OPEN ACCESS

University of BRISTOL

Wang, Q., Holmes, M., Davey Smith, G., & Ala-Korpela, M. (2017). Genetic support for a causal role of insulin resistance on circulating branched-chain amino acids and inflammation. *Diabetes Care*, *40*(12), 1779-1786. https://doi.org/10.2337/dc17-1642

Peer reviewed version

Link to published version (if available):
10.2337/dc17-1642

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

Genetic Support for a Causal Role of Insulin Resistance on Circulating Branched-Chain Amino Acids and Inflammation

Journal:	*Diabetes Care*
Manuscript ID	DC17-1642.R1
Manuscript Type:	Original Article: Cardiovascular and Metabolic Risk
Date Submitted by the Author:	10-Sep-2017
Complete List of Authors:	Wang, Qin; University of Oulu Holmes, Michael; University of Oxford, Nuffield Department of Population Health Davey Smith, George; University of Bristol Ala-Korpela, Mika; Oulun Yliopisto Laaketieteellisen tiedekunta, Computational Medicine

SCHOLARONE™
Manuscripts

Genetic Support for a Causal Role of Insulin Resistance on Circulating Branched-Chain Amino Acids and Inflammation

Qin Wang, PhD[1]*, Michael V Holmes, MD, PhD[2,3,4,5]*,
George Davey Smith, MD, DSc[5,6], Mika Ala-Korpela, PhD[1,5,6,7]**

1. Computational Medicine, Faculty of Medicine, University of Oulu and Biocenter Oulu, Oulu, Finland;
2. Medical Research Council Population Health Research Unit at the University of Oxford, Oxford, UK;
3. Clinical Trial Service Unit & Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK;
4. National Institute for Health Research, Oxford Biomedical Research Centre, Oxford University Hospital, Oxford, UK;
5. Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK;
6. School of Social and Community Medicine, University of Bristol, Bristol, UK;
7. NMR Metabolomics Laboratory, School of Pharmacy, University of Eastern Finland, Kuopio, Finland.

The authors wish to note equal contributions for Dr Wang and Dr Holmes

Short running title: Causal Role of Insulin Resistance on BCAAs
Subject codes: Biomarkers, Mechanisms, Risk Factors, Type 2 Diabetes, Genetics

Word counts: Abstract 248 (/ 250), Main text 3,942 (/ 4,000)
References 40/40, Figures 4 (/ 4); Online Supplemental Material (16 pages)

** Correspondence:
Professor Mika Ala-Korpela
Computational Medicine, Faculty of Medicine
University of Oulu, P.O.Box 5000
FI-90014 University of Oulu, Oulu, Finland
E-mail: mika.ala-korpela@oulu.fi
Mobile: +358 40 19 77 657

Abstract

Objective: Insulin resistance has deleterious effects on cardiometabolic disease. We used Mendelian randomization analyses to clarify the causal relationships of insulin resistance on circulating blood-based metabolites to shed light on potential mediators of the insulin resistance to cardiometabolic disease relationship.

Research Design and Methods: We used 53 single nucleotide polymorphisms associated with insulin resistance from a recent genome-wide association study to explore their effects on circulating lipids and metabolites. We used published summary-level data from two genome-wide association studies (GWASs) of European individuals; data on the exposure (insulin resistance) were obtained from meta-GWASs of 188,577 individuals and data on the outcomes (58 metabolic measures assessed by NMR) were taken from a GWAS of 24,925 individuals.

Results: One standard deviation (SD) genetically elevated insulin resistance (equivalent to 55% higher geometric mean of fasting insulin, 0.89 mmol/L higher triglycerides and 0.46 mmol/L lower HDL-C) was associated with higher concentrations of all branched-chain amino acids, isoleucine (0.56 SD; 95%CI: 0.43, 0.70), leucine (0.42 SD; 95%CI: 0.28, 0.55) and valine (0.26 SD; 95%CI: 0.12, 0.39) as well as with higher glycoprotein acetyls (an inflammation marker; 0.47 SD; 95%CI: 0.32, 0.62) (P<0.0003 for each). Results were broadly consistent when using multiple sensitivity analyses to account for potential genetic pleiotropy.

Conclusions: We provide robust evidence that insulin resistance causally impacts on each individual branched-chain amino acid and inflammation. Taken together with existing studies, this implies that branched-chain amino acid metabolism lies on a causal pathway from adiposity and insulin resistance to type 2 diabetes.

Non-standard Abbreviations and Acronyms: BCAA, branched-chain amino acid; IR, insulin resistance; MR, Mendelian randomization; IVW, inverse variance weighted; GlycA, glycoprotein acetyls; BCKD, branched-chain α-ketoacid dehydrogenase.

The obesity pandemic is a public health crisis leading to a dramatic surge in the incidence of type 2 diabetes mellitus (T2DM) and related diseases (e.g., cardiovascular diseases) (1). Adiposity, particularly visceral adiposity (2), is associated with insulin resistance (IR) and subsequent T2DM. Recent genetic studies employing the Mendelian randomization approach have shown adiposity traits (such as general adiposity, indexed by body mass index, and central adiposity, indexed by waist-to-hip ratio) to show causal relationships with blood pressure, lipids, coronary heart disease, stroke and diabetes (3-6). Furthermore, such studies have demonstrated that adiposity traits causally impact on insulin resistance (3,4,6). Insulin resistance is the clinical state of a reduced sensitivity to insulin, typically manifested as elevated levels of fasting insulin and often accompanied with higher levels of circulating triglycerides and lower levels of high-density lipoprotein cholesterol (HDL-C) (7).

Exploring the molecular mechanism by which IR leads to T2DM may help to identify biomarkers that could mediate the relationship, and provide novel opportunities for disease prevention. Recent studies have suggested that branched-chain amino acids (BCAAs) might play a role in the development of T2DM. Prospective observational studies show that higher levels of circulating BCAAs are positively associated with markers of insulin resistance (8) and risk of incident T2DM (9,10). Recent genetic studies have also implicated the metabolism of BCAAs in the development of diabetes (11).

Insulin resistance is a complex trait, which can be assessed by different metrics, including clamp/insulin suppression test (gold standard), insulin sensitivity test (based on OGTT), HOMA-IR and fasting insulin. The GENESIS consortium has published a GWAS of insulin sensitivity measured by clamp/insulin suppression test in a modest number of subjects (N = 5624) (12). However, the statistical power limits the findings of this study. Other metrics which can be more easily measured, such as fasting insulin or HOMA-IR, are often used in large-scale genetic and epidemiological studies. In the GWAS of fasting insulin conducted by Scott et al (up to 108,557 individuals) (13), they also tested the associations of insulin-associated SNPs with lipid traits. They found that majority of the insulin-associated SNPs were associated with HDL-C and/or TG and this pattern was not observed for those SNPs associated with fasting glucose or 2hour glucose. Subsequently, a genetic instrument was built for insulin resistance that used the 19 SNPs associated with fasting insulin, and restricted the instrument to those SNPs that were also associated with triglycerides and HDL-C (14). This instrument was recently adopted by Mahendran et al, and the results suggest that insulin resistance might be causal for circulating concentrations of BCAAs (15). More recently, Lotta et al considerably expanded the set of SNPs associated with three components of insulin resistance (higher fasting insulin, higher triglycerides and lower HDL-cholesterol), identifying 53 such SNPs and found that the SNPs in aggregate also associated with risks of CHD and T2DM (7).

Here we aim to: 1) assess the causal effects of insulin resistance employing these 53 SNPs recently identified from across the genome that associated with higher fasting

insulin, higher triglycerides and lower HDL-C (7); 2) use multiple instruments and multiple sensitivity analysis as a means to detect and correct for potential genetic pleiotropy in order to ensure reliable findings; 3) expand the outcome measures from BCAAs to a comprehensive panel of amino acids (including alanine, glutamine, tyrosine and phenylalanine), lipoprotein subclasses, fatty acids, glycolysis-related measures and one inflammatory marker, which are established or emerging biomarkers for T2DM and cardiovascular diseases; 4) provide an overview of the potential causal pathways and mediator roles that insulin resistance places in the underlying association of adiposity with T2DM, by incorporating our findings into multiple strands of genetic evidence.

Research Design and Methods

We used published summary-level data from two GWA studies of European individuals (7,16). Data on the exposure (insulin resistance) were obtained from a meta-GWASs (meta-analysis of genome-wide association studies) of up to 188,577 individuals (7) and data on the outcome (58 circulating metabolic measures) were taken from a GWAS of up to 24,925 individuals (16). Characteristics of these GWASs are reported in Supplemental Tables 1 & 2.

Generation of Genetic Instruments

We used the 53 SNPs associated with an insulin resistance phenotype from Lotta et al (7). In brief, Lotta et al. conducted a meta-GWAS to identify SNPs that associated with an insulin resistance phenotype of: (i) higher fasting insulin adjusted for BMI; (ii) higher triglycerides; and, (iii) lower HDL-C at P<0.005 for each trait. The combined association with the triad of phenotypes have been proposed as a means to characterising the genetic architecture of insulin resistance (7). This meta-GWAS identified 53 SNPs, of which a subset of 25 loci had been previously associated with triglycerides or HDL-C at genome-wide significance, whereas the remaining 28 had not.

We used the 53 SNPs to generate a genetic instrument for IR. To conduct the Mendelian randomization (MR) analyses (17), we needed to obtain the association of SNPs with the exposure (insulin resistance) and also the associations with outcomes (metabolic measures). Lotta et al. (7) did not provide beta or SE for the associations of individual SNPs with the insulin resistance phenotype. To generate our own SNP to exposure estimate, we took the absolute value of the standardized beta coefficient for each of the 53 SNP associations with the individual components of the composite IR phenotype (i.e. fasting insulin adjusted for BMI, triglycerides and HDL-C) and meta-analysed the estimates together using a fixed-effect inverse-variance weighted method (data sources provided in Supplemental Table 3). We used this meta-analysed value as the SNP-exposure estimate for the summary-level MR analyses. Supplemental Fig. 1 shows the associations of the 53 individual SNPs for our insulin

resistance trait with the three individual components. Most of the SNPs fell in a straight line (with a slope equal to 1), suggesting a similar contribution of the three traits to the 'composite' insulin resistance phenotype with the exception of rs1011685 (near *LPL*), which had a much weaker effect on insulin adjusted for BMI. We therefore conducted sensitivity analyses in which rs1011685 was excluded from the instrument.

Two-Sample Mendelian Randomization Analysis

We used data from Kettunen et al. (16) to obtain SNP-associations with metabolic measures. Summary data for 58 measures were used in this study, including 14 lipoprotein subclasses, 3 lipoprotein size measures, 9 total lipids, ApoA1, ApoB, 10 fatty acids related measures, 9 amino acids, one inflammation marker - glycoprotein acetyls and several other measures. These metabolic measures were quantified by a high-throughput NMR metabolomics platform using primarily fasting serum samples with an approximately 1:1 male-to-female ratio and age span of 20-60 years (Supplemental Table 2). We used a conventional inverse variance weighted (IVW) MR analysis, in which the SNP to outcome estimate is regressed on the SNP to exposure, with the y-axis intercept forced through the origin. The data used for the MR analyses are presented in Supplemental Tables 3 & 4.

Sensitivity Analyses

As the conventional IVW MR approach can be vulnerable to unbalanced horizontal pleiotropy (18), we conducted MR-Egger, weighted median and weighted mode-based MR analyses, which allow relaxation of some of the instrumental variable assumptions. The characteristics of these different MR methods are summarized in Supplemental Table 5. Overall, use of several MR methods that each makes different assumptions on the amount and type of genetic confounding is a useful strategy to assess the robustness of findings to potential violations of the instrumental variable assumptions (19).

In addition to the 53 SNP instrument, we: (i) removed the rs1011685 (near *LPL*), which, as described above, did not show consistent associations across individual phenotypes of insulin resistance; (ii) used the 28 SNPs reported in Lotta et al. (7) that were not in loci previously associated with triglycerides or HDL-C at genome-wide significance, and (iii) used 12 SNPs associated with fasting insulin (BMI adjusted) reported by MAGIC consortium (13). As fasting insulin is another marker of insulin resistance, consistent results of the primary analysis and sensitivity analysis (iii) would provide further confidence in concluding the causal role of insulin resistance on the circulating metabolites. Further, sensitivity analysis (ii & iii) are helpful in assessing the contribution of primarily lipid-associated SNPs on the casual effect estimates. In addition, to quantify whether the genetic instruments for IR associated with BMI, we regressed the associations of SNPs with insulin resistance against the

associations of SNPs with BMI using summary-level data from GIANT consortium (20). A final step was to remove SNPs from the 53 SNP instrument that individually associated with BMI at P<0.001 using GIANT summary statistics (20) in order to clarify whether this materially altered the MR effect estimates.

Genetic effect estimates are presented as standard deviation (SD) differences in metabolite concentrations per 1-SD genetically higher insulin resistance. To gain insight into the association of the genetic instrument with its individual components, we quantified the association of a 1-SD higher genetically higher insulin resistance on fasting insulin from the MAGIC consortium (13), and the blood lipids HDL-C and triglycerides from the Global Lipids Genetics Consortium (21). We used a two-sided P<0.001 (= 0.05/58; multiple testing correction) to denote evidence of an association.

All analyses were conducted in R.

Results

The associations of the 53 SNPs with each of the metabolic measures are shown in **Figure 1**. As expected, all the SNPs associated with higher BMI-adjusted insulin and triglycerides and lower HDL-C. There were also general trends of the SNPs to associate with higher VLDL and lower HDL traits.

Causal effect estimates of insulin resistance, proxied by the 53-SNP instrument, on the individual metabolic traits are illustrated in **Figure 2**. The association magnitudes (betas), standard errors and corresponding P-values are reported in Supplemental Table 6. A one-SD genetically higher insulin resistance was associated with 55% (95%CI: 50%, 60%) higher fasting insulin adjusted for BMI, 0.89 mmol/l (95%CI: 0.85, 0.93) higher triglycerides and 0.46 mmol/l (95%CI: 0.44, 0.48) lower HDL-C. In addition, there were clear associations of the genetic instrument for insulin resistance with higher concentrations of all VLDL subclasses with more moderate associations with IDL and LDL subclasses. In contrast, the associations were inverse for most HDL subclasses. The genetic instrument was positively associated with VLDL and negatively with HDL particle size. These findings corroborate the characteristics of the instrument as devised by Lotta et al (7). Similarly, we identified positive associations of the genetic instrument with circulating fatty acids, including monounsaturated and omega-3 fatty acids. Interestingly, the genetic instrument only weakly associated with an increase in NMR-quantified glucose, a finding in keeping with the observation by Lotta et al. using the MAGIC consortium data (7). As reported in the original study (7), the genetic instruments were negatively associated with BMI (Supplemental Table 7).

We identified strong positive associations of genetically higher insulin resistance with the BCAAs, isoleucine, leucine and valine. These estimates correspond to a 0.56 SD (95%CI: 0.43, 0.70) higher isoleucine, 0.42 SD (95%CI: 0.28, 0.55) higher leucine and 0.26 SD (95%CI: 0.12, 0.39) higher valine per 1-SD higher insulin resistance. Weaker associations were noticed with the other amino acids. In addition, genetically

higher insulin resistance was positively associated with glycoprotein acetyls, an inflammation marker (0.47 SD; 95%CI: 0.32, 0.62).

Sensitivity Analyses

Most of the associations identified for the 53-SNP instrument were replicated with the 28-SNP instrument (limited to those SNPs that were not in loci of prior GWAS hits for triglycerides or HDL-C; **Figure 2**) as well as the 12-SNP instrument (identified in a GWAS of fasting insulin adjusted for BMI; Supplemental Fig. 2). The associations were also consistent when rs1011685 near *LPL* was removed from the 53-SNP instrument (**Figure 2**). Removal of 6 SNPs associated with BMI (P<0.001) had no material effect on the MR estimates (data not shown).

To investigate the robustness of these MR estimates to potential confounding by genetic pleiotropy, we also investigated the association of the 53-SNP instrument with the BCAAs and glycoprotein acetyls (GlycA) using MR-Egger, weighted median and weighted mode-based estimators. Discordance of the point estimates was noticed across the methods, predominantly due to the inclusion of the rs1011685 variant that had minimal effects on insulin adjusted for BMI (Supplemental Table 8). Since MR approaches can be vulnerable to the inclusion of such outliers, we repeated the sensitivity analyses excluding rs1011685, which led to estimates across all MR methods that were comparable to the IVW approach (**Figure 3** and Supplemental Table 8).

The intercepts of MR Egger were of generally small magnitude (absolute values ≤ 0.01, far smaller than the corresponding beta coefficients) with little or no evidence that they departed from zero, providing little evidence for the presence of genetic pleiotropy (Supplemental Table 9). Because the MR-Egger estimate of the causal effect (obtained from the slope of the regression line) can be underestimated when the assumption of no measurement error of the exposure (NOME) is violated, the heterogeneity index (I^2) was used to detect the extent of this potential violation (22). Results remained consistent when SIMEX-adjusted MR-Egger was used to correct potential errors of the SNP to exposure estimates (Supplemental Table 8).

Pathways

Figure 4 and Supplemental Table 10 illustrates the current evidence base for various pathways leading from adiposity to T2DM. Prior MR studies have shown that general adiposity (measured by BMI) and central adiposity (measured by waist-to-hip ratio adjusted for BMI) causally influence fasting insulin, HDL-C and triglycerides (4). BMI has been previously shown to influence BCAAs (23) and in this study we show that both BMI and waist-to-hip ratio (adjusted for BMI) impact these traits (Supplemental Fig. 3 & 4). Both BMI and waist-to-hip ratio causally affect diabetes (4). Our study shows that insulin resistance impacts on BCAAs, and together with a prior study providing genetic support of BCAAs metabolism in T2DM (11), the various sources of

data support a causal pathway that is from adiposity to insulin resistance to BCAAs to diabetes.

Conclusions

Our study provides genetic evidence in support of higher levels of insulin resistance leading to an elevation in circulating branched-chain amino acids. Within the context of other studies, our findings support the hypothesis that the metabolism of BCAAs may be a mediator that is downstream of adiposity and insulin resistance on the causal pathway to T2DM. If true, then this not only has important aetiological relevance, but also could point towards potential novel opportunities for disease treatment and prevention.

These findings for insulin resistance and BCAAs are consistent with a recent paper by Mahendran et al. (15) in which ten IR associated SNPs were used to quantify the association with a composite measure of BCAAs in a one-sample MR setting of ~1300 individuals. However, the selection of SNPs into the 10-SNP instrument may induce bias as the instrument was enriched for GWAS hits of fasting insulin that were also associated with triglycerides and HDL-C (18). In this study, a more robust approach was taken to instrument derivation by selecting >50 SNPs across the genome, which have recently been identified using a hypothesis-free approach to show directionally consistent associations with a triad of phenotypes that mark insulin resistance; this 53-SNP instrument was used to infer the causality of insulin resistance using a two-sample MR design with little overlap between datasets and with data on ~180,000 individuals for the SNP to exposure (IR) estimates and data on ~25,000 individuals for the SNP to outcome (metabolic markers) estimates. The consistent results that we report derived from multiple genetic instruments and multiple MR sensitivity analyses provide robust evidence that insulin resistance impacts on BCAAs in a cause and effect manner. Particularly, as insulin resistance can be measured by various metrics (e.g. a triad of the phenotypes as defined here and also by fasting insulin alone), the consistent results of the 53-SNP instrument (a genetic proxy for the insulin resistance triad) and 12-SNP instrument (a genetic proxy for fasting insulin alone) across the metabolic profile strengthens the evidence base for a causal role of insulin resistance and potentially validates the biological meaning of insulin resistance as defined by a complex phenotype characterised by higher insulin, higher triglycerides and lower HDL-C.

Interventional studies provide orthogonal support for our findings that obesity and insulin resistance causally impact on circulating BCAAs. Multiple longitudinal studies have shown that BCAA levels were reduced after various insulin-sensitising interventions, including weight loss surgery through gastric bypass, pioglitazone therapy or physical exercise (24-26). Also, a reduction in BCAA concentrations was observed following the secretion of insulin during oral glucose tolerance test (OGTT), with individuals with insulin resistance showing less BCAA suppression (i.e. higher BCAA concentrations) following OGTT (27). Prospective studies have identified

circulating BCAAs to be predictive of incident T2DM and a recent genetic study found that the metabolism of BCAAs is likely causally linked to T2DM (11). Triangulating these sources of evidence provides support for the hypothesis that circulating BCAAs may mediate the relation from adiposity and insulin resistance to T2DM. On the other hand, observational studies have reported that higher dietary intake of BCAAs is associated with an improved cardiometabolic risk profile including a lower risk of T2DM (28,29). However, dietary BCAAs, both measured in absolute terms or as a percentage of total protein, are only weakly correlated with circulating concentrations of BCAAs (28,29). There is also evidence that the expression of enzymes involved in BCAA catabolism (e.g. branched-chain α-ketoacid dehydrogenase, BCKD) is reduced in obese and diabetic individuals (11,30). BCKD is responsible for the rate-limiting step of BCAA catabolism and BCKD can be activated by its regulatory phosphatase encoded by *PPM1K*. Individuals with T2DM have reduced up-regulation of *PPM1K* in skeletal muscle during OGTT (11). Consistent with this, after weight loss surgery, BCKD concentrations are increased leading to a commensurate reduction in BCAAs (24,30). Thus, elevated circulating BCAA levels observed in obese and diabetic individuals could arise from impaired BCAA catabolism (11). Putting these strands of evidence together, it is plausible that pharmacotherapies to improve or restore the function of BCAA catabolism may represent a means to prevent T2DM. However, further studies are required to understand the exact role of BCAA metabolism in the aetiology of T2DM.

In contrast to the strong effects of IR on the BCAAs, we noticed a generally weaker effect of IR on alanine, glutamine, and aromatic amino acids (phenylalanine and tyrosine). Each of these biomarkers has been associated with the risk of insulin resistance, hyperglycaemia, type 2 diabetes and cardiovascular diseases (8,10,31). Although imprecise estimates were observed for these measures in the MR analyses reported herein, the consistent results from different instruments on these traits merits further investigation in larger datasets to clarify whether these represent causal relationships.

The association that we identify of insulin resistance with GlycA is novel. GlycA is a marker of both acute phase and chronic inflammation and has been linked to neutrophil activity (32). GlycA reflects circulating levels of various inflammatory glycoproteins (primarily alpha-1-acid glycoprotein and haptoglobin), and is also associated with a wide range of inflammatory cytokines (32). Prospective observational studies have identified positive associations of GlycA with cardiovascular disease, T2DM and premature mortality (33,34). A role for inflammation in the development of T2DM has been proposed for many years on account of the observational associations between higher concentrations of biomarkers of inflammation, such as C-reactive protein, interleukin-1, interleukin-6, and the risk of T2DM (35,36). Although recent MR studies have so far failed to provide evidence in support of this hypothesis (36), it remains plausible that such causal pathways (from inflammation to T2DM) exist, and that larger studies and/or investigations of other inflammatory markers and pathways may identify a causal role

of inflammation in T2DM. Therefore, GlycA could represent a biomarker either involved in, or correlated to an inflammation pathway involved in the aetiology of T2DM. A causal role of inflammation in vascular disease is gaining traction given recent findings from genetic studies in humans of the interleukin-6 receptor and CHD (37), and more recently in phase III clinical trials of anti-inflammatory drugs for the treatment of CHD (CANTOS trial of canakinumab, a monoclonal antibody to interleukin-1 beta) (38). Of note, a previous study has suggested that BMI has a causal impact on circulating concentrations of GlycA (23) (as we also report in Supplemental Fig. 3 for BMI and Fig.4 for WHR) and our study here provides clarification on the potential causal pathway, showing that insulin resistance is also causal for GlycA. However, elucidating the causal role of GlycA in cardiometabolic disease remains challenging using an MR approach since at present the identified genetic variants associated with GlycA are limited in number (16), thus hindering our ability to answer this important question. Larger GWAS of GlycA may facilitate this endeavor.

Strengths of this study include: (i) a comprehensive genetic instrument for an insulin resistance phenotype using findings from a recent GWAS (7); (ii) characterizing and validating the genetic instrument for IR with a repertoire of biomarkers of triglyceride and HDL-C metabolism; (iii) use of multiple sensitivity analyses (both in the derivation of the genetic instruments and their application to state of the art MR methodologies) which provided robust and consistent evidence; (iv) quantifying the causal effects of insulin resistance on each of the three BCAAs individually; (iv) adding important new information on the effect of insulin resistance on an inflammation marker; and, (v) a data summation that provides evidence of a causal pathway from adiposity through IR and BCAA to T2DM.

Limitations include: (i) analyses were conducted at the summary level and we could not investigate associations by subgroups, e.g., of age or sex, meaning that it is not possible to test whether these associations are modified by age; (ii) our analyses were conducted using European datasets which may hamper their translational relevance to non-Europeans, however risk factors for disease tend to show similar relationships across geographical regions (39) and emerging studies are providing evidence that shows the genetic architecture for common diseases is likely similar across ethnic groups (40) (iii) a meta-GWAS of 3 traits was used to proxy IR, which may not include other traits related to IR and may have limited clinical relevance, although in the original paper by Lotta et al. (7), associations were identified for diabetes and heart disease; (iv) meta-GWAS may select SNPs on the basis of pleiotropy (i.e., by their very selection, they associate with higher fasting insulin, higher triglycerides and lower HDL-C) and thus SNPs may tag heterogeneous pathways, some of which may result in unbalanced horizontal pleiotropy (18). Against this are the consistent associations across the different genetic instruments, their stability to various MR sensitivity analyses (with each MR approach having its own assumptions on the amount and type of genetic pleiotropy, see Supplemental Table 5), and the general consistency with a prior study that used a weaker instrument in a

much smaller dataset (15). Finally, the instruments were derived from a meta-GWAS that included fasting insulin adjusted for BMI; conditioning on a trait in discovery GWAS can induce collider bias, as evidenced by the negative association of the instruments with BMI. However, this negative association with BMI would be expected to diminish the association of the genetic instruments with BCAA that we report (and also diminish the association with T2DM and CHD reported by Lotta et al. (7) rather than augment it, and therefore is unlikely to result in major bias in the MR estimates we report. Further, removal of six SNPs that were associated with BMI (at $P<0.001$ using GIANT summary statistics) had no material impact on the causal estimates derived from MR.

In conclusion, our findings provide new information in support of a causal role of insulin resistance on branched chain amino acids and inflammation. Taken together with recent findings from complimentary studies, these data suggest BCAA metabolism may lie on a causal pathway from adiposity and insulin resistance to type 2 diabetes.

Duality of Interest
None of the authors report any conflicts of interest.

Author Contributions
Q.W., M.V.H., and M.A.K. designed the study. Q.W. conducted the statistical analyses. Q.W. and M.V.H. wrote the manuscript. All authors interpreted the data and contributed to the discussion, reviewed and edited the manuscript. M.A.K. is the guarantor of this work and, as such, had full access to all the data in the study and takes the responsibility for the integrity of the data and the accuracy of the data analyses.

References

1.   The GBD 2015 Obesity Collaborators. Health Effects of Overweight and Obesity in 195 Countries over 25 Years. N Engl J Med 2017;377:13–27.

2.   Hardy OT, Czech MP, Corvera S. What causes the insulin resistance underlying obesity? Current opinion in endocrinology, diabetes, and obesity. Curr Opin Endocrinol Diabetes Obes 2012;19:81–7.

3.   Holmes MV, Lange LA, Palmer T, et al. Causal effects of body mass index on cardiometabolic traits and events: a Mendelian randomization analysis. Am J Hum Genet 2014;94:198–208.

4.   Dale CE, Fatemifar G, Palmer TM, et al. Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. Circulation 2017;135:2373–88.

5.   Lyall DM, Celis-Morales C, Ward J, et al. Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank: A Mendelian Randomization Study. JAMA Cardiol 2017;2:882–9.

6.   Emdin CA, Khera AV, Natarajan P, et al. Genetic Association of Waist-to-Hip Ratio With Cardiometabolic Traits, Type 2 Diabetes, and Coronary Heart Disease. JAMA 2017;317:626–34.

7.   Lotta LA, Gulati P, Day FR, et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. Nat Genet 2017;49:17–26.

8.   Würtz P, Soininen P, Kangas AJ, et al. Branched-chain and aromatic amino acids are predictors of insulin resistance in young adults. Diabetes Care 2013;36:648–55.

9.   Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. Nat Med 2011;17:448–53.

10. Guasch-Ferré M, Hruby A, Toledo E, et al. Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis. Diabetes Care 2016;39:833–46.

11. Lotta LA, Scott RA, Sharp SJ, et al. Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. PLoS Med 2016;13:e1002179.

12. Knowles JW, Xie W, Zhang Z, et al. Identification and validation of N-acetyltransferase 2 as an insulin sensitivity gene. J Clin Invest 2015;125:1739–51.

13. 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.

14. Scott RA, Fall T, Pasko D, et al. Common genetic variants highlight the role of insulin resistance and body fat distribution in type 2 diabetes, independent of obesity. Diabetes 2014;63:4378–87.

15. Mahendran Y, Jonsson A, Have CT, et al. Genetic evidence of a causal effect of insulin resistance on branched-chain amino acid levels. Diabetologia 2017;60:873–8.

16. Kettunen J, Demirkan A, Würtz P, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commun 2016;7:11122.

17. Smith GD, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 2003;32:1–22.

18. Holmes MV, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. Nat Rev Cardiol 2017; doi: 10.1038/nrcardio.2017.78.

19. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol 2017; https://doi.org/10.1093/ije/dyx102.

20. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197–206.

21. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013;45:1274–83.

22. Bowden J, Del Greco M F, Minelli C, et al. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. Int J Epidemiol 2016;45:1961–74.

23. Würtz P, Wang Q, Kangas AJ, et al. Metabolic Signatures of Adiposity in Young Adults: Mendelian Randomization Analysis and Effects of Weight Change. PLoS Med 2014;11:e1001765.

24. Laferrère B, Reilly D, Arias S, et al. Differential metabolic impact of gastric bypass surgery versus dietary intervention in obese diabetic subjects despite identical weight loss. Sci Transl Med 2011;3:80re2.

25. Kakazu E, Kondo Y, Ninomiya M, et al. The influence of pioglitazone on the plasma amino acid profile in patients with nonalcoholic steatohepatitis (NASH). Hepatol Int 2013;7:577–85.

26. Kujala UM, Mäkinen V-P, Heinonen I, et al. Long-term leisure-time physical activity and serum metabolome. Circulation 2013;127:340–8.

27. Shaham O, Wei R, Wang TJ, et al. Metabolic profiling of the human response to a glucose challenge reveals distinct axes of insulin sensitivity. Mol Syst Biol 2008;4:214.

28.   Nagata C, Nakamura K, Wada K, Tsuji M, Tamai Y, Kawachi T. Branched-chain amino acid intake and the risk of diabetes in a Japanese community: the Takayama study. Am J Epidemiol 2013;178:1226–32.

29.   Jennings A, MacGregor A, Pallister T, Spector T, Cassidy A. Associations between branched chain amino acid intake and biomarkers of adiposity and cardiometabolic health independent of genetic factors: A twin study. Int J Cardiol 2016;223:992–8.

30.   She P, Van Horn C, Reid T, Hutson SM, Cooney RN, Lynch CJ. Obesity-related elevations in plasma leucine are associated with alterations in enzymes involved in branched-chain amino acid metabolism. Am J Physiol Endocrinol Metab 2007;293:E1552–63.

31.   Würtz P, Havulinna AS, Soininen P, et al. Metabolite Profiling and Cardiovascular Event Risk: A Prospective Study of Three Population-Based Cohorts. Circulation 2015;131:774–85.

32.   Ritchie SC, Würtz P, Nath AP, et al. The Biomarker GlycA Is Associated with Chronic Inflammation and Predicts Long-Term Risk of Severe Infection. Cell Systems 2015;1:293–301.

33.   Connelly MA, Gruppen EG, Wolak-Dinsmore J, et al. GlycA, a marker of acute phase glycoproteins, and the risk of incident type 2 diabetes mellitus: PREVEND study. Clinica Chimica Acta 2016;452:10–7.

34.   Fischer K, Kettunen J, Würtz P, et al. Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17,345 persons. PLoS Med 2014;11:e1001606.

35.   Wang X, Bao W, Liu J, et al. Inflammatory Markers and Risk of Type 2 Diabetes. Diabetes Care 2013;36:166–75.

36.   Swerdlow DI. Mendelian Randomization and Type 2 Diabetes. Cardiovasc Drugs Ther 2016;30:51–7.

37.   Swerdlow DI, Holmes MV, Kuchenbaecker KB, et al. The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. Lancet 2012 31;379:1214–24.

38.   Ridker PM, Everett BM, Thuren T, et al. Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. N Engl J Med 2017;doi: 10.1056/NEJMoa1707914.

39.   Yusuf S, Hawken S, Ounpuu S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet 2004;364:937–52.

40.   Gan W, Walters RG, Holmes MV, et al. Evaluation of type 2 diabetes genetic risk variants in Chinese adults: findings from 93,000 individuals from the China Kadoorie Biobank. Diabetologia 2016;59:1446–57.

Figure Legends

Figure 1. Heat map of the 53 SNPs and their associations with 58 circulating biomarkers. The units are reported as an SD-difference in metabolic measure per insulin resistance increasing allele. Abbreviations: VLDL, very-low-density lipoprotein; IDL, intermediate-density lipoprotein; LDL, low-density lipoprotein; HDL, high-density lipoprotein; lipoprotein measures without further specification refer to total lipid concentrations; apoA1, apolipoprotein A-I; apoB, apolipoprotein B; w79s FA, omega 7, 9 and saturated fatty acids; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acids; DHA, docosahexaenoic acid; DBinFA, the average number of double bonds in fatty acids; FALen, the average fatty acid chain length. Insulin resistance was defined as a triad of higher fasting insulin (BMI adjusted), higher triglycerides and lower HDL-C. Metabolic measures were quantified by the high-throughput NMR metabolomics platform using primarily fasting serum samples.

Figure 2. **Forrest plot of the causal effect estimates of insulin resistance on circulating metabolic measures.** Estimates are derived from inverse variance weighted (IVW) Mendelian randomization analyses. The three instruments are: 53 SNPs identified from Lotta et al. (7) (black diamonds); 52 SNPs removing an outlier variant rs1011685 (near *LPL*) (red circles); 28 SNPs in loci not previously associated with high-density lipoprotein cholesterol and triglycerides at genome-wide significance (blue circles). Open and close symbols indicate $P \geq 0.001$ and $P < 0.001$, respectively. Units are given as SD difference in metabolic measures per 1-SD genetically higher insulin resistance. Abbreviations are as listed in the caption for Figure 1.

Figure 3. **Funnel plots for the three branched-chain amino acids and glycoprotein acetyls showing the causal effect estimates.** IVW refers to the conventional inverse variance weighted method (using 53 SNPs; red vertical lines), Egger to the MR-Egger (using 52 SNPs; green lines), WME to the weighted median estimator (using 52 SNPs; blue lines) and WMBE for the weighted mode-based estimator (using 52 SNPs; purple line). For the results shown for MR-Egger, WME and WMBE, the outlier SNP rs1011685 near *LPL* was removed. The 95% CIs for each method are shown as the colourful horizontal lines. Each individual black circle describes the causal effect estimate using the individual SNP as the instrument.

Figure 4. Strands of evidence from multiple genetic studies supporting a causal pathway from adiposity, through insulin resistance and branched-chain amino acids, to diabetes. Sources of evidence: 1. Holmes et al. (3), Dale et al. (4); 2. Dale et al. (4), Emdin et al. (6); 3. Würtz et al. (23), this study (Supplemental Fig. 3); 4. This study (Supplemental Fig. 4); 5. Holmes et al. (3), Dale et al. (4), Lyall et al.

(5); 6. Emdin et al. (6), Dale et al. (4); 7. This study (Figure 2), Mahendran et al. (15); 8. Lotta et al. (11). For details of these studies and the MR estimates provided, please see Supplemental Table 10. *refers to adjustment with BMI.

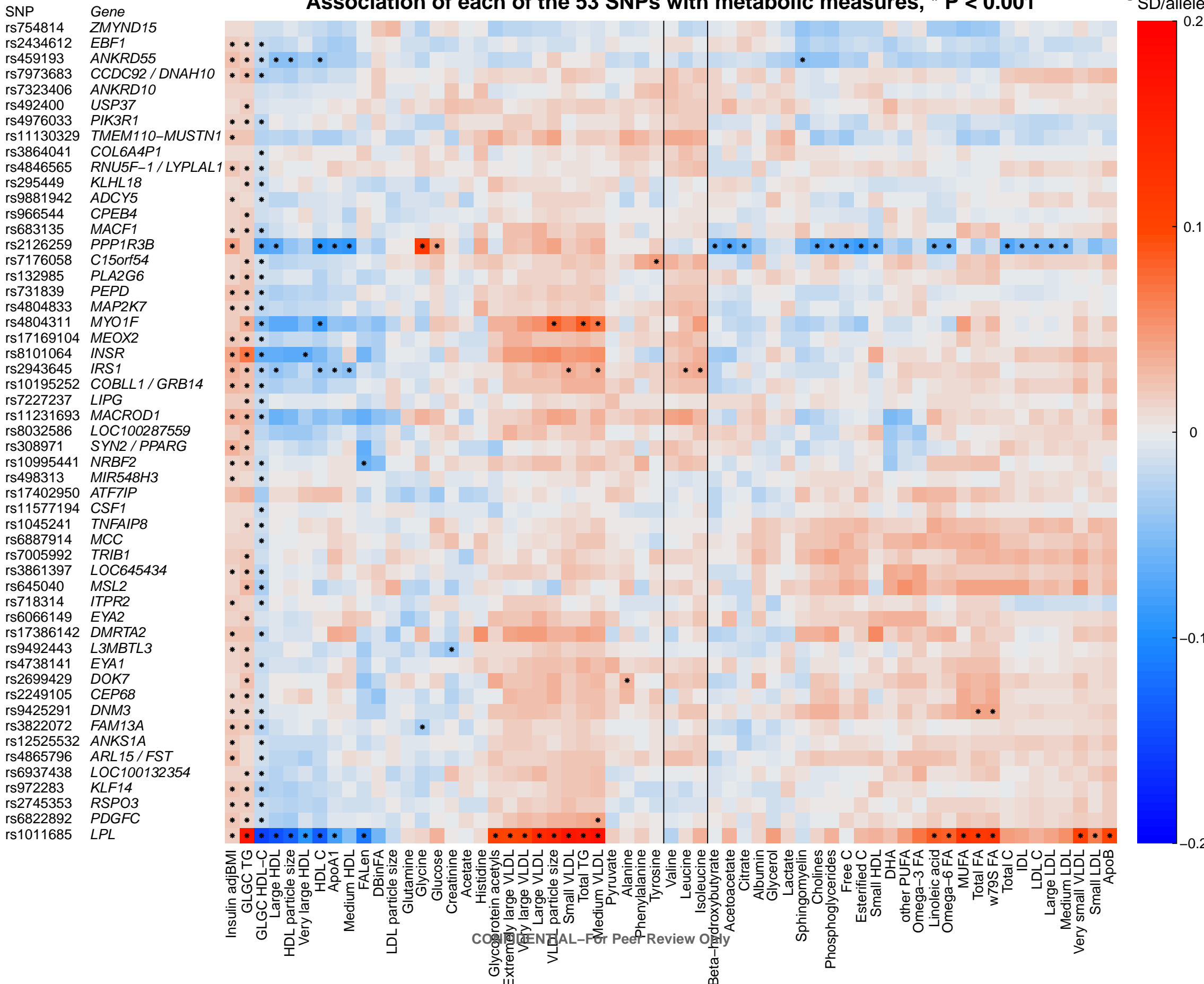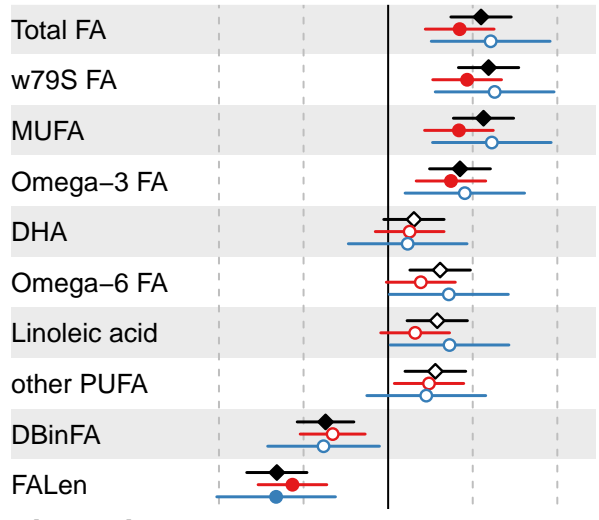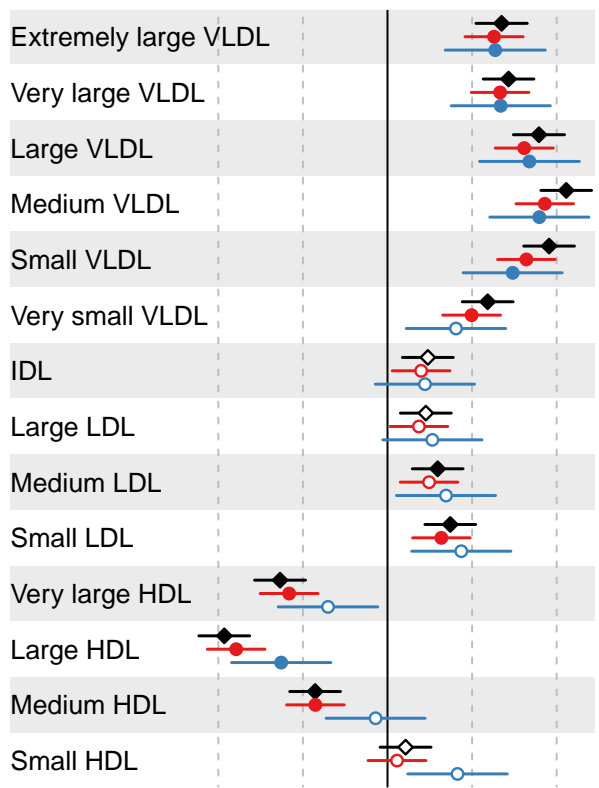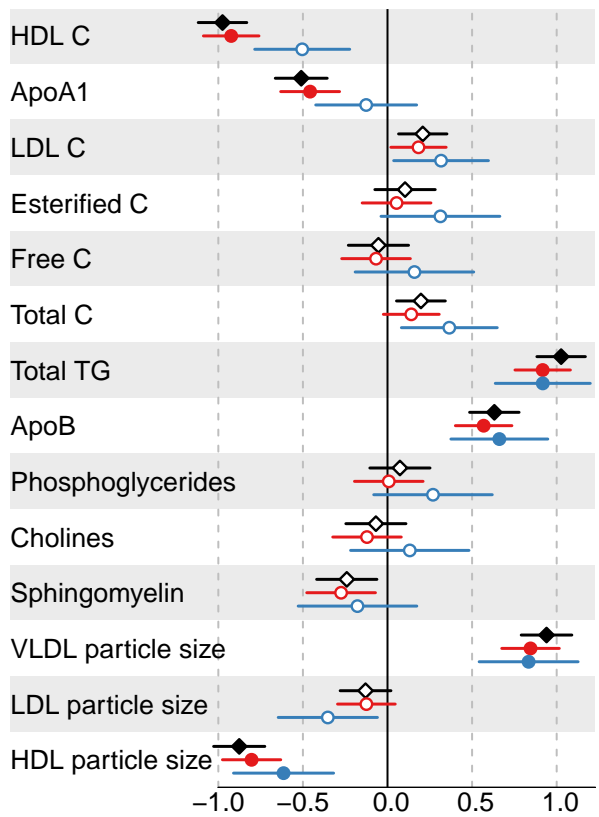# Association of each of the 53 SNPs with metabolic measures, * P < 0.001

Total lipids in lipoprotein subclasses
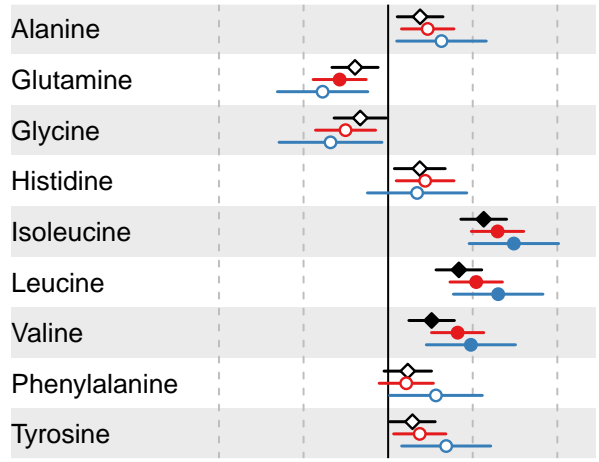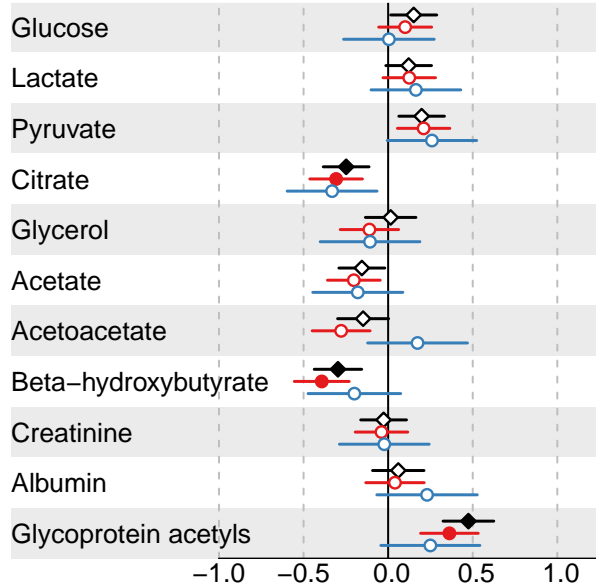
Causal effect estimates; SD metabolite (95% CI) per 1−SD insulin resistance

IVW 53−SNP : ◆ P < 0.001   ◇ P ≥ 0.001
IVW 52−SNP : ● P < 0.001   ○ P ≥ 0.001
IVW 28−SNP : ● P < 0.001   ○ P ≥ 0.001
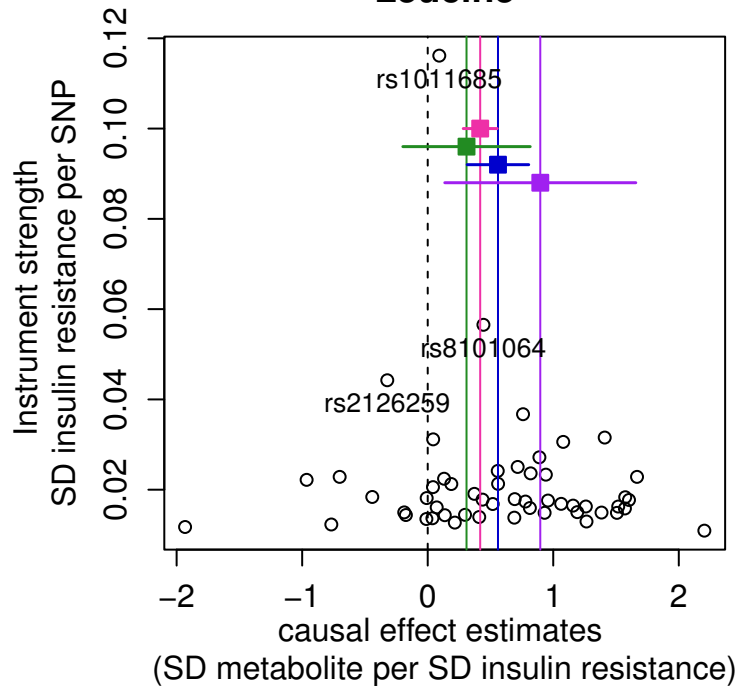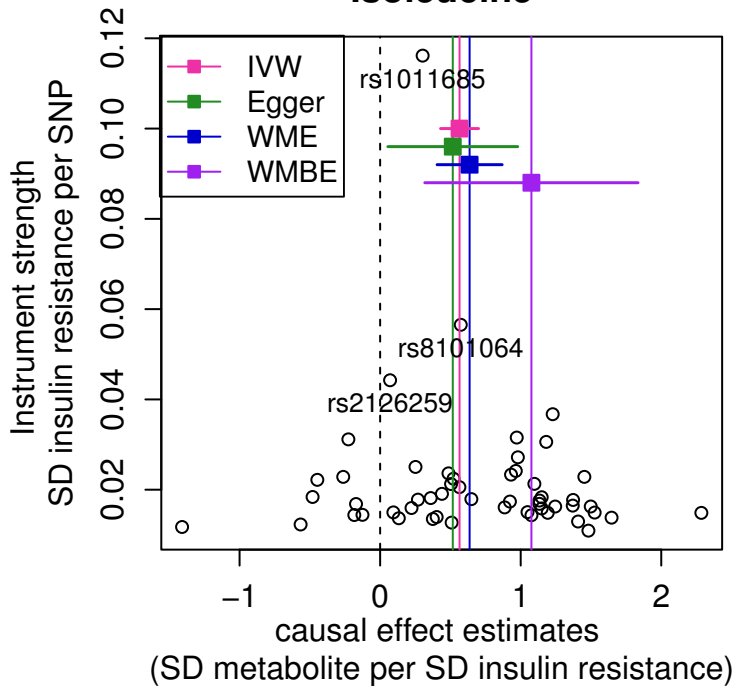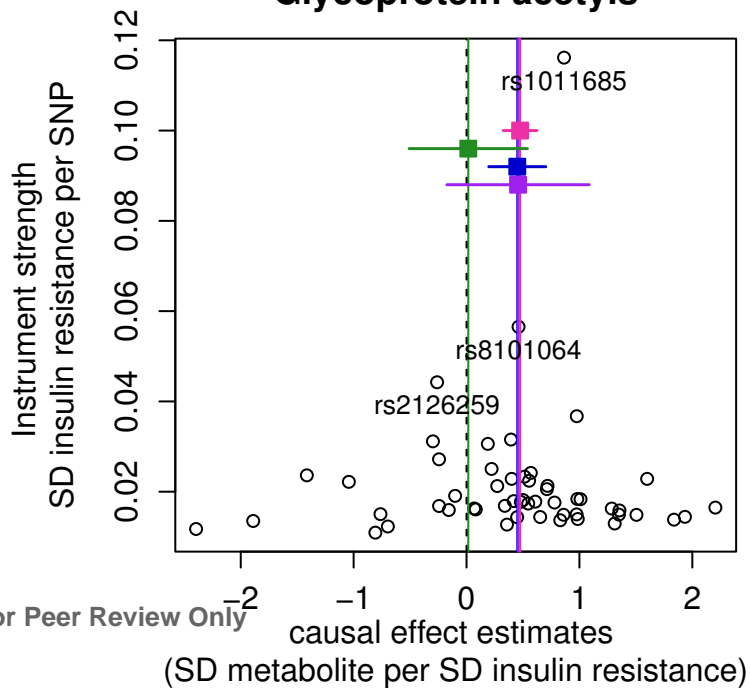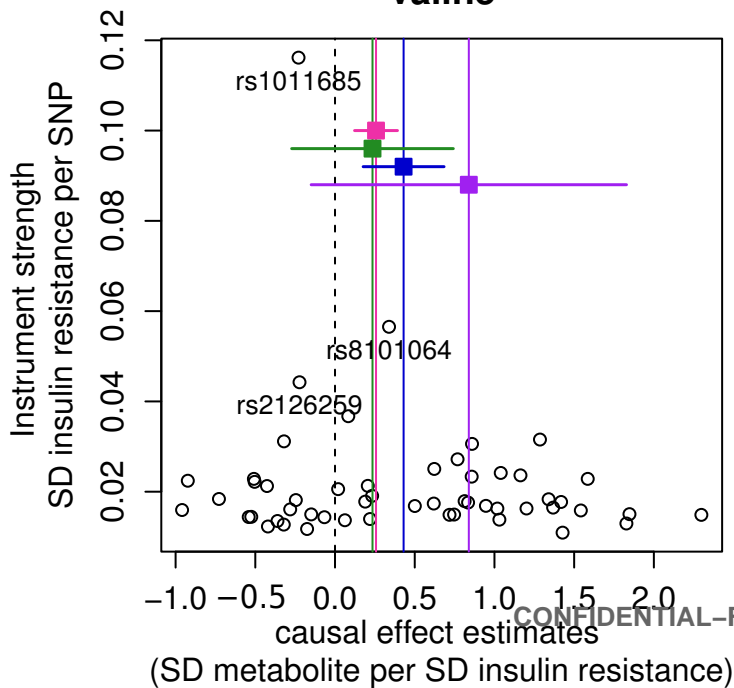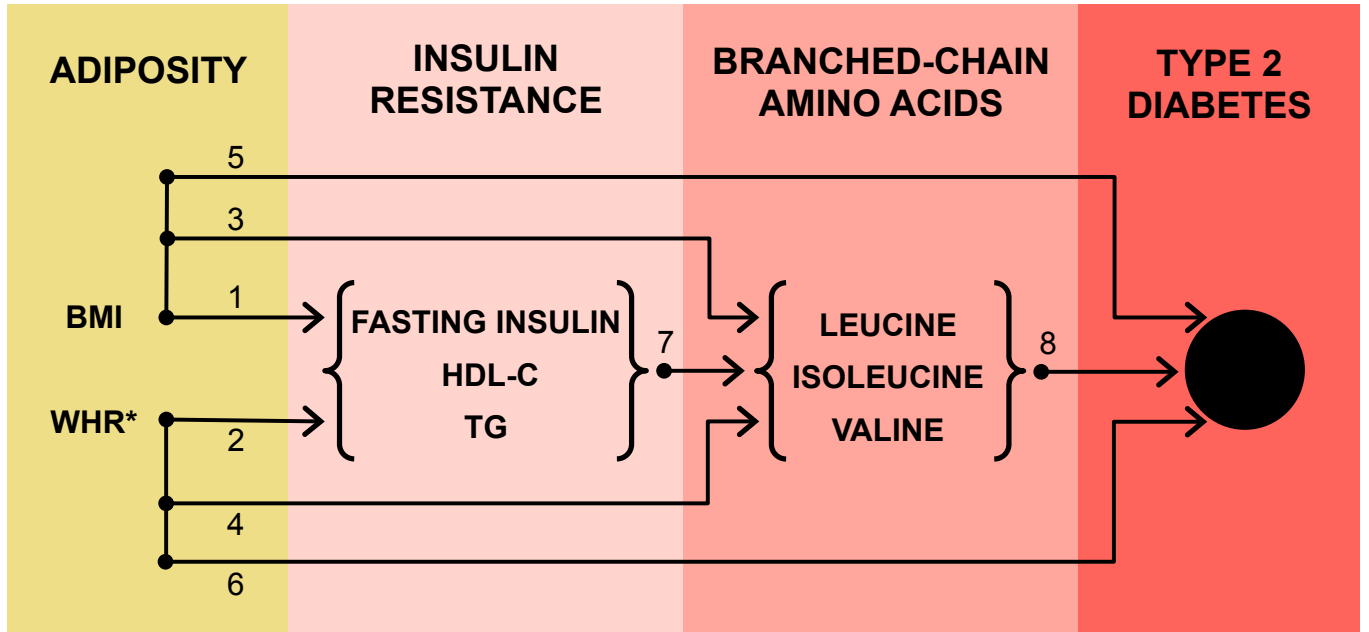
SUPPLEMENTAL MATERIAL

Genetic support for a causal role of insulin resistance on circulating branched-chain amino acids and inflammation
Qin Wang, Michael V Holmes, George Davey Smith, Mika Ala-Korpela

Content
Supplemental Tables

Supplemental Figures

Supplemental Table 1. Characteristics of the GWASs studies used in this study.

Characteristics	Exposure GWAS* (insulin resistance)		Outcome GWAS
Phenotype	Fasting insulin adjusted for BMI	Triglycerides and HDL cholesterol	Circulating metabolic measures profiled by high-throughput quantitative serum NMR metabolomics
Consortium	MAGIC	GLGC	MAGNETIC
PubMed ID	22885924, 22581228	24097068	27005778
N	Up to 108,557 participants	Up to 188,577 participants	Up to 24,925 participants
Ethnicity	European	European	European
Genotype data	GWAS array and metabochip array	GWAS array and metabochip array	GWAS array

*: Lotta et al (1) built a genetic instrument for insulin resistance by combining published GWAS results for fasting insulin (adjusted for BMI), HDL cholesterol and triglycerides. Here, the characteristics of the GWAS for each individual phenotype of insulin resistance are listed.

Supplemental Table 2. Characteristics of the individual cohorts used in the metabolite-GWAS.

Study		Age		BMI		Female%
	N	**Mean**	**s.d.**	**Mean**	**s.d.**	
EGCUT	3,287	46.3	19.5	26.4	5.4	58
ERF	2,118	48.2	14.7	26.7	4.7	58
FTC	664	23.9	2.1	23.1	3.7	50
FR97	3,661	45.3	12.8	26.3	4.5	55
COROGENE	828	53.2	13.2	26.6	4.1	54
GenMets	572	55.8	7.3	27.2	4.5	57
HBCS	708	61.3	2.9	27.1	4.1	60
KORA	1,745	60.9	8.8	28.2	4.8	52
LLS	2,227	59.2	6.8	25.4	3.5	54
NTR	1,192	38.8	12.8	24.6	4.2	64
NFBC 1966	4,709	31.2	0.4	24.6	4.1	51
PredictCVD	374	47.5	14.6	26.6	4.4	37
PROTE	597	38.3	16	25.2	4.6	51
YFS	2,390	37.7	5	26	4.7	54

Supplemental Table 3 and 4 presented as Excel files due to their large size.

Supplemental Table 5. Characteristics of different MR methods[*].

Method	Principle	Breakdown level †	Stability to presence of horizontal pleiotropy	Analytical code for the methods
IVW (2)	Combing the ratio estimates from individual SNPs via fixed-effect meta-analysis; alternatively, interpreted as the regression of SNP-outcome onto SNP-exposure forced through original.	0%	Stable if the sum of horizontal pleiotropic effects across all variants is zero.	R and Stata codes for IVW and MR-Egger regression methods are available at Web Appendix of reference (3)
MR-Egger regression ‡ (3)	The regression of SNP-outcome onto SNP-exposure allows the y-intercept to float. The y-intercept is used to infer the presence and extent of unbalanced horizontal pleiotropy and the slope indicates the corrected causal effect estimate.	100%	Stable even if all variants are invalid provided that the associations of individual variants with the exposure are unrelated to the corresponding pleiotropic effects.	
Weighted median estimator (4)	The median of the ratio estimates obtained from individual SNPs	50%	Stable if more than 50% of the weights is contributed by valid variants, regardless of the type of horizontal pleiotropy.	R code is provided in Supporting Information Appendix 2 from reference (4)
Weighted mode-based estimator (5)	The highest density of the ratio estimates across all SNPs	Depends on the weights of individual SNP	Stable if the largest number of similar individual-instrument causal effect estimates arise from valid instruments, regardless of the type of horizontal pleiotropy	R code is provided in supplementary materials from reference (5)

* This table is based on multiple previous publications (3-5).

† Breakdown refers to the proportion of information that can come from invalid instrumental variables before the method gives biased estimates.

‡ The MR-Egger estimate of the causal effect can be prone to be underestimated when the assumption of no-measurement error (NOME) of the SNP to exposure is violated; the heterogeneity index (I^2) was used to detect the extent of the violation and MR-Egger with SIMEX was used to correct such violation (6).

Supplemental Table 6. The causal effect estimates using 53-SNP, 52-SNP and 28-SNP instrument via inverse-variance weighted method as shown in Figure 2.

Metabolites	53-SNP instrument			52-SNP instrument			28-SNP instrument		
	Beta	SE	Pvalue	Beta	SE	Pvalue	Beta	SE	Pvalue
Extremely large VLDL	0.67	0.08	4E-12	0.63	0.09	1E-09	0.64	0.10	1E-04
Very large VLDL	0.72	0.07	4E-13	0.67	0.08	3E-10	0.67	0.10	7E-05
Large VLDL	0.90	0.08	2E-16	0.81	0.09	8E-13	0.84	0.10	3E-06
Medium VLDL	1.06	0.07	P<2E-16	0.93	0.08	5E-15	0.90	0.10	1E-06
Small VLDL	0.96	0.07	P<2E-16	0.82	0.08	4E-13	0.74	0.10	2E-05
Very small VLDL	0.59	0.07	2E-10	0.50	0.09	4E-07	0.41	0.10	9E-03
IDL	0.24	0.07	2E-03	0.20	0.08	2E-02	0.22	0.10	1E-01
Large LDL	0.23	0.07	4E-03	0.19	0.08	3E-02	0.27	0.10	7E-02
Medium LDL	0.30	0.07	2E-04	0.25	0.08	6E-03	0.35	0.10	2E-02
Small LDL	0.37	0.07	8E-06	0.32	0.08	4E-04	0.44	0.10	5E-03
Very large HDL	-0.64	0.08	2E-11	-0.58	0.09	1E-08	-0.35	0.10	2E-02
Large HDL	-0.96	0.07	P<2E-16	-0.90	0.08	2E-14	-0.63	0.10	2E-04
Medium HDL	-0.43	0.07	5E-07	-0.43	0.08	7E-06	-0.07	0.10	6E-01
Small HDL	0.11	0.08	2E-01	0.06	0.09	5E-01	0.41	0.10	8E-03
HDL C	-0.98	0.07	P<2E-16	-0.92	0.08	2E-15	-0.50	0.10	1E-03
ApoA1	-0.51	0.08	2E-08	-0.46	0.09	3E-06	-0.13	0.10	4E-01
LDL C	0.21	0.07	5E-03	0.18	0.08	3E-02	0.32	0.10	3E-02
Esterified C	0.10	0.09	2E-01	0.05	0.10	6E-01	0.31	0.20	8E-02
Free C	-0.05	0.09	5E-01	-0.07	0.10	5E-01	0.16	0.20	4E-01
Total C	0.20	0.07	8E-03	0.14	0.08	9E-02	0.37	0.10	1E-02
Total TG	1.03	0.07	P<2E-16	0.92	0.08	2E-15	0.92	0.10	3E-07
ApoB	0.63	0.07	1E-11	0.57	0.08	1E-08	0.66	0.10	6E-05
Phosphoglycerides	0.07	0.09	4E-01	0.01	0.10	9E-01	0.27	0.20	1E-01
Cholines	-0.07	0.09	4E-01	-0.12	0.10	2E-01	0.13	0.20	4E-01
Sphingomyelin	-0.24	0.09	9E-03	-0.27	0.10	9E-03	-0.18	0.20	3E-01
VLDL particle size	0.94	0.07	P<2E-16	0.85	0.08	1E-13	0.83	0.10	3E-06
LDL particle size	-0.13	0.07	9E-02	-0.13	0.08	1E-01	-0.35	0.10	2E-02
HDL particle size	-0.88	0.08	4E-16	-0.80	0.09	1E-12	-0.61	0.10	2E-04
Total FA	0.55	0.09	1E-07	0.42	0.10	1E-04	0.61	0.20	1E-03
w79S FA	0.59	0.09	2E-08	0.47	0.10	3E-05	0.63	0.20	1E-03
MUFA	0.56	0.09	5E-08	0.42	0.10	1E-04	0.61	0.20	1E-03
Omega-3 FA	0.42	0.09	2E-05	0.37	0.10	7E-04	0.45	0.20	1E-02
DHA	0.15	0.09	9E-02	0.13	0.10	2E-01	0.12	0.20	5E-01
Omega-6 FA	0.31	0.09	1E-03	0.19	0.10	6E-02	0.36	0.20	5E-02
Linoleic acid	0.29	0.09	2E-03	0.16	0.10	1E-01	0.36	0.20	4E-02
other PUFA	0.28	0.09	3E-03	0.24	0.10	2E-02	0.23	0.20	2E-01
DBinFA	-0.37	0.08	5E-05	-0.33	0.10	1E-03	-0.38	0.20	3E-02
FALen	-0.66	0.09	1E-09	-0.57	0.10	9E-07	-0.66	0.20	6E-04
Alanine	0.19	0.07	7E-03	0.23	0.08	4E-03	0.32	0.10	2E-02
Glutamine	-0.20	0.07	6E-03	-0.29	0.08	6E-04	-0.39	0.10	6E-03

Glycine	-0.17	0.08	4E-02	-0.25	0.09	6E-03	-0.34	0.10	3E-02
Histidine	0.19	0.07	2E-02	0.22	0.08	1E-02	0.17	0.10	2E-01
Isoleucine	0.56	0.07	3E-11	0.65	0.08	4E-11	0.74	0.10	4E-06
Leucine	0.42	0.07	9E-08	0.52	0.08	1E-08	0.65	0.10	3E-05
Valine	0.26	0.07	3E-04	0.41	0.08	2E-06	0.49	0.10	7E-04
Phenylalanine	0.12	0.07	1E-01	0.11	0.08	2E-01	0.28	0.10	4E-02
Tyrosine	0.14	0.07	4E-02	0.19	0.08	2E-02	0.34	0.10	1E-02
Glucose	0.15	0.07	3E-02	0.10	0.08	2E-01	0.00	0.10	1E+00
Lactate	0.12	0.07	8E-02	0.13	0.08	1E-01	0.16	0.10	2E-01
Pyruvate	0.20	0.07	5E-03	0.21	0.08	9E-03	0.26	0.10	5E-02
Citrate	-0.25	0.07	6E-04	-0.31	0.08	2E-04	-0.33	0.10	2E-02
Glycerol	0.01	0.07	8E-01	-0.11	0.09	2E-01	-0.11	0.10	5E-01
Acetate	-0.16	0.07	3E-02	-0.20	0.08	1E-02	-0.18	0.10	2E-01
Acetoacetate	-0.15	0.08	5E-02	-0.28	0.09	2E-03	0.17	0.10	2E-01
Beta-hydroxybutyrate	-0.30	0.07	9E-05	-0.39	0.08	1E-05	-0.20	0.10	1E-01
Creatinine	-0.03	0.07	7E-01	-0.04	0.08	6E-01	-0.02	0.10	9E-01
Albumin	0.06	0.08	4E-01	0.04	0.09	6E-01	0.23	0.10	1E-01
Glycoprotein acetyls	0.47	0.07	5E-08	0.36	0.08	8E-05	0.25	0.10	9E-02

The three instruments are: 53 SNPs identified from Lotta et al. (1); 52 SNPs removing an outlier variant rs1011685 (near *LPL*); 28 SNPs not previously associated with high-density lipoprotein cholesterol and triglycerides at genome-wide significance. The unit for Betas is reported as: SD metabolite per 1-SD higher insulin resistance. Insulin resistance is defined as a triad of higher fasting insulin (BMI adjusted), higher triglycerides and lower HDL cholesterol.

Supplemental Table 7. Associations of insulin resistance genetic instrument with body mass index.

Instrument	Estimate	Std. Error	CI 2.50%	CI 97.50%	Pvalue
Association of insulin resistance genetic instrument with BMI (SD BMI per SD insulin resistance)					
53 SNPs	-0.11	0.022	-0.16	-0.07	3.15E-06
52 SNPs	-0.16	0.025	-0.21	-0.11	4.29E-08
28 SNPs	-0.12	0.042	-0.21	-0.04	6.86E-03
Association of fasting insulin (BMI adjusted) genetic instrument with BMI (SD BMI per SD insulin)					
12 SNPs	-0.33	0.035	-0.41	-0.26	1.35E-06

Two-sample MR via inverse variance weighted method was used to assess the causal effect estimates of insulin resistance (and fasting insulin) on body mass index (BMI). The associations of the instrument SNPs with BMI were obtained from GIANT consortium for up to 339,224 individuals (7). The four instruments are: 53 SNPs identified from Lotta et al. (1); 52 SNPs removing an outlier variant rs1011685 (near *LPL*); 28 SNPs not previously associated with high-density lipoprotein cholesterol and triglycerides at genome-wide significance; 12 SNPs associated with fasting insulin adjusted for BMI reported by MAGIC consortium (8).

Supplemental Table 8. Sensitivity analyses using different methods across all instruments.

	Isoleucine Beta (95%CI); Pvalue	Leucine Beta (95%CI); Pvalue	Valine Beta (95%CI); Pvalue	GlycA Beta (95%CI); Pvalue
53 SNPs				
IVW	0.56 (0.43, 0.70) p=3e-11	0.42 (0.28, 0.55) p=9e-08	0.26 (0.12, 0.39) p=3e-04	0.47 (0.32, 0.62) p=5e-08
MR-Egger				
Standard approach	0.32 (0.07, 0.58) p=1e-02 I^2=0.95	0.10 (-0.17, 0.37) p=5e-01 I^2=0.95	-0.17 (-0.44, 0.10) p=2e-01 I^2=0.95	0.63 (0.34, 0.92) p=3e-05 I^2=0.95
SIMEX adjusted	0.34 (0.02, 0.65) p=4e-2	0.11 (-0.25, 0.47) p=5e-01	-0.18 (-0.55, 0.19) p=4e-01	0.67 (0.26, 1.08) p=2e-03
Weighted Median	0.42 (0.21, 0.64) p=1e-04	0.23 (0.00, 0.45) p=5e-02	-0.02 (-0.27, 0.23) p=9e-01	0.56 (0.32, 0.79) p=3e-06
Weighted MBE	0.32 (0.05, 0.59) p=2e-02	0.09 (-0.19, 0.37) p=5e-01	-0.25 (-0.51, 0.02) p=7e-02	0.85 (0.53, 1.17) p=4e-06
52 SNPs (removing outlier SNP rs1011685)				
IVW	0.65 (0.49, 0.80) p=4e-11	0.52 (0.37, 0.68) p=1e-08	0.41 (0.26, 0.57) p=2e-06	0.36 (0.19, 0.53) p=8e-05
MR-Egger				
Standard approach	0.52 (0.05, 0.99) p=3e-02 I^2=0.81	0.31 (-0.18, 0.79) p=2e-01 I^2=0.81	0.24 (-0.25, 0.72) p=3e-01 I^2=0.81	0.02 (-0.49, 0.52) p=9e-01 I^2=0.82
SIMEX adjusted	0.60 (-0.02, 1.21) p=6e-02	0.35 (-0.35, 1.05) p=3e-01	0.27 (-0.45, 0.98) p=5e-01	0.02 (-0.71, 0.75) p=9e-01
Weighted Median	0.64 (0.39, 0.88) p=4e-07	0.56 (0.32, 0.80) p=5e-06	0.43 (0.17, 0.69) p=1e-03	0.45 (0.21, 0.69) p=3e-04
Weighted MBE	1.08 (0.29, 1.86) p=9e-03	0.90 (0.17, 1.62) p=2e-02	0.84 (-0.17, 1.85) p=1e-01	0.46 (-0.19, 1.10) p=2e-01
28 SNPs (not previously associated HDL-C or TGs at genome-wide significance)				
IVW	0.74 (0.48, 1.01) p=4e-06	0.65 (0.39, 0.92) p=3e-05	0.49 (0.23, 0.75) p=7e-04	0.25 (-0.04, 0.54) 9e-02
MR-Egger				
Standard approach	0.70 (-0.44, 1.83) p=2e-01 I^2=0.10	1.16 (0.04, 2.27) p=4e-2 I^2=0.09	0.62 (-0.53, 1.77) p=3e-01 I^2=0.10	0.30 (-1.03, 1.63) p=7e-01 I^2=0.12
SIMEX adjusted	1.01 (-0.54, 2.57) p=2e-01	1.63 (0.00, 3.25) p=5e-02	0.86 (-1.03, 2.74) p=4e-01	0.44 (-1.63, 2.51) p=7e-01
Weighted Median	0.92 (0.56, 1.29) p=6e-07	0.77 (0.39, 1.15) p=8e-05	0.55 (0.16, 0.94) p=6e-03	0.42 (0.05, 0.80) p=3e-02
Weighted MBE	1.08 (0.15, 2.00) p=3e-02	0.06 (-0.90, 1.02) p=9e-01	-0.28 (-1.50, 0.93) p=7e-01	0.50 (-0.44,1.44) p=3e-01

The results were broadly consistent after removing the outlier SNP (rs1011685). MR-Egger (3) suggests there is little directional pleiotropy, as the intercepts were of generally small magnitudes (MR-Egger intercepts ≤ 0.01) in comparison with the betas. As MR-Egger estimate of causal effect is prone to be underestimated when the assumption of no-measurement error (NOME) is violated, the heterogeneity index (I^2) was used to detect the extent of the violation and SIMEX-method was used to

correct such violation (6). The lower value of I^2 (value lies between 0 and 1) indicates stronger NOME violation and thus stronger dilution of the estimates. For mode-based estimator (5), the smoothing parameter for the empirical density function was set to 0.5, in order to increase the precision to detect potential pleiotropy. Weighted median estimates are valid as long as the majority of weights contributing to the analyses are valid (4). For all the methods, inverse-variance of the ratio estimators were used as the weights. IVW: inverse-variance weighted; GlycA: glycoprotein acetyls.

Supplemental Table 9. The intercepts for MR-Egger.

|	Egger-Intercepts	SE	P value
53 SNPs			
Ile	0.0061	0.003	0.08
Leu	0.008	0.004	0.04
Val	0.0109	0.004	0.01
GlycA	-0.0039	0.004	0.37
52 SNPs			
Ile	0.0027	0.005	0.62
Leu	0.0043	0.006	0.48
Val	0.0036	0.006	0.57
GlycA	0.0071	0.006	0.26
28 SNPs			
Ile	0.0008	0.009	0.93
Leu	-0.0083	0.01	0.41
Val	-0.0022	0.012	0.85
GlycA	-0.0008	0.013	0.95

The intercepts of MR Egger were of generally small magnitudes (intercepts ≤ 0.01, far less than the betas) with little significance level, suggesting that there is little directional pleiotropy of the genetic instruments.

Supplemental Table 10. Details of Mendelian randomization studies and estimates to accompany Figure 4.

| Ref in Figure 4	Exposure	Outcome	Mendelian randomization estimate (per-1SD higher exposure)	References
1	BMI	Fasting insulin HDL-cholesterol Triglycerides	logFI: 0.39 SD (0.29, 0.49); HDL-C: -0.22 SD (-0.29, -0.14); logTG: 0.19 SD (0.12, 0.28).	Holmes et al. (9) and Dale et al.*(10)
2	WHRadjBMI	Fasting insulin HDL-cholesterol Triglycerides	logFI: 0.33 SD (0.15, 0.51); HDL-C: -0.38 SD (-0.49, -0.28); logTG: 0.45 SD (0.28, 0.62).	Dale et al.* (10) and Emdin et al.(11)
3	BMI	BCAAs	Isoleucine: 0.20 SD (0.12, 0.28); Leucine: 0.17 SD (0.09, 0.25); Valine: 0.19 SD (0.10, 0.27).	Würtz et al. (12) and Supplemental Figure 3*
4	WHRadjBMI	BCAAs	Isoleucine: 0.21 SD (0.09, 0.33); Leucine: 0.15 SD (0.03, 0.27); Valine: 0.10 SD (-0.02, 0.22).	Supplemental Figure 4
5	BMI	T2D	OR 1.96 (1.41, 2.78)	Holmes et al. (9); Dale et al. *(10); Lyall et al. (13)
6	WHRadjBMI	T2D	OR 1.82 (1.38, 2.42)	Emdin et al. (11); Dale et al.*(10)
7	Insulin resistance[#]	BCAAs	Isoleucine: 0.56SD (0.43, 0.70); Leucine: 0.42SD (0.28, 0.55); Valine: 0.26SD (0.12, 0.39)	Figure 2*, Mahendran et al. (14)
8	BCAAs	T2D	The study provided genetic evidence for a causal link between BCAA metabolism and risk of T2D. For further details see ref (15)	Lotta et al. (16)

Where more than one reference is provided, the reference followed by * is the one used for the data reported under the "MR estimate" column. #: Insulin resistance is a composite trait proxied by higher fasting insulin (adjusted for BMI), triglycerides and lower HDL-cholesterol. logFI: logarithm of fasting insulin; SD: standard deviation

Supplemental Figure 1. The associations of the individual SNPs for the composite insulin resistance with the individual insulin resistance phenotypes.
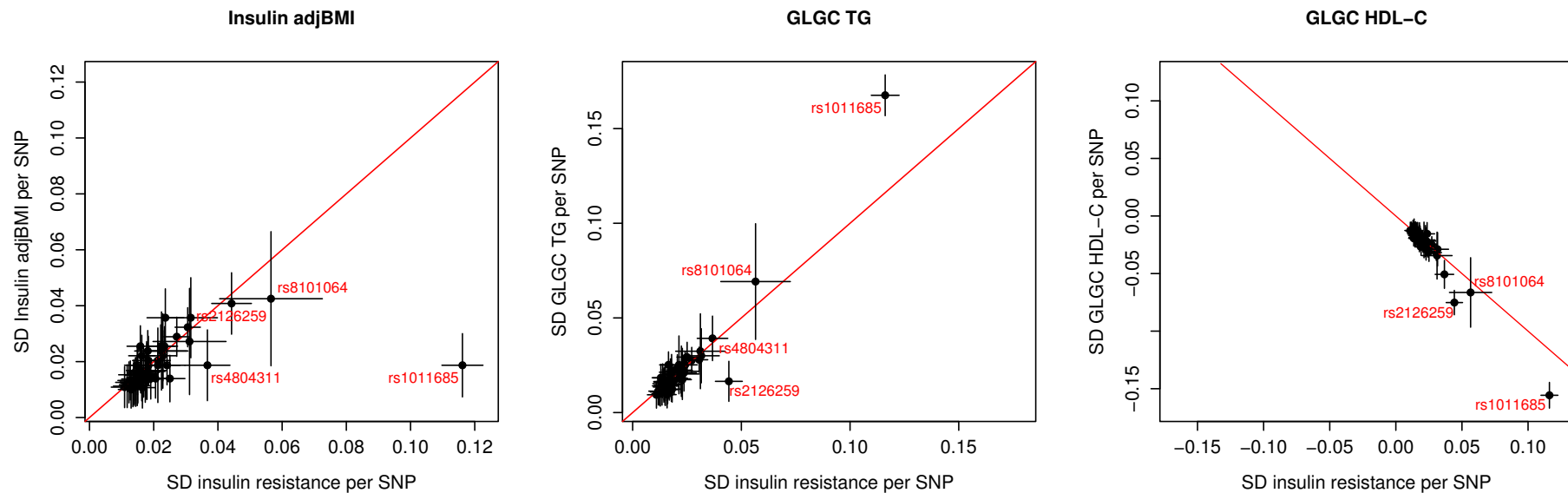

The SNP associations with insulin (BMI adjusted) was obtained from MAGIC Consortium (8). The SNP associations with TG and HDL-C were obtained from GLGC consortium (17). The SNP association with composite insulin resistance was defined as the meta-analyzed results of SNP associations with fasting insulin (BMI adjusted), TG and HDL-C. Among the 53 SNPs, most of the SNPs fell in a straight line (with a slope equal to 1), suggesting a similar contribution of the three traits to the 'composite' insulin resistance phenotype with the exception of rs1011685 (near *LPL*), which had a much weaker effect on insulin adjusted for BMI. Each dot indicates one SNP. The horizontal lines denotes the CIs for the associations of SNPs with composite insulin resistance while the vertical lines denotes the CIs for the associations of SNPs with the individual phenotypes.

Supplemental Figure 2. Causal effect estimates of fasting insulin on circulating metabolic measures.
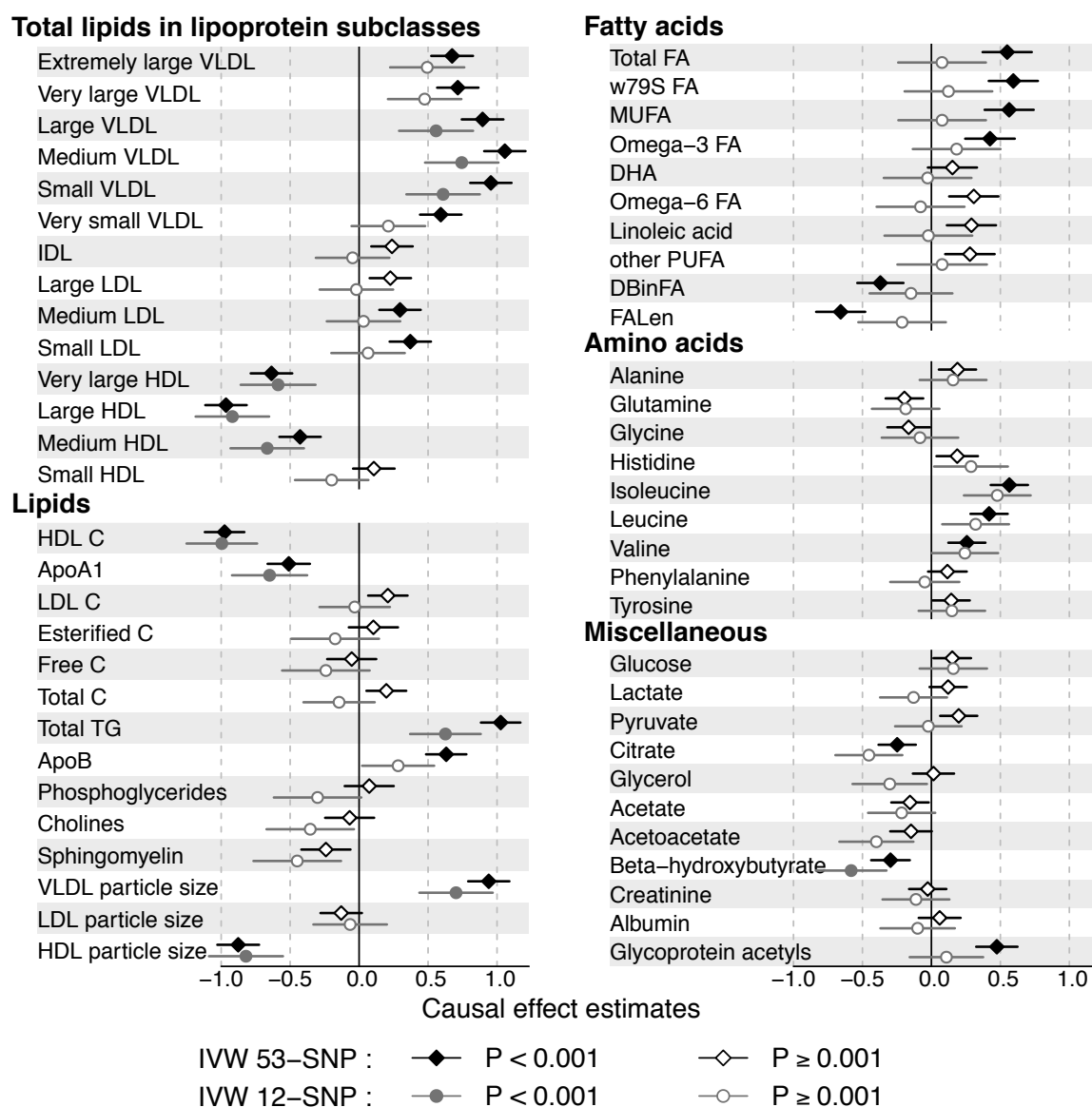
Black diamond: Causal effect estimates of insulin resistance (a composite trait denoted as increased insulin adjusted for BMI, increased TGs and decreased HDL-C) on circulating metabolic measures, using insulin resistance associated 53 SNPs as the instrument. The unit is SD difference in metabolite concentration per SD higher insulin resistance. These are also shown in Figure 2. **Grey circle:** Causal effect estimates of fasting insulin (adjusted for BMI) on the circulating metabolic measures, using the 12 loci reported by a previous GWAS as the instrument (8). The unit is SD difference in metabolite concentration per SD higher fasting insulin (adjusted for BMI). Both estimates were calculated via two-sample Mendelian randomization analysis using inverse-variance weighted method (2).

Supplemental Figure 3. Causal effect estimates of body mass index on branched-chain amino acids and glycoprotein acetyls.

BMI ⟶ Amino acids and glycoprotein acetyls
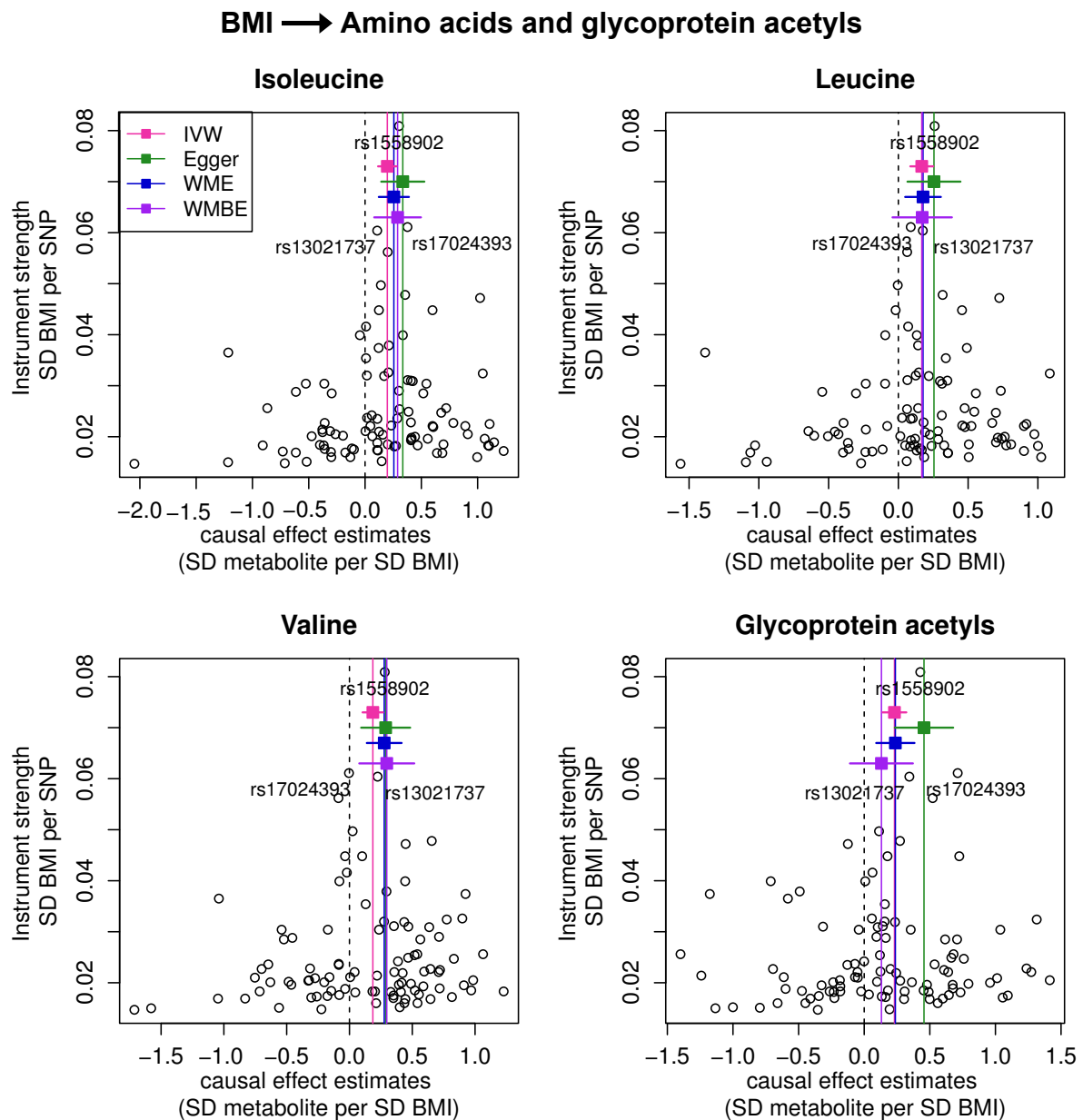
The 97 BMI-associated loci reported by previous GWA study (GIANT consortium) were used as the instrument to assess the causal effects of BMI on circulating metabolic measures. Summary statistics from this BMI-GWAS (up to 322,154 individuals, primarily European) (7) and metabolic-GWAS (up to 24,925 participants, European) (18) were used in the two-sample Mendelian randomization analysis. The causal effect estimates using 97 SNPs were calculated by inverse-variance weighted method (red vertical line), MR-Egger (green line) weighted median estimator (WME; blue line) and weighted mode-based estimator (WMBE; purple line). The 95% CIs for each method are depicted as the colourful horizontal lines. The x-axis shows the causal effect estimates using individual SNP as the instrument. The y-axis describes the association strength (beta coefficients) of individual SNPs with BMI (rs1885988 was used as the genetic proxy for rs12016871; $R^2 > 0.8$).

Supplemental Figure 4. Causal effect estimates of waist-to-hip ratio on branched-chain amino acids and glycoprotein acetyls.

Waist-to-hip ratio (adjusted for BMI) ⟶ Amino acids and glycoprotein acetyls
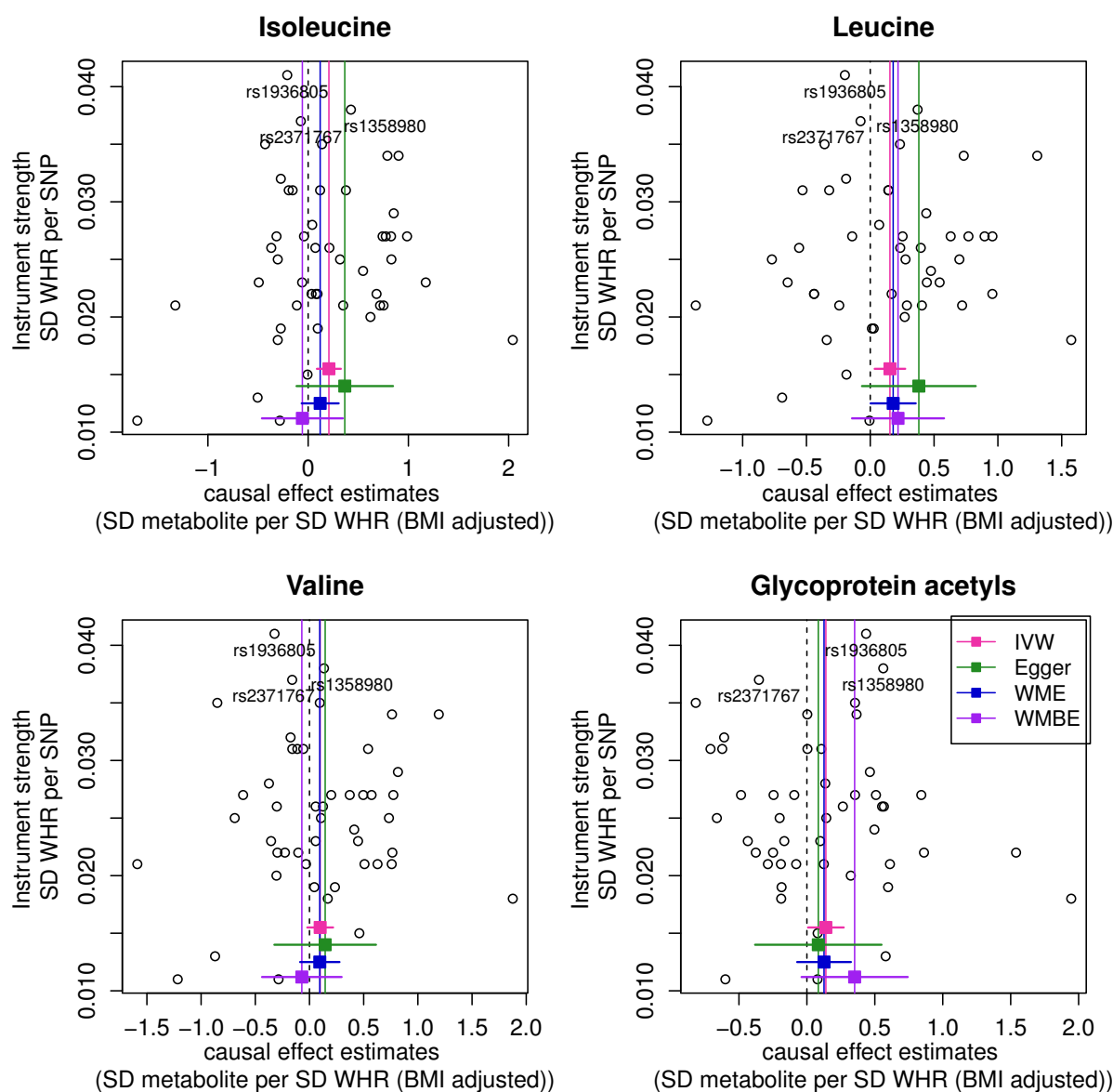
The 49 waist-to-hip ratio (WHR, BMI-adjusted)-associated loci reported by previous GWA study (GIANT consortium) were used as the instrument to assess the causal effects of WHR on circulating metabolic measures. Summary statistics from this WHR-GWAS (up to 224,459 individuals, almost all were European) (19) and metabolic-GWAS (up to 24,925 participants, European) (18) were used in the two-sample Mendelian randomization analysis. The causal effect estimates using 49 SNPs were calculated by inverse-variance weighted method (red vertical line), MR-Egger (green line), weighted median estimator (WME; blue line) and weighted mode-based estimator (WMBE; purple line). The x-axis shows the causal effect estimates using individual SNPs as the instrument. The y-axis describes the association strength (beta coefficients) of individual SNP with WHR. SNP rs7759742 was missing from the metabolic-GWAS and was therefore not included in the analysis.

References

1. Lotta LA, Gulati P, Day FR, et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. Nat Genet 2017;49(1):17–26.

2. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013;37(7):658–65.

3. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44(2):512–25.

4. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 2016;40(4):304–14.

5. Hartwig FP, Smith GD, Bowden J. Robust Inference In Two-Sample Mendelian Randomisation Via The Zero Modal Pleiotropy Assumption. Int J Epidemiol 2017; https://doi.org/10.1093/ije/dyx102.

6. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. Int J Epidemiol 2016;45:1961-74.

7. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197–206.

8. 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(9):991–1005.

9. Holmes MV, Lange LA, Palmer T, et al. Causal effects of body mass index on cardiometabolic traits and events: a Mendelian randomization analysis. Am J Hum Genet 2014;94(2):198–208.

10. Dale CE, Fatemifar G, Palmer TM, et al. Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. Circulation 2017;135(24):2373–88.

11. Emdin CA, Khera AV, Natarajan P, et al. Genetic Association of Waist-to-Hip Ratio With Cardiometabolic Traits, Type 2 Diabetes, and Coronary Heart Disease. JAMA 2017;317(6):626–34.

12. Würtz P, Wang Q, Kangas AJ, et al. Metabolic Signatures of Adiposity in Young Adults: Mendelian Randomization Analysis and Effects of Weight Change. PLoS Med 2014;11:e1001765.

13. Lyall DM, Celis-Morales C, Ward J, et al. Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank: A Mendelian Randomization Study. JAMA Cardiol 2017;2:882-889.

14. Mahendran Y, Jonsson A, Have CT, et al. Genetic evidence of a causal effect of insulin resistance on branched-chain amino acid levels. Diabetologia 2017;60(5):873–8.

15. Holmes MV, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. Nat Rev Cardiol. 2017; doi: 10.1038/nrcardio.2017.78.

16. Lotta LA, Scott RA, Sharp SJ, et al. Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. PLoS Med 2016;13:e1002179.

17. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013;45(11):1274–83.

18. Kettunen J, Demirkan A, Würtz P, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commun 2016;7:11122.

19. Shungin D, Winkler TW, Croteau-Chonka DC, et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature 2015;518(7538):187–96.