Published in final edited form as:

Hum Genet. 2015 March ; 134(3): 343–358. doi:10.1007/s00439-014-1524-3.

A Systematic Heritability Analysis of the Human Whole Blood Transcriptome

Tianxiao Huan^{1,2}, Chunyu Liu^{1,2}, Roby Joehanes^{1,2,3}, Xiaoling Zhang¹, Brian H. Chen^{1,2}, Andrew D. Johnson¹, Chen Yao^{1,2}, Paul Courchesne^{1,2}, Christopher J. O'Donnell¹, Peter J. Munson³, and Daniel Levy^{1,2}

¹The National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, MA

²Population Sciences Branch, Division of Intramural Research, National Heart, Lung and Blood Institute, Bethesda, MD

³Mathematical and Statistical Computing Laboratory, Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, MD

Abstract

Genome-wide expression quantitative trait locus (eQTL) mapping may reveal common genetic variants regulating gene expression. In addition to mapping eQTLs, we systematically evaluated the heritability of the whole blood transcriptome in 5626 participants from the Framingham Heart Study. Of all gene expression measurements, about 40% exhibit evidence of being heritable

($h^2_{geneExp} > 0$, ($p < 0.05$)), the average heritability was estimated to be 0.13, and 10% display

$h^2_{geneExp} > 0.2$. In order to identify the role of eQTLs in promoting phenotype differences and disease susceptibility, we investigated the proportion of *cis/trans* eQTLs in different heritability categories and discovered that genes with higher heritability are more likely to have *cis* eQTLs that explain large proportions of variance in the expression of the corresponding genes. Single *cis* eQTLs explain 0.33–0.53 of variance in transcripts on average, whereas single *trans* eQTLs only explain 0.02–0.07. The top *cis* eQTLs tend to explain more variance in the corresponding gene

when its $h^2_{geneExp}$ is greater. Taking body mass index (BMI) as a case study, we cross-linked *cis/trans* eQTLs with both GWAS SNPs and differentially expressed genes for BMI. We discovered that BMI GWAS SNPs in 16p11.2 (e.g., rs7359397) are associated with several BMI differentially expressed genes in a *cis* manner (e.g. *SULT1A1*, *SPNS1*, and *TUFM*). These BMI signature genes explain a much larger proportion of variance in BMI than do the GWAS SNPs. Our results shed light the impact of eQTLs on the heritability of the human whole blood transcriptome and its relations to phenotype differences.

Correspondence should be addressed to: Daniel Levy, MD, Framingham Heart Study, Population Sciences Branch, National Heart, Lung, and Blood Institute, 73 Mt. Wayte Avenue, Suite 2, Framingham, MA 01702, Levyd@nih.gov, Phone: 508-935-3458, Fax: 508-872-2678.

Disclosure declaration

The authors declare no conflict of interest.

Keywords

heritability; eQTL; transcriptome; gene expression

Introduction

Genome-wide association studies (GWAS) have identified numerous loci associated with complex phenotypes and disease traits, such as coronary heart disease (Lotta and Peyvandi 2011; Schunkert et al. 2011), blood pressure (Ehret et al. 2011; Levy et al. 2009; Newton-Cheh et al. 2009), lipids (Teslovich et al. 2010), body mass index (Guo et al. 2013; Monda et al. 2013; Speliotes et al. 2010), and type 2 diabetes (Morris et al. 2012; Voight et al. 2010). Although GWAS have provided valuable insights into the genetic basis of complex traits, they explain only a small proportion of the heritability of most traits (Tenesa and Haley 2013). For the majority of loci from GWAS, the mechanistic underpinnings for association with the corresponding traits are unknown. Gene expression may mediate genotype-phenotype association, and as such may be considered an intermediate molecular phenotype (Cheung and Spielman 2009). Dissecting the genetic determinants of gene expression can provide insight into mechanisms of disease and the functional consequences of genetic variation. Genome-wide expression quantitative trait locus (eQTL) mapping has been widely used to identify common genetic variants regulating gene expression (Myers et al. 2007; Schadt et al. 2008; Stranger et al. 2005; Stranger et al. 2007; Zeller et al. 2010). Although numerous studies (Myers et al. 2007; Schadt et al. 2008; Stranger et al. 2005; Stranger et al. 2007; Westra et al. 2013; Zeller et al. 2010) have revealed tens of thousands of *cis* and *trans* eQTLs, many *cis* and *trans* eQTLs remain unrecognized due to sample size limitations and tissue specificity (Grundberg et al. 2012; Price et al. 2011). Further investigations are needed to answer additional questions, such as, are these *cis* or *trans* components related to the heritability of gene expression and what is the proportion of variance in gene expression that can be explained by single *cis* or *trans* eQTLs?

In this study, we systematically investigated the heritability of the human whole blood transcriptome using Framingham Heart Study (FHS) pedigrees. Our goals were three fold. First, we explored the distribution of heritability of genome-wide gene expression levels using a large sample with well-defined extended pedigrees. Second, we investigated how *cis* and *trans* eQTLs are related to transcript heritability levels. Third, we sought to understand how the genetic basis of gene expression relates to phenotype differences and disease susceptibility via *cis* and *trans* components. To accomplish these aims, we estimated the overall heritability of approximately 18,000 genes in 5626 participants from the FHS, and assessed the proportion of variance in gene expression that is attributable to *cis* and *trans* eQTLs. By cross-linking the eQTLs with GWAS results of metabolic traits including body mass index (BMI), blood pressure (BP), and lipid traits in the NHGRI GWAS Catalog (Hindorff et al. 2009), we discovered trait-associated single nucleotide polymorphisms (SNPs) that explain relatively large proportions of the genetic variance of multiple gene transcripts, despite the fact that these SNPs only explain a small proportion of phenotypic variance for the same metabolic traits. Last, taking BMI as an example, by cross-linking the eQTLs with trait-associated SNPs (GWAS SNPs (Hindorff et al. 2009)), we discovered that

some GWAS SNPs are *cis* eQTLs of certain gene transcripts and can explain large proportions of variance in expression of these transcripts. In addition, some of the corresponding *cis* eQTL gene transcripts show differential expression for BMI. Although the GWAS SNPs only explain a small proportion of phenotypic variance in BMI, these differentially expressed *cis* eQTL related gene transcripts explain a larger proportion of variance in BMI.

Materials and Methods

Study population description

In 1971, the offspring and offspring spouses (offspring cohort, N=5124) of the original FHS cohort participants were recruited and have been examined approximately every four years (except for the interval between examinations 1 and 2 with an intervening 8 years) (Feinleib et al. 1975). From 2002 to 2005, the adult children (third generation cohort, N=4095) of the offspring cohort participants were recruited and examined (Splansky et al. 2007). A total of 5626 participants from the offspring (N=2446) and third generation (N=3180) cohorts were included in this study. Whole blood samples were collected at the eighth examination of the offspring cohort and the second examination of the third generation cohort. All participants provided written consent for genetic research.

Gene expression profiling

Fasting peripheral whole blood samples (2.5ml) in PAXgene™ tubes (PreAnalytiX, Hombrechtikon, Switzerland) were collected and the Affymetrix Human Exon Array ST 1.0 (Affymetrix, Inc., Santa Clara, CA) was utilized to measure mRNA expression levels genome wide (N~18,000 genes). Details of the design, sampling, RNA isolation, and mRNA measurement have been described previously (Huan et al. 2013; Joeheanes et al. 2013). All data used herein are available online in dbGaP (<http://www.ncbi.nlm.nih.gov/gap>; accession number phs000007).

Genotyping and quality control

DNA isolation, and genotyping with the Affymetrix 500K mapping array and the Affymetrix 50K gene-focused MIP array have been described previously (Levy et al. 2009). A total of 503,551 SNPs with successful call rate >0.95 and Hardy-Weinberg Equilibrium (HWE) $P > 10^{-6}$ were retained. Imputation of ~36.3 million SNPs in 1000 Genomes Phase 1 SNP data was conducted using MACH (Li et al. 2010). In this study, we used the 1000-genome resource imputed SNPs with minor allele frequency (MAF) >0.01 and imputation ratio >0.3, yielding approximately 8 million SNPs for eQTL analysis.

eQTL identification

The eQTL list was generated using the 5257 individuals that had genome-wide genotypes and gene expression profiling. A *cis* eQTL was defined as an eQTL within a 1 megabase (Mb) flanking the gene. A *trans* eQTL was defined as an eQTLs in a different chromosome from the gene. The remaining eQTLs residing in the same chromosome but extending 1Mb distance were defined as *long-range cis* eQTLs. The imputed SNPs were coded in an additive genetic model of allele dosage that was continuous and ranged from 0 to 2. Linear

mixed effects (LME) regression models were used to determine the association between gene expression (adjusting for age, sex, 10 technical covariates, cell types, the first principle component [PC1], and familial relatedness) and the imputed SNP dosage from the 1000-Genomes resource (Abecasis et al. 2012). Supplementary Table S1 shows the 10 technical covariates and PC1. Cell types included white blood cells [WBC], red blood cells [RBC], platelets, neutrophils, lymphocytes, monocytes, eosinophils, and basophils, which were imputed by partial least squares (PLS) regression (Joehanes R, PhD, unpublished data). Genomic coordinates were based on NCBI human reference genome build 37/hg19. The Benjamini-Hochberg method (BH) (Benjamini and Hochberg 1995) was used to calculate false discovery rates (FDR) of *cis*, *trans*, and *long-range cis* eQTLs separately.

We evaluated if the effect sizes of *cis* and *trans* eQTLs in independent individual sets showed concordance of allele effect. We split the overall samples into the discovery and replication sets (1:1 ratio) that included independent families between the two sets. Using the same methods described above, we identified eQTLs in the discovery and replication sets respectively and compared the effect sizes of eQTLs from the discovery and replication set.

To determine if differences in cell type distributions in different samples affected the identification of *cis* and *trans* eQTLs, we applied regression models without including cell types as covariates to repeat the eQTL analysis for the *cis* and *trans* eQTLs at FDR < 1e-4. We then compared the t-test statistics of *cis* and *trans* eQTLs with and without adjusting for cell types.

By design, the Affymetrix Human Exon 1.0 ST array platform contains SNPs in some probes for targeting mRNA sequence; this SNP-on-probe phenomenon can affect the binding affinity of microarray probe sequences and may lead to false-positive and false-negative results (Ramasamy et al. 2013). To address this problem, we further limited eQTLs to the 6059 transcripts without any SNPs-on-probes. We reported the eQTL results for all 18,000 transcripts in the main body of the manuscript and the results for the 6059 transcripts without SNP-on-probes in the Supplementary Files.

Estimation of the additive heritability for a gene expression trait

The narrow-sense heritability estimate of a gene expression trait (denoted as $h_{geneExp}^2$) was the proportion of the additive polygenic genetic variance of the total phenotypic variance of a gene expression trait: $h_{geneExp}^2 = \sigma_A^2 / \sigma_{geneExp}^2$, here σ_A^2 denotes the additive polygenic genetic variance and $\sigma_{geneExp}^2$ denotes the total phenotypic variance of a gene expression trait. In order to obtain both variance components and *p*-values for heritability estimates, the $h_{geneExp}^2$ estimate was obtained using variance-component methodology implemented in SOLAR (Almasy and Blangero 1998) and in the *lmekin()* function of Kinship Package (Abecasis et al. 2001) (<http://cran.r-project.org/web/packages/kinship/>).

Heritability estimation for all gene expression traits was performed using the combined sample (N=5626). To determine whether heritability estimates were consistent in individual

cohorts, we conducted analysis in the offspring cohort (N=2446) and the third generation cohort (N=3180) separately. The average heritability estimate was calculated as

$$\left(\sum_1^n h_{geneExp}^2 \right) / n, \text{ where } n = \text{number of gene expression traits.}$$

We further investigated if the heritability estimates in relation to cohorts, sample size and blood cell type portions (see Supplementary Files).

The proportion of variance in gene expression attributable to peak eQTLs

The proportion of variance in a single gene expression trait that is attributable to a single eQTL (or eSNP) was denoted as h_{eQTL}^2 and was calculated as follows:

$$h_{eQTL}^2 = \max \left(0, \frac{\sigma_{QTL.null}^2 + \sigma_{err.null}^2 - \sigma_{QTL.full}^2 - \sigma_{err.full}^2}{\sigma_{geneExp}^2} \right) \quad (\text{Equation 1})$$

where $\sigma_{geneExp}^2$ was the total phenotypic variance of a gene expression trait; $\sigma_{QTL.full}^2$ and $\sigma_{err.full}^2$ were the polygenic variance and the residual error, respectively, when modeling with the tested eQTL; $\sigma_{QTL.null}^2$ and $\sigma_{err.null}^2$ were the polygenic variance and the residual error when modeling without the tested eQTL. The relative proportion of variance explained by a peak eQTL with respect to the heritability of a gene expression trait was calculated by $rh_{eQTL}^2 = h_{eQTL}^2 / h_{geneExp}^2$. The *lmekin()* function in the Kinship package was used to estimate h_{eQTL}^2 .

The proportion of variance in a clinical trait attributable to a signature gene

The proportion of the variance in a clinical trait attributable to a signature gene was calculated as follows:

$$h_{geneExp_trait}^2 = \max \left(0, \frac{\sigma_{geneExp.null}^2 + \sigma_{err.full}^2 - \sigma_{geneExp.full}^2 - \sigma_{err.full}^2}{\sigma_{clinical\ trait}^2} \right) \quad (\text{Equation 2})$$

where $\sigma_{clinical\ trait}^2$ was the total phenotypic variance of the clinical trait, $\sigma_{geneExp.full}^2$ and $\sigma_{err.full}^2$ were the polygenic variance and the residual error when modeling with the tested signature gene, $\sigma_{geneExp.null}^2$ and $\sigma_{err.null}^2$ were the polygenic variance and the residual error when modeling without the tested signature gene. The *lmekin()* function in the Kinship package was used to estimate $h_{geneExp_trait}^2$.

GWAS resources for metabolic traits

We cross-linked the eQTLs (or eSNPs) identified in this study with top SNPs reported from the GWAS for several metabolic traits, including hypertension/blood pressure (BP), obesity/body mass index (BMI) and lipid traits (High-density lipoprotein [HDL], Low-density lipoprotein [LDL], Total cholesterol [TC] and Triglycerides [TG]), in the NHGRI GWAS

Catalog (accessed Dec 2013) (Hindorff et al. 2009). We used a p value $5e-8$ to define a trait-specific significant result.

Identification of differentially expressed genes for BMI

To determine the proportion of variance that can be explained by differentially expressed genes or by eQTLs of these differentially expressed genes for a clinical trait, we used BMI as an example for proof-of-concept. We identified differentially expressed genes for BMI which was measured at the same examinations of the offspring and the third generation cohorts at which gene expression was assessed. Residuals for genes (after adjusting for technical covariates) were used to identify differentially expressed genes associated with BMI after accounting for age, sex, and family structure in LME models implemented in *lme4*() (Abecasis et al. 2001). We applied Benjamini–Hochberg’s FDR<0.05 to denote significant signals and calculated FDR for selected BMI signature genes (Benjamini and Hochberg 1995). The proportion of the variance in BMI attributable to multiple differentially expressed signature genes was calculated using Equation (2).

Results

Study participants

Table 1 summarizes the characteristics of the 5626 FHS participants used in this study. At the baseline examination, the offspring cohort is, on average ~20 years older than the third generation cohort (66 versus 46 years). Supplementary Table S2 displays the family sizes in the 5626 individuals. Among the participants, 399 are unrelated individuals, and 5227 are from 704 families (family size ≥ 2). Among the 5227 individuals, 2352 are from 570 small families (family size 2–9), 1283 are from 101 medium size families (family size 10–19), 940 are from 29 relatively large families (family size 20–99), and 651 are from large families (size ≥ 100).

Heritability estimation of genome-wide gene expression measurements

The narrow sense heritability of ~18,000 transcripts was estimated using the FHS pedigrees. For simplicity, we refer to “narrow sense heritability” as “heritability” in the rest of this manuscript. Figure 1 displays the distribution of heritability estimates for 18,000 genes based on all 5626 participants (Supplementary Figure S1 displays the results for 6059 genes without SNPs-on-probes [see Methods]). The average heritability of global gene expression was estimated to be 0.072 (the median heritability estimate was 0.037). Among all genes, 7161 (~40% of the genome) were estimated to have $h^2_{geneExp} > 0$ ($p < 0.05$), 1,730 (~10%) with $h^2_{geneExp} > 0.2$ ($p < 1.8e-6$), 286 (~2%) with $h^2_{geneExp} > 0.4$ ($p < 6.3e-41$), and 50 (~0.3%) with $h^2_{geneExp} > 0.6$ ($p < 3.7e-105$).

We further investigated if the heritability estimates differed in relation to cohort (offspring vs. third generation), sample size, and blood cell type portions. We discovered that the overall heritability distribution is similar in the FHS offspring and third generation cohorts. In addition, differences in cell type proportions are not likely to significantly affect the heritability of gene expression levels. We also discovered that larger sample size and more

comprehensive family structure provide more accurate heritability estimates (see Supplementary Files, Supplementary Figure S2–S3, and Supplementary Table S3).

Effects of *cis* and *trans* eQTLs' on gene expression

We performed a genome-wide association analysis of global gene expression levels and identified *cis*, *trans* and *long-range cis* eQTLs. The proportion of gene transcripts with significant *cis*, *trans* and *long-range cis* eQTLs is much smaller at $h^2_{\text{geneExp}} < 0.2$ than at $h^2_{\text{geneExp}} > 0.2$. At $\text{FDR} < 1\text{e-}4$, 43% of genes have *cis* eQTLs at $h^2_{\text{geneExp}} < 0.2$, compared to 93% of genes at $h^2_{\text{geneExp}} > 0.2$. Similarly, 3% of genes have *trans* eQTLs at $h^2_{\text{geneExp}} < 0.2$, compared to 21% of genes at $h^2_{\text{geneExp}} > 0.2$; and 1% of genes have *long-range cis* eQTLs at $h^2_{\text{geneExp}} < 0.2$, compared to 10% of genes at $h^2_{\text{geneExp}} > 0.2$.

We further computed the proportion of variance explained by a single *cis* or *trans* eQTL (h^2_{eQTL}) in the study samples for genes with expression heritability levels $h^2_{\text{geneExp}} > 0.2$ (1,730 genes or 10% of the total gene expression traits). By splitting the overall samples into independent discovery and replication sets (see Methods), we discovered that *cis* and *trans* eQTLs show consistent allelic directional associations both in the discovery and replication sets (Supplementary Figure S4). We evaluated if cell types affected the identification of eQTLs. The t-test statistics remained very similar with or without adjusting for cell types for both *cis* (Pearson $r = 0.99$) and *trans* (Pearson $r = 0.98$) eQTLs, suggesting that the differences of cell type distributions in different samples did not have large impact on the identification of either *cis* or *trans* eQTLs (Supplementary Figure S5). A similar finding was also reported by Westra et al. in a *trans* eQTL study (Westra et al. 2013).

We further applied a series of FDR thresholds ($\text{FDR} < 1\text{e-}4$, $1\text{e-}6$, $1\text{e-}8$ and $1\text{e-}10$) to select significant eQTLs. Figure 2 summarizes the proportions of gene transcripts with significant *cis*, *trans* or *long-range cis* eQTLs at different $h^2_{\text{geneExp}} > 0$ levels = (0.2, 0.3), (0.3, 0.4), (0.4, 0.5) and (0.5, 0.6). We observed that gene expression traits with higher heritability estimates tend to have higher proportions of significant *cis* eQTLs. For example, at $\text{FDR} < 1\text{e-}4$ (corresponding $p = 3.2\text{e-}6$), 49 of the 50 genes at $h^2_{\text{geneExp}} > 0$ of (0.6, 1) are associated with *cis* eQTLs. For the other four heritability categories, i.e. $h^2_{\text{geneExp}} = (0.2, 0.3)$, (0.3, 0.4), (0.4, 0.5), and (0.5, 0.6), the proportions of genes with *cis* eQTLs are 92%, 91%, 96%, and 100%, respectively (Figure 2A). Compared to the proportions of significant *cis* eQTLs, in general, the proportions of significant *trans* eQTL are much smaller across all heritability categories. The transcripts of *trans* eQTLs are associated with heritability values in an opposite manner. For FDR thresholds, gene expression traits with higher heritability estimates tend to have lower proportions of significant *trans* eQTLs (Figure 2B). For example, at FDR threshold $< 1\text{e-}4$ (corresponding $p = 1.6\text{e-}10$), the proportions of *trans* eQTLs genes are 22%, 24%, 16%, 16%, and 2%, respectively, for the 5 heritability categories. The long-range *cis* eQTLs exhibit characteristics similar to those *cis* eQTLs rather than *trans* eQTLs (Figure 2C). For all FDR thresholds, gene expression traits with higher heritability estimates tend to have higher proportions of significant *long-range cis* eQTLs.

Through investigating the peak *cis*, *trans*, or *long-range cis* eQTL for every gene, we discovered that the peak *cis* eQTLs explain a large proportion of variance in their respective gene expression level (Supplementary Table S4 and Figure 3A). On average the proportion of variance in the transcript with $h^2_{\text{geneExp}} > 0.2$ explained by a peak *cis* eQTL is about one third. Moreover, when the heritability of gene expression increases from h^2_{geneExp} of (0.2, 0.3) to (0.6, 1), the average $h^2_{\text{cis-eQTL}}$ increases from 0.33 to 0.53 (Figure 3A). In contrast, the peak *trans* eQTLs only explain a small proportion of variance in their respective gene expression levels. The average proportion of variance in the transcript explained by a single peak *trans* eQTL ranged from 0.02 to 0.07 in all of the five h^2_{geneExp} categories (Supplementary Table S5 and Figure 3B). For *long-range cis* eQTLs, when the heritability of gene expression increases from h^2_{geneExp} of (0.2, 0.3) to (0.6, 1), the average $h^2_{\text{long-cis-eQTL}}$ increases from 0.07 to 0.18 (Supplementary Table S6 and Figure 3C).

In order to avoid the SNP-on-probe effect that may lead false-positive and false-negative results, we repeated analyses limited to 6,059 gene transcripts without any SNPs-on-probes (see Methods). Supplementary Figures S6–S7 present results for these 6,059 genes. The results are similar for these 6,059 genes vs. all 18,000 genes (Figures 2–3).

Metabolic traits GWAS SNPs that explained a large proportion of variability in eQTL genes

We cross-linked the GWAS results for several metabolic traits with the *cis* and *trans* eQTLs for genes with $h^2_{\text{geneExp}} > 0.2$ (the threshold for choosing eQTLs is $\text{FDR} < 1\text{e-}3$). Table 2 summarizes a list of *cis* eQTLs that explain relatively large proportions of heritability in the corresponding gene expression levels ($rh^2_{\text{eQTL}} \geq 0.1$). These *cis* eQTLs are among the peak signals for the corresponding genes (eQTL $p < 1\text{e-}40$). Although these significant GWAS SNPs (that also are *cis* eQTLs) explain only very small proportions ($rh^2_{\text{trait-SNP}} \leq 0.001$) of genetic variance in the metabolic traits with which they are associated, these SNPs (*cis* eQTLs) explain relatively large proportions of the genetic variance in gene expression levels, and more strikingly, several *cis* eQTLs display large rh^2_{eQTL} estimates (> 0.5).

Although some *cis* eQTLs display higher values of rh^2_{eQTL} for a transcript, the transcript do not necessary correspond to the closest gene to the *cis* eQTL. For example, rs486416 (associated with total cholesterol and triglycerides in GWAS (Teslovich et al. 2010)) is located in the intron of *EHMT2*, but this SNP is associated with expression of *SLC44A4* in a *cis* manner with $rh^2_{\text{eQTL}} = 0.87$ for the *SLC44A4* transcript. Additionally, we discovered that some *cis* eQTLs are associated with multiple gene expression traits and display relatively high rh^2_{eQTL} . For example, rs12936231 (associated with HDL in GWAS (Teslovich et al. 2010)) is a *cis* eSNP of both *GSDMB* ($rh^2_{\text{eQTL}} = 0.92$, 28KB away from rs12936231) and *ORMDL3* ($rh^2_{\text{eQTL}} = 0.34$, 48KB away from rs12936231) (Figure 4A).

Table 3 summarizes *trans* eQTLs for genes with $h_{geneExp}^2 > 0.2$. In contrast to *cis* eQTLs, most *trans* eQTLs display smaller rh_{eQTL}^2 (< 0.1) with only one exception *trans* eQTL rs289754 for *BTN3A1* ($rh_{eQTL}^2 = 0.19$). Although having smaller average rh_{eQTL}^2 , some *trans* eQTLs are also associated with multiple gene transcripts. Table 3 displays two such *trans* eSNPs: one example is rs3184504, which is nonsynonymous in *SH2B3* and is associated in GWAS with SBP, DBP (Levy et al. 2009) and LDL (Teslovich et al. 2010). rs3184504 is a *trans* eQTL associated with five gene transcripts including *FCGR1A* ($rh_{eQTL}^2 = 0.04$), *ARHGEF40* ($rh_{eQTL}^2 = 0.03$), *MYADM* ($rh_{eQTL}^2 = 0.03$), *IDO1* ($rh_{eQTL}^2 = 0.02$) and *IFIT3* ($rh_{eQTL}^2 = 0.02$). Another SNP, rs289754, is a *trans* eSNP associated with three genes located in the MHC Cluster I – *BTN3A1* ($rh_{eQTL}^2 = 0.19$), *HCP5* ($rh_{eQTL}^2 = 0.10$) and *HLA-F* ($rh_{eQTL}^2 = 0.04$) (Figure 4B).

Identification of “master” eQTLs associated with BMI signatures genes

Based on our finding that some eQTLs explain large proportions of variation in gene expression levels (h_{eQTL}^2) -- although they only explain small proportions of variance in metabolic traits -- we investigated the relations between the trait-specific differential expression signature genes and the genes with the eQTLs that also are GWAS SNPs for the same trait. Using BMI as a proof-of-concept example, at first, we identified BMI differentially expressed genes at $FDR < 0.05$ (i.e. BMI signature genes, Supplementary Table S7). Second, we cross-linked these signature genes with significant BMI GWAS SNPs as well as with eQTLs. By doing so we discovered that 9 BMI signature genes are associated with 6 BMI GWAS SNPs (Speliotes et al. 2010) in a *cis* manner (Table 4). Among a total of 18,000 genes measured by the Affymetrix Exon Array, 18 genes had *cis* eQTLs with $p < 5 \times 10^{-8}$ in BMI GWAS. Comparison of the two ratios $9/5409$ and $18/18000$ by hypergeometric test yielded $p = 0.02$, indicating the BMI transcriptomic signatures were enriched for BMI GWAS SNP-associated transcripts at $p = 0.02$ by hypergeometric test. rs7359397 is a *cis* eQTL for three BMI signature genes – *SULT1A1* (FDR adjusted $p = 0.05$), *SPNS1* (FDR adjusted $p = 5.6 \times 10^{-4}$), and *TUFM* (FDR adjusted $p = 0.03$). Other two genes located near rs7359397, *CD19* and *SH2B1*, also are BMI signature genes at FDR adjusted $p = 0.04$ and 1×10^{-4} , however, rs7359397 was not recognized as a *cis* eQTL for these two genes (Figure 5A–B). We further discovered groups of differentially expressed genes that are associated with the same GWAS SNPs and jointly explain greater proportions of the variance in BMI than do the SNPs associated with the trait in GWAS. For instance, rs7359397 only explained 0.0086% of phenotypic variance in BMI (calculated using FHS data alone), but the three BMI signature genes (*SULT1A1*, *SPNS1*, and *TUFM*) that are associated with this SNP in *cis* manner, jointly explained 0.5% of the interindividual variability in BMI (Figure 5C).

Discussion

We systematically investigated the heritability of global gene expression using whole blood samples from 5626 FHS participants. To our knowledge, this is the largest heritability study of global gene expression. We observed that approximately 40% of the genome (7161 gene)

display $h^2_{geneExp} > 0$ and ~10% displayed $h^2_{geneExp} > 0.2$. We confirmed that transcripts with higher heritability estimates are more likely to be associated with *cis* eQTLs, but this is not the case for *trans* eQTL (Dixon et al. 2007; Goring et al. 2007; Grundberg et al. 2012). Moreover, we found that single *cis* eQTLs explain on average 0.33–0.53 of variance in expression levels for genes with $h^2_{geneExp} > 0.2$. Single *trans* eQTLs, however, only explain 0.02–0.07 of variance in transcripts with $h^2_{geneExp} > 0.2$. Interestingly we discovered that the top *cis* eQTLs tend to explain more variance in the respective transcript when $h^2_{geneExp}$ increased, but there is no obvious trend for the top *trans* eQTLs. Taking BMI as a case study, by cross-linking the *cis/trans* eQTLs with BMI GWAS SNPs (Hindorff et al. 2009; Speliotes et al. 2010) and BMI signature genes, we observed that some BMI GWAS SNPs are actually the *cis* eQTLs for multiple BMI signature genes. Although these SNPs (also *cis* eQTLs) explain only very small proportions of variance in BMI, they explain relatively larger proportions of variance in gene expression levels. These BMI signature genes -- driven by the BMI GWAS SNPs -- jointly explain a larger proportion of the variability in BMI than do the GWAS SNPs.

We found that 40% of gene transcript levels are heritable and the average heritability of global gene expression genome-wide is estimated to be 0.07 (for all 18,000 genes) and 0.13 (for 7,161 genes with $h^2_{geneExp} > 0$). Several studies reported that 40–70% of gene transcripts are heritable with $h^2_{geneExp} > 0$ (Emilsson et al. 2008; Goring et al. 2007; Price et al. 2011; Stranger et al. 2007). Dixon *et al* reported that the average heritability for global gene expression is 0.2 (Dixon et al. 2007), and the average heritability for transcripts with $h^2_{geneExp} > 0$ range between 0.15 and 0.30 (Emilsson et al. 2008; Price et al. 2011; Stranger et al. 2007). These heritability estimates are considerable larger than our findings. By sub-sampling FHS families, we discovered that small sample sizes and relatively simple family structures display a wider range of average heritability estimates (Supplementary Figures S3) than large and complex family structures, which may partly explain the different heritability estimates reported in several previous studies and in our sample. Heritability is the proportion of phenotypic variation attributable to genetic variation; varying levels of pedigree structures are needed to obtain accurate variance parameters from genetic and environmental components in order to get more accurate estimation of heritability. The current study sample (N=5626) that consists of extended families of two-generations of FHS participants is five times larger than the sample sizes of prior studies (Emilsson et al. 2008; Goring et al. 2007; Price et al. 2011; Stranger et al. 2007). In addition, the heritability estimates are further compared between the offspring cohort (N=2446) and the third generation cohort (N=3180), and between samples with higher vs. lower WBC, neutrophil, or lymphocyte proportions. With a large sample size, complex family structure, and careful study design, we believe that the heritability estimates of global gene expression in the current study are more stable and reliable than previous findings.

In this study, we confirmed that *cis* components explain a larger proportion of variance in expression levels of genes than *trans* components (Grundberg et al. 2012; Price et al. 2011). In addition, we discovered that the proportion of genetic variance in a transcript explained

by the peak eQTL is different for transcripts with different heritability estimates. On average, approximately 53% ($rh_{eQTL}^2=0.53$) of genetic variance for transcripts with $h_{geneExp}^2>0.6$ is explained by a single *cis* eQTL, and the proportion decrease to 33% ($rh_{eQTL}^2=0.33$) for transcripts with $h_{geneExp}^2=0.2$ to 0.3. This finding suggests that, in general, transcripts with higher heritability estimates are affected by fewer nearby genetic loci compared to transcripts with lower heritability. In contrast, the most significant *trans* eQTLs explain, on average, much smaller proportions of genetic variance in transcripts and do not show obvious trends in relation to heritability estimates, (h_{eSNP}^2 ranges from 0.02 to 0.07 for transcripts on average in all $h_{geneExp}^2$ levels). Grundberg *et al.* estimated that at least 0.12 of total heritability of transcripts is “missed” when only considering common *cis* eQTL effects, and they proposed that the “missing heritability” might be due to low-frequency and rare variants (Grundberg et al. 2012). However, our findings suggest that the “missing heritability” may also be attributable in part to *trans* eQTLs with modest genetic effects that are not detected easily in smaller sized eQTL studies.

By cross-linking GWAS results for metabolic traits with transcripts and with *cis* and *trans* eQTLs, we identified a number of GWAS SNPs that explain large proportions of the variation in several genes (Table 2–3). For example, rs12936231, a *cis* eQTL for *GSDMB*, explains up to 92% of genetic variance in *GSDMB*. We also discovered that some GWAS SNPs, that we consider to represent master eQTLs, are associated with multiple transcripts. For example, rs3184504 is a *trans* eQTL for 5 genes, *FCGR1A*, *ARHGEF40*, *MYADM*, *IDO1*, and *IFIT3*. The mechanism underlying the associations of multiple transcripts with a single eQTL (a master eSNP) has not been elucidated to date. One possible mechanism is that transcripts sharing common *cis/trans* eQTLs may be tightly co-expressed and involved in the same biological process or pathway (Rotival et al. 2011). Another possible explanation is that the *cis/trans* eQTL and its proxy SNPs may have an effect on a regulatory component, such as transcription factors (TFs) or microRNAs (miRNAs) that regulate multiple transcripts, and this eQTL is thereby associated with multiple transcripts targeted by the TFs or the miRNAs. Figure 4B illustrates a “master” effect by a *trans* eQTL that is located in or near a TF or a key regulator. Different alleles of the *trans* eQTL may affect the transcription or translation of the TF or key regulator, giving rise to different isoforms of the TF or key regulation. Multiple genes whose expression levels are regulated by those TFs or key regulators may show association with this SNP. For example, rs289754 is located in the intron of *NLRC5* (16q13) and is a *trans* eQTL for three genes: *BTN3A1* (6p22), *HCP5* (6p21.3) and *HLA-F* (6p21.3). Rs289754 explained 0.19, 0.10, and 0.04 of the heritability of *BTN3A1*, *HCP5* and *HLA-F*, respectively (Table 4). *NLRC5* has been newly recognized as a key regulator of MHC class I related immune functions (Kobayashi and van den Elsen 2012). However, *NLRC5* itself does not have a DNA-binding domain and must interact with multiple proteins to activate the MHC class I dependent immune response (Kobayashi and van den Elsen 2012). The third possible explanation is that the common *cis* eQTLs in the nearby region display allelic effects on the nucleosome, chromatin structure, or TF binding affinity, and in this way govern the transcriptional activity of nearby genes. Figure 4A shows a mechanism by which a “master” *cis* eQTL affects transcription of

multiple nearby genes. Supportive evidence exists for rs12936231 (17q21), which displays CTCF promoter binding in eight cell lines (Verlaan et al. 2009). CTCF is a well-characterized protein involved in chromatin looping and chromosome repositioning. The same study (Verlaan et al. 2009) reported that rs12936231 is a significant *cis* eQTL for *GSDMB* (distance ~28KB from transcription start site [TSS]) and *ORMDL3* (distance ~48 KB from TSS) gene expression. We were able to confirm that rs12936231 is significantly associated with transcripts from the nearby genes *GSDMB* (p value=0) and *ORMDL3* (p value = 1.34e-92). In addition, we discovered that rs12936231 explain 0.94 of heritability of *GSDMB* and 0.34 of the heritability of *ORMDL3* expression levels (Table 3), thereby providing further evidence that rs12936231 regulates both genes. However, a functional study is needed to further elucidate the role of the eQTL in relation to gene regulation. We further linked the trait-GWAS SNPs and trait-signature genes via their eQTLs. Taking BMI as proof-of-concept case study, we found several BMI signature genes that are associated with BMI GWAS SNPs in a *cis* manner. For example, BMI GWAS SNP rs7359397 ($p=1.64e-14$ in BMI GWAS (Speliotes et al. 2010)) is a *cis* eQTL for 3 BMI signature genes (*SULT1A1*, *SPNS1*, and *TUFM*, Figure 4 and Table 4) among its six *cis* eQTL genes. These three BMI signature genes extend the proportion of phenotypic variance of BMI explained by rs7359397.

Conclusions

We systematically evaluated the heritability of global whole blood gene expression using a large sample size of FHS extended pedigrees. We discovered that with increasing heritability of transcripts, there is an increase in the contribution of *cis*, but not *trans* regulation. Future studies are needed to investigate if this finding holds true across different tissues. By cross-linking the eQTLs with trait-associated SNPs (i.e. BMI GWAS SNPs (Hindorff et al. 2009)), we discovered that some GWAS SNPs are *cis* eQTLs of certain gene transcripts and can explain large proportions of variance in expression of these transcripts. In addition, some of the corresponding *cis* eQTL gene transcripts show differential expression in relation to the same trait (i.e. BMI). Although the GWAS SNPs only explain a small proportion of phenotypic variance in BMI, the differentially expressed *cis* eQTL related gene transcripts explained a larger proportion of variance in BMI. Overall, our results shed light on the impact of eQTLs on the heritability of the human blood transcriptome, and the role of eQTLs in promoting phenotype differences and disease susceptibility.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

D. L., C. L. and T.H. designed, directed, and supervised the project. D. L. was responsible for funding of the project. T. H., C. L. and D. L. drafted the manuscript. P. C. organized the experiment material and data exchange. All authors participated in revising and editing the manuscripts. All authors have read and approved the final version of the manuscript.

References

- Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012; 491:56–65. [PubMed: 23128226]
- Abecasis GR, Cardon LR, Cookson WO, Sham PC, Cherny SS. Association analysis in a variance components framework. *Genet Epidemiol*. 2001; 21(Suppl 1):S341–S346. [PubMed: 11793695]
- Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet*. 1998; 62:1198–1211. [PubMed: 9545414]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995:289–300.
- Cheung VG, Spielman RS. Genetics of human gene expression: mapping DNA variants that influence gene expression. *Nature Reviews Genetics*. 2009; 10:595–604.
- Dixon AL, Liang L, Moffatt MF, Chen W, Heath S, Wong KC, Taylor J, Burnett E, Gut I, Farrall M, Lathrop GM, Abecasis GR, Cookson WO. A genome-wide association study of global gene expression. *Nat Genet*. 2007; 39:1202–1207. [PubMed: 17873877]
- Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, Smith AV, Tobin MD, Verwoert GC, Hwang SJ, Pihur V, Vollenweider P, O'Reilly PF, Amin N, Bragg-Gresham JL, Teumer A, Glazer NL, Launer L, Zhao JH, Aulchenko Y, Heath S, Sober S, Parsa A, Luan J, Arora P, Dehghan A, Zhang F, Lucas G, Hicks AA, Jackson AU, Peden JF, Tanaka T, Wild SH, Rudan I, Igl W, Milaneschi Y, Parker AN, Fava C, Chambers JC, Fox ER, Kumari M, Go MJ, van der Harst P, Kao WH, Sjogren M, Vinay DG, Alexander M, Tabara Y, Shaw-Hawkins S, Whincup PH, Liu Y, Shi G, Kuusisto J, Tayo B, Seielstad M, Sim X, Nguyen KD, Lehtimäki T, Matullo G, Wu Y, Gaunt TR, Onland-Moret NC, Cooper MN, Platou CG, Org E, Hardy R, Dahgam S, Palmen J, Vitart V, Braund PS, Kuznetsova T, Uitterwaal CS, Adeyemo A, Palmas W, Campbell H, Ludwig B, Tomaszewski M, Tzoulaki I, Palmer ND, Aspelund T, Garcia M, Chang YP, O'Connell JR, Steinle NI, Grobbee DE, Arking DE, Kardina SL, Morrison AC, Hernandez D, Najjar S, McArdle WL, Hadley D, Brown MJ, Connell JM, Hingorani AD, Day IN, Lawlor DA, Beilby JP, Lawrence RW, Clarke R, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*. 2011; 478:103–109. [PubMed: 21909115]
- Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, Carlson S, Helgason A, Walters GB, Gunnarsdottir S, Mouy M, Steinthorsdottir V, Eiríksdóttir GH, Björnsdóttir G, Reynisdóttir I, Gudbjartsson D, Helgadóttir A, Jonasdóttir A, Styrkarsdóttir U, Gretarsdóttir S, Magnusson KP, Stefansson H, Fossdal R, Kristjansson K, Gislason HG, Stefansson T, Leifsson BG, Thorsteinsdóttir U, Lamb JR, Gulcher JR, Reitman ML, Kong A, Schadt EE, Stefansson K. Genetics of gene expression and its effect on disease. *Nature*. 2008; 452:423–428. [PubMed: 18344981]
- Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med*. 1975; 4:518–525. [PubMed: 1208363]
- Goring HHH, Curran JE, Johnson MP, Dyer TD, Charlesworth J, Cole SA, Jowett JBM, Abraham LJ, Rainwater DL, Comuzzie AG, Mahaney MC, Almasy L, MacCluer JW, Kissebah AH, Collier GR, Moses EK, Blangero J. Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nature Genetics*. 2007; 39:1208–1216. [PubMed: 17873875]
- Grundberg E, Small KS, Hedman AK, Nica AC, Buil A, Keildson S, Bell JT, Yang TP, Meduri E, Barrett A, Nisbett J, Sekowska M, Wilk A, Shin SY, Glass D, Travers M, Min JL, Ring S, Ho KR, Thorleifsson G, Kong A, Thorsteindóttir U, Ainali C, Dimas AS, Hassanali N, Ingle C, Knowles D, Krestyaninova M, Lowe CE, Di Meglio P, Montgomery SB, Parts L, Potter S, Surdulescu G, Tsaprouni L, Tsoka S, Bataille V, Durbin R, Nestle FO, O'Rahilly S, Soranzo N, Lindgren CM, Zondervan KT, Ahmadi KR, Schadt EE, Stefansson K, Smith GD, McCarthy MI, Deloukas P, Dermitzakis ET, Spector TD, R MTHE. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nature Genetics*. 2012; 44 1084–+
- Guo Y, Lanktree MB, Taylor KC, Hakonarson H, Lange LA, Keating BJ, Fairfax BP, Elbers CC, Barnard J, Farrall M. Gene-centric meta-analyses of 108 912 individuals confirm known body mass index loci and reveal three novel signals. *Human Molecular Genetics*. 2013; 22:184–201. [PubMed: 23001569]

- Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A*. 2009; 106:9362–9367. [PubMed: 19474294]
- Huan T, Zhang B, Wang Z, Joehanes R, Zhu J, Johnson AD, Ying S, Munson PJ, Raghavachari N, Wang R, Liu P, Courchesne P, Hwang SJ, Assimes TL, McPherson R, Samani NJ, Schunkert H, Meng Q, Suver C, O'Donnell CJ, Derry J, Yang X, Levy D. A systems biology framework identifies molecular underpinnings of coronary heart disease. *Arterioscler Thromb Vasc Biol*. 2013; 33:1427–1434. [PubMed: 23539213]
- Joehanes R, Ying S, Huan T, Johnson AD, Raghavachari N, Wang R, Liu P, Woodhouse KA, Sen SK, Tanriverdi K, Courchesne P, Freedman JE, O'Donnell CJ, Levy D, Munson PJ. Gene expression signatures of coronary heart disease. *Arterioscler Thromb Vasc Biol*. 2013; 33:1418–1426. [PubMed: 23539218]
- Kobayashi KS, van den Elsen PJ. NLRC5: a key regulator of MHC class I-dependent immune responses. *Nat Rev Immunol*. 2012; 12:813–820. [PubMed: 23175229]
- Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, Glazer NL, Morrison AC, Johnson AD, Aspelund T, Aulchenko Y, Lumley T, Kottgen A, Vasan RS, Rivadeneira F, Eiriksdottir G, Guo X, Arking DE, Mitchell GF, Mattace-Raso FU, Smith AV, Taylor K, Scharpf RB, Hwang SJ, Sijbrands EJ, Bis J, Harris TB, Ganesh SK, O'Donnell CJ, Hofman A, Rotter JI, Coresh J, Benjamin EJ, Uitterlinden AG, Heiss G, Fox CS, Witteman JC, Boerwinkle E, Wang TJ, Gudnason V, Larson MG, Chakravarti A, Psaty BM, van Duijn CM. Genome-wide association study of blood pressure and hypertension. *Nat Genet*. 2009; 41:677–687. [PubMed: 19430479]
- Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic epidemiology*. 2010; 34:816–834. [PubMed: 21058334]
- Lotta LA, Peyvandi F. Addressing the complexity of cardiovascular disease by design. *Lancet*. 2011; 377:356–358. [PubMed: 21239053]
- Monda KL, Chen GK, Taylor KC, Palmer C, Edwards TL, Lange LA, Ng MC, Adeyemo AA, Allison MA, Bielak LF. A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. *Nature genetics*. 2013
- Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V, Strawbridge RJ, Khan H, Grallert H, Mahajan A. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature genetics*. 2012; 44:981–990. [PubMed: 22885922]
- Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, Marlowe L, Kaleem M, Leung D, Bryden L, Nath P, Zismann VL, Joshipura K, Huentelman MJ, Hu-Lince D, Coon KD, Craig DW, Pearson JV, Holmans P, Heward CB, Reiman EM, Stephan D, Hardy J. A survey of genetic human cortical gene expression. *Nat Genet*. 2007; 39:1494–1499. [PubMed: 17982457]
- Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, Najjar SS, Zhao JH, Heath SC, Eyheramendy S, Papadakis K, Voight BF, Scott LJ, Zhang F, Farrall M, Tanaka T, Wallace C, Chambers JC, Khaw KT, Nilsson P, van der Harst P, Polidoro S, Grobbee DE, Onland-Moret NC, Bots ML, Wain LV, Elliott KS, Teumer A, Luan J, Lucas G, Kuusisto J, Burton PR, Hadley D, McArdle WL, Brown M, Dominiczak A, Newhouse SJ, Samani NJ, Webster J, Zeggini E, Beckmann JS, Bergmann S, Lim N, Song K, Vollenweider P, Waeber G, Waterworth DM, Yuan X, Groop L, Orho-Melander M, Allione A, Di Gregorio A, Guarrera S, Panico S, Ricceri F, Romanazzi V, Sacerdote C, Vineis P, Barroso I, Sandhu MS, Luben RN, Crawford GJ, Jousilahti P, Perola M, Boehnke M, Bonnycastle LL, Collins FS, Jackson AU, Mohlke KL, Stringham HM, Valle TT, Willer CJ, Bergman RN, Morken MA, Doring A, Gieger C, Illig T, Meitinger T, Org E, Pfeuffer A, Wichmann HE, Kathiresan S, Marrugat J, O'Donnell CJ, Schwartz SM, Siscovick DS, Subirana I, Freimer NB, Hartikainen AL, McCarthy MI, O'Reilly PF, Peltonen L, Pouta A, de Jong PE, Snieder H, van Gilst WH, Clarke R, Goel A, Hamsten A, Peden JF, et al. Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet*. 2009; 41:666–676. [PubMed: 19430483]
- Price AL, Helgason A, Thorleifsson G, McCarroll SA, Kong A, Stefansson K. Single-Tissue and Cross-Tissue Heritability of Gene Expression Via Identity-by-Descent in Related or Unrelated Individuals. *Plos Genetics*. 2011; 7

- Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010; 26:2336–2337. [PubMed: 20634204]
- Ramasamy A, Trabzuni D, Gibbs JR, Dillman A, Hernandez DG, Arepalli S, Walker R, Smith C, Illori GP, Shabalín AA. Resolving the polymorphism-in-probe problem is critical for correct interpretation of expression QTL studies. *Nucleic acids research*. 2013; 41:e88–e88. [PubMed: 23435227]
- Rotival M, Zeller T, Wild PS, Maouche S, Szymczak S, Schillert A, Castagne R, Deiseroth A, Proust C, Brocheton J, Godefroy T, Perret C, Germain M, Eleftheriadis M, Sinning CR, Schnabel RB, Lubos E, Lackner KJ, Rossmann H, Munzel T, Rendon A, Erdmann J, Deloukas P, Hengstenberg C, Diemert P, Montalescot G, Ouwehand WH, Samani NJ, Schunkert H, Tregouet DA, Ziegler A, Goodall AH, Cambien F, Tiret L, Blankenberg S. Integrating genome-wide genetic variations and monocyte expression data reveals trans-regulated gene modules in humans. *PLoS Genet*. 2011; 7:e1002367. [PubMed: 22144904]
- Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, Kasarskis A, Zhang B, Wang S, Suver C, Zhu J, Millstein J, Sieberts S, Lamb J, GuhaThakurta D, Derry J, Storey JD, Avila-Campillo I, Kruger MJ, Johnson JM, Rohl CA, van Nas A, Mehrabian M, Drake TA, Lusk AJ, Smith RC, Guengerich FP, Strom SC, Schuetz E, Rushmore TH, Ulrich R. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol*. 2008; 6:e107. [PubMed: 18462017]
- Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardisino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Muhleisen TW, Muhlestein JB, Munzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nothen MM, Olivieri O, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet*. 2011; 43:333–338. [PubMed: 21378990]
- Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Lango Allen H, Lindgren CM, Luan J, Magi R, Randall JC, Vedantam S, Winkler TW, Qi L, Workalemahu T, Heid IM, Steinthorsdottir V, Stringham HM, Weedon MN, Wheeler E, Wood AR, Ferreira T, Weyant RJ, Segre AV, Estrada K, Liang L, Nemesh J, Park JH, Gustafsson S, Kilpelainen TO, Yang J, Bouatia-Naji N, Esko T, Feitosa MF, Kutalik Z, Mangino M, Raychaudhuri S, Scherag A, Smith AV, Welch R, Zhao JH, Aben KK, Absher DM, Amin N, Dixon AL, Fisher E, Glazer NL, Goddard ME, Heard-Costa NL, Hoesel V, Hottenga JJ, Johansson A, Johnson T, Ketkar S, Lamina C, Li S, Moffatt MF, Myers RH, Narisu N, Perry JR, Peters MJ, Preuss M, Ripatti S, Rivadeneira F, Sandholt C, Scott LJ, Timpson NJ, Tyrer JP, van Wingerden S, Watanabe RM, White CC, Wiklund F, Barlassina C, Chasman DI, Cooper MN, Jansson JO, Lawrence RW, Peltikka N, Prokopenko I, Shi J, Thiering E, Alavere H, Alibrandi MT, Almgren P, Arnold AM, Aspelund T, Atwood LD, Balkau B, Balmforth AJ, Bennett AJ, Ben-Shlomo Y, Bergman RN, Bergmann S, Biebermann H, Blakemore AI, Boes T, Bonnycastle LL, Bornstein SR, Brown MJ, Buchanan TA, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet*. 2010; 42:937–948. [PubMed: 20935630]
- Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, D'Agostino RB Sr, Fox CS, Larson MG, Murabito JM, O'Donnell CJ, Vasan RS, Wolf PA, Levy D. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol*. 2007; 165:1328–1335. [PubMed: 17372189]

- Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, Lyle R, Hunt S, Kahl B, Antonarakis SE, Tavare S, Deloukas P, Dermitzakis ET. Genome-wide associations of gene expression variation in humans. *PLoS Genet.* 2005; 1:e78. [PubMed: 16362079]
- Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, Ingle CE, Dunning M, Flicek P, Koller D, Montgomery S, Tavare S, Deloukas P, Dermitzakis ET. Population genomics of human gene expression. *Nature Genetics.* 2007; 39:1217–1224. [PubMed: 17873874]
- Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. *Nat Rev Genet.* 2013; 14:139–149. [PubMed: 23329114]
- Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, Johansen CT, Fouchier SW, Isaacs A, Peloso GM, Barbalic M, Ricketts SL, Bis JC, Aulchenko YS, Thorleifsson G, Feitosa MF, Chambers J, Orho-Melander M, Melander O, Johnson T, Li X, Guo X, Li M, Shin Cho Y, Jin Go M, Jin Kim Y, Lee JY, Park T, Kim K, Sim X, Twee-Hee Ong R, Croteau-Chonka DC, Lange LA, Smith JD, Song K, Hua Zhao J, Yuan X, Luan J, Lamina C, Ziegler A, Zhang W, Zee RY, Wright AF, Witteman JC, Wilson JF, Willemssen G, Wichmann HE, Whitfield JB, Waterworth DM, Wareham NJ, Waeber G, Vollenweider P, Voight BF, Vitart V, Uitterlinden AG, Uda M, Tuomilehto J, Thompson JR, Tanaka T, Surakka I, Stringham HM, Spector TD, Soranzo N, Smit JH, Sinisalo J, Silander K, Sijbrands EJ, Scuteri A, Scott J, Schlessinger D, Sanna S, Salomaa V, Saharinen J, Sabatti C, Ruukonen A, Rudan I, Rose LM, Roberts R, Rieder M, Psaty BM, Pramstaller PP, Pichler I, Perola M, Penninx BW, Pedersen NL, Pattaro C, Parker AN, Pare G, Oostra BA, O'Donnell CJ, Nieminen MS, Nickerson DA, Montgomery GW, Meitinger T, McPherson R, McCarthy MI, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature.* 2010; 466:707–713. [PubMed: 20686565]
- Verlaan DJ, Berlivet S, Hunninghake GM, Madore AM, Lariviere M, Moussette S, Grundberg E, Kwan T, Ouimet M, Ge B, Hoberman R, Swiatek M, Dias J, Lam KC, Koka V, Harmsen E, Soto-Quiros M, Avila L, Celedon JC, Weiss ST, Dewar K, Sinnett D, Laprise C, Raby BA, Pastinen T, Naumova AK. Allele-specific chromatin remodeling in the ZBP2/GSDMB/ORMDL3 locus associated with the risk of asthma and autoimmune disease. *Am J Hum Genet.* 2009; 85:377–393. [PubMed: 19732864]
- Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, Zeggini E, Huth C, Aulchenko YS, Thorleifsson G, McCulloch LJ, Ferreira T, Grallert H, Amin N, Wu G, Willer CJ, Raychaudhuri S, McCarroll SA, Langenberg C, Hofmann OM, Dupuis J, Qi L, Segre AV, van Hoek M, Navarro P, Ardlie K, Balkau B, Benediktsson R, Bennett AJ, Blagieva R, Boerwinkle E, Bonnycastle LL, Bostrom KB, Bravenboer B, Bumpstead S, Burt NP, Charpentier G, Chines PS, Cornelis M, Couper DJ, Crawford G, Doney ASF, Elliott KS, Elliott AL, Erdos MR, Fox CS, Franklin CS, Ganser M, Gieger C, Grarup N, Green T, Griffin S, Groves CJ, Guiducci C, Hadjadj S, Hassanali N, Herder C, Isomaa B, Jackson AU, Johnson PRV, Jorgensen T, Kao WHL, Klopp N, Kong A, Kraft P, Kuusisto J, Lauritzen T, Li M, Lieverse A, Lindgren CM, Lyssenko V, Marre M, Meitinger T, Midthjell K, Morken MA, Narisu N, Nilsson P, Owen KR, Payne F, Perry JRB, Petersen A-K, Platou C, Proenca C, Prokopenko I, Rathmann W, Rayner NW, Robertson NR, Rocheleau G, Roden M, Sampson MJ, Saxena R, Shields BM, Shrader P, Sigurdsson G, Sparso T, Strassburger K, Stringham HM, Sun Q, Swift AJ, Thorand B, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet.* 2010; 42:579–589. [PubMed: 20581827]
- Westra H-J, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, Christiansen MW, Fairfax BP, Schramm K, Powell JE. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nature genetics.* 2013; 45:1238–1243. [PubMed: 24013639]
- Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, Maouche S, Germain M, Lackner K, Rossmann H, Eleftheriadis M, Sinning CR, Schnabel RB, Lubos E, Mennerich D, Rust W, Perret C, Proust C, Nicaud V, Loscalzo J, Hubner N, Tregouet D, Munzel T, Ziegler A, Tiret L, Blankenberg S, Cambien F. Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PLoS One.* 2010; 5:e10693. [PubMed: 20502693]

A **B**

h^2	Number of Genes	% of Transcriptome (18,000 transcripts)
$h^2 > 0.6$	50	0.3
$0.6 \geq h^2 > 0.5$	77	0.4
$0.5 \geq h^2 > 0.4$	98	0.5
$0.4 \geq h^2 > 0.3$	384	2
$0.3 \geq h^2 > 0.2$	1060	6

Figure 1. Heritability Distribution in Framingham Cohorts

A) Heritability ($h^2_{geneExp}$) distribution of transcripts; B) Summary of genes in different $h^2_{geneExp}$ sub-categories.

Figure 2. Relationship Between Heritability Estimates and Proportion of Genes with *cis*, *trans* or long-range *cis* eQTLs

A) Proportion of transcripts having *cis* eQTLs in different heritability levels; B) Proportion of transcripts having *trans* eQTLs in different heritability levels; C) Proportion of transcripts having long-range *cis* eQTLs in different heritability levels.

Figure 3. Proportion of Transcript Variance Explained by Peak eQTLs

A) Variance proportion of a transcript explained by a single peak *cis* eQTL; B) Variance proportion of a transcript explained by a single peak *trans* eQTL; C) Variance proportion of a transcript explained by a single peak *long-range cis* eQTL.

A

B

Figure 4. Genotype-specific Gene Expression and a Proposed Schematic of Master *cis*- and *trans*-eQTL Regulation of Transcripts

A) An example of master *cis* eQTL model; B) An example of master *trans* eQTL model.

The number above the directed curve line indicated the heritability proportion of transcripts explained by the SNP. TF indicated transcription factor. The y-axis geneExp_resid indicated the gene expression values were adjusted for age, sex and technique covariates.

A

B

C

Figure 5. *cis/trans* eQTLs linking BMI GWAS SNPs and BMI signature genes

A) Regional association plots (drawn by LocusZoom (Pruim et al. 2010)) of BMI GWAS SNPs (peak SNP rs7359397) and its proxy genes. The y-axis is $-\log_{10}$ -transformed BMI GWAS p values from GIANT (Speliotes et al. 2010); the genes marked by green rectangle indicating eQTL genes, by blue rectangle indicating BMI signature genes and by red rectangle indicating the overlap of eQTL genes and BMI signature genes. The genes marked by green rectangle indicating eQTL genes and by blue rectangle indicating BMI signature genes. B) Schematic figure depicting BMI GWAS SNP (peak SNP rs7359397) is associated with multiple genes in *cis*-manner, and some of those genes are differentially expressed for BMI; C) rs7359397 driving BMI differential gene expression signatures. The number above the directed curved line indicated the phenotypic proportion of transcripts explained by the SNP, the phenotypic proportion of traits explained by the SNP, or the phenotypic proportion of traits explained by trait signatures.

Table 1

Characteristics of the 5,626 Framingham Heart Study Participants

	Offspring Cohort N=2,446 (examination cycle 8: 2005–2008)	Third Generation Cohort N=3,180 (examination cycle 2: 2008–2011)
Phenotypes/Covariates	Mean (SD)	Mean (SD)
Male (%)	45	47
Age (yrs)	66 (9)	46 (9)
Body mass index (kg/m ²)	28.4 (5.4)	28.0 (5.9)

Table 2

cis eQTLs among GWAS Results for Metabolic Traits

Gene	Chr.	eQTL	SNP location	$\#h^2_{geneExp}$	SNP -- Gene		SNP -- Trait		GWAS-Trait
					eSNP-Pval	$\#h^2_{eQTL}$	$\#GWAS-p$ value	$\#h^2_{trait-SNP}$	
GSTM4	chr1	rs4970843	Intronic (SORT1)	0.55	7.26E-63	0.11	9.77E-23	0.002	LDL/TG
RHD	chr1	rs926438	Intronic (RHCE)	0.52	5.19E-289	0.47	1.68E-09	5.01E-7	LDL/TG
DOCK7	chr1	rs1748195	Intronic (DOCK7)	0.23	2.73E-162	0.60	2.43E-42	0.004	LDL/TG/TG
GSTM4	chr1	rs4970843	Intronic (SORT1)	0.55	7.26E-63	0.11	9.77E-23	0.001	LDL/TG
CLCN6	chr1	rs1256713 6	Intronic (CLCN6)	0.21	6.68E-52	0.21	3.41E-10	0.007	SBP/DBP
C4B	chr6	rs2072633	Intronic (CFB)	0.21	1.56E-40	0.16	2.67E-08	0	TC
HLA-B	chr6	rs437179	Coding(SKI V2L)	0.46	1.38E-59	0.11	6.98E-12	5.60E-4	TC
HLA-B	chr6	rs2524095	Unknown	0.46	4.88E-113	0.22	3.02E-08	0.001	TG
LOC10051 0059	chr6	rs2894254	Intronic (C6orf10)	0.74	1.65E-218	0.24	8.92E-10	0.004	TG
HCP5	chr6	rs2596565	Unknown	0.29	5.63E-28	0.13	1.55E-10	0.004	TG/TG
NCRNA00 243	chr6	rs3130544	Unknown	0.26	3.70E-100	0.37	4.84E-10	0.003	TG/TG
FLOT1				0.25	1.39E-59	0.23			
SLC44A4	chr6	rs486416	Intronic (EHMT2)	0.56	0	0.87	1.99E-14	0.006	TG/TG
CSGALNA CT1	chr8	rs1967103	Intronic (CSGALNA CT1)	0.55	5.97E-218	0.35	5.78E-09	0.006	HDL/TG
GRINA	chr8	rs6985603	5' upstream (SPATC1)	0.26	1.10E-104	0.35	3.39E-10	0.002	LDL
C11orf10	chr11	rs174534	Intronic (C11orf9)	0.32	2.01E-255	0.65	2.68E-18	0.005	HDL/LDL/T G/TG
FADS2	chr11	rs174549	Intronic (FADS1)	0.74	0	0.54	2.08E-20	0.004	HDL/LDL/T G/TG

Gene	Chr.	eQTL	SNP location	SNP -- Gene		SNP -- Trait	
				eSNP-Pval	$r^2 h^2_{eQTL}$	$\ddagger$ GWAS-p value	GWAS-Trait
SIDT2	chr11	rs6589597	Unknown	2.17E-297	0.46	7.22E-14	TC
SIDT2	chr11	rs7107152	Intronic (SIDT2)	0	0.68	7.13E-10	TG
POC1B	chr12	rs11105328	Intergenic	2.53E-41	0.11	1.08E-09	SBP/DBP
MAP2K5	chr15	rs12902812	Intronic (MAP2K5)	1.40E-198	0.73	1.33E-08	BMI
APOB48R	chr16	rs12443881	Intronic (ATXN2L)	0	0.72	5.32E-10	BMI
GSDMB	chr17	rs12936231	Intronic (ZPBP2)	0	0.92	2.47E-09	HDL
ORMDL3				1.34E-92	0.34		
LILRB3	chr19	rs431420	Unknown	3.50E-160	0.33	7.51E-13	HDL
LILRB2				5.93E-124	0.31		
SLC44A2	chr19	rs2288904	Coding (SLC44A2)	2.37E-48	0.16	5.72E-10	LDL
PVRL2	chr19	rs8104483	Intronic (PVRL2)	1.13E-271	0.40	1.94E-20	LDL/TC

$*h^2_{geneExp}$ is the proportion of total phenotypic variance that can be explained by the additive genetic components ($h^2_{geneExp} = \sigma^2_A / \sigma^2_{geneExp}$)

$\dagger r^2 h^2_{eQTL}$ is the ratio of the proportion of variance explained by a peak eQTL to the heritability of a gene expression trait ($r^2 h^2_{eQTL} = h^2_{eQTL} / h^2_{geneExp}$)

$\ddagger$ GWAS-p value: the minimum trait-associated SNP p value in NHGRI GWAS Catalog (Hindorff et al. 2009)

$\S r^2 h^2_{trait-SNP}$ is the ratio of the proportion of variance explained by a SNP to the heritability of a clinical trait ($r^2 h^2_{trait-SNP} = h^2_{SNP} / h^2_{trait}$). $r^2 h^2_{trait-SNP}$ is calculated from FHS samples along.

Table 3

trans eQTLs among GWAS for Metabolic Traits

Gene	Chr. for gene	eQTL	SNP Location	$h^2_{geneExp}$	SNP -- Gene		SNP - Trait		
					eSNP-Pval	$\hat{\sigma}^2 h^2_{eQTL}$	$\hat{\sigma}^2 GWAS-p$ value	$\hat{\sigma}^2 h^2_{trait-SNP}$	GWAS-Trait
SELP	chr1			0.37	4.84E-09	0.02			
AQP10	chr1	rs10761741	Intronic (JMJD1C)	0.21	9.04E-14	0.06	1.81E-11	0.002	TG
ITGA2B	chr17			0.4	1.22E-10	0.02			
CXCL5	chr4			0.29	2.37E-13	0.04			
PKHD1L1	chr8	rs10761779	Intergenic	0.32	3.60E-10	0.03	4.13E-11	0.002	TG
PPBP	chr4			0.30	1.28E-08	0.02			
EBF1	chr5	rs11078927	Intronic (GSDMB)	0.28	2.76E-08	0.02	1.28E-08	1.88E-4	HDL
CPA3	chr3	rs1178977	Intronic (BAZ1B)	0.39	1.44E-08	0.01	2.90E-55	0.03	TG
ALAS2	chrX	rs1800562	Coding (HFE)	0.37	2.73E-13	0.03	6.07E-10	0.005	LDL/TC
ALPL	chr1	rs2872507	Intergenic	0.25	5.10E-09	0.05	2.83E-08	1.89E-4	HDL
BTN3A1	chr6			0.23	2.22E-49	0.19			
HCP5	chr6	rs289754	Intronic (NLR5)	0.29	2.57E-35	0.1	8.98E-10	1.17E-4	HDL
HLA-F	chr6			0.26	8.53E-14	0.04			
FCGR1A	chr1			0.27	9.82E-14	0.04			
ARHGEF40	chr14			0.28	8.82E-11	0.03			
MYADM	chr19	rs3184504	Coding (SH2B3)	0.25	3.88E-09	0.03	2.33E-14	0.002	SBP/DBP/LDL/TC
IDO1	chr8			0.26	2.02E-08	0.02			
IFIT3	chr10			0.22	2.82E-08	0.02			
PLA2G7	chr6	rs3808456	Intronic (TRPS1)	0.45	2.18E-10	0.03	1.28E-09	0	HDL
ESPN	chr1	rs634869	Intergenic	0.32	2.29E-11	0.03	4.94E-08	4.17E-4	HDL
GUCY1A3	chr4			0.27	2.31E-08	0.03			
ITGB3	chr17	rs7077256	Intronic (REEP3)	0.54	4.64E-12	0.02	1.70E-10	0.001	TG
CLU	chr8			0.47	3.84E-09	0.02			
GPR109A	chr12	rs7832357	Intergenic	0.22	3.83E-11	0.03	1.84E-11	0.004	TG
MME	chr3			0.23	4.34E-09	0.03			

h_{genExp}^2 is the heritability of the transcript

r_{eQTL}^2 is the heritability proportion of the transcript explained by a single eSNP

$\hat{\tau}$ GWAS-p value: the minimum trait-associated SNP p value in NHGRI GWAS catalog (Hindorff et al. 2009)

$h_{trait-SNP}^2$ is the ratio of the proportion of variance explained by a SNP to the heritability of a clinical trait (r_{trait}^2/h_{SNP}^2). $r_{trait-SNP}^2$ is calculated from FHS samples along.

Table 4

BMI GWAS SNPs for BMI Differentially Expressed Genes linking by eQTLs

Gene	Chr. for gene	eQTL	SNP Location	$*h^2_{\text{geneExp}}$	Gene -- Trait			SNP -- Gene			SNP -- Trait	
					Signature P value	Signature FDR	Phenotypic Proportion (Gene for Trait)	eSNP p value	Cis Trans	Phenotypic Proportion (SNP for Gene)	#GWAS p value	Phenotypic Proportion (SNP for Trait)
ETV5	chr3	rs9816226	Intronic (DGKG)	0.01	0.0002	0.001	0.003	4.44E-16	cis	0.012	7.61E-14	1.5E-4
AAGAB	chr15	rs2278076	3' downstream (MAP2K5)	0.04	0.0003	0.002	0.001	6.08E-06	cis	0.004	1.46E-10	0
CLN6	chr15	rs3784707	Intronic (MAP2K5)	0.17	1.35E-10	2.58E-09	0.008	9.66E-06	cis	0.004	1.46E-10	0
SULT1A1				0.21	0.01	0.05	0.001	5.54E-07	cis	0.006		
SPNS1	chr16	rs7498665	Coding (SH2B1)	0.05	8.25E-05	5.57E-04	0.004	5.77E-15	cis	0.012	1.75E-10	8.6E-5
TUFM				0.19	0.007	0.03	0.002	5.16E-08	cis	0.148		
PTPRJ	chr11	rs2290851	Intronic (FNBP4)	0.19	1.28E-17	5.79E-16	0.014	3.88E-07	cis	0.005		
MTCH2				0.08	1.13E-05	9.26E-05	0.003	3.60E-14	cis	0.011	2.77E-10	4.5E-4
LIN7C	chr11	rs1519480	3' downstream (BDNF)	0.08	2.55E-06	2.39E-05	0.004	2.22E-14	cis	0.012	2.12E-09	4.4E-4

[‡] GWAS-p value: for BMI, the p values are from GIANT (Speliotes et al. 2010);