

## **Evaluating the role of amino acids in type 2 diabetes risk: a Mendelian randomization study**

Jacky Man Yuen Mo,<sup>1</sup> Baoting He,<sup>1</sup> Tommy Hon Ting Wong,<sup>1</sup> Ying Liang,<sup>1</sup> Shan Luo,<sup>1</sup>  
Kenneth Lo,<sup>2</sup> Jimmy Chun Yu Louie,<sup>3</sup> Shiu Lun Au Yeung<sup>1\*</sup>

### **Affiliation**

1. School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China
2. Department of Food Science and Nutrition, The Hong Kong Polytechnic University, Hong Kong SAR, China
3. Department of Nursing and Allied Health, School of Health Sciences, Swinburne University of Technology, Melbourne, Australia

### **\*Corresponding author**

Dr. Au Yeung Shiu Lun

School of Public Health, LKS Faculty of Medicine, The University of Hong Kong  
1/F, Patrick Manson Building, 7 Sassoon Road, Hong Kong SAR, China

Phone: 852-3917 6740

Fax: 852-3520 1945

Email: [ayslryan@hku.hk](mailto:ayslryan@hku.hk)

Number of words: 245 (Abstract); 2642 (Main text)

Number of tables: 0

Number of figures: 3

## Abstract

**Background:** Previous observational and Mendelian randomization studies suggested different amino acids associated with type 2 diabetes (T2D). However, these studies may suffer from confounding or the use of invalid instruments, respectively.

**Methods:** We extracted strong ( $p < 5 \times 10^{-8}$ ), independent ( $r^2 < 0.001$ ) genetic variants associated with nine amino acids (alanine, glutamine, glycine, histidine, phenylalanine, tyrosine, isoleucine, leucine, and valine) from summary statistics of UK Biobank ( $N \leq 115,075$ ), with exclusion of potentially pleiotropic variants. We then applied them to T2D summary statistics from DIAMANTE Consortium (without UK Biobank participants) ( $N = 455,313$ ) and FinnGen study ( $N = 365,950$ ), and glycemic traits (MAGIC consortium,  $N \leq 209,605$ ). Inverse variance weighed (IVW) method was the main analysis, with multiple sensitivity analyses to assess robustness of findings.

**Results:** Alanine was associated with higher T2D risk, correcting for multiple testing (Odds Ratio (OR) 1.50 per SD; 95% CI 1.16 to 1.95). At nominal significance, isoleucine was associated with higher T2D risk (OR 1.13; 95% CI 1.00 to 1.27) and tyrosine was associated with lower T2D risk (OR 0.89; 95% CI 0.80 to 0.99). Alanine was also associated with lower insulin, higher glycated hemoglobin and glucose whereas isoleucine and leucine were associated with lower insulin. These associations were consistent in most sensitivity analyses.

**Conclusion:** Alanine likely contributed to higher T2D risk whilst the associations for isoleucine and tyrosine requires further verification. Whether these findings explain health effects of sources of amino acids, such as diet, should be further explored.

## Introduction

Type 2 diabetes (T2D) is projected to affect over 570 million adults worldwide by 2025, with an increasing prevalence and associated burden on healthcare systems across the globe.<sup>1</sup> As such, identifying modifiable etiologic factors related to T2D would aid prevention and early diagnosis of T2D. Recent research have increasingly suggested certain amino acids, in particular branched chain amino acids, are potential risk factors of T2D. Earlier prospective cohort studies suggest alanine, tyrosine and isoleucine may increase T2D whilst glycine and serine inversely associated with T2D risk.<sup>2-4</sup> Although the exact mechanistic pathways were not fully elucidated, impairment of insulin secretion and insulin resistance may be one of these pathways.<sup>5,6</sup> However, whether these findings were causal were unclear given the possibility of confounding by socioeconomic position and lifestyle factors in these observational studies.

Mendelian randomization (MR), a design that is less vulnerable to confounding than conventional observational studies,<sup>7</sup> have been increasingly used to ascertain the causal role of metabolomic signatures associated with T2D. Previous MR studies have shown branched amino acids (BCAAs), including leucine, isoleucine and valine, may increase T2D risk whereas alanine and tyrosine, an aromatic amino acid, may decrease T2D.<sup>8-10</sup> However, these MR studies may be flawed due to the use of limited number of instruments derived from small genome wide association studies (GWAS) of amino acids,<sup>11,12</sup> and the use of a single instrument (e.g. variants in protein phosphatase Mg<sup>2+</sup>/Mn<sup>2+</sup> dependent 1K [*PPMIK*] locus) to estimate the causal relationship between multiple amino acids and T2D and hence constitute horizontal pleiotropy.<sup>9,13</sup> In order to better evaluate the role of these amino acids in T2D risk taking into account of potentially invalid instruments in early MR studies, we conducted one of the largest MR studies to date by leveraging large GWAS summary

statistics of nine amino acids covered in the Nuclear Magnetic Resonance (NMR) spectrometry platform in UK Biobank,<sup>14</sup> with exclusion of pleiotropic instruments based on NMR measured metabolites,<sup>15 16</sup> and respective GWAS of T2D and glycaemic traits in Europeans.

## Methods

A valid MR study should fulfill three instrumental variable assumptions.<sup>17</sup> First, relevance – genetic instruments are strongly associated with the exposure of interest. Second, independence – the instruments are free from confounding between the exposure and outcome. Third, exclusion restriction – the effect of instruments on the outcome must be exerted via the exposure with no alternative pathways.

### *Genetic instruments of amino acids*

The metabolic profiling of plasma samples from a randomly subset of participants (N = 118,461) from the UK Biobank was quantified using high-throughput NMR spectroscopy metabolomics (Nightingale Health Ltd., Helsinki, Finland, <https://nightingalehealth.com/>).<sup>18</sup> We obtained strong ( $p$  value  $< 5 \times 10^{-8}$ ), independent ( $r^2 < 0.001$ ) genetic predictors, single nucleotide polymorphism (SNPs), of nine circulating amino acids including BCAAs (isoleucine, leucine, and valine), aromatic amino acids (phenylalanine and tyrosine) and other amino acids (alanine, glutamine, glycine, and histidine) (**Supplementary Table 1**), generated by Integrative Epidemiology Unit, University of Bristol (<https://gwas.mrcieu.ac.uk/>).<sup>19</sup> Measurements of amino acids were inverse rank-normalized and the genetic associations of amino acids were obtained using a linear mixed model adjusted for age, sex, fasting status, and genotyping array. To reduce the impact of horizontal pleiotropy which biases MR

estimates, we removed instruments that are associated with more than five NMR metabolites (including non-amino acid metabolites) in the same UK Biobank summary statistics, as per previous studies (**Supplementary Figure 1**).<sup>15 16</sup>

### ***Genetic association of type 2 diabetes***

We obtained genetic association from the DIABetes Meta-Analysis of Trans-Ethnic association studies (DIAMANTE) consortium, which is the largest genetic consortium of T2D at the time of analysis. To avoid possible bias arising from overlap of samples used in the amino acid GWAS and the T2D GWAS, we restricted the analyses without the UK Biobank samples, which included 55,005 cases and 400,308 controls of European.<sup>20</sup> T2D was defined differently across studies, such as blood sugar level (glycated hemoglobin [HbA<sub>1c</sub>] ≥ 6.5%, fasting glucose [FG] ≥ 7.0 mmol/L or 2-hour glucose [2hGlu] ≥ 11.1 mmol/L), the use of anti-diabetic medication and physician diagnosis. The genetic associations were obtained using multivariable linear regression, adjusted for age, sex, and the top 10 genetic principal components (PC).

We also obtained genetic associations of T2D from FinnGen study (Round 9), a combination of Finnish Biobanks, which consists of 57,698 T2D cases and 308,252 controls participants.<sup>21</sup> T2D was defined based on hospital record linkage and the genetic associations of T2D were obtained using multivariable linear regression adjusted for age, sex and the top 10 PCs.

### ***Genetic association of glycemic traits***

We obtained genetic associations of FI (log-transformed), HbA<sub>1c</sub> (%), FG (mmol/L) and 2hGlu (mmol/L) from the Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC), restricted to participants of European descent (N ≤ 209,605).<sup>22</sup> Participants with type 1 or type 2 diabetes, who reported taking anti-diabetic medications or had FG ≥ 7

mmol/L, 2hGlu  $\geq$  11.1 mmol/L, or HbA<sub>1c</sub>  $\geq$  6.5% were excluded.<sup>22</sup> The genetic associations of glycemic traits were obtained using multivariable linear regression adjusted for age, sex, body mass index (except for HbA<sub>1c</sub>), study-specific covariates and PCs. Although adjustment for covariate in GWAS may introduce collider bias, the original GWAS investigators did not find strong evidence of such bias.<sup>22 23</sup>

### ***Main analysis***

We assessed the instrument strength using F statistic via approximation, where F statistic  $>10$  indicated low evidence for weak instrument bias.<sup>24</sup> We calculated the variance explained ( $R^2$ ) of each amino acid by summing up  $R^2$  of individual instrument, which was calculated using effect allele frequencies, effect estimates and corresponding standard errors, as per a previous study.<sup>25 26</sup> We obtained the causal estimates of amino acids with T2D using inverse-variance weighted (IVW) method with multiplicative random effects, which assumed no overall horizontal pleiotropy.<sup>27</sup> We assessed  $I^2$  for instrument heterogeneity where a high  $I^2$  indicated the possibility of invalid instruments. To increase the statistical power, we meta-analyzed the IVW estimates from DIAMANTE and FinnGen using random effect meta-analysis.

### *Sensitivity analyses*

To triangulate the evidence from MR methods which relied on different assumptions, we also conducted MR-Egger, which allows for genetic pleiotropy but requires the instrument strength independent of direct effect (InSIDE) assumption;<sup>28</sup> weighted median method, which assumes at least half of weights were derived from valid instruments,<sup>29</sup> and MR-Pleiotropy RESidual Sum and Outlier (PRESSO), which removed possible invalid instruments via the MR-PRESSO outlier test.<sup>30</sup> We assessed overall horizontal pleiotropy using MR-Egger intercept test.<sup>28</sup>

To correct for multiple testing whilst taking into account the correlated nature of the amino acid, we calculated the principal components (PCs) needed to explain 99% of the variance of the nine amino acid exposures, using individual level data of our ongoing UK Biobank project (Application number: 14864), which gave 7 PCs. Statistical significance was therefore set as  $p < 0.0071$  ( $0.05/7$ ).

All analyses were performed using “TwoSampleMR”, “MendelianRandomization” and “MR-PRESSO” package on R (version 4.3.1).

### *Ethics approval*

This study only used publicly available summary data and hence no ethics approval is required. Details of ethical approval and participant consent for each of the studies that contributed to the GWAS can be found in the original publications.<sup>19-22</sup>

## Results

**Supplementary Table 2** shows the list of instruments for each amino acid. All instruments had an F statistic of  $>10$ , indicating low level of evidence for weak instruments bias. The F-statistics and  $R^2$  for each amino acid are as follow: alanine (30-250, 1.91%), glutamine (28-2209, 4.51%), glycine (26-377, 3.16%), histidine (35-157, 0.77%), phenylalanine (30-664, 1.13%), tyrosine (30-1027, 2.63%), isoleucine (33-267, 0.60%), leucine (32-488, 0.93%) and valine (30-673, 1.58%). Variants which were excluded due to associations with more than 5 metabolites were shown in **Supplementary Table 3**.

### *Association of genetically predicted circulating amino acids with risk of type 2 diabetes*

**Figure 1** shows the association between genetically predicted amino acids and T2D risk. After combining the IVW estimates from DIAMANTE and FinnGen, per standard deviation (SD) increase in alanine (Odds ratio (OR) 1.50; 95% CI 1.16 to 1.95) was associated with higher T2D risk, correcting for multiple testing. At nominal significance, histidine (OR 1.32; 95% CI 1.04 to 1.66) and isoleucine (OR 1.13; 95% CI 1.00 to 1.27) were associated with higher T2D risk. Conversely, genetically predicted tyrosine (OR 0.89; 95% CI 0.80 to 0.99) was associated with lower T2D risk. These findings were generally consistent across both DIAMANTE and FinnGen (**Supplementary Table 4**).

**Supplementary Table 5** and **6** show the MR analyses using DIAMANTE and FinnGen, respectively. There were signs of heterogeneity based on  $I^2$  for most analyses. Notably, horizontal pleiotropy was evident for histidine, implying possible biases in IVW. Sensitivity analyses gave directionally similar estimates apart from MR-Egger regarding alanine- and histidine-associated analyses.



### *Association of genetically predicted circulating amino acids with glycemc traits*

**Figures 2** and **3** show the association of genetically predicted amino acids with glycemc traits, including FI, FG, HbA<sub>1c</sub> and 2hGlu. Using inverse-variance weighted, alanine (Beta [ $\beta$ ] -0.06; 95% CI -0.09 to -0.03) and two BCAAs (isoleucine [ $\beta$ ] -0.07; 95% CI -0.11 to -0.03] and leucine [ $\beta$ ] -0.05; 95% CI -0.08 to -0.02]) were associated with lower FI, (**Figure 2**, left panel). Alanine was also associated with higher HbA<sub>1c</sub> ( $\beta$  0.06; 95% CI 0.02 to 0.11) (**Figure 2**, right panel) and FG ( $\beta$  0.14; 95% CI 0.06 to 0.23) (**Figure 3**, left panel). There were signs of heterogeneity although evidence for horizontal pleiotropic effects was limited. Sensitivity analyses generally gave directionally similar estimates (**Supplementary Table 7-10**).

### **Discussion**

To the best of our knowledge, this is one of the first MR studies which comprehensively assessed the role of nine amino acids in T2D risk and glycemc traits using respective large genetic consortia where earlier MR studies have only investigated BCAAs and tyrosine, and may suffer from methodological issues such as the use of smaller amino acid GWAS, the use of a single variant for investigation of multiple BCAAs.<sup>8-10</sup> We showed, for the first-time using MR, that alanine likely increases T2D risk, and hence confirmed the positive findings in previous observational studies.<sup>31-33</sup> Consistent with previous studies, our study showed that isoleucine (one of the BCAAs) likely increases T2D risk.<sup>9,34</sup> However, the positive association between histidine and T2D risk was likely due to horizontal pleiotropy. In addition, we found tyrosine, an aromatic amino acid, to be inversely associated with T2D risk, which was consistent with a previous MR study but not observational studies.<sup>8,35</sup>

The positive association between alanine and glycemc traits concurred with the findings from earlier observational studies,<sup>31-33</sup> and our study extended by showing an increased risk of T2D

with higher alanine using a design less vulnerable to confounding. This is in contrary to an earlier study which showed an inverse association of alanine on T2D risk based on a single variant<sup>10</sup> which is vulnerable to horizontal pleiotropy. The potential harm of alanine in glucose homeostasis was also consistent with its association with increased fasting glucose and reduced insulin, consistent with a previous prospective cohort study.<sup>2</sup> Although how alanine impacts T2D risk has not been thoroughly explored, alanine is involved in the glucose-alanine cycle where alanine is converted into pyruvate by alanine aminotransferase in the liver and then to glucose via the gluconeogenic pathway.<sup>36,37</sup> Hence, it is postulated that elevated alanine may lead to hyperglycemia and reduce insulin secretion, subsequently induces insulin resistance,<sup>38</sup> a hallmark feature of T2D. Alternatively, other studies suggested elevated alanine amongst people with non-alcoholic fatty liver disease,<sup>32,39</sup> which in itself likely increases T2D risk.<sup>40</sup> Additional studies would be necessary to clarify the role of alanine in these two related diseases. Our study also helped clarify that phenylalanine unlikely increases T2D risk, in contrary to a small prospective cohort study.<sup>41</sup>

The positive association of BCAAs with T2D risk was well studied in observational studies.<sup>3</sup><sup>35,41,42</sup> With the inclusion of more genetic instruments, our study clarified that amongst the BCAAs, isoleucine may be more relevant to the development of T2D, consistent with findings from various analyses and data sources,<sup>9</sup> although it is noted that the association did not pass correction for multiple testing and hence requires replications. Nevertheless, previous studies suggested that BCAAs could lead to persistent mammalian target of rapamycin complex 1 (mTORC1) activation, causing insulin resistance.<sup>5</sup> On the other hand, others speculated that elevated BCAAs reflects impaired BCAA metabolism and hence the accumulation of mitotoxic metabolites (BCAA dysmetabolism model).<sup>5</sup> Although an earlier MR study found no association between BCAAs and FI, our study suggests an inverse

association for isoleucine and leucine.<sup>43</sup> Such inconsistency in findings might be explained by the use of smaller sample size and insufficient considerations on potential pleiotropy when constructing BCAAs instruments in the MR analysis.<sup>43</sup> Conversely, other MR studies suggested insulin resistance leads to elevated BCAAs.<sup>44</sup> Additional studies, such as studying metabolic profiles of T2D risk variants related to BCAAs (e.g. *PPMIK*) or interactions with lipids,<sup>5 45</sup> would help disentangle the etiologic role of BCAAs in T2D risk.

The inverse association between tyrosine and T2D risk was inconsistent with earlier observational studies,<sup>35 41</sup> but in line with another MR study.<sup>8</sup> According to observational study, tyrosine is associated with reduced insulin sensitivity and insulin secretion,<sup>2</sup> and hence contributes to increased T2D risk. However, our MR study suggested little evidence for an association of tyrosine in insulin. This is somewhat also consistent with one observational study showing little impact on the association of tyrosine in T2D upon adjustment for insulin despite the reported positive association with T2D risk.<sup>41</sup> Nevertheless, these inconsistent findings may either imply confounding and biases in earlier studies, or issues with external validity to humans from animal studies as commented elsewhere.<sup>46</sup> Having said that, animal studies may help identify possible pathways, such as postulated dopamine pathway in animal study,<sup>47</sup> to be tested and may inform the mechanisms concerning tyrosine in T2D risk and reconcile the discrepancies between observational and MR studies.

From the public health perspective, our study provides potentially more credible evidence that alanine and isoleucine may increase T2D risk. Although it remains unclear whether increase in circulating amino acids is the consequence of high amino acid intake or a dysfunction in amino acid metabolism, or both, a prospective study indicated that BCAAs may partially mediate the association of meat consumption in T2D risk.<sup>48</sup> Future dietary

guidelines and recommendations may look into alanine and isoleucine intake and supplements given their potential to increase T2D risk. However, since our study focused on the general population, more research is needed to explore any potential effect modification by baseline characteristics, such as individuals with BCAA deficiency and obese individuals.

Although our study is less vulnerable to residual confounding than conventional observational studies, there are some limitations. First, MR studies have three main assumptions for valid causal inference. Although we used strong and independent genetic instruments from large GWAS to reduce the risk of weak instrument bias and genetics were less likely to be confounded, violation of exclusion restriction assumption was possible and hence may have explained the discrepancies between the MR and observational studies.

However, we excluded instruments exhibiting horizontal pleiotropy by screening the instruments' association with other NMR metabolomic markers, as well as performing sensitivity analyses which relied on other assumptions and showed similar findings.

Second, we did not use the most updated T2D GWAS from DIAMANTE Consortium as this most updated version did not provide summary statistics without UK Biobank,<sup>49</sup> and hence could introduce possible sample overlap between the GWAS of amino acid and T2D and induce biases. Third, we did not investigate sex specific effects due to the use of summary statistics. As such, follow up studies using individual level data can be conducted to explore the shape of the associations for better targeted interventions, such as sex specific dietary guidelines. Lastly, we only used data derived from European populations. Whether these associations apply to other populations, such as Asians, require further investigations.

## **Conclusion**

Overall, this study highlighted the causal role of alanine, and potentially, isoleucine in increased T2D risk, which may help inform corresponding dietary recommendations. The paradoxical inverse association of tyrosine with T2D risk requires further investigation.

**Acknowledgment:** We would like to thank MRC Integrative Epidemiology Unit at the University of Bristol for providing the summary statistics of NMR metabolites on IEU GWAS database, and Dr Carolina Borges for the discussion on the respective data. We thank the participants and investigators of DIAMANTE consortium and FinnGen study for providing the summary statistics for T2D. We thank MAGIC for providing summary statistics for FG, 2hGlu, HbA<sub>1c</sub> and FI. Part of this research (Calculation of principal components) has been conducted using the UK Biobank Resource (Application number: 14864).

**Availability of data and materials:**

The summary statistics of NMR metabolites are available on IEU GWAS database ([https://gwas.mrcieu.ac.uk/datasets/?gwas\\_id\\_icontains=met-d](https://gwas.mrcieu.ac.uk/datasets/?gwas_id_icontains=met-d)). The summary statistics for T2D of DIAMANTE consortium and FinnGen study are available on (<https://diagram-consortium.org/downloads.html>) and ([https://risteys.finregistry.fi/endpoints/E4\\_DM2](https://risteys.finregistry.fi/endpoints/E4_DM2)), respectively. The summary statistics of glyceic traits are available from the MAGIC (<https://www.magicinvestigators.org>).

**Conflict of interests:** none.

**Funding:** This study was partly funded by the Health and Medical Research Fund, Health Bureau, Hong Kong Special Administrative Region, China (CFS-HKU-1).

**Authors' contributions:** SLAY conceptualized the study. JMYM conducted the analysis with feedback from SLAY. YL cross checked the analyses. SL calculated the PCs required to correct for multiple testing using UK Biobank data. JMYM and SLAY interpreted the results. JMYM wrote the first draft of the manuscript with critical feedback and revisions from BH,

TWHT, SL, KL, JCYL, and SLAY. All authors gave final approval of the version to be published. JMYM is the guarantor.

## REFERENCES

1. Lin X, Xu Y, Pan X, et al. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. *Scientific Reports* 2020;10(1):14790. doi: 10.1038/s41598-020-71908-9
2. Vangipurapu J, Stancáková A, Smith U, et al. Nine Amino Acids Are Associated With Decreased Insulin Secretion and Elevated Glucose Levels in a 7.4-Year Follow-up Study of 5,181 Finnish Men. *Diabetes* 2019;68(6):1353-58. doi: 10.2337/db18-1076 [published Online First: 20190318]
3. Ahola-Olli AV, Mustelin L, Kalimeri M, et al. Circulating metabolites and the risk of type 2 diabetes: a prospective study of 11,896 young adults from four Finnish cohorts. *Diabetologia* 2019;62(12):2298-309. doi: 10.1007/s00125-019-05001-w [published Online First: 20191004]
4. Wilson PW, Meigs JB, Sullivan L, et al. Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Arch Intern Med* 2007;167(10):1068-74. doi: 10.1001/archinte.167.10.1068
5. Lynch CJ, Adams SH. Branched-chain amino acids in metabolic signalling and insulin resistance. *Nat Rev Endocrinol* 2014;10(12):723-36. doi: 10.1038/nrendo.2014.171 [published Online First: 20141007]
6. Yan-Do R, MacDonald PE. Impaired “Glycine”-mia in Type 2 Diabetes and Potential Mechanisms Contributing to Glucose Homeostasis. *Endocrinology* 2017;158(5):1064-73. doi: 10.1210/en.2017-00148
7. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *Bmj* 2018;362:k601. doi: 10.1136/bmj.k601 [published Online First: 2018/07/14]
8. Jäger S, Cuadrat R, Wittenbecher C, et al. Mendelian Randomization Study on Amino Acid Metabolism Suggests Tyrosine as Causal Trait for Type 2 Diabetes. *Nutrients* 2020;12(12) doi: 10.3390/nu12123890 [published Online First: 20201219]
9. 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(11):e1002179. doi: 10.1371/journal.pmed.1002179 [published Online First: 20161129]
10. Yuan S, Larsson SC. An atlas on risk factors for type 2 diabetes: a wide-angled Mendelian randomisation study. *Diabetologia* 2020;63(11):2359-71. doi: 10.1007/s00125-020-05253-x [published Online First: 20200908]
11. Shin SY, Fauman EB, Petersen AK, et al. An atlas of genetic influences on human blood metabolites. *Nat Genet* 2014;46(6):543-50. doi: 10.1038/ng.2982 [published Online First: 20140511]
12. 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. doi: 10.1038/ncomms11122 [published Online First: 20160323]
13. Holmes MV, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat Rev Cardiol* 2017;14(10):577-90. doi: 10.1038/nrcardio.2017.78 [published Online First: 20170601]
14. Julkunen H, Cichońska A, Tiainen M, et al. Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nature Communications* 2023;14(1):604. doi: 10.1038/s41467-023-36231-7
15. Zheng J, Haberland V, Baird D, et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nature Genetics* 2020;52(10):1122-31. doi: 10.1038/s41588-020-0682-6



16. Karjalainen MK, Karthikeyan S, Oliver-Williams C, et al. Genome-wide characterization of circulating metabolic biomarkers reveals substantial pleiotropy and novel disease pathways. *medRxiv* 2022:2022.10.20.22281089. doi: 10.1101/2022.10.20.22281089
17. Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nature Reviews Methods Primers* 2022;2(1):6. doi: 10.1038/s43586-021-00092-5
18. Würtz P, Kangas AJ, Soininen P, et al. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies. *Am J Epidemiol* 2017;186(9):1084-96. doi: 10.1093/aje/kwx016
19. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562(7726):203-09. doi: 10.1038/s41586-018-0579-z
20. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nature Genetics* 2018;50(11):1505-13. doi: 10.1038/s41588-018-0241-6
21. Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* 2023;613(7944):508-18. doi: 10.1038/s41586-022-05473-8
22. Chen J, Spracklen CN, Marenne G, et al. The trans-ancestral genomic architecture of glycemic traits. *Nat Genet* 2021;53(6):840-60. doi: 10.1038/s41588-021-00852-9 [published Online First: 20210531]
23. 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. *Nature Genetics* 2012;44(9):991-1005. doi: 10.1038/ng.2385
24. Pierce BL, Ahsan H, Vanderweele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int J Epidemiol* 2011;40(3):740-52. doi: 10.1093/ije/dyq151 [published Online First: 20100902]
25. Yarmolinsky J, Bonilla C, Haycock PC, et al. Circulating Selenium and Prostate Cancer Risk: A Mendelian Randomization Analysis. *JNCI: Journal of the National Cancer Institute* 2018;110(9):1035-38. doi: 10.1093/jnci/djy081
26. Lee SH, Goddard ME, Wray NR, et al. A better coefficient of determination for genetic profile analysis. *Genet Epidemiol* 2012;36(3):214-24. doi: 10.1002/gepi.21614
27. Bowden J, Del Greco MF, Minelli C, et al. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med* 2017;36(11):1783-802. doi: 10.1002/sim.7221 [published Online First: 20170123]
28. 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. doi: 10.1093/ije/dyv080 [published Online First: 20150606]
29. Bowden J, Davey Smith G, Haycock PC, et al. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* 2016;40(4):304-14. doi: 10.1002/gepi.21965 [published Online First: 2016/04/07]
30. Verbanck M, Chen CY, Neale B, et al. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet* 2018;50(5):693-98. doi: 10.1038/s41588-018-0099-7 [published Online First: 20180423]
31. Long J, Yang Z, Wang L, et al. Metabolite biomarkers of type 2 diabetes mellitus and pre-diabetes: a systematic review and meta-analysis. *BMC Endocrine Disorders* 2020;20(1):174. doi: 10.1186/s12902-020-00653-x

32. Würtz P, Tiainen M, Mäkinen VP, et al. Circulating metabolite predictors of glycemia in middle-aged men and women. *Diabetes Care* 2012;35(8):1749-56. doi: 10.2337/dc11-1838 [published Online First: 20120504]
33. Semba RD, Gonzalez-Freire M, Moaddel R, et al. Altered Plasma Amino Acids and Lipids Associated With Abnormal Glucose Metabolism and Insulin Resistance in Older Adults. *J Clin Endocrinol Metab* 2018;103(9):3331-39. doi: 10.1210/jc.2018-00480
34. Flores-Guerrero JL, Osté MCJ, Kieneker LM, et al. Plasma Branched-Chain Amino Acids and Risk of Incident Type 2 Diabetes: Results from the PREVENTD Prospective Cohort Study. *J Clin Med* 2018;7(12) doi: 10.3390/jcm7120513 [published Online First: 20181204]
35. Chen S, Akter S, Kuwahara K, et al. Serum amino acid profiles and risk of type 2 diabetes among Japanese adults in the Hitachi Health Study. *Scientific Reports* 2019;9(1):7010. doi: 10.1038/s41598-019-43431-z
36. Felig P. The glucose-alanine cycle. *Metabolism - Clinical and Experimental* 1973;22(2):179-207. doi: 10.1016/0026-0495(73)90269-2
37. Felig P, Pozefsk T, Marlis E, et al. Alanine: Key Role in Gluconeogenesis. *Science* 1970;167(3920):1003-04. doi: doi:10.1126/science.167.3920.1003
38. Giri B, Dey S, Das T, et al. Chronic hyperglycemia mediated physiological alteration and metabolic distortion leads to organ dysfunction, infection, cancer progression and other pathophysiological consequences: An update on glucose toxicity. *Biomedicine & Pharmacotherapy* 2018;107:306-28. doi: <https://doi.org/10.1016/j.biopha.2018.07.157>
39. Gaggini M, Carli F, Rosso C, et al. Altered amino acid concentrations in NAFLD: Impact of obesity and insulin resistance. *Hepatology* 2018;67(1):145-58. doi: <https://doi.org/10.1002/hep.29465>
40. Au Yeung SL, Borges MC, Wong THT, et al. Evaluating the role of non-alcoholic fatty liver disease in cardiovascular diseases and type 2 diabetes: a Mendelian randomization study in Europeans and East Asians. *Int J Epidemiol* 2023;52(3):921-31. doi: 10.1093/ije/dyac212
41. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. *Nature Medicine* 2011;17(4):448-53. doi: 10.1038/nm.2307
42. Cuomo P, Capparelli R, Iannelli A, et al. Role of Branched-Chain Amino Acid Metabolism in Type 2 Diabetes, Obesity, Cardiovascular Disease and Non-Alcoholic Fatty Liver Disease. *Int J Mol Sci* 2022;23(8) doi: 10.3390/ijms23084325 [published Online First: 20220413]
43. 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):873-78. doi: 10.1007/s00125-017-4222-6 [published Online First: 20170210]
44. Wang Q, Holmes MV, Davey Smith G, et al. Genetic Support for a Causal Role of Insulin Resistance on Circulating Branched-Chain Amino Acids and Inflammation. *Diabetes Care* 2017;40(12):1779-86. doi: 10.2337/dc17-1642 [published Online First: 20171018]
45. Newgard CB. Interplay between lipids and branched-chain amino acids in development of insulin resistance. *Cell Metab* 2012;15(5):606-14. doi: 10.1016/j.cmet.2012.01.024
46. Pound P, Bracken MB. Is animal research sufficiently evidence based to be a cornerstone of biomedical research? *BMJ : British Medical Journal* 2014;348:g3387. doi: 10.1136/bmj.g3387

47. Korner J, Cline GW, Slifstein M, et al. A role for foregut tyrosine metabolism in glucose tolerance. *Molecular Metabolism* 2019;23:37-50. doi: <https://doi.org/10.1016/j.molmet.2019.02.008>
48. Isanejad M, LaCroix AZ, Thomson CA, et al. Branched-chain amino acid, meat intake and risk of type 2 diabetes in the Women's Health Initiative. *Br J Nutr* 2017;117(11):1523-30. doi: 10.1017/s0007114517001568
49. Mahajan A, Spracklen CN, Zhang W, et al. Multi-ancestry genetic study of type 2 diabetes highlights the power of diverse populations for discovery and translation. *Nature Genetics* 2022;54(5):560-72. doi: 10.1038/s41588-022-01058-3

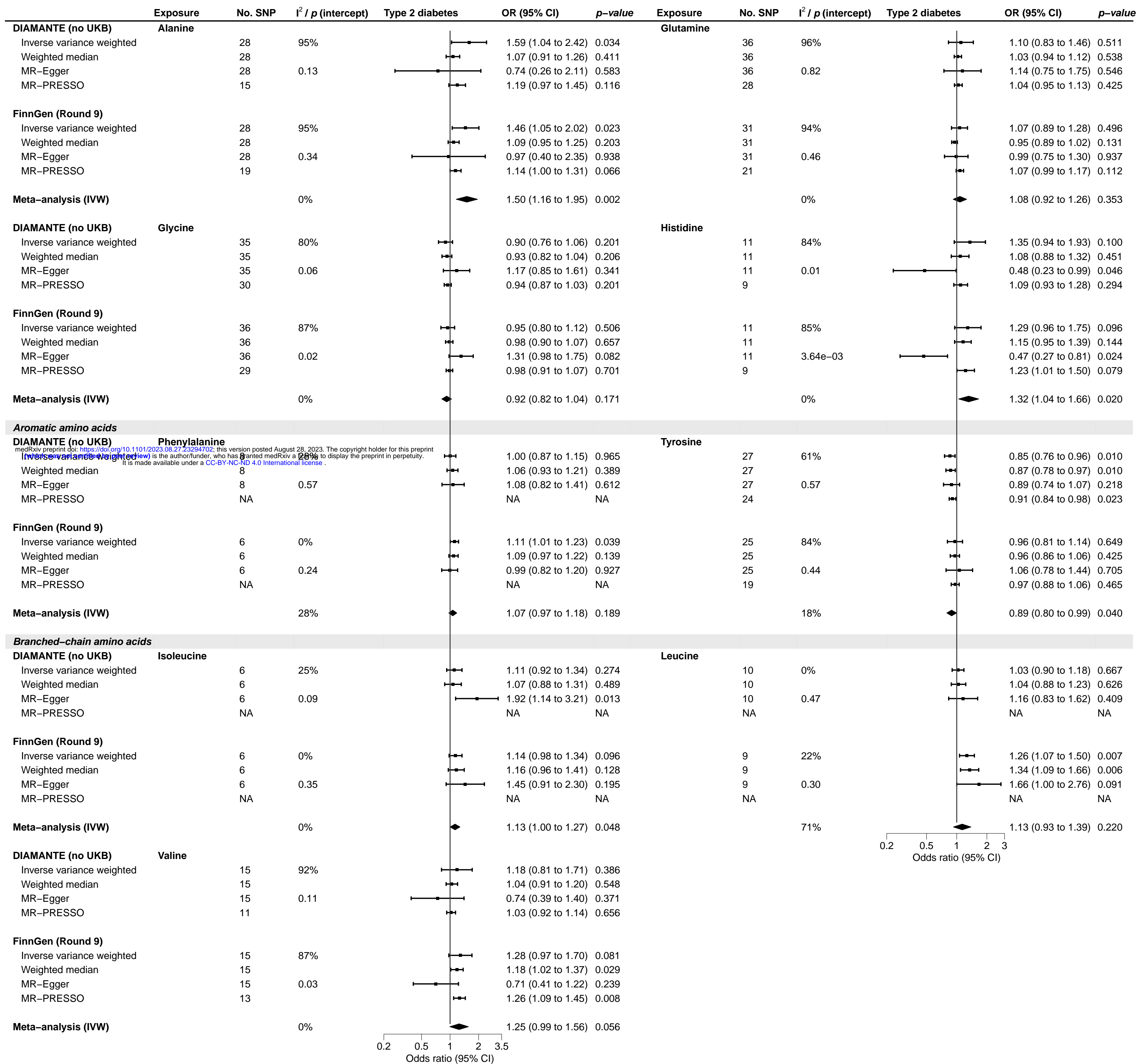
## **Figure legend**

Figure 1: The association of genetically predicted amino acids (per SD) with type 2 diabetes (T2D) risk

Figure 2: The association of genetically predicted amino acids (per SD) with fasting insulin and glycated hemoglobin

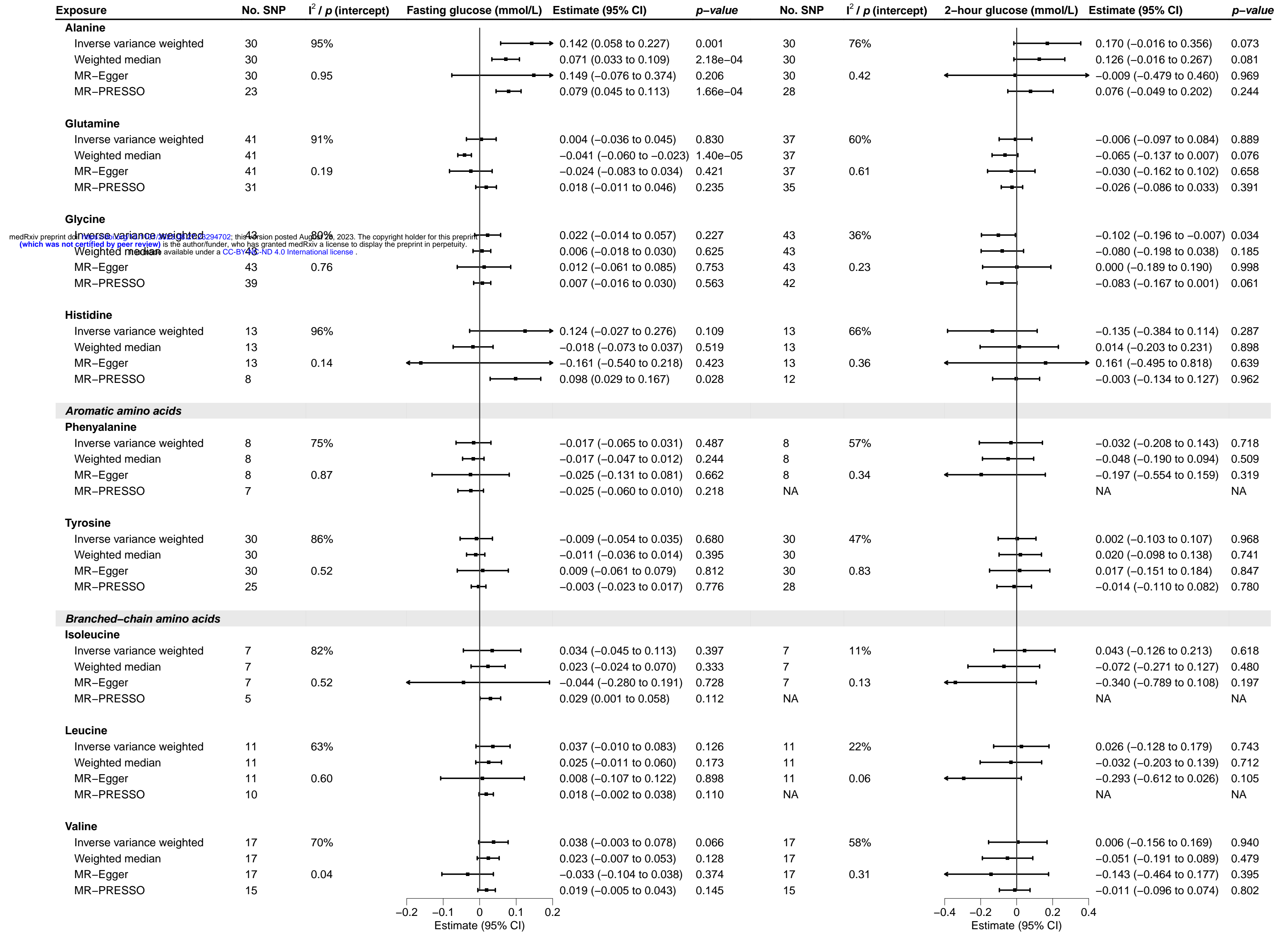
Figure 3: The association of nine genetically predicted amino acids (per SD) with fasting glucose and 2-hour glucose

Supplementary Figure 1: Summary of study design



Exposure	No. SNP	I <sup>2</sup> / p (intercept)	Fasting insulin (log-transformed)	Estimate (95% CI)	p-value	No. SNP	I <sup>2</sup> / p (intercept)	Glycated hemoglobin (%)	Estimate (95% CI)	p-value
<b>Alanine</b>										
Inverse variance weighted	30	52%		-0.059 (-0.092 to -0.027)	3.78e-04	30	89%		0.065 (0.019 to 0.110)	0.005
Weighted median	30			-0.060 (-0.095 to -0.024)	0.001	30			0.019 (-0.010 to 0.047)	0.195
MR-Egger	30	0.55		-0.035 (-0.121 to 0.051)	0.436	30	0.16		-0.016 (-0.134 to 0.101)	0.791
MR-PRESSO	29			-0.051 (-0.079 to -0.024)	0.001	22			0.042 (0.021 to 0.063)	0.001
<b>Glutamine</b>										
Inverse variance weighted	41	82%		0.002 (-0.030 to 0.034)	0.905	41	83%		0.013 (-0.009 to 0.035)	0.251
Weighted median	41			0.020 (0.001 to 0.039)	0.037	41			0.004 (-0.009 to 0.017)	0.528
MR-Egger	41	0.13		0.028 (-0.018 to 0.074)	0.239	41	0.89		0.015 (-0.018 to 0.047)	0.387
MR-PRESSO	34			-0.005 (-0.038 to 0.027)	0.753	36			0.004 (-0.010 to 0.018)	0.575
<b>Glycine</b>										
Inverse variance weighted	43	86%		-0.027 (-0.076 to 0.022)	0.273	43	60%		-0.004 (-0.023 to 0.015)	0.690
Weighted median	43			-0.011 (-0.040 to 0.018)	0.465	43			-0.006 (-0.023 to 0.012)	0.507
MR-Egger	43	0.17		0.034 (-0.064 to 0.131)	0.504	43	0.19		0.019 (-0.019 to 0.056)	0.339
MR-PRESSO	39			-0.016 (-0.038 to 0.007)	0.179	41			-0.008 (-0.024 to 0.008)	0.315
<b>Histidine</b>										
Inverse variance weighted	13	93%		0.044 (-0.085 to 0.174)	0.504	13	66%		0.028 (-0.011 to 0.067)	0.160
Weighted median	13			-0.001 (-0.059 to 0.058)	0.979	13			0.010 (-0.027 to 0.047)	0.591
MR-Egger	13	0.37		-0.109 (-0.452 to 0.234)	0.547	13	0.06		-0.061 (-0.153 to 0.031)	0.217
MR-PRESSO	9			0.034 (-0.025 to 0.094)	0.294	11			0.008 (-0.021 to 0.036)	0.604
<b>Aromatic amino acids</b>										
<b>Phenylalanine</b>										
Inverse variance weighted	8	44%		-0.032 (-0.068 to 0.004)	0.084	8	48%		0.023 (-0.001 to 0.047)	0.065
Weighted median	8			-0.040 (-0.073 to -0.007)	0.017	8			0.021 (-0.000 to 0.041)	0.054
MR-Egger	8	0.50		-0.056 (-0.132 to 0.020)	0.200	8	0.63		0.011 (-0.042 to 0.063)	0.706
MR-PRESSO	NA			NA	NA	NA			NA	NA
<b>Tyrosine</b>										
Inverse variance weighted	30	62%		-0.004 (-0.035 to 0.026)	0.782	30	71%		-0.014 (-0.037 to 0.009)	0.225
Weighted median	30			-0.005 (-0.033 to 0.023)	0.727	30			-0.013 (-0.031 to 0.005)	0.154
MR-Egger	30	0.60		-0.014 (-0.063 to 0.034)	0.561	30	0.36		-0.027 (-0.063 to 0.008)	0.146
MR-PRESSO	26			-0.012 (-0.034 to 0.010)	0.308	28			-0.010 (-0.027 to 0.007)	0.251
<b>Branched-chain amino acids</b>										
<b>Isoleucine</b>										
Inverse variance weighted	7	0%		-0.069 (-0.107 to -0.031)	3.80e-04	7	24%		-0.008 (-0.036 to 0.020)	0.572
Weighted median	7			-0.061 (-0.109 to -0.013)	0.012	7			-0.018 (-0.049 to 0.013)	0.267
MR-Egger	7	0.85		-0.080 (-0.187 to 0.028)	0.207	7	0.26		-0.054 (-0.131 to 0.023)	0.225
MR-PRESSO	NA			NA	NA	NA			NA	NA
<b>Leucine</b>										
Inverse variance weighted	11	0%		-0.051 (-0.083 to -0.019)	0.002	11	23%		-0.003 (-0.027 to 0.021)	0.806
Weighted median	11			-0.047 (-0.087 to -0.007)	0.022	11			-0.010 (-0.036 to 0.016)	0.447
MR-Egger	11	0.73		-0.038 (-0.114 to 0.038)	0.348	11	0.08		-0.048 (-0.098 to 0.001)	0.089
MR-PRESSO	NA			NA	NA	NA			NA	NA
<b>Valine</b>										
Inverse variance weighted	17	92%		0.008 (-0.079 to 0.095)	0.861	17	59%		-0.002 (-0.028 to 0.024)	0.881
Weighted median	17			-0.032 (-0.065 to 0.002)	0.063	17			-0.011 (-0.033 to 0.011)	0.325
MR-Egger	17	0.05		-0.136 (-0.293 to 0.020)	0.108	17	0.13		-0.036 (-0.084 to 0.012)	0.166
MR-PRESSO	15			-0.007 (-0.052 to 0.037)	0.758	16			-0.007 (-0.027 to 0.013)	0.496

medRxiv preprint doi: <https://doi.org/10.1101/2023.08.27.23294705>; this version posted August 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.



medRxiv preprint doi: <https://doi.org/10.1101/2023.08.26.23294702>; this version posted August 26, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

