Brian H. Chen,[1,2,3] Marie-France Hivert,[4,5,6] Marjolein J. Peters,[7,8] Luke C. Pilling,[9] John D. Hogan,[10] Lisa M. Pham,[10] Lorna W. Harries,[11] Caroline S. Fox,[2,3] Stefania Bandinelli,[12] Abbas Dehghan,[13] Dena G. Hernandez,[14] Albert Hofman,[13] Jaeyoung Hong,[15] Roby Joehanes,[2,3,16] Andrew D. Johnson,[2,3] Peter J. Munson,[17] Denis V. Rybin,[18] Andrew B. Singleton,[14] André G. Uitterlinden,[7,8,13] Saixia Ying,[17] MAGIC Investigators, David Melzer,[9] Daniel Levy,[2,3] Joyce B.J. van Meurs,[7,8] Luigi Ferrucci,[1] Jose C. Florez,[5,19,20,21] Josée Dupuis,[2,15] James B. Meigs,[20,21,22] and Eric D. Kolaczyk[10,23]

Peripheral Blood Transcriptomic Signatures of Fasting Glucose and Insulin Concentrations

Diabetes 2016;65:3794–3804 | DOI: 10.2337/db16-0470

Genome-wide association studies (GWAS) have successfully identified genetic loci associated with glycemic traits. However, characterizing the functional significance of these loci has proven challenging. We sought to gain insights into the regulation of fasting insulin and fasting glucose through the use of gene expression microarray data from peripheral blood samples of participants without diabetes in the Framingham Heart Study (FHS) (n = 5,056), the Rotterdam Study (RS) (n = 723), and the InCHIANTI Study (Invecchiare in Chianti) (n = 595). Using a false discovery rate q <0.05, we identified three transcripts associated with fasting glucose and 433 transcripts associated with fasting insulin levels after adjusting for age, sex, technical covariates, and complete blood cell counts. Among the findings, circulating *IGF2BP2* transcript levels were positively associated with fasting insulin in both the FHS and RS. Using 1000 Genomes–imputed genotype data, we identified 47,587 *cis*-expression quantitative trait loci (eQTL) and 6,695 *trans*-eQTL associated with the 433 significant insulin-associated transcripts. Of note, we identified a *trans*-eQTL (rs592423), where the A allele was associated with higher *IGF2BP2* levels and with fasting insulin in an independent genetic meta-analysis comprised of 50,823 individuals. We conclude that integration of

[1]Longitudinal Studies Section, Translational Gerontology Branch, Intramural Research Program, National Institute on Aging, National Institutes of Health, Baltimore, MD
[2]Framingham Heart Study, National Heart, Lung, and Blood Institute, Framingham, MA
[3]Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, Bethesda, MD
[4]Department of Population Medicine, Harvard Pilgrim Health Care Institute, Harvard Medical School, Boston, MA
[5]Diabetes Research Center, Department of Medicine, Massachusetts General Hospital, Boston, MA
[6]Department of Medicine, Université de Sherbrooke, Sherbrooke, Quebec, Canada
[7]Department of Internal Medicine, Erasmus University Medical Center Rotterdam, Rotterdam, the Netherlands
[8]Netherlands Genomics Initiative–sponsored Netherlands Consortium for Healthy Aging, Leiden and Rotterdam, the Netherlands
[9]Epidemiology and Public Health Group, Institute of Biomedical and Clinical Sciences, University of Exeter Medical School, Exeter, U.K.
[10]Program in Bioinformatics, Boston University, Boston, MA
[11]Institute of Biomedical and Clinical Sciences, University of Exeter Medical School, Exeter, U.K.
[12]Geriatric Rehabilitation Unit, Azienda Sanitaria di Firenze, Florence, Italy
[13]Department of Epidemiology, Erasmus University Medical Center Rotterdam, Rotterdam, the Netherlands
[14]Laboratory of Neurogenetics, Intramural Research Program, National Institute on Aging, National Institutes of Health, Bethesda, MD
[15]Department of Biostatistics, Boston University School of Public Health, Boston, MA
[16]Hebrew SeniorLife, Harvard Medical School, Boston, MA
[17]Mathematical and Statistical Computing Laboratory, Center for Information Technology, National Institutes of Health, Bethesda, MD
[18]Data Coordinating Center, Boston University, Boston, MA
[19]Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA
[20]Metabolism Program and Program in Medical and Population Genetics, Broad Institute, Cambridge, MA
[21]Department of Medicine, Harvard Medical School, Boston, MA
[22]Division of General Internal Medicine, Massachusetts General Hospital, Boston, MA
[23]Department of Mathematics and Statistics, Boston University, MA

Corresponding authors: James B. Meigs, jmeigs@partners.org, and Eric D. Kolaczyk, kolaczyk@bu.edu.

Received 22 April 2016 and accepted 4 September 2016.

This article contains Supplementary Data online at http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db16-0470/-/DC1.

B.H.C. and M.-F.H. contributed equally to this work.

J.B.M. and E.D.K. jointly directed this work.

genomic and transcriptomic data implicate circulating *IGF2BP2* mRNA levels associated with glucose and insulin homeostasis.

Genome-wide association studies (GWAS) using arrays containing hundreds of thousands of single nucleotides polymorphisms (SNPs) have revealed multiple genetic variants associated with fasting glucose or fasting insulin in humans (1–5). Yet, all together, those SNPs explained only a small percentage of the total variation in fasting glucose (4.8%) and fasting insulin (1.2%) (6). Transcriptomic profiling provides a high-throughput platform to expand genomic associations and reveal how gene expression complements studies on genetic variations.

To date, most transcriptomic studies of fasting glucose and fasting insulin have examined a limited number of genes. Transcriptomic response to insulin treatment has been reported, but the sample sizes have been relatively small (7,8). These studies have been instrumental in testing hypothesis-driven studies on the acute molecular effects of insulin. However, few studies have comprehensively investigated the genetic regulation of steady-state fasting glucose or fasting insulin levels, particularly using blood transcript levels.

To deepen our understanding of the regulation of fasting glucose and fasting insulin, we performed a transcriptome-wide association study (TWAS) in three well-characterized cohort studies: Framingham Heart Study (FHS), the Rotterdam Study (RS), and the InCHIANTI Study (Invecchiare in Chianti). Using a hypothesis-free approach, we applied stringent criteria for cross-replication across cohorts and applied pathway analyses to provide an integrated view of our findings. We further used expression quantitative trait loci (eQTL) to link TWAS and GWAS findings to identify associated transcripts that may be under genetic control.

RESEARCH DESIGN AND METHODS

Overview of Approach

As depicted in Fig. 1, we conducted a TWAS in three independent cohorts (described below). We then focused on highly reproducible transcripts across cohorts, which we defined as having a false discovery rate (FDR) of $q < 0.05$ in both sets of results, separated by array platform (i.e., Affymetrix vs. Illumina). Multiple approaches were used to assess the reproducibility and biological relevance of our transcript associations. First, we examined the transcriptomic associations of published genes assigned to genetic variants reported in prior fasting glucose and fasting insulin GWAS. Next, we used eQTL analysis to comprehensively assess any convergence of findings from associations between genetic variants and transcripts identified by our TWAS as well as with insulin and glucose levels. Signals confirmed by both approaches represent highly reproducible findings that span several large populations. Last, we conducted gene set enrichment analysis (GSEA) to provide insights into biological pathways that may be involved in the regulation of transcripts associated with fasting glucose or insulin levels.

Gene Expression Correlations Across Tissues

RNA sequencing data from the Genotype-Tissue Expression (GTEx) Project (http://www.gtexportal.org/static/datasets/gtex_analysis_v6/rna_seq_data/GTEx_Analysis_v6_RNA-seq_RNA-SeQCv1.1.8_gene_rpkm.gct.gz, accessed on 29 June 2016) (9). Analysis was restricted to tissues determined a priori to be of relevance to glycemic traits, including visceral fat, kidney, liver, muscle, and pancreas. Values with reads per kilobases of transcript per million mapped reads <1 were excluded. Replicate samples were combined by taking the median (or mean, if even number of replicates) value for each transcript. For each pairwise tissue comparison, Spearman correlations were computed for each individual that had transcript levels available in both tissues. Correlation coefficients for each tissue pair were obtained by taking the mean across individuals with transcript data in both tissues. To estimate the sample size needed in nonblood tissue to achieve equivalent statistical power as our study, we multiplied the sample size from our study with the squared correlation coefficient obtained from our GTEx analysis, following the approach described in Pritchard et al. (10).

Study Populations

Detailed descriptions of the three population-based cohorts that were included in the current analysis can be found in the Supplementary Data. Briefly, the first cohort (FHS) included participants from the FHS Offspring Study's 8th examination cycle (n = 2,049) and the Third Generation's 2nd examination cycle (n = 3,007). The second cohort (RS) included participants from the third recruitment cohort of the RS (n = 881). The third cohort (InCHIANTI Study) included participants from the third follow-up visit (n = 698). Participants were excluded if they were missing data on glucose, insulin, or blood cell counts or had type 2 diabetes. Informed consent was obtained from each FHS participant and the study protocol was approved under Boston University Medical Center's institutional review board protocol (H-27984). RS has been approved by the Medical Ethics Committee of the Erasmus University Medical Center Rotterdam and by the Ministry of Health, Welfare and Sport of the Netherlands, which implemented the Wet Bevolkingsonderzoek: ERGO (Population Studies Act: Rotterdam Study). All RS participants provided written informed consent to participate in the study and to obtain information from their treating physicians. Ethics approval of the InCHIANTI Study was granted by the Istituto Nazionale di Riposo e Cura per Anziani institutional review board in Italy, and participants gave informed consent to participate.

Outcome Definitions

FHS participants fasted overnight to provide blood specimens, which were frozen in EDTA tubes at −80°C until assayed. Fasting insulin levels were quantified using the

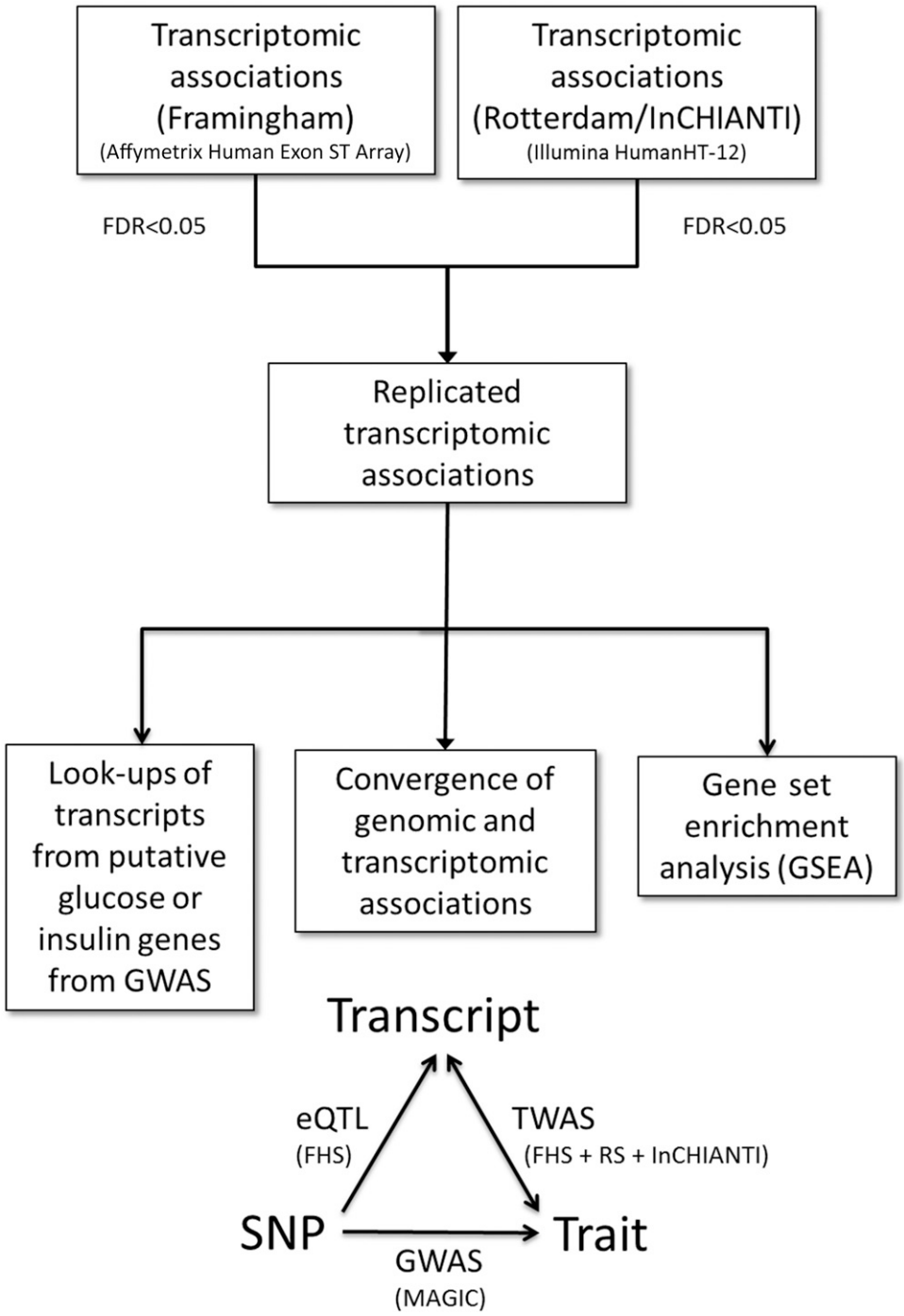

Figure 1—Overview of analytic approach.

Roche e411 immunoanalyzer (Roche Diagnostics, Risch-Rotkreuz, Switzerland). Ten percent of our samples were run in duplicate. The interassay coefficients of variation (CV) for fasting insulin were 4.5% in the Third Generation and 3.8% in the Offspring cohorts. Fasting glucose samples were run on fresh EDTA plasma samples using a hexokinase assay on the Roche Hitachi 911 chemistry analyzer (Roche Diagnostics). All samples were measured in duplicate. The interassay CV for glucose was 1.8%. Glycated hemoglobin (HbA_{1c}) was measured in whole blood using a turbidimetric immunoassay on the Roche Hitachi 911 chemistry analyzer following a hemolysis step. A total of 15.3% of samples were run in duplicate to provide an interassay CV of 2.7%.

For the RS samples, venous blood samples were obtained after an overnight fast (at least 8 h). The glucose samples were stored in −20°C serum samples for <1 week before processing. Insulin was measured in serum samples stored

at −80°C. Serum glucose was assessed using the hexokinase method (Boehringer Mannheim, Mannheim, Germany). Serum insulin levels were quantified using the cobas Roche electrochemiluminescence immunoassay (12017547 122) on a Modular Analytics E170 analyzer (Roche Diagnostics). The interassay CV was <4.9% for insulin and 1.7% for glucose.

InCHIANTI participants fasted overnight (12 h) and were sedentary for at least 15 min before providing blood samples. Aliquots of plasma and serum were prepared and frozen at −80°C. Fasting glucose was determined by an enzymatic colorimetric assay using a modified glucose oxidase-peroxidase method (Roche Diagnostics GmbH, Mannheim, Germany) and a Modular P800 Hitachi analyzer (Hitachi Chemical, Tokyo, Japan). The intra-assay CV was 0.9%, and the interassay CV was 1.8%. Fasting insulin was not quantified in InCHIANTI at this examination cycle.

Transcriptomic Profiling

Detailed descriptions of the transcriptomic profiling can be found elsewhere (11). Briefly, whole-blood samples were collected in PAXgene tubes. Following RNA amplification, global transcript levels were quantified using the Affymetrix Human Exon1.0 ST Array for FHS, Illumina HumanHT-12 v4 Expression BeadChip for RS, and Illumina Human HT-12 v3 BeadChip for InCHIANTI, as described previously (11). Transcriptomic data were RMA normalized (12) for FHS or quantile normalized and $\log_2$-transformed for RS and InCHIANTI. Data are accessible to the public for FHS at dbGaP (accession "phs000363.v7.p8"), for RS at the Gene Expression Omnibus (GEO) (GSE33828), and for InCHIANTI at GEO (GSE48152).

Analysis of Transcript Associations With Glycemic Traits

Fasting insulin levels were natural log-transformed due to their skewed distributions. Associations between the transcript levels (independent variables) and glycemic traits (dependent variables) were adjusted for age, age^2, and sex using linear regression in RS and InCHIANTI and with mixed-effect models in FHS to account for familial correlation. Separate models were further adjusted for BMI to assess its contribution to the transcript associations with glucose or insulin levels. Because fasting insulin levels were not available in InCHIANTI, the fasting insulin results reported are from FHS and RS only, but the fasting glucose results include all three cohorts. Statistically significant transcript associations with glucose (or insulin) levels were defined as FDR q <0.05 in FHS and the meta-analysis of RS and InCHIANTI in a cross-replication manner (13). The glucose results from the RS and InCHIANTI were meta-analyzed because they used similar microarray platforms and had substantially smaller sample sizes compared with the FHS data set. To test for enrichment of known biological pathways in our top gene lists we used GSEA preranked feature using 1,000 permutations (14). We used the Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets restricted to those containing between 15 and 500 genes from our analysis.

Integration of Genomic, Transcriptomic, and Glycemic Trait Data

To infer causal direction of effect, we tested for consistency of associations across genomic, transcriptomic, and trait data (Fig. 1, bottom). First, we identified eQTL for transcripts that were associated with fasting glucose or insulin levels at FDR q <0.05 across cohorts (FHS and RS for insulin, FHS and meta-analysis of RS and InCHIANTI for glucose). We tested associations between each identified cross-replicated transcript and genotypes with minor allele frequencies greater than 1% and 1000 Genomes imputation R^2 >0.3 in FHS. In total 1.52×10^{11} tests were conducted; significant eQTL were identified as having FDR q <0.05. After identifying these eQTL in FHS samples, we examined the eQTL's association with the respective trait (glucose or insulin) in the HapMap-imputed data from the Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) (15) after excluding data from FHS, RS, and InCHIANTI samples (n = 50,823 after exclusion of the three cohorts). This list of SNPs was then pruned based on linkage disequilibrium (r^2 >0.8 in HapMap) and limited to SNPs with at least 10 cohorts contributing results. We considered an eQTL to be significantly associated with glucose or insulin levels in the MAGIC data set using an FDR q <0.05. We used corroborating evidence from transcript-to-trait, SNP-to-transcript (i.e., eQTL), and SNP-to-trait to support hypotheses of directionality and biological relevance.

RESULTS

Study Population Characteristics

Study population characteristics are presented in Table 1. All three cohorts were from European descent, participants were middle-aged at the time of blood draw, and about 55% were women. The InCHIANTI Study participants were older on average than FHS or RS participants. BMI levels were comparable across cohorts. Although fasting glucose levels were comparable between FHS (median 96.0 mg/dL) and RS (95.4 mg/dL), InCHIANTI participants were substantially lower (87 mg/dL). Fasting insulin levels in RS were substantially higher (median 76.0 pmol/L) than that of FHS participants (56.1 pmol/L).

Gene Expression Correlations Across Tissues

Using publicly available RNA sequencing data from the GTEx Project, the correlations of transcript levels from whole blood to other tissues were assessed. Spearman correlations with whole blood ranged from r = 0.39 in liver to r = 0.54 in kidney (Table 2 and Supplementary Fig. 1). Thus, if the transcript associations that we observed in our analysis with fasting insulin hold in other tissues, a sample size of 110 liver samples would be sufficient to achieve the same statistical power as RS, and a sample size of 769 liver samples would be sufficient to achieve the same power as FHS. For kidney samples, samples sizes of 211 and 1,474 would be sufficient to achieve the same power as RS and FHS, respectively.

Table 1—Characteristics of the study participants

	Discovery cohort	Replication cohorts	
	FHS	RS	InCHIANTI
n	5,056	723	595
Age (years), mean ± SD	54.0 ± 13.1	59.3 ± 7.9	71.5 ± 15.9
Sex (male), *n* (%)	2,261 (44.7)	320 (44.3)	262 (44)
BMI (kg/m^2), mean ± SD	27.7 ± 5.3	27.4 ± 4.5	26.9 ± 4.2
Microarray platform	Affymetrix GeneChip Human Exon 1.0 ST	Illumina HumanHT-12 v4 BeadChip	Illumina HumanHT-12 v3 BeadChip
Biomarker levels, median (IQR)			
Fasting glucose (mg/dL)	96.0 (90.0–102.0)	95.4 (90.0–102.6)	87 (79–95)
Fasting insulin (pmol/L)	56.1 (38.2–83.1)	76.0 (55.0–106.7)	NA

Fasting glucose conversion from mg/dL to mmol/L, multiply by 0.0555. Fasting insulin conversion from pmol/L to μIU/mL, multiply by 0.144.

Transcriptomic Associations With Fasting Glucose

To identify transcriptomic signatures that were replicated in multiple data sets, we focused on transcripts that were significantly associated (FDR $q < 0.05$) with fasting glucose in FHS and in the meta-analysis of RS and InCHIANTI. Using this criterion, we identified three out of 12,051 unique transcripts that were significantly associated with fasting glucose in both FHS and the meta-analyzed results from RS and InCHIANTI (Fig. 2). The three identified gene transcripts were *MARCH8* (FHS $q = 4.5 \times 10^{-15}$; RS + InCHIANTI $q = 8.4 \times 10^{-3}$), *OSBP2* (FHS $q = 2.0 \times 10^{-11}$; RS + InCHIANTI $q = 1.3 \times 10^{-2}$), and *TNS1* (FHS $q = 4.3 \times 10^{-12}$; RS + InCHIANTI $q = 4.6 \times 10^{-2}$). The full set of fasting glucose TWAS results can be found in Supplementary Table 1 (all Supplementary Tables can be accessed at http://sites.bu.edu/fhspl/publications/pbtsfgic_supp/). After BMI adjustment, no transcripts met our stringent significance criteria (Supplementary Table 2). However, the direction and magnitude of *t*-statistics with and without BMI adjustment were highly correlated ($r = 0.90$ in FHS and $r = 0.92$ in RS + InCHIANTI) (Supplementary Fig. 2).

Transcriptomic Associations With Fasting Insulin

We applied a similar strategy for cross-cohort reproducibility of transcripts significantly associated with fasting insulin (Fig. 2). Fasting insulin was associated with 433 transcripts with FDR $q < 0.05$ in both FHS and the analysis of RS. The full set of fasting insulin TWAS results can be found in Supplementary Table 3. Further adjustment for BMI yielded 117 transcripts that were significant in both sets of results, of which 112 (95.7%) were also significant in the BMI-unadjusted results (Supplementary Table 4). Among significant findings, two of the transcripts from our fasting glucose analyses (*OSBP2* and *TNS1*) were also associated with fasting insulin (with or without adjustment for BMI). The *t*-statistics for the fasting insulin associations across all transcripts with and without BMI adjustment were highly correlated ($r = 0.89$ in FHS and $r = 0.87$ in RS), suggesting minimal influence on the transcript-insulin associations by BMI (Supplementary Fig. 2). Fasting insulin and fasting glucose associations across all transcripts were highly consistent in FHS (*t*-statistic correlations $r = 0.83$) but less so in RS ($r = 0.48$) (Supplementary Fig. 3).

Table 2—Spearman correlation coefficients between tissue transcript levels from the GTEx Project

	Visceral	Kidney	Liver	Muscle	Pancreas	Blood
Visceral	1.00	0.77	0.66	0.61	0.71	0.50
Kidney		1.00	0.72	0.62	0.78	0.54
Liver			1.00	0.58	0.72	0.39
Muscle				1.00	0.62	0.41
Pancreas					1.00	0.40
Blood						1.00

Transcriptomic Associations in Gene Regions From Prior GWAS

Genetic variants have been previously shown to be associated with fasting glucose and fasting insulin in GWAS (1,6,16). Variants may influence levels of gene expression and circulating transcripts of closely positioned genes (i.e., *cis*-eQTL), so we investigated whether the transcript levels of putative genes assigned to genetic variants identified in GWAS of fasting glucose and fasting insulin (1,6,16) were associated with their respective trait (fasting glucose or fasting insulin). We performed separate analyses in FHS and in the combined data set of RS + InCHIANTI for each trait. Among 22 putative fasting insulin genes and 45 putative fasting glucose genes reported in prior GWAS, we were able to test 11 fasting insulin transcripts and 27 fasting glucose transcripts that were available on both microarray platforms used in our study (1,6,16). Among loci identified by fasting glucose GWAS, circulating *IGF2BP2* transcript levels were most strongly associated with fasting glucose in the current analysis ($P = 1.32 \times 10^{-18}$ in FHS and $P = 3.01 \times 10^{-5}$ in RS +

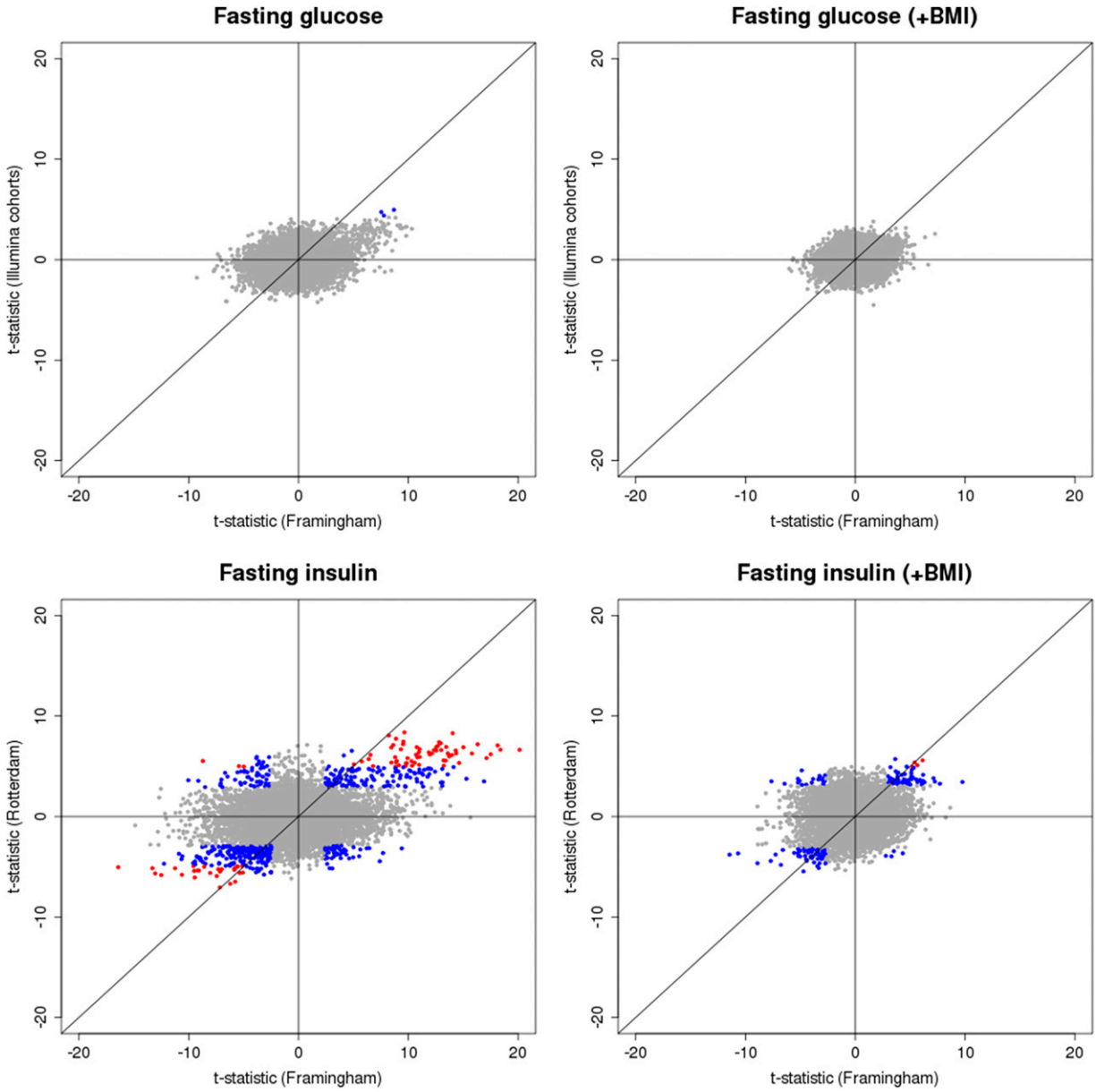

Figure 2—Plots of transcriptomic associations with fasting glucose and fasting insulin in discovery (n = 5,056) vs. replication cohorts (n = 1,318). Significant associations in both discovery and replication cohorts (Bonferroni correction) are highlighted in red. FDR q <0.05 in both discovery and replication cohorts are highlighted in blue. All models adjusted for age, sex, technical covariates, blood counts, and family structure (FHS only).

InCHIANTI) (Table 3); this association remained the same after adjustment for BMI. Among loci identified by fasting insulin GWAS, we found that fasting insulin levels were strongly associated with circulating levels of *TCF7L2* transcripts ($P = 4.66 \times 10^{-19}$) and *ARL15* transcripts ($P = 6.51 \times 10^{-14}$) in FHS, but these findings were not replicated in RS (Table 4).

eQTL Analysis of Glucose- or Insulin-Associated Transcripts and Integration With Genetic Association of Glucose and Insulin Levels

Consistency of evidence across multiple levels of biology (e.g., between SNP, transcript, and phenotypic trait) may help identify transcripts whose levels alter fasting glucose (or insulin) levels rather than the converse. To this end, we identified SNPs (i.e., eQTL) that were significantly associated with transcript levels for the genes identified in our primary TWAS (three transcripts associated with fasting glucose and 433 transcripts with fasting insulin). Using 1000 Genomes–imputed SNPs in FHS, we identified *cis*- and *trans*-eQTL at FDR q <0.05. In total, we identified 43 *cis*-eQTL and 357 *trans*-eQTL for fasting glucose transcripts, 47,567 *cis*-eQTL and 6,695 *trans*-eQTL for fasting insulin transcripts, and 9,815 *cis*-eQTL and 1,017 *trans*-eQTL for BMI-adjusted fasting insulin

Table 3—Transcriptomic associations with fasting glucose levels for genes identified in GWAS of fasting glucose

	FHS (n = 5,056)		RS and InCHIANTI (n = 1,318)	
Gene	t	P value	Z	P value
IGF2BP2	8.84	1.32E-18	4.17	3.01E-05
ARAP1	−5.44	5.53E-08	−2.55	1.08E-02
TOP1	3.94	8.29E-05	0.03	9.78E-01
TCF7L2	3.01	2.66E-03	0.39	7.00E-01
PDX1	−2.18	2.91E-02	−0.10	9.22E-01
OR4S1	−2.18	2.92E-02	0.32	7.45E-01
MTNR1B	2.16	3.12E-02	0.19	8.53E-01
SLC30A8	2.11	3.50E-02	0.04	9.68E-01
P2RX2	−1.78	7.58E-02	1.63	1.04E-01
FOXA2	−1.47	1.41E-01	0.03	9.74E-01
GLIS3	−1.28	2.00E-01	0.42	6.73E-01
CREB3L1	−1.27	2.03E-01	−0.40	6.86E-01
DPYSL5	1.12	2.63E-01	2.09	3.70E-02
KL	1.08	2.80E-01	−0.23	8.21E-01
CDKAL1	0.63	5.29E-01	−1.12	2.62E-01
IKBKAP	0.58	5.65E-01	−1.26	2.07E-01
CRY2	0.55	5.82E-01	1.13	2.59E-01
PCSK1	0.55	5.83E-01	−1.08	2.80E-01
GRB10	0.45	6.52E-01	1.79	7.38E-02
SLC2A2	−0.44	6.59E-01	−0.06	9.52E-01
G6PC2	−0.42	6.77E-01	0.22	8.22E-01
FADS1	0.22	8.24E-01	−2.32	2.02E-02
GCKR	−0.21	8.32E-01	0.58	5.61E-01
ADRA2A	−0.20	8.45E-01	2.05	4.06E-02
GIPR	−0.12	9.08E-01	0.02	9.86E-01
PROX1	0.10	9.19E-01	0.84	4.01E-01
GCK	0.05	9.58E-01	−0.91	3.61E-01

transcripts. These eQTL included many SNPs that were in linkage disequilibrium with one another, so we pruned the list to select independent eQTL from each genomic region.

On the basis of these lists of *cis*- and *trans*-eQTL, we associated each independent eQTL with its respective phenotypic trait (fasting glucose or fasting insulin) in the MAGIC data set. None of the *cis*-eQTL of the significant insulin transcripts were associated with fasting insulin (FDR <0.05), but 92 *trans*-eQTL were significantly associated with fasting insulin. These 92 *trans*-eQTL were associated with three transcripts identified in our fasting insulin TWAS—*RAB36* (lowest q = 0.003 for the eQTL association with fasting insulin), *IGF2BP2* (q = 0.05), and *PLEK2* (q = 0.05) (Table 5 and Supplementary Table 5). Among these three genes, the *RAB36* SNP-to-transcript association was highly significant ($q = 1.28 \times 10^{-35}$), but the transcript association with fasting insulin was relatively weak (*trans*-eQTL for rs1049256 $t = -2.98$; q = 0.04). In contrast, the *IGF2BP2* and *PLEK2* *trans*-eQTL associations were modest (q = 0.002 and q = 0.05, respectively), but the transcript associations with fasting insulin were highly significant ($q = 3.46 \times 10^{-8}$ and $q = 1.76 \times 10^{-7}$, respectively). It is notable that the same SNP on chromosome 6 (rs592423) was associated with *IGF2BP2* and *PLEK2* transcripts in *trans*. *IGF2BP2* and *PLEK2* transcript levels were moderately correlated in FHS (Spearman r = 0.53, $P < 2.2 \times 10^{-16}$). None of the eQTL identified for fasting glucose transcripts were associated with fasting glucose levels in the MAGIC meta-analyses. Despite the moderate correlations of global transcript levels across tissues, none of the cross-tissue correlations were significant for *IGF2BP2* (ENSG00000073792.11) (Supplementary Fig. 4), *PLEK2* (ENSG00000100558.4) (Supplementary Fig. 5), or *RAB36* (ENSG00000100228.8) (Supplementary Fig. 6) in GTEx.

GSEA

Fasting glucose and fasting insulin are complex traits that result from the effect of multiple genetic influences and regulation. As a complementary approach to examining associations for single gene transcripts individually, we tested whether gene sets representing biological pathways were enriched among fasting glucose– or fasting insulin–associated transcripts. To accomplish this task, we conducted GSEA using FHS results selected as the cohort with the largest sample size in our study. Association results from RS and InCHIANTI were underpowered and were not used in this analysis (Supplementary Fig. 7). Systemic lupus erythematosus–related pathway genes were significantly enriched for fasting glucose (with or without BMI adjustment) and fasting insulin GSEA analyses (all q <0.05) (Table 6). Gene sets that were significantly enriched in both fasting glucose and BMI-adjusted fasting insulin levels included tryptophan metabolism, porphyrin metabolism, proteasome, fatty acid metabolism, butanoate metabolism, lysine degradation, and propanoate metabolism. Branched-chain amino

Table 4—Transcriptomic associations with fasting insulin levels for genes identified in GWAS of fasting insulin

	FHS (n = 5,056)		RS (n = 723)	
Gene	t	P value	t	P value
TCF7L2	8.96	4.66E-19	1.31	1.92E-01
ARL15	−7.52	6.51E-14	−0.84	3.99E-01
UHRF1BP1	5.14	2.89E-07	−0.29	7.71E-01
PDGFC	4.49	7.28E-06	−0.29	7.68E-01
LYPLAL1	−3.85	1.18E-04	2.18	2.97E-02
FTO	2.19	2.86E-02	−0.46	6.45E-01
GRB14	−1.41	1.58E-01	0.14	8.88E-01
PEPD	−0.84	4.01E-01	−2.93	3.45E-03
RSPO3	0.82	4.10E-01	−1.23	2.19E-01
HIP1	−0.60	5.48E-01	−1.70	8.99E-02
GCKR	−0.53	5.98E-01	1.21	2.25E-01

acid (BCAA) degradation for valine, leucine, and isoleucine were identified as significantly enriched pathways that emerged from the GSEA analyses of BMI-adjusted fasting insulin–associated transcripts (q = 0.03).

DISCUSSION

We examined whole blood mRNA transcript associations with fasting glucose and fasting insulin levels in individuals without diabetes from three independent studies to identify genes that may be involved in the regulation of glycemic traits. We found 433 transcripts associated with fasting insulin and 3 with fasting glucose that were significant (FDR q <0.05) in independent sets of data. In contrast to GWAS of fasting glucose and insulin levels (1), we found higher number of significant associations with fasting insulin than with fasting glucose. We also identified genetic variants that were associated with our insulin/glucose-associated transcripts (eQTL), most notably a *trans*-eQTL (rs592423) that was associated with circulating transcript levels of *IGF2BP2*. The same eQTL was also associated with fasting insulin in a large independent sample from MAGIC. Thus, the convergence of genomic and transcriptomic associations suggest a role for *IGF2BP2* and its circulating transcript in the regulation of fasting insulin.

Although the correlations between transcript levels from blood and other tissues in the GTEx Project were generally moderate, our study was able to provide sample size estimates for future studies that may choose to investigate nonblood tissues. Although these studies will prove difficult logistically and ethically in healthy individuals, these estimates serve as a step forward in bringing forth such studies.

We found that higher circulating transcript levels of *IGF2BP2* were associated with higher fasting insulin. In animal models, overexpression of *IGF2BP2* causes β-cell damage in islets (17). Animal models also showed that *IGF2BP2* knockout mice are leaner than controls (18). In our human study, we found that the association between fasting insulin and *IGF2BP2* transcript levels remained significant after adjustment for BMI, suggesting that obesity may not be the predominant driver of this observation. SNPs rs4402960 and rs1470579 located in *IGF2BP2* (on chromosome 2) have been associated with type 2 diabetes in multiple studies (19–22), but our *cis*-eQTL analyses did not highlight these two variants.

Surprisingly, our eQTL analyses revealed that rs592423 (chromosome 6) was associated with *IGF2BP2* (chromosome 3) transcript levels in *trans*. This eQTL association was reported previously in an independent eQTL analysis of three large, population-based cohorts (23). The same variant rs592423 (chromosome 6) was also associated in *trans* with transcript levels of *PLEK2* (chromosome 14), another transcript that we found significantly associated with insulin levels in both FHS and RS. Proxy SNPs with rs592423 (i.e., rs628751 at r^2 = 0.90 and rs643381 at r^2 = 0.74) were associated with *PLEK2* transcript levels

Table 5—Genomic, transcriptomic, and fasting insulin associations that reached FDR q <0.05 in all three comparisons

SNP	SNP Chr	Allele	MAF§	Gene	Gene Chr	SNP-transcript*			SNP-insulin†		Transcript-insulin‡	
						t	R^2	FDR	t	FDR	t	FDR
rs1049256	3	C	0.46	*RAB36*	22	14.44	0.04	1.28×10^{-35}	3.92	0.003	−2.98	0.04
rs592423	6	A	0.45	*IGF2BP2*	3	−7.20	0.01	0.002	2.72	0.05	6.60	3.46×10^{-8}
rs592423	6	A	0.45	*PLEK2*	14	−6.70	0.008	0.05	2.72	0.05	6.29	1.76×10^{-7}

For each significant transcript, the SNP with the smallest P value for fasting insulin levels is shown. Chr, chromosome; MAF, minor allele frequency. *SNP-transcript associations based on FHS eQTL (n = 5,056). †SNP-insulin associations based on MAGIC GWAS results (n = 50,823; FHS, RS, and InCHIANTI results not included in meta-analysis). ‡Transcript-insulin associations based on the FHS and RS meta-analysis. §Based on HapMap CEU.

Table 6—GSEA of transcriptomic associations for fasting glucose and fasting insulin in FHS Offspring cohort

KEGG	Enrichment score	FDR *q* value
Fasting glucose		
Systemic lupus erythematosus	−2.15	0.001
Porphyrin and chlorophyll metabolism	2.05	0.007
Fatty acid metabolism	1.91	0.010
Tryptophan metabolism	1.93	0.012
Butanoate metabolism	1.88	0.012
Lysine degradation	1.86	0.012
Proteasome	1.95	0.013
Propanoate metabolism	1.82	0.018
Fasting glucose (BMI-adjusted)		
Systemic lupus erythematosus	−1.90	0.015
Asthma	−2.02	0.017
Intestinal immune network for IgA production	−1.92	0.018
Fasting insulin		
Systemic lupus erythematosus	−1.97	0.008
Fasting insulin (BMI-adjusted)		
Tryptophan metabolism	2.12	0.001
Proteasome	2.01	0.003
Lysine degradation	1.95	0.007
Aminoacyl tRNA biosynthesis	1.92	0.008
Alanine aspartate and glutamate metabolism	1.86	0.010
Fatty acid metabolism	1.87	0.011
Citrate cycle TCA cycle	1.88	0.012
Porphyrin and chlorophyll metabolism	1.81	0.013
Glyoxylate and dicarboxylate metabolism	1.82	0.014
Butanoate metabolism	1.81	0.015
Valine leucine and isoleucine degradation	1.73	0.030
Propanoate metabolism	1.72	0.031
Cysteine and methionine metabolism	1.70	0.035
Primary immunodeficiency	1.69	0.035
Huntington disease	1.66	0.042
RNA polymerase	1.64	0.048

in a prior eQTL meta-analysis that included RS and InCHIANTI samples (24). Little is known currently about *PLEK2*, but it may play an integral role in erythropoiesis (25). Although the *PLEK2* association may be suggestive of a tissue-specific artifact, a biological role for *IGF2BP2* cannot be ruled out. We found that the A allele at rs592423 was nominally associated with higher fasting insulin (β = 0.009, P = 0.004) in the MAGIC GWAS meta-analysis. The A allele of rs592423 has also been previously associated with adiponectin levels in GWAS (MANTRA Bayes Factor = 6.5; n = 37,430) (26). Our observation that the same eQTL (rs592423) was associated with two different insulin-associated transcripts (*IGF2BP2* and *PLEK2*) in *trans* suggests that this genetic variant may influence a transcription factor that may affect expression of multiple genes. The rs592423 variant is located 145 kb from the closest coding sequence, *CITED2*, which is a transcriptional coactivator of the peroxisome proliferative activated receptor γ coactivator 1α (PGC-1α) and peroxisome proliferator–activated receptor α pathways (27,28). Furthermore, *CITED2* expression was downregulated in response to insulin infusion in vivo, which may be explained by negative feedback due to the elevated insulin levels (29). In addition, located within 50 kb of rs592423 is a long noncoding RNA (*LOC645434*) that is expressed highly in breast tissue and moderately in adipose and brain tissues. Future studies are needed to test whether *CITED2* or *LOC645434* influences levels of *IGF2BP2*, *PLEK2*, insulin, or adiponectin, as well as elucidating their specific roles in regulating fasting insulin levels. Studies of model organisms may be used to provide further insights into the mechanisms between rs592423 (or its equivalent in other species) and expression of *IGF2BP2*, *PLEK2*, and *CITED2*. Inbred knockouts of the *CITED2* gene ortholog may elucidate its role in relation to *IGF2BP2* and/or *PLEK2* expression. Outbred animal models may also be used to confirm the utility of our approach of genetic variation altering gene expression in more controlled settings.

In addition to *IGF2BP2*, our TWAS findings identified several known genes associated with various metabolic traits. Notable associations with fasting insulin transcripts included genes involved in insulin signaling (*IRS2*, *FOXO4*) and adipocyte biology or adipokine regulation (*ITLN1*, *PID1*, *ADIPOR1*). *ITLN1* plays a role in insulin-stimulated glucose uptake in adipocytes. *PID1* participates in adipocyte proliferation. *ADIPOR1* is a receptor for adiponectin. The eQTL for *ITLN1* (rs4656953) was nominally associated with fasting insulin (adjusted for BMI) in the MAGIC GWAS meta-analysis (β = 0.006, P = 0.04), implying genetic regulation of *ITLN1* that also influences insulin regulation. Our TWAS also identified other transcripts of genes near loci that were previously identified in GWAS of type 2 diabetes (*IGF2BP2*, *JAZF1*) (21), HbA_{1c} levels (*PIEZO1*) (30), and lipid profiles (*MYLIP*, *AKT1*) (31). Significant cross-replicated insulin transcripts associations also identified genes that are known BCAA transporters (*SLC16A3*, *SLC43A2*, *SLC1A5*, *SLC7A5*), whose related metabolites were implicated with type 2 diabetes (32). These confirmatory findings serve as a positive control for the ability for transcriptomic data in whole blood to screen for trait-related genes.

Among the gene sets that were enriched in BMI-adjusted insulin associations was degradation of valine, leucine, and isoleucine, three major BCAAs. Increased levels of BCAA are associated with insulin resistance and higher risk of type 2 diabetes incidence (32–34). From the list of individual transcripts that were significant in both FHS and RS for fasting insulin, *SLC16A3*, *SLC43A2*, *SLC1A5*, and *SLC7A5* are known BCAA transporters. We also found the significant insulin-associated transcript *SLC22A4*, which was previously identified in a GWAS of the ratio of valine to isovalerylcarnitine (35) This GWAS also implicated a variant near *SLC7A5* (BCAA transporter) in association with plasma levels of kynurenine, a metabolite of tryptophan. The tryptophan metabolism gene set emerged from our

GSEA of both glucose- and BMI-adjusted insulin-associated transcripts. BCAA and tryptophan pathways have strong biological evidence for a role in insulin resistance. In muscle, the mTOR growth pathway receives signals via three pathways: growth factors including insulin, glucose levels via AMPK, and BCAAs via mTORC1 (36). This is of particular relevance due to the link between type 2 diabetes and cancer (36) and the role of skeletal muscle in insulin resistance and the development of type 2 diabetes (37). Overall, our results support the notion that BCAA are likely implicated in insulin resistance pathways. Functional studies and intervention trials are still needed to clarify their direction of effect and exact biological role.

Our findings may lead to identification of novel genes implicated in glucose/insulin regulation. For example, we found that multiple eQTL on chromosome 3 were associated with both *RAB36* transcript levels and with fasting insulin in the MAGIC GWAS meta-analysis. *RAB36* is a member RAS oncogene family and is suspected to be involved in vesicular traffic (38), which is a crucial step in insulin secretion.

Our main limitation is our inability to infer direction of effect of observed transcript associations with glycemic traits, given the cross-sectional nature of our study. Although further studies are needed to confirm the exact biological roles of our transcriptomic findings in relation to glucose/insulin regulation, anchoring transcriptomic results to eQTL and eQTL-to-trait may provide some indication of whether the transcriptomic signature is under genetic control or is a response to levels of the glycemic trait. We hypothesized that transcript levels under genetic control would have an eQTL that is associated both with a glucose/insulin-associated transcript and with the respective trait. A causal inference test for *IGF2BP2* could not be conducted because the identified eQTL was a weak instrument (based on a small R^2) and thus would have biased results in a formal causal test (39).

Additional limitations should be noted when interpreting the results of our study. First, extrapolating transcriptomic signatures from one tissue to another is difficult. Although tissues may share a portion of their transcriptomic signatures (40,41), it remains unclear how to distinguish the shared and unshared signatures. These tissue-specific transcript levels may partly account for the lack of associations in our whole-blood transcriptomic data for many known fasting glucose and fasting insulin biosynthesis and metabolism genes. Next, we only examined mRNA transcript levels present in the Affymetrix and Illumina gene expression microarrays, but a more comprehensive quantification of the transcriptome (e.g., RNA sequencing) and splicing isoforms may be able to provide additional insights (42). Moreover, array-specific differences might have caused us to miss significant and reproducible transcript associations with the glycemic traits. The Illumina array probes hybridize to the 3′ ends of mRNA transcripts, and the Affymetrix Exon array probes hybridize to the exons. Thus, the Affymetrix array may detect total mRNA transcripts (i.e., multiple isoforms), while the Illumina array detects only a subset of isoforms. Last, a large number of comparisons were conducted study-wide. To reduce the risk of false positive findings, we used several highly stringent approaches. In addition to computing FDRs and validation in independent samples, we further required transcript-to-phenotype associations to reach this level of significance in both discovery and replication samples. Furthermore, we used complementary approaches using genetic variation to provide further support for our main findings.

In conclusion, using three well-characterized, population-based cohorts, we identified hundreds of mRNA transcripts whose levels in whole blood were associated with fasting glucose or fasting insulin levels. We identified genes and pathways that have been discovered using other biological assays, supporting the use of circulating mRNA transcript levels as a discovery tool. For instance, our GSEA identified BCAA regulation, which was primarily related to type 2 diabetes through metabolomic studies. Furthermore, we identified an eQTL associated with fasting insulin and acting as a *trans*-eQTL for the insulin-associated transcript *IGF2BP2*, one of the most reproducible type 2 diabetes and glycemic trait GWAS findings. Taken as a whole, transcriptomics, particularly when combined with genomic data, was a tractable modality for highlighting important genes and pathways that may be involved in the regulation of fasting insulin levels.

Acknowledgments. The authors thank all study participants and staff from the RS, the participating general practitioners, and the pharmacists. The authors thank Julie McGrath of the Harvard Pilgrim Health Care Institute for help in the preparation of the manuscript.

Funding. This work was funded by the European Commission (HEALTH-F2-2008-201865, GEFOS, HEALTH-F2-2008 35627, and TREAT-OA 200800), the Netherlands Organisation for Scientific Research (175.010.2005.011 and 911-03-012), the Research Institute for Diseases in the Elderly (014-93-015, RIDE2), the Netherlands Genomics Initiative/Netherlands Consortium for Healthy Ageing (050-060-810), and Netherlands Organisation for Scientific Research VIDI grant (917103521). The RS is funded by Erasmus University Medical Center Rotterdam and Erasmus University; Netherlands Organisation for Health Research and Development; the Research Institute for Diseases in the Elderly; the Netherlands Ministry of Education, Culture and Science; the Netherlands Ministry for Health, Welfare and Sports; the European Commission (DG XII); and the Municipality of Rotterdam.

The FHS is funded by the National Institutes of Health (NIH) contract N01-HC-25195. The laboratory work for this investigation was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, NIH. The analytical component of this project was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, and the Center for Information Technology, NIH. This study utilized the high-performance computational capabilities of the Biowulf Linux cluster (https://hpc.nih.gov/) and Helix Systems (http://helix.nih.gov) at the NIH. L.M.P., J.D., J.B.M., and E.D.K. were partially supported by NIH grant R01 DK078616. J.B.M. is partly supported by NIH grant K24 DK080140.

This research was supported in part by the Intramural Research Program, National Institute on Aging, NIH (Z01-AG000947 and Z01-AG000185). The InCHIANTI study baseline (1998–2000) was supported as a targeted project (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the National Institute on Aging (contracts 263 MD 9164 and 263 MD 821336). L.W.H. was funded by a grant from the U.K. Medical Research Council (MR/J006777/1).

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. B.H.C., M.-F.H., M.J.P., and L.C.P. researched the data and wrote the manuscript. J.D.H. and J.D. researched the data, contributed to discussion, and reviewed and edited the manuscript. L.M.P., J.H., R.J., A.D.J., P.J.M., D.V.R., S.Y., and D.M. researched the data and reviewed and edited the manuscript. L.W.H., J.C.F., and E.D.K. contributed to discussion and reviewed and edited the manuscript. C.S.F., S.B., A.D., D.G.H., A.H., A.B.S., A.G.U., and J.B.J.v.M. acquired the data and reviewed and edited the manuscript. MAGIC Investigators acquired the data and researched the data. D.L. acquired the data, researched the data, and reviewed and edited the manuscript. L.F. contributed to the discussion, acquired the data, and reviewed and edited the manuscript. J.B.M. provided support and supervision, contributed to discussion, and reviewed and edited the manuscript. B.H.C. and M.-F.H. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet 2010;42:105–116
2. Rasmussen-Torvik LJ, Guo X, Bowden DW, et al. Fasting glucose GWAS candidate region analysis across ethnic groups in the Multiethnic Study of Atherosclerosis (MESA). Genet Epidemiol 2012;36:384–391
3. Hayes MG, Urbanek M, Hivert MF, et al. Identification of HKDC1 and BACE2 as genes influencing glycemic traits during pregnancy through genome-wide association studies. Diabetes 2013;62:3282–3291
4. Chen G, Bentley A, Adeyemo A, et al. Genome-wide association study identifies novel loci association with fasting insulin and insulin resistance in African Americans. Hum Mol Genet 2012;21:4530–4536
5. Kim YJ, Go MJ, Hu C, et al. Large-scale genome-wide association studies in East Asians identify new genetic loci influencing metabolic traits. Nat Genet 2011;43:990–995
6. Scott RA, Lagou V, Welch RP, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat Genet 2012;44:991–1005
7. Wu X, Wang J, Cui X, et al. The effect of insulin on expression of genes and biochemical pathways in human skeletal muscle. Endocrine 2007;31:5–17
8. Rome S, Clement K, Rabasa-Lhoret R, et al. Microarray profiling of human skeletal muscle reveals that insulin regulates approximately 800 genes during a hyperinsulinemic clamp. J Biol Chem 2003;278:18063–18068
9. Lonsdale J, Thomas J, Salvatore M, et al. The Genotype-Tissue Expression (GTEx) Project. Nat Genet 2013;45:580–585
10. Pritchard JK, Przeworski M. Linkage disequilibrium in humans: models and data. Am J Hum Genet 2001;69:1–14
11. Peters MJ, Joehanes R, Pilling LC, et al. The transcriptional landscape of age in human peripheral blood. Nat Commun 2015;6:8570
12. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249–264
13. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 2003;100:9440–9445
14. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–15550
15. Prokopenko I, Langenberg C, Florez JC, et al. Variants in MTNR1B influence fasting glucose levels. Nat Genet 2009;41:77–81
16. Manning AK, Hivert MF, Scott RA, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. Nat Genet 2012;44:659–669
17. Casellas A, Mallol C, Salavert A, et al. Insulin-like growth factor 2 overexpression induces beta-cell dysfunction and increases beta-cell susceptibility to damage. J Biol Chem 2015;290:16772–16785
18. Dai N, Zhao L, Wrighting D, et al. IGF2BP2/IMP2-deficient mice resist obesity through enhanced translation of Ucp1 mRNA and other mRNAs encoding mitochondrial proteins. Cell Metab 2015;21:609–621
19. 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, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nat Genet 2014;46:234–244
20. Saxena R, Saleheen D, Been LF, et al. Genome-wide association study identifies a novel locus contributing to type 2 diabetes susceptibility in Sikhs of Punjabi origin from India. Diabetes 2013;62:1746–1755
21. Voight BF, Scott LJ, Steinthorsdottir V, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010;42:579–589
22. Unoki H, Takahashi A, Kawaguchi T, et al. SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. Nat Genet 2008;40:1098–1102
23. Schramm K, Marzi C, Schurmann C, et al. Mapping the genetic architecture of gene regulation in whole blood. PLoS One 2014;9:e93844
24. Westra HJ, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 2013;45:1238–1243
25. Zhao B, Keerthivasan G, Mei Y, et al. Targeted shRNA screening identified critical roles of pleckstrin-2 in erythropoiesis. Haematologica 2014;99:1157–1167
26. Dastani Z, Hivert MF, Timpson N, et al. 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 2012;8:e1002607
27. Sakai M, Matsumoto M, Tujimura T, et al. CITED2 links hormonal signaling to PGC-1alpha acetylation in the regulation of gluconeogenesis. Nat Med 2012; 18:612–617
28. Tien ES, Davis JW, Vanden Heuvel JP. Identification of the CREB-binding protein/p300-interacting protein CITED2 as a peroxisome proliferator-activated receptor alpha coregulator. J Biol Chem 2004;279:24053–24063
29. Coletta DK, Balas B, Chavez AO, et al. Effect of acute physiological hyperinsulinemia on gene expression in human skeletal muscle in vivo. Am J Physiol Endocrinol Metab 2008;294:E910–E917
30. Chen P, Takeuchi F, Lee JY, et al. Multiple nonglycemic genomic loci are newly associated with blood level of glycated hemoglobin in East Asians. Diabetes 2014;63:2551–2562
31. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013;45:1274–1283
32. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. Nat Med 2011;17:448–453
33. Wang TJ, Ngo D, Psychogios N, et al. 2-Aminoadipic acid is a biomarker for diabetes risk. J Clin Invest 2013;123:4309–4317
34. Newgard CB, An J, Bain JR, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. Cell Metab 2009;9:311–326
35. Shin SY, Fauman EB, Petersen AK, et al. An atlas of genetic influences on human blood metabolites. Nat Genet 2014;46:543–550
36. O'Connell TM. The complex role of branched chain amino acids in diabetes and cancer. Metabolites 2013;3:931–945
37. DeFronzo RA, Tripathy D. Skeletal muscle insulin resistance is the primary defect in type 2 diabetes. Diabetes Care 2009;32(Suppl. 2):S157–S163
38. Grosshans BL, Ortiz D, Novick P. Rabs and their effectors: achieving specificity in membrane traffic. Proc Natl Acad Sci U S A 2006;103:11821–11827
39. Burgess S, Thompson SG. Avoiding bias from weak instruments in Mendelian randomization studies. Int J Epidemiol 2011;40:755–764
40. Grundberg E, Small KS, Hedman AK, et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat Genet 2012;44:1084–1089
41. Nica AC, Parts L, Glass D, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. PLoS Genet 2011;7: e1002003
42. Zhang X, Joehanes R, Chen BH, et al. Identification of common genetic variants controlling transcript isoform variation in human whole blood. Nat Genet 2015;47:345–352