Methods and Applications

Genetic Insights into Glycine's Protective Role Against CAD — European and East Asia, 2015 and 2020

Jiaying Yu[1,2,&]; Zhuolin Zhu[3,&]; Ting Wang[1,2]; Yuanhao Wei[1,2]; Lianjie Huang[1,2]; Qianru Zhang[1,2]; Yuting Zhang[1,2]; Yiran Wang[1,2]; Guiyou Liu[4]; Xiang Shu[5,#]; Rennan Feng[1,2,#]

ABSTRACT

Introduction: The purpose of this study is to examine the potential causal relationship between levels of circulating glycine and coronary artery disease (CAD) using a two-step Mendelian randomization (MR) analysis.

Methods: We analyzed data from genome-wide association studies (GWAS) conducted on European and East Asian populations. To assess the causal effects of circulating glycine levels on the risk of CAD. We used the inverse-variance weighting (IVW), weighted median (WM), MR-Egger, and Mendelian Randomization Pleiotropy RESidual Sum and Outlier (MR-PRESSO) methods. Furthermore, we conducted mediation analysis to investigate the contribution of blood pressure and other cardiovascular disease-related traits.

Results: The two-step Mendelian randomization analysis revealed that higher levels of glycine in the blood were associated with a reduced risk of CAD in Europeans [odds ratio (*OR*)=0.84, 95% confidence interval (*CI*): 0.72, −0.98; *P*=0.029] and East Asians: (*OR*=0.76, 95% *CI*: 0.66, −0.89; $P=3.57\times10^{-4}$). Sensitivity analysis confirmed the robustness of these findings. Additionally, our results suggest that about 6.06% of the observed causal effect is mediated through genetically predicted systolic blood pressure (SBP) in the European population.

Discussion: Our results contribute to the current knowledge regarding the involvement of glycine in the progression of CAD, and provide valuable methodological insights for the prevention and treatment of this condition.

Coronary artery disease (CAD), also known as atherosclerosis or coronary heart disease, is the leading cause of global mortality (*1*). Glycine, a non-essential amino acid, plays a critical role in cell growth, immune function, antioxidant response, and anti-inflammatory processes (*2*). Previous studies have shown positive effects of glycine on cardiovascular health (*3*). The therapeutic potential of glycine for metabolic disorders and cardiovascular diseases has been proposed. However, studies investigating the association between circulating glycine levels and CAD risk have yielded inconsistent results (*4*–*6*). Therefore, the causal relationship between glycine and CAD remains controversial, and if such a relationship exists, it may be influenced by metabolic factors such as blood pressure.

Mendelian randomization (MR) is a statistical technique that uses genetic variants as instrumental variables to assess the causal impact of an exposure on an outcome (*7*). MR leverages the fact that genetic variants are randomly assigned at conception, making them immune to confounding factors typically found in observational studies. In order to explore the causal relationship between circulating glycine levels and the risk of CAD, as well as to uncover the underlying mechanisms, we conducted a comprehensive study using a two-step MR approach. This study aimed to investigate the potential causal effects of circulating glycine on CAD risk in individuals of European ancestry and East Asians.

METHODS

In the study, we analyzed the relationship between specific genetic instruments and glycine levels in the UK Biobank (UKB), which included 114,972 individuals of European descent (Nightingale Health Plc; Biomarker Quantification Version 2020) and the study conducted by Wittemans et al. (*4*), which included 30,118 individuals of European ancestry.

All details regarding the GWAS summary-level data are presented in Supplementary Table S1 (available at https://weekly.chinacdc.cn/). In order to address

potential weak instrumental bias, instrumental variables (IVs) should significantly associate with the exposure ($P<5\times10^{-8}$) and exhibit minimal linkage-disequilibrium (LD) with other single nucleotide polymorphisms (SNPs) ($R^2<0.001$) within a clump distance of 1,000 kb (Supplementary Table S2, available at https://weekly.chinacdc.cn/). By utilizing the PhenoScanner database, we identified and subsequently excluded pleiotropic SNPs that are correlated with confounding factors (Supplementary Table S2).

Our study utilized a two-step MR analysis design (Figure 1). The primary analysis was conducted using the inverse variance-weighted (IVW) method. In instances where heterogeneity was detected, we employed the IVW method with random effects. To further explore the robustness of our findings, we conducted sensitivity analyses using alternative approaches, including MR-Egger regression, the weighted median method, and Mendelian Randomization Pleiotropy RESidual Sum and Outlier (MR-PRESSO) analysis, accounting for multiple genetic variants (*8*–*9*). To assess pleiotropy, we utilized the MR-Egger intercept test and the MR-PRESSO global test. Additionally, we evaluated heterogeneity of the MR findings using Cochran's Q-statistic and the I^2 index (*10*). The MR analyses were conducted using the 'dplyr' and 'TwoSampleMR' packages in R (version 4.0.5, R foundation for statistical computing, Vienna, Austria) The threshold of statistical significance was $P<0.05$.

RESULTS

The 19 identified SNPs collectively accounted for approximately 6.5% of the variation in circulating glycine levels (R^2=6.5%). Furthermore, the F-statistic, which exceeds 18.66, indicates a low probability of weak instrument bias occurring in this study.

We utilized a panel of 18 SNPs, excluding 1 palindromic SNP with intermediate allele frequencies, to assess the correlation between genetically predicted higher circulating glycine levels and decreased risk of CAD in Europeans. The analysis revealed a significant correlation [odds ratio (*OR*)=0.84, 95% confidence interval (*CI*): (0.72, 0.98), P=0.029, $P_{\text{Cochran's Q}}$=0.018, $P_{\text{MR-PRESSO global test}}$=0.03, $P_{\text{MR-Egger intercept test}}$=0.069]. However, when using instruments consisting of 4, 3, and 1 SNPs, no significant associations were observed, despite consistently indicating the same direction of association (Figure 2A).

Out of the initial 19 SNPs, 14 were included in our East Asian-focused MR analysis using data from Biobank Japan. Five SNPs were excluded due to missing data or palindromic status. Our analysis showed a consistent protective relationship between genetically predicted glycine levels and the risk of CAD (*OR*=0.76, 95% *CI*: 0.66, 0.89; $P=3.57\times10^{-4}$, $P_{\text{Cochran's Q}}$=0.186, $P_{\text{MR-PRESSO global test}}$=0.199, $P_{\text{MR-Egger intercept test}}$=0.038), However, other IVs sets did not show a significant association, likely due to limited statistical power (Figure 2B).

We found a significant relation between higher genetically predicted circulating glycine levels and lower genetically predicted SBP (β=−0.74, 95% *CI*: −1.28, −0.20; P=0.007, $P_{\text{Cochran's Q}}$=0.196, $P_{\text{MR-PRESSO global test}}$=0.287], $P_{\text{MR-Egger intercept test}}$=0.759). Using the CPS1 and GLDC instruments, we observed consistent effects of glycine on SBP (β=−0.62, 95% *CI*: −1.19, −0.06; P=0.03, $P_{\text{Cochran's Q}}$=0.551, $P_{\text{MR-Egger intercept test}}$=0.786] (Supplementary Figure S1, available at https:// weekly.chinacdc.cn/).

Furthermore, our analysis revealed that there was a negative relationship between genetically predicted

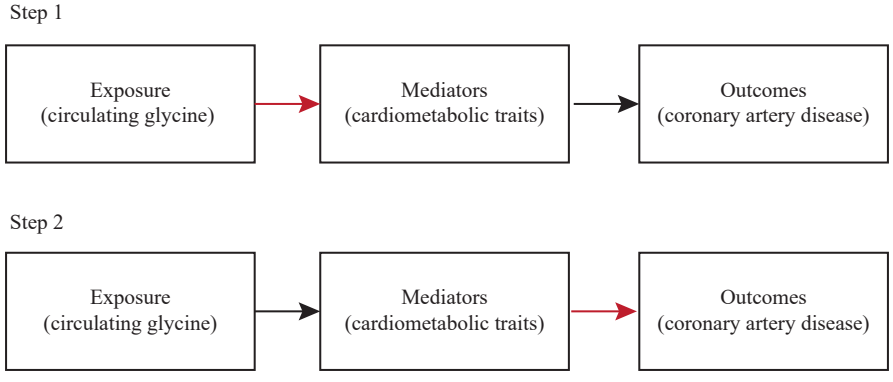


FIGURE 1. Schematic diagram of a two-step Mendelian randomization.

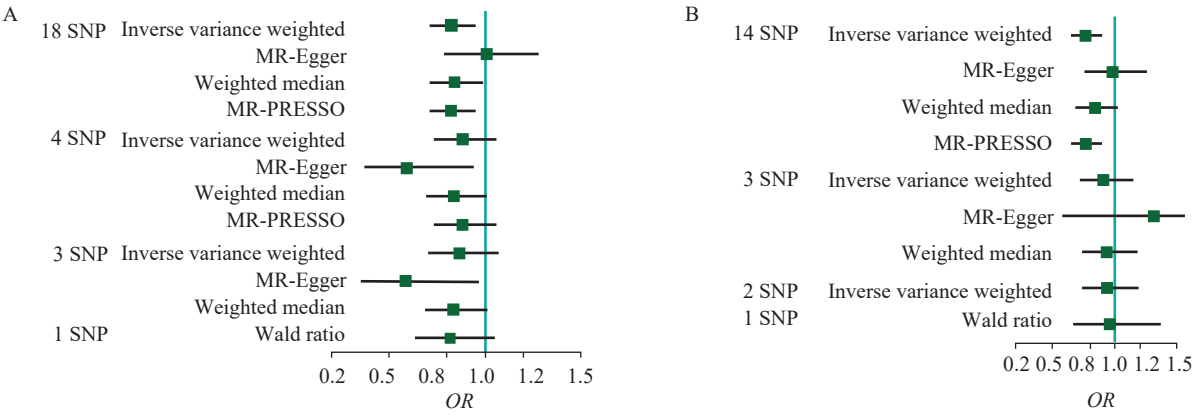


FIGURE 2. Forest plots showing effect sizes ± 95% confidence intervals for the association between genetically predicted circulating glycine levels and CAD. (A) European; (B) East Asian.
Note: Four sets of IVs were used to estimate the association between circulating glycine levels and CAD risk: 1) significant glycine-related SNPs (19 SNPs) identified in the GWAS of circulating glycine; 2) loci near genes encoding enzymes-related to glycine metabolism (GLDC, PHGDH, PSPH, ALDH1L1, and CPS1, 4 SNPs); 3) loci near genes encoding enzymes related to glycine metabolism except for the pleiotropic CPS1 locus (which showed significant associations with multiple metabolites, 3 SNPs); and 4) the loci at GCSH and GLDC encoding enzymes of the glycine cleavage system (1 SNP).
Abbreviation: SNP=single nucleotide polymorphism; MR-PRESSO=Mendelian Randomization Pleiotropy RESidual Sum and Outlier; GWAS=genome-wideassociation studies; ALDH1L1=aldehyde dehydrogenase 1 family member L1; CAD=coronary artery disease; CPS1=carbamoyl-phosphate synthase; GLDC=glycine decarboxylase; GCSH=glycine cleavage system protein H; PHGDH=phosphoglycerate dehydrogenase; PSPH=phosphoserine phosphatase; SBP=systolic blood pressure.

circulating glycine levels and genetically predicted DBP (β=−0.30, 95% *CI*: −0.56, −0.05; P=0.02, $P_{\text{Cochran's Q}}$=0.356, $P_{\text{MR-PRESSO global test}}$=0.455, $P_{\text{MR-Egger intercept test}}$=0.527). This association remained consistent when using alternative IV sets (Supplementary Figure S1, available at https://weekly.chinacdc.cn/). However, we did not find any association between predicted glycine levels and anthropometric, glycemic, inflammatory, or blood lipid traits (Supplementary Figures S2–S5, available at https://weekly.chinacdc.cn/).

Out of the 461 SNPs related to SBP, we included 412 in our MR analysis. Four SNPs were excluded due to insufficient data, and 45 SNPs were removed after an outlier test, using MR-PRESSO. We detected significant heterogeneity ($P_{\text{Cochran's Q}}$=2.17×10^{-8}), and the presence of horizontal pleiotropy was confirmed by using the MR-PRESSO global test (P<1×10^{-4}). Our results demonstrated a positive association between SBP and the risk of CAD, with each unit increase in SBP associated with a 3% increase in CAD risk (*OR*=1.03, 95% *CI*: 1.02, 1.03; P_{WM}=2.33×10^{-19}) (Table 1).

The potential mediation of systolic blood pressure (SBP) on the association between circulating glycine and CAD risk was investigated. The mediation effect involving SBP was found that 6.06% of the effect of circulating glycine, which was genetically predicted using 19 SNPs, was mediated through the genetically predicted SBP (Supplementary Table S3, available at https://weekly.chinacdc.cn/).

TABLE 1. Association of genetically predicted SBP with CAD risk in the Mendelian randomization analysis.

Method	*OR*	95% *CI*	*P*
WM	1.03	1.02 to 1.03	2.33×10^{-19}
IVW	1.03	1.03 to 1.04	1.76×10^{-56}
MR-Egger	1.04	1.03 to 1.05	1.73×10^{-11}
MR-PRESSO	1.03	1.03 to 1.04	1.96×10^{-44}

Note: *OR*>1 indicates that increased SBP was associated with an increased risk of CAD; The Cochran's Q=589.23 (P=2.17×10^{-8}), and I^2=30.24%, indicating that there was heterogeneity. The MR-PRESSO global test (P<1×10^{-4}) and MR-Egger intercept test (P=0.481) indicated that there was horizontal pleiotropy for the selected instruments.
Abbreviation: CAD=coronary artery disease; *CI*=confidence interval; MR-PRESSO=Mendelian Randomization Pleiotropy RESidual Sum and Outlier; IVW=inverse variance weighted; *OR*=odds ratio; SBP=systolic blood pressure; WM=weighted median.

DISCUSSION

The two-step MR analysis demonstrated a significant causal association between decreased levels of genetically predicted circulating glycine and CAD. Our estimation revealed that approximately 6.06% of this potential causal effect is mediated through

genetically predicted SBP, suggesting that the protective influence of circulating glycine may be attributed to its effect on reducing SBP. However, our findings suggest that other risk factors associated with CAD, such as glycemic characteristics, lipid profiles, and inflammatory markers, may not play a considerable role as mediators in this relationship.

Previous studies conducted on European white populations have produced inconsistent results regarding the association between glycine levels and the risk of developing CAD. Specifically, one study on European whites did not provide strong evidence for a causal relationship between glycine and CAD risk (*6*). Additionally, the relationship between circulating glycine levels and the CAD risk in different racial groups remains uncertain. There is only one previous study that investigated this relationship in Singaporean Chinese individuals, and reported a similar protective effect (*5*). However, this study had limitations such as a relatively small sample size and the inclusion of only two SNPs, one of which was not validated. Consequently, the study design may have compromised the validity of their findings. In our research, we have cross-validated IVs using two independent GWAS datasets, which strengthens the reliability and robustness of our conclusions.

Previous epidemiological studies have suggested that dietary glycine may have a protective effect on blood pressure regulation (*11*). Our MR analysis provides further support for this association. We found an inverse genetic correlation between circulating glycine levels and SBP, which accounted for approximately nearly 6.06% of the genetic association between glycine and CAD. In rat models of metabolic syndrome, diets rich in glycine have been shown to reduce hypertension by reducing free radical production and enhancing nitric oxide utilization (*12*). However, our study did not uncover any significant correlations between glycine and lipid traits or inflammatory markers (*13*). Further comprehensive investigations are needed to explore other potential mechanisms that may explain the genetic link between glycine and CAD.

Our study had several strengths. First, we implemented a thorough process to select valid genetic instruments for MR analysis. This procedure reduced the potential bias caused by weak instruments and improved the statistical power of our study. We used different MR methodologies to ensure the reliability of our estimates regarding the causal relationship between circulating glycine and the risk of CAD. Additionally, we conducted MR analyses on two separate populations, obtaining consistent results.

This study was subject to some limitations. First, the genetic instruments used in our analysis were derived from datasets consisting only of individuals of European descent. To date, there have been no large-scale GWAS studies investigating the relationship between circulating glycine and genetic instruments in East Asian populations. Additionally, our mediation analysis is subject to potential bias, as accurately establishing causal relationships can be challenging and distinguishing between mediation and confounding can be statistically complex.

The present study utilized MR analysis to investigate the possible causal link between serum glycine levels and CAD. These findings contribute to our understanding of the role of glycine in the development of CAD and provide methodological insights for the prevention and treatment of the disease.

Conflicts of interest: No conflicts of interest.

Acknowledgments: The study was based on the data provided by the Medical Research Council Integrative Epidemiology Unit. We thank the investigators of the previous studies.

Funding: Supported by the National Natural Science Foundation of China (82273612), and by Open Project of Key Laboratory of Science and Engineering for the Multi-Modal Prevention and Control of Major Chronic Diseases, Ministry of Industry and Information Technology (Grant No. MCD-2023-1-09).

doi: 10.46234/ccdcw2024.034

[#] Corresponding authors: Rennan Feng, fengrennan@yeah.net; Xiang Shu, shux@mskcc.org.

[1] Department of Nutrition and Food Hygiene, School of Public Health, Harbin Medical University, Harbin City, Heilongjiang Province, China; [2] Key Laboratory of Precision Nutrition and Health of Ministry of Education, School of Public Health, Harbin Medical University, Harbin City, Heilongjiang Province, China; [3] Songyang County Center for Disease Prevention and Control, Songyang City, Zhejiang Province China; [4] Department of Neurology, Xuanwu Hospital, Capital Medical University, Beijing, China; [5] Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York City, NY, USA.
[&] Joint first authors.

Submitted: February 07, 2024; Accepted: February 22, 2024

REFERENCES

1. GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for

the Global Burden of Disease Study 2013. Lancet 2015;385(9963): 117 – 71. https://doi.org/10.1016/S0140-6736(14)61682-2.
2. Ballevre O, Cadenhead A, Calder AG, Rees WD, Lobley GE, Fuller MF, et al. Quantitative partition of threonine oxidation in pigs: effect of dietary threonine. Am J Physiol 1990;259(4 Pt 1):E483-91. http://dx.doi.org/10.1152/ajpendo.1990.259.4.E483.
3. Razak MA, Begum PS, Viswanath B, Rajagopal S. Multifarious beneficial effect of nonessential amino acid, glycine: a review. Oxid Med Cell Longev 2017;2017:1716701. https://doi.org/10.1155/2017/1716701.
4. Wittemans LBL, Lotta LA, Oliver-Williams C, Stewart ID, Surendran P, Karthikeyan S, et al. Assessing the causal association of glycine with risk of cardio-metabolic diseases. Nat Commun 2019;10(1):1060. https://doi.org/10.1038/s41467-019-08936-1.
5. Chang XL, Wang L, Guan SP, Kennedy BK, Liu JJ, Khor CC, et al. The association of genetically determined serum glycine with cardiovascular risk in East Asians. Nutr Metab Cardiovasc Dis 2021;31(6):1840 – 4. https://doi.org/10.1016/j.numecd.2021.03.010.
6. Jia Q, Han Y, Huang P, Woodward NC, Gukasyan J, Kettunen J, et al. Genetic determinants of circulating glycine levels and risk of coronary artery disease. J Am Heart Assoc 2019;8(10):e011922. https://doi.org/10.1161/JAHA.119.011922.
7. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations: update for summer 2023. Wellcome Open Res 2019;4:186. https://doi.org/10.12688/wellcomeopenres.15555.1.
8. Bowden J, Smith GD, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44(2):512 – 25. https://doi.org/10.1093/ije/dyv080.
9. 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. https://doi.org/10.1002/gepi.21965.
10. Del Greco MF, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. Stat Med 2015;34(21):2926 – 40. https://doi.org/10.1002/sim.6522.
11. Stamler J, Brown IJ, Daviglus ML, Chan Q, Miura K, Okuda N, et al. Dietary glycine and blood pressure: the international study on macro/micronutrients and blood pressure. Am J Clin Nutr 2013;98(1):136 – 45. https://doi.org/10.3945/ajcn.112.043000.
12. Slomowitz LA, Gabbai FB, Khang SJ, Satriano J, Thareau S, Deng AH, et al. Protein intake regulates the vasodilatory function of the kidney and NMDA receptor expression. Am J Physiol Regul Integr Comp Physiol 2004;287(5):R1184 – 9. https://doi.org/10.1152/ajpregu.00169.2003.
13. Ding YP, Svingen GFT, Pedersen ER, Gregory JF, Ueland PM, Tell GS, et al. Plasma glycine and risk of acute myocardial infarction in patients with suspected stable angina pectoris. J Am Heart Assoc 2015;5(1):e002621. https://doi.org/10.1161/JAHA.115.002621.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLE S1. Description of GWAS summary-level data used in this study.

Trait	Year	GWAS ID	Population	Sample size	Number of SNPs	Source (Pubmed ID)
Circulating glycine levels	2020	met-d-Gly	European	114,972	12,321,875	Nightingale Health Plc; Biomarker Quantification Version 2020
CAD	2015	ebi-a-GCST003116	European	141,217	8,597,751	26343387
CAD	2020	bbj-a-159	East Asian	212,453	8,881,048	–
BMI	2018	ieu-b-40	European	681,275	2,336,260	30124842
WHRadjBMI	2015	ieu-a-79	European	210,082	210,082	25673412
FG	2012	ebi-a-GCST005186	European	58,074	2,599,409	22581228
FI	2012	ebi-a-GCST005185	European	51,750	2,598,774	22581228
DBP	2018	ieu-b-39	European	757,601	7,160,619	30224653
SBP	2018	ieu-b-38	European	757,601	7,088,083	30224653
CRP	2018	ieu-b-35	European	204,402	2,414,379	30388399
IL-6	2019	prot-c-4673_13_2	European	1,338	501,428	28240269
TG	2013	ebi-a-GCST002216	European	94,595	2,410,057	24097068
TC	2013	ebi-a-GCST002221	European	94,595	2,418,562	24097068

Note: "–" means not applicable. ID means GWAS ID on the MRbase website, which is a website of large-scale public GWAS datasets; PMID, the ID number of the articles published in PubMed.

Abbreviation: SNP=single nucleotide polymorphism; GWAS=genome-wideassociation studies; CAD=coronary artery disease; BMI=body mass index; WHRadjBMI=waist-to-hip ratio adjusted for BMI; FG=fasting blood glucose; FI=fasting blood insulin; DBP=diastolic blood pressure; SBP=systolic blood pressure; IL-6=interleukin-6; CRP=C-reactive protein; TG=triglycerides; TC=total cholesterol.

SUPPLEMENTARY TABLE S2. Selection of valid genetic instruments based on UKB and Wittemans's study.

UKB									
SNP	**Chr**	**Pos**	**EA**	**NEA**	**EAF**	**β**	**Se**	***P***	**Loci**
rs10190808*,†	2	211993631	C	G	0.136421	0.0346173	0.0056719	2.40×10^{-9}	*CPS1-IT1*
rs10934753*,†	3	125906179	A	G	0.417116	0.0690031	0.0039623	3.40×10^{-74}	*ALDH1L1*
rs11045886	12	21386493	C	A	0.165117	0.0306817	0.0052791	8.30×10^{-9}	*SLCO1B1*
rs11172190	12	57766305	T	C	0.503709	0.0236924	0.0039114	1.00×10^{-10}	*R3HDM2*
rs112247225†	16	81154900	T	C	0.046621	−0.128232	0.0092851	5.20×10^{-45}	*PKD1L2*
rs11242109	5	131677047	T	G	0.47905	0.0226052	0.003901	2.00×10^{-8}	*SLC22A4*
rs149181595	15	43685807	C	A	0.027601	0.0942587	0.0119122	3.90×10^{-17}	*TUBGCP4*
rs17722201	2	209645912	C	T	0.219135	0.0275746	0.0047091	9.80×10^{-9}	*PTH2R*
rs192322963†	8	17445955	A	G	0.025449	0.0829439	0.0125376	4.10×10^{-11}	*PDGFRL*
rs1965869†	4	89677537	C	T	0.715726	−0.0242087	0.0043317	9.00×10^{-9}	*FAM13A*
rs2026972*,†	9	6538279	C	G	0.308331	−0.0821551	0.0042321	2.20×10^{-87}	*GLDC*
rs2608913†	6	131870261	C	T	0.217028	−0.0243496	0.0047367	2.30×10^{-8}	*ARG1*
rs2657879†	12	56865338	G	A	0.182536	0.0478885	0.0050385	5.20×10^{-22}	*GLS2*
rs2711697†	12	47265729	C	A	0.369926	0.0265213	0.0040341	1.10×10^{-11}	*SLC38A4*
rs28435239	9	5989087	A	G	0.769224	−0.0289595	0.0046491	8.60×10^{-10}	*KIAA2026*
rs34945403†	15	58430763	G	A	0.067631	−0.0634669	0.0080365	1.70×10^{-15}	*AQP9*
rs35034344†	2	211026796	T	A	0.273031	−0.0653462	0.0045911	1.90×10^{-46}	*KANSL1L*
rs4380169	2	212145768	C	T	0.490602	−0.0421118	0.0038951	3.60×10^{-27}	*ENSAP3*
rs4889229	16	81113672	T	C	0.917408	0.117726	0.0071008	4.20×10^{-64}	*RP11-303E16.10*
rs561931*,†	1	120254506	G	A	0.580975	0.0290991	0.0039588	4.50×10^{-14}	*PHGDH*
rs56819961†	12	47137673	C	T	0.212573	0.0426129	0.0047614	1.40×10^{-20}	*SLC38A4*
rs6587644	1	151994458	A	G	0.304982	0.0221024	0.0042589	2.40×10^{-8}	*NBPF18P*
rs67523949	12	348506	T	C	0.536231	0.0263338	0.0039096	3.30×10^{-11}	*SLC6A13*
rs7188156	16	79938114	G	T	0.14474	0.0317192	0.0055593	1.20×10^{-8}	*LINC01228*
rs75604103†	2	211692010	G	A	0.121625	−0.0989874	0.0060071	1.10×10^{-60}	*ENSAP3*
rs7704653†	5	90255685	G	A	0.724033	−0.0316456	0.0044087	3.70×10^{-13}	*ADGRV1*
rs7800001†	7	56072010	C	T	0.755284	0.0723549	0.0045524	5.50×10^{-61}	*GBAS*
rs79687284†	1	214150821	C	G	0.034645	0.0679817	0.0106777	7.80×10^{-11}	*PROX1-AS1*
rs9532939†	13	42440496	A	T	0.345473	0.0252027	0.0042637	8.70×10^{-9}	*VWA8*
rs11666281	19	18234588	T	C	0.25424	0.0044789	−0.0370833	1.30×10^{-17}	*MAST3*
rs6601302†	8	9239458	G	T	0.750342	0.004516	0.0276846	2.10×10^{-10}	–
rs2657879†	12	56865338	G	A	0.182536	0.0050385	0.0478885	5.20×10^{-22}	*GLS2*
rs28601761†	8	126500031	G	C	0.420161	0.003998	0.060914	1.40×10^{-55}	*TRIB1*
rs4240624	8	9184231	A	G	0.90923	0.0067925	−0.126861	2.00×10^{-82}	*LOC157273*
Wittemans's study									
rs4646961	1	76217169	A	G	0.297	0.048	0.006	8.41×10^{-19}	intronic variant in ACADM
rs561931*,†	1	120254506	G	A	0.593	0.033	0.006	7.57×10^{-14}	5' UTR variant of PHGDH
rs10184004†	2	165508389	T	C	0.4	0.036	0.006	1.53×10^{-9}	Intergenic variant near COBLL1 (28 kb) and GRB14 (30 kb)
rs715*,†	2	211543055	C	T	0.313	0.444	0.006	3.00×10^{-1632}	3'UTR variant of CPS1
rs9862438*	3	125910381	T	C	0.416	0.058	0.006	1.13×10^{-30}	ncRNA intronic variant in ALDH1L1-AS2

Continued

UKB									
SNP	**Chr**	**Pos**	**EA**	**NEA**	**EAF**	***β***	**Se**	***P***	**Loci**
rs148685782	4	155533035	G	C	0.996	0.309	0.049	2.01×10^{-10}	Synonymous variant in FGG
rs71640034	4	187161048	A	G	0.511	0.034	0.006	5.57×10^{-8}	intronic variant in KLKB1
rs156380	5	53378450	C	T	0.807	0.031	0.007	4.50×10^{-8}	intronic variant in ARL15
rs3105793	5	90226061	A	G	0.273	0.028	0.006	4.04×10^{-8}	intronic variant in ADGRV1
rs10900807	5	131757480	G	C	0.805	0.036	0.007	1.26×10^{-9}	ncRNA intronic variant in C5orf56
rs2545801	5	176841339	C	T	0.747	0.042	0.007	7.23×10^{-14}	intergenic variant near F12 (5 kb) and GRK6 (12 kb)
rs543159	6	160776017	A	C	0.482	0.035	0.006	4.20×10^{-10}	intronic variant in SLC22A3
rs4947534*	7	56079094	C	T	0.76	0.072	0.007	7.12×10^{-34}	3' UTR variant of PSPH
rs9987289†	8	9183358	A	G	0.1	0.124	0.01	1.74×10^{-49}	ncRNA intronic variant in LOC157273
rs28601761†	8	126500031	G	C	0.416	0.063	0.006	8.49×10^{-30}	intergenic variant near TRIB1 (49kb) and LINC00861 (435 kb)
rs17591030*	9	6550024	C	T	0.715	0.08	0.006	1.88×10^{-40}	intron variant in GLDC
rs676996†	9	136146077	T	G	0.668	0.04	0.006	4.39×10^{-15}	intron variant in ABO
rs190595610	10	32274880	A	G	0.997	0.253	0.056	8.96×10^{-9}	Intergenic variant near ARHGAP12 (57 kb) and KIF5B (23 kb)
rs10740134†	10	65315433	T	C	0.515	0.038	0.006	1.18×10^{-12}	intron variant in REEP3
rs12297321	12	47109387	T	C	0.152	0.048	0.008	7.41×10^{-13}	Intergenic variant near SLC38A4 (38 kb) and LOC100288798 (630 kb)
rs2638314	12	56866334	A	T	0.182	0.042	0.007	1.52×10^{-8}	intronic variant in GLS2
rs9514191	13	104520138	C	G	0.312	0.034	0.006	3.10×10^{-8}	intergenic variant near LINC01309 (440 kb) and DAOA-AS1(159 kb)
rs201393666	15	43677979	A	C	0.029	0.097	0.017	2.64×10^{-8}	intronic variant in TUBGCP4
rs2280195	15	58467095	A	G	0.441	0.028	0.006	3.15×10^{-9}	intronic variant in AQP9
rs9923732*	16	81110903	A	G	0.914	0.119	0.011	1.22×10^{-41}	Upstream variant of C16orf46, 9 kb downstream of GCSH
rs8078686†	17	45735706	C	T	0.509	0.035	0.006	3.66×10^{-11}	intron variant in KPNB1
rs273510†	19	18223350	A	G	0.708	0.034	0.006	3.57×10^{-9}	intron variant in MAST3

Note: "–" means Gene loci was not found. We excluded rs1047891 (homocysteine levels), rs11666281 (body mass index), rs13107325 (diastolic blood pressure), rs1801133 (homocysteine levels), rs2657879 (fasting blood glucose), rs28601761 (triglycerides), rs36105243 (type 2 diabetes), rs4240624 (total cholesterol), rs56113850 (smoking status: current), rs6601302 (body mass index), and rs79687284 (diabetes diagnosed by doctors) for pleiotropic effects (from PhenoScanner).

Abbreviation: UKB=UK Biobank; Chr=chromosome; Pos=position; EA=equal alleles; NEA=non-equal alleles; EAF=equal allele frequencies; Se=standard error; SNP=single nucleotide polymorphism; PSPH=phosphoserine phosphatase; GLDC=glycine decarboxylase; GCSH=glycine cleavage system protein H.

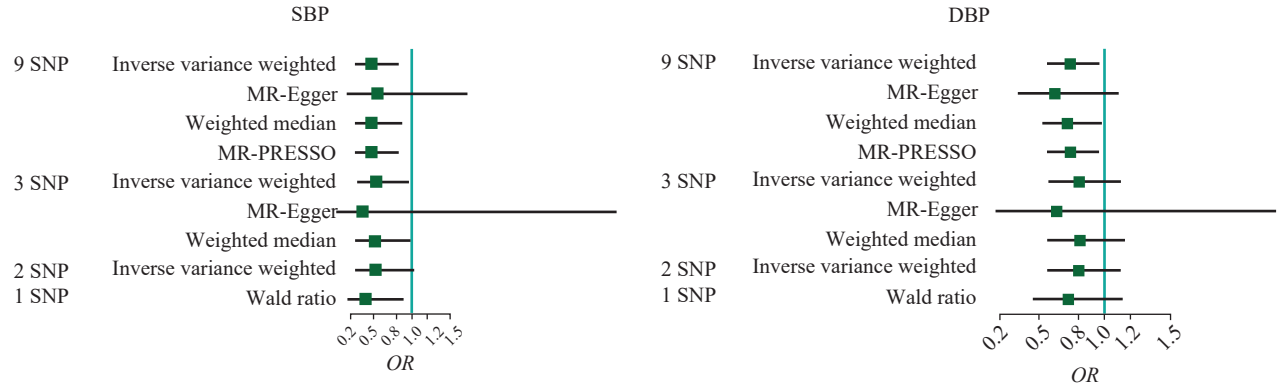
* means loci related to circulating glycine.

† means the SNPs used in this study.

SUPPLEMENTARY TABLE S3. The mediation effect of circulating glycine on CAD via SBP.

Mediator	**Total effect**	**Direct effect A**	**Direct effect B**	**Mediation effect**		**Mediated proportion (%)**
	***β* (95% *CI*)**	***β* (95% *CI*)**	***β* (95% *CI*)**	***β* (95% *CI*)**	***P***	**(95% *CI*)**
SBP	−0.20 (−0.34, −0.05)	−0.74 (−1.28, −0.20)	0.03 (0.02, 0.03)	-0.02 (-0.04, -4.94×10^{-3})	9.93×10^{-3}	6.06 (3.62, 10.64)

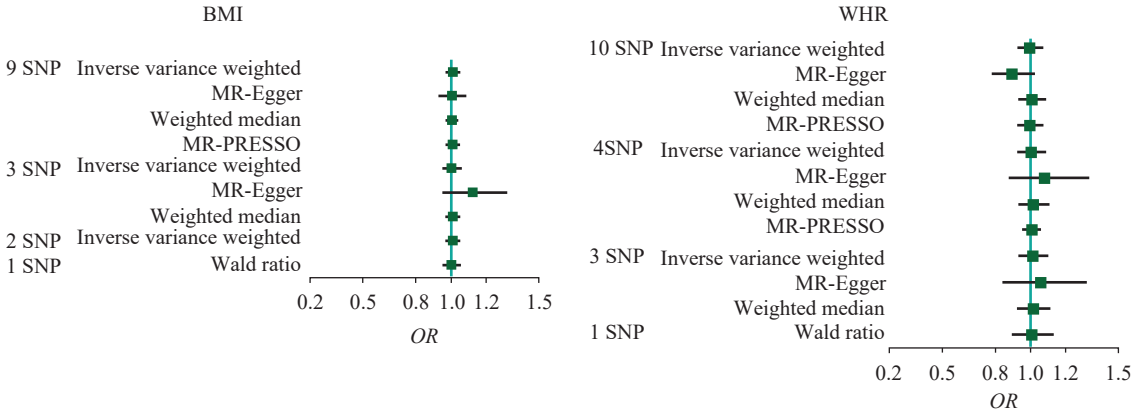
Abbreviation: *CI*=confidence interval; CAD=coronary artery disease; SBP=systolic blood pressure.



SUPPLEMENTARY FIGURE S1. Forest plots of the effect sizes ± 95% confidence intervals of the association between genetically predicted circulating glycine levels and SBP/DBP.

Note: The association effect size was estimated based on one standard deviation change of genetically predicted circulating glycine levels (9 SNP: in 19 SNPs,16 SNPs were available for SBP dataset; 1 SNP was palindromic with intermediate allele frequencies, and 6 SNPs were outliers removed by MR-PRESSO outlier test; 3 SNP: in 4 SNPs, 3 SNPs were available for SBP dataset; 2 SNP: in 3 SNPs, 2 SNPs were available for SBP dataset; 1 SNP: 1 SNP was available for SBP dataset; 9 SNP: in 19 SNPS,16 SNPs were available for DBP dataset; 1 SNP was palindromic with intermediate allele frequencies, and 6 SNPs were outliers removed by MR-PRESSO outlier test; 3 SNP: in 5 SNPs, 3 SNPs were available for DBP dataset; 2 SNP: in 3 SNPs, 2 SNPs were available for DBP dataset; 1 SNP: 1 SNP was available for DBP dataset).

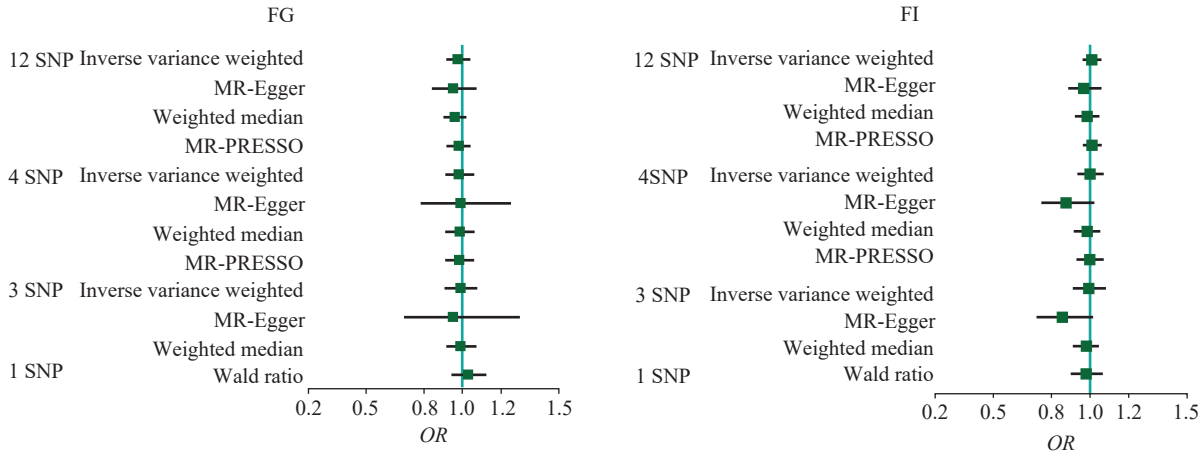
Abbreviation: MR-PRESSO=Mendelian Randomization Pleiotropy RESidual Sum and Outlier; SNP=single nucleotide polymorphism; DBP=diastolic blood pressure; SBP=systolic blood pressure.



SUPPLEMENTARY FIGURE S2. Forest plots of the effect sizes ± 95% confidence intervals of the association between genetically predicted circulating glycine levels and BMI/WHR.

Note: The association effect size was estimated based on one standard deviation change of genetically predicted circulating glycine levels (9 SNP: in 19 SNPs, 12 SNPs were available for BMI dataset, and 3 SNPs were outliers removed by MR-PRESSO outlier test; 3 SNP: in 4 SNPs, 3 SNPs were available for BMI dataset; 2 SNP: in 3 SNPs, 2 SNPs were available for BMI dataset; 1 SNP: 1 SNP was available for BMI dataset; 10 SNP: in 19 SNPs, 14 SNPs were available for WHR dataset, and 4 SNPs were outliers removed by MR-PRESSO outlier test; 4 SNP: 4 SNPs were available for WHR dataset; 3 SNP: 3 SNPs were available for WHR dataset; 1 SNP: 1 SNP was available for WHR dataset).

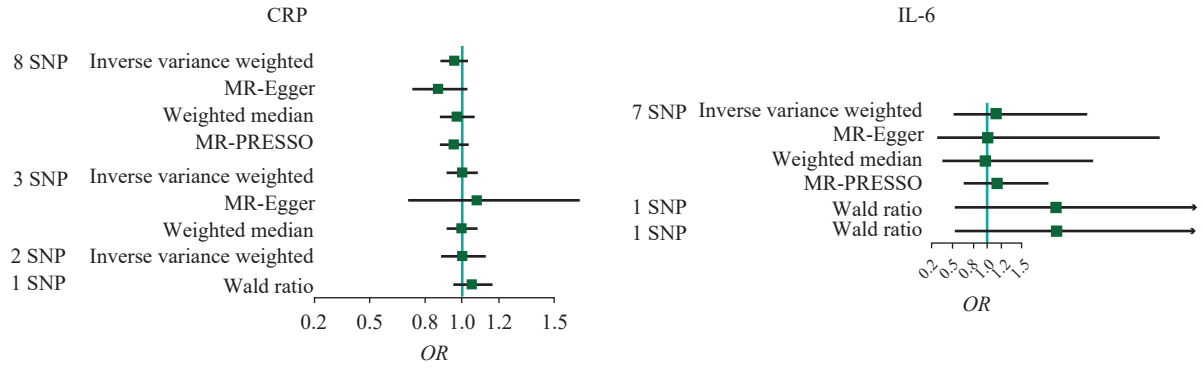
Abbreviation: SNP=single nucleotide polymorphism; MR-PRESSO=Mendelian Randomization Pleiotropy RESidual Sum and Outlier; BMI=body mass index; WHR=waist-to-hip ratio adjusted for body mass index.



SUPPLEMENTARY FIGURE S3. Forest plots of the effect sizes ± 95% confidence intervals of the association between genetically predicted circulating glycine levels and FG/FI.

Note: The association effect size was estimated based on one standard deviation change of genetically predicted circulating glycine levels (12 SNP: in 19 SNPs,14 SNPs were available for FG dataset, and 2 SNPs were outliers removed by MR-PRESSO outlier test; 4 SNP: 4 SNPs were available for FG dataset; 3 SNP: 3 SNPs were available for FG dataset; 1 SNP: 1 SNP was available for FG dataset; 12 SNP: in 19 SNPS,14 SNPs were available for FI dataset, and 2 SNPs were outliers removed by MR-PRESSO outlier test; 4 SNP: 4 SNPs were available for FI dataset; 3 SNP: 3 SNPs were available for FI dataset; 1 SNP: 1 SNP was available for FI dataset).

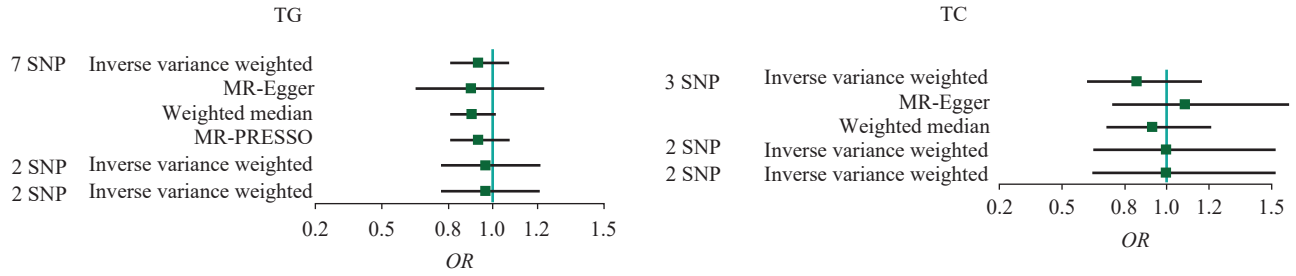
Abbreviation: SNP=single nucleotide polymorphism; MR-PRESSO=Mendelian Randomization Pleiotropy RESidual Sum and Outlier; FG=fasting blood glucose; FI=fasting blood insulin.



SUPPLEMENTARY FIGURE S4. Forest plots of the effect sizes ± 95% confidence intervals of the association between genetically predicted circulating glycine levels and CRP/IL-6.

Note: The association effect size was estimated based on one standard deviation change of genetically predicted circulating glycine levels (8 SNP: in 19 SNPS,13 SNPs were available for CRP dataset; and 5 SNPs were outliers removed by MR-PRESSO outlier test; 3 SNP: in 4 SNPs, 3 SNPs were available for CRP dataset; 2 SNP: in 3 SNPs, 2 SNPs were available for CRP dataset; 1 SNP: 1 SNP was available for CRP dataset. 7 SNP: in 19 SNPs, 8 SNPs were available for IL-6 dataset and 1 SNP was removed for being palindromic with intermediate allele frequencies; 1 SNP: in 4 SNPs, 2 SNPs were available for IL-6 dataset, and 1 SNP was removed for being palindromic with intermediate allele frequencies; 1 SNP: in 3 SNPS1 SNP was available for IL-6 dataset; 1SNP: 1 SNP was removed for being palindromic with intermediate allele frequencies).

Abbreviation: SNP=single nucleotide polymorphism; MR-PRESSO=Mendelian Randomization Pleiotropy RESidual Sum and Outlier; CRP=C-reactive protein; IL-6=Interleukin-6.



SUPPLEMENTARY FIGURE S5. Forest plots of the effect sizes ± 95% confidence intervals of the association between genetically predicted circulating glycine levels and TG/TC.

Note: The association effect size was estimated based on one standard deviation change of genetically predicted circulating glycine levels (7 SNP: in 19 SNPS 14 SNPs were available for TG dataset, 2 SNPs were removed for being palindromic with intermediate allele frequencies, and 5 SNPs were outliers removing by MR-PRESSO outlier test; 2 SNP: in 4 SNPs, 4 SNPs were available for TG dataset, 2 SNPs were removed for being palindromic with intermediate allele frequencies; 2 SNP: in 3 SNPS, 1 SNP was removed for being palindromic with intermediate allele frequencies 1 SNP: was removed for being palindromic with intermediate allele frequencies. 3 SNP: in 19 SNPs, 14 SNPs were available for TC dataset 2 SNPs were removed for being palindromic with intermediate allele frequencies, and 9 SNPs were outliers removing by MR-PRESSO outlier test; 2 SNP: in 4 SNPs, 4 SNPs were available for TC dataset, 2 SNPs were removed for being palindromic with intermediate allele frequencies; 2 SNP: in 3 SNPs, 1 SNP was removed for being palindromic with intermediate allele frequencies; 1 SNP: was removed for being palindromic with intermediate allele frequencies).

Abbreviation: SNP=single nucleotide polymorphism; MR-PRESSO=Mendelian Randomization Pleiotropy RESidual Sum and Outlier; TG=Triglyceride; TC=Total cholesterol.