Tzu Chi Medical Journal 2025; 37 (2): 181-188

Original Article

3-Hydroxy-3-methylglutaryl-CoA reductase variants strongly associated with low-density lipoprotein cholesterol levels and diabetes mellitus risk in a Taiwanese population: A Mendelian randomization study

*Lung-An Hsu[a†], Ming-Sheng Teng[b†], De-Min Duan[c], Kuan-Hung Yeh[c,d], Semon Wu[e], Yu-Lin Ko[b,c,d]**

[a]The First Cardiovascular Division, Department of Internal Medicine, Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Taoyuan, Taiwan, [b]Department of Research, Taipei Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, New Taipei, Taiwan, [c]Division of Cardiology, Department of Internal Medicine and The Cardiovascular Medical Center, Taipei Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, New Taipei, Taiwan, [d]School of Medicine, Tzu Chi University, Hualien, Taiwan, [e]Department of Life Science, Chinese Culture University, Taipei, Taiwan

[†]Both authors contributed equally to this work.

Submission : 20-Sep-2024
Revision : 08-Nov-2024
Acceptance : 13-Dec-2024
Web Publication : 07-Apr-2025

Abstract

Objectives: 3-Hydroxy-3-methylglutaryl-CoA reductase (HMGCR) is a rate-limiting enzyme involved in cholesterol synthesis. The common *HMGCR* variants are associated with low-density lipoprotein cholesterol (LDL-C) levels. We aimed to identify novel *HMGCR* variants influencing the lipid profiles of Taiwanese and assess the causal links between LDL-C levels and diabetic risk based on HMGCR genotypes. **Materials and Methods:** Genome-wide genotyping of 108,880 participants from Taiwan Biobank was used for the association studies and Mendelian randomization (MR) analysis. **Results:** Regional association and stepwise linear regression analyses showed *HMGCR* rs3064191, rs150454634, and rs13354746 variants were independently associated with total cholesterol (TC), LDL-C, and non-high-density lipoprotein cholesterol (non-HDL-C) levels with the former two variants in strong linkage disequilibrium with *HMGCR* rs3846662, a variant influencing exonal alternative splicing, and *HMGCR* rs191835914 (p.Y311S), an Asian-specific nonsynonymous mutation, respectively. Multivariate MR analyses showed significant associations between weighted genetic risk scores using LDL-C-determining *HMGCR* variants and using genome-wide association study identifying LDL-C-determining 47 variants and the prevalence of diabetes mellitus (DM) ($P = 0.0011$ and $P = 1.66 \times 10^{-8}$, respectively). **Conclusion:** The *HMGCR* variants exhibited significant associations with TC, LDL-C, and non-HDL-C levels as well as causally with DM risk in our Taiwanese population. *HMGCR* genotypes may play an important role and serve as a reference for the prevention and treatment of cardiovascular diseases in the clinical settings.

Keywords: *3-Hydroxy-3-methylglutaryl-CoA reductase, Diabetes mellitus, Low-density lipoprotein cholesterol, Mendelian randomization, Nonsynonymous mutation*

Introduction

Elevated levels of circulating low-density lipoprotein cholesterol (LDL-C) are recognized as a substantial causal factor for cardiovascular atherogenesis, vascular mortality, and all-cause mortality [1,2]. Studies have consistently shown that reducing LDL-C levels correlates with a decreased incidence of major vascular events across diverse populations, encompassing various risk profiles, ethnicities, and medication usage [3-7]. Mendelian randomization (MR) studies have further reinforced the causal links between LDL-C levels and adverse vascular events and mortality [8-12]. Conversely, intriguing inverse causal associations have been observed between elevated LDL-C and the hazard of diabetes mellitus (DM) in the individuals of diverse ethnic backgrounds [8-10,13-17]. However, certain studies suggest that these associations may be mechanism-or gene-specific [11,18]. Therefore, understanding the causality between genetic variants, LDL-C levels, DM, and their potential molecular mechanisms acting on LDL-C levels is crucial for the prevention and treatment of cardiovascular diseases.

The therapeutic target of statins, a class of LDL-C-lowering drugs, is 3-Hydroxy-3-methylglutaryl-CoA

Supplementary material available online

Access this article online

Quick Response Code:

Website: www.tcmjmed.com

DOI: 10.4103/tcmj.tcmj_247_24

*Address for correspondence: Dr. Yu-Lin Ko, Division of Cardiology, Department of Internal Medicine and The Cardiovascular Medical Center, Taipei Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, 289 Jianguo Road, Xindian, New Taipei, Taiwan. E-mail: yulinkotw@yahoo.com.tw

How to cite this article: Hsu LA, Teng MS, Duan DM, Yeh KH, Wu S, Ko YL. 3-Hydroxy-3-methylglutaryl-CoA reductase variants strongly associated with low-density lipoprotein cholesterol levels and diabetes mellitus risk in a Taiwanese population: A Mendelian randomization study. Tzu Chi Med J 2025;37(2):181-8.

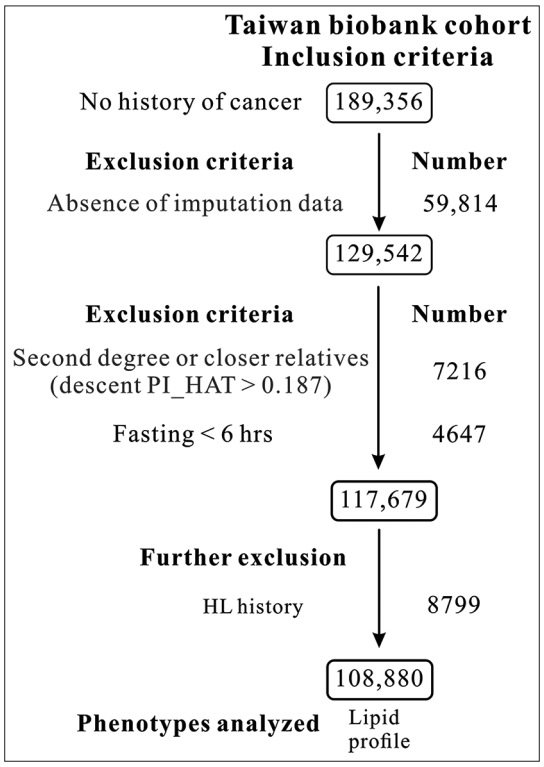


Figure 1: Flow chart of study population inclusion. Out of 129,542 participants from TWB, 108,880 were included in the genotype-phenotype analysis after excluding individuals who met the exclusion criteria, such as QC for GWAS (n = 7,216) meaning no imputation data and sec-ond-degree or closer relatives (descent PI_HAT > 0.187), fasting for < 6 h (n = 4,647), and a history of HL (n = 8,799). QC: Quality control, GWAS: Genome-wide association study, HL: Hyperlipidemia

reductase (HMGCR). This enzyme, a bottleneck factor in cholesterol biosynthesis across mammalian cells, plays a crucial role in lipid metabolism [19,20]. Recently, the functional *HMGCR* intron 13 variant, rs3846662, has been identified for its influence on alternative splicing of *HMGCR* [21-23]. Similar to variants in *PCSK9* and *ACYL*, *HMGCR* variants, as determined by genetic risk scores, have demonstrated protective effects against cardiovascular events [8,24]. In our prior investigations utilizing regional association analysis, we successfully uncovered the pivotal impact of ethnicity-specific variants on the genetic factors influencing LDL-C levels among the individuals from the Taiwan Biobank (TWB), a comprehensive database derived from a population-based cohort study spanning Taiwan [9,25-27]. In addition, MR studies, as outlined in our previous work, have convincingly revealed the causal association between LDL-C and the susceptibility of DM in connection with *PCSK9* and *APOB* variants [9,25]. In the current study, we have delved into understanding the role played by *HMGCR* variants in determining the levels of circulating lipids. Furthermore, our investigation has extended to assessing the causal relationships between LDL-C levels and the susceptibility of DM based on the distinct genotypes of *HMGCR* within the Taiwanese population.

Materials and methods

Study population

Detailed information on the participants from the TWB involved in this study cohort has been previously outlined [9,25,26]. In brief, the TWB cohort study, which is population-based, recruited participants from centers across Taiwan between February 2008 and December 2020. The initial enrolment comprised 129,542 individuals who underwent whole-genome genotyping array analysis. Following the application of exclusion criteria, which included stringent quality control (QC) for array data, fasting duration of <6 h, absence of imputation data, second degree or closer relatives (descent PI_HAT > 0.187), and a history of hyperlipidemia, 108,880 eligible individuals from the TWB were selected for the current regional association studies and genotype–phenotype analyses. The participant enrolment process is visually depicted in Figure 1.

This study received approval from the Institutional Review Board of Taipei Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, with the approval number 08-XD-005, and the Ethical Governance Committee of the TWB, with the approval number TWBR11107-03. This study was conducted in accordance with the Declaration of Helsinki. Before the collection of original data, each participant provided written informed consent, adhering to the ethical standards and guidelines.

Lipid profile detection

The serum lipid profile, including total cholesterol (TC), LDL-C, high-density lipoprotein cholesterol (HDL-C), and triglyceride (TG), was assessed as previously described [9,25,26]. Non-HDL-C levels were derived by deducting HDL-C from TC, while remnant cholesterol (RC)

levels were ascertained by subtracting the combined sum of HDL-C and LDL-C from TC.

DNA extraction and genotyping

We conducted the single-nucleotide variation genotyping using the Axiom Genome-Wide Array Plate System (Affymetrix, Santa Clara, CA, USA) with custom TWB chips. Our study analyzed genotyping data from two genome-wide association study (GWAS) arrays: the GWAS CHB-1 Array (TWBv1.0) and CHB-2 Array (TWBv2.0) which included 27,685 and 103,316 individuals, respectively. Whole-genome sequencing (WGS) was employed to confirm the genotyping data from TWBv1.0 and TWBv2.0 arrays (n = 1,459), showing the consistency rate 99.92% and 99.80% for TWBv1.0 and TWBv2.0, respectively.

Genome-wide association study analysis

For GWAS of LDL-C levels, the PLINK software (version 1.09; https://www.cog-genomics.org/plink/; Shaun Purcell, Cambridge, MA, USA; accessed on March 6, 2023) was employed. GWAS analyses were conducted under an additive genetic model, using linear regression to evaluate the association between each single-nucleotide polymorphism and log-transformed LDL-C levels. Covariates included age, sex, body mass index (BMI), and smoking status. Genome-wide significance was defined as $P < 5 \times 10^{-8}$.

Regional association study with conditional analysis

In our investigation, we conducted a regional association study incorporating conditional analysis to identify polymorphisms within the *HMGCR* region influencing LDL-C levels, utilizing a methodology established in previous studies [9,25,26]. Genome-wide genotype imputation was performed using SHAPEIT and IMPUTE2, with the 1000 Genomes Project Phase 3 East Asian populations served as the reference. Subsequent QC measures, including stringent filtering of single-nucleotide polymorphisms (SNPs) based on IMPUTE2 imputation quality scores >0.3, call rate ≥97%, missing rate <3%, minor allele frequency (MAF) ≥0.01, and P value for the Hardy–Weinberg equilibrium of $\geq 10^{-6}$, were applied. Following QC, 108,880 participants were selected for the regional association study, focusing on 1007 SNPs near the *HMGCR* gene on chromosome 5q13.3, specifically at positions 74.53–74.76 Mb. A $P < 5 \times 10^{-8}$ was considered to indicate genome-wide significance.

Statistical analysis

Continuous variables, presented as medians (interquartile ranges), underwent the analysis for normal distribution using the Kolmogorov–Smirnov test. The categorical data, expressed as percentages, were assessed for distribution using the Chi-squared test. To meet normality assumptions, a logarithmic transformation was applied to study the parameters before analysis. Linear regression analysis was employed to analyze regional association in the study. The regression models were adjusted for age, sex, BMI, and smoking status. The independent factors associated with the studied phenotype were assessed utilizing the stepwise linear regression. Employing the assumption of an additive genetic effect and adjusting for gender, age, BMI, and smoking, general linear regression analysis was conducted to explore the associations between the genotypes and target phenotypes. Genotype–phenotype association analysis was conducted using SPSS (version 23; SPSS, Chicago, IL, USA).

Mendelian randomization analysis for 3-Hydroxy-3-methylglutaryl-CoA reductase variants and weighted genetic risk scores for causal association between low-density lipoprotein cholesterol levels and risk of diabetes mellitus

A two-stage least squares (2SLS) regression analysis was conducted to examine the causal relationship between LDL-C levels and DM risk using two separate sets of instrumental variables (IVs) in MR analyses. The first MR analysis utilized *HMGCR* variants with known functional relevance, selected for their strong linkage disequilibrium (LD; $R^2 > 0.8$) with LDL-C GWAS-significant lead SNPs identified through regional plotting and conditional association analyses; for example, although rs3064191 was a lead SNP in the LDL-C GWAS, it was excluded due to its intronic, nonfunctional nature. This analysis included a weighted genetic risk score (WGRS) derived from these *HMGCR* variants. The second MR analysis employed a WGRS based on 47 lead SNPs identified in the LDL-C GWAS, with each SNPs effect allele weighted by its effect size on LDL-C levels. Each analysis was conducted separately to assess the potential causal link between LDL-C levels and DM risk. The analysis involved two steps: first, regressing *HMGCR* variants and WGRSs to predict LDL-C levels by linear regression analysis, and second, regressing the predicted LDL-C levels on DM risk by logistic regression analysis. We computed the WGRSs by summing the risk alleles at each locus, multiplied by their effect sizes derived from our regional association analysis. The included covariates adjusted for gender, age, smoking, BMI, and other possible confounders (e.g. estimated glomerular filtration rate, platelet count, hemoglobin, TG, aspartate aminotransferase, and RC). We considered the risk allele with consistently correlated associations with the target variables. The three core IV assumptions for MR in this study, including relevance, independence, and exclusion restriction, were described as follows: (1) Relevance: *HMGCR* variants and WGRSs are strongly associated with LDL-C levels. An F statistic >10 confirmed the strength of these IVs [28]. (2) Independence: IVs are independent of confounders affecting both LDL-C levels and DM risk. (3) Exclusion Restriction: IVs influence DM risk only through LDL-C levels, not through other pathways. To confirm these IVs as appropriate instruments, we further assessed potential pleiotropy. Although rs191835914 showed nominal associations with DM (P = 0.0127), it did not reach genome-wide significance and lost significance after adjusting for LDL-C, indicating vertical rather than horizontal pleiotropy.

Sensitivity analysis for Mendelian randomization studies

To perform sensitivity analyses, we used various statistical methods, such as funnel plots, inverse variance weighting (IVW), simple median, weighted median, and MR-Egger regression. Heterogeneity was assessed using scatter plots and Cochran's or Rücker's Q tests. Supplementary Method 1 provided more details.

Table 1: Associations of 3-Hydroxy-3-methylglutaryl-CoA reductase variants with serum lipid profile

Genetic variants	Genotypes			Beta	SE	*P*
***HMGCR* rs191835914**	**AA (105,617)**	**AC+CC (2140)**				
TC (mg/dL)	193.0 (171.0–217.0)	187.0 (166.0–209.0)		−0.0142	0.0017	1.41×10^{-17}
LDL-C (mg/dL)	119.0 (100.0–141.0)	131.0 (95.0–133.0)		−0.0232	0.0025	1.36×10^{-20}
Non-HDL-C (mg/dL)	138.0 (116.0–162.0)	131.0 (111.3–153.8)		−0.0203	0.0023	2.10×10^{-19}
HDL-C (mg/dL)	53.0 (45.0–63.0)	54.0 (45.0–64.0)		0.0003	0.0020	0.8611
TG (mg/dL)	91.0 (64.0–133.0)	88.0 (62.0–128.8)		−0.0095	0.0047	0.0415
***HMGCR* rs150454634**	**CC (105,642)**	**CA+AA (2142)**		**Beta**	**SE**	***P***
TC (mg/dL)	193.0 (171.0–217.0)	187.0 (166.0–209.0)		−0.0143	0.0017	9.11×10^{-18}
LDL-C (mg/dL)	119.0 (100.0–141.0)	131.0 (95.0–133.0)		−0.0234	0.0025	5.34×10^{-21}
Non-HDL-C (mg/dL)	138.0 (116.0–162.0)	131.0 (111.0–153.0)		−0.0204	0.0022	1.10×10^{-19}
HDL-C (mg/dL)	53.0 (45.0–63.0)	54.0 (45.0–64.0)		0.0004	0.0020	0.8352
TG (mg/dL)	91.0 (64.0–133.0)	88.0 (62.0–128.3)		−0.0093	0.0046	0.0447
***HMGCR* rs13354746**	**CC (61,923)**	**CT (39,750)**	**TT (6505)**	**Beta**	**SE**	***P***
TC (mg/dL)	192.0 (170.0–215.0)	194.0 (172.0–218.0)	196.0 (173.0–221.0)	0.0052	0.0004	1.30×10^{-42}
LDL-C (mg/dL)	118.0 (98.0–139.0)	120.0 (100.0–142.0)	123.0 (102.0–145.0)	0.0084	0.0006	4.08×10^{-49}
Non-HDL-C (mg/dL)	137.0 (115.0–160.0)	139.0 (117.0–163.0)	141.0 (119.0–165.0)	0.0075	0.0005	6.55×10^{-48}
HDL-C (mg/dL)	53.0 (45.0–63.0)	53.0 (45.0–63.0)	53.0 (45.0–63.0)	−0.0004	0.0005	0.3305
TG (mg/dL)	91.0 (64.0–133.0)	90.0 (64.0–134.0)	91.0 (64.0–135.0)	0.0022	0.0011	0.0432
***HMGCR* rs3064191**	**CC (27,657)**	**CTTGTAC (54,373)**	**CTTGTACTTGTA (26,471)**	**Beta**	**SE**	***P***
TC (mg/dL)	195.0 (173.0–219.0)	193.0 (171.0–217.0)	191.0 (169.0–214.0)	−0.0052	0.0003	2.39×10^{-57}
LDL-C (mg/dL)	122.0 (101.0–143.0)	119.0 (99.0–141.0)	117.0 (98.0–137.0)	−0.0082	0.0005	2.36×10^{-63}
Non-HDL-C (mg/dL)	140.0 (118.0–164.0)	138.0 (116.0–162.0)	135.0 (114.0–159.0)	−0.0075	0.0004	1.91×10^{-64}
HDL-C (mg/dL)	53.0 (45.0–63.0)	53.0 (45.0–63.0)	53.0 (45.0–63.0)	0.0005	0.0004	0.2107
TG (mg/dL)	91.0 (64.0–135.0)	90.0 (64.0–133.0)	91.0 (64.0–133.0)	−0.0028	0.0009	0.0020
***HMGCR* rs3846662**	**GG (31,122)**	**GA (54,298)**	**AA (23,460)**	**Beta**	**SE**	***P***
TC (mg/dL)	195.0 (173.0–219.0)	193.0 (171.0–217.0)	190.0 (169.0–213.0)	−0.0050	0.0003	6.54×10^{-53}
LDL-C (mg/dL)	121.0 (101.0–143.0)	119.0 (99.0–141.0)	117.0 (98.0–137.0)	−0.0078	0.0005	4.92×10^{-57}
Non-HDL-C (mg/dL)	140.0 (118.0–164.0)	138.0 (116.0–162.0)	135.0 (114.0–159.0)	−0.0072	0.0004	4.27×10^{-59}
HDL-C (mg/dL)	53.0 (45.0–63.0)	53.0 (45.0–63.0)	53.0 (45.0–63.0)	0.0004	0.0004	0.2651
TG (mg/dL)	91.0 (64.0–134.0)	90.0 (64.0–133.0)	91.0 (64.0–133.0)	−0.0025	0.0009	0.0061

P: Adjusted for gender, age, BMI, and smoking and determined using the linear regression. TC: Total cholesterol, LDL-C: Low-density lipoprotein cholesterol, HDL-C: High-density lipoprotein cholesterol, TG: Triglyceride, *HMGCR*: 3-Hydroxy-3-methylglutaryl-CoA reductase, SE: Standard error, BMI: Body mass index

Results

Regional association study for genetic variants around the 3-Hydroxy-3-methylglutaryl-CoA reductase gene region and linkage disequilibrium between the variants

We conducted comprehensive regional association analyses to discover the associations of genetic variations around the *HMGCR* region at positions 74.53–74.76 Mb on chromosome 5q13.3 with LDL-C levels [Supplementary Figure 1]. Our data disclosed the lead polymorphism, rs3064191, an insertion-deletion variant (ATTGT/) situated at the intron region of *HMGCR*, is associated with LDL-C levels ($P = 2.36 \times 10^{-63}$) with genome-wide association significance. Subsequent sequential conditional analysis, adjusting for rs3064191 genotypes, revealed rs150454634 as the second lead SNP ($P = 2.39 \times 10^{-14}$). Further adjustment for genotypes of both rs3064191 and rs150454634 identified rs13354746 as the third lead SNP ($P = 5.38 \times 10^{-11}$). We further investigated whether functional *HMGCR* variants could be identified as candidates associated with LDL-C levels. Supplementary Figure 2 and Supplementary Tables 1-3 showed that rs3846662, considered functional as it has been reported to influence exon 13 alternative splicing [23,29], was found in strong LD ($r^2 = 0.82$) with the lead SNP rs3064191, whereas one noteworthy variant rs191835914 (p.Y311S), an Asian-specific nonsynonymous mutation on exon 9, was found in nearly complete LD ($r^2 = 1.00$) with the second lead SNP rs150454634.

Association of nonsynonymous mutations in 3-Hydroxy-3-methylglutaryl-CoA reductase with serum lipid profile

We extended our investigation to explore the potential links of the five candidate variants with the serum lipid profile. The findings demonstrated strong associations, with all five variants exhibiting significant correlations with TC, LDL-C, and non-HDL-C levels [Table 1]. Due to the very low frequencies of the homozygous variants of rs191835914 and rs150454634, we combined heterozygous and homozygous forms of these two variants for analysis. After adjusting for age, sex, BMI, and current smoking, significant genome-wide associations were noted between rs191835914, rs150454634, rs13354746, rs3064191, and rs3846662 genotypes and TC, LDL-C, and non-HDL-C levels [Table 1].

Stepwise linear regression analysis results

We employed stepwise linear regression to investigate the association of selected *HMGCR* variants with serum lipid levels, adjusting for age, gender, BMI, and smoking. The

Table 2: Results of stepwise linear regression analysis for triglyceride, low-density lipoprotein cholesterol, and nonhigh-density lipoprotein cholesterol levels, including 3-Hydroxy-3-methylglutaryl-CoA reductase region polymorphisms

	TC			LDL-C			Non-HDL-C		
	Beta (SE)	R^2	*P*	Beta (SE)	R^2	*P*	Beta (SE)	R^2	*P*
Age (years)	0.0013 (0.00002)	0.0337	$<10^{-307}$	0.0014 (0.00003)	0.0166	$<10^{-307}$	0.0019 (0.00003)	0.0351	$<10^{-307}$
Sex (male vs. female)	0.0139 (0.0005)	0.0046	3.28×10^{-168}	--	--	--	−0.0055 (0.0008)	0.0008	2.61×10^{-13}
BMI (kg/m^2)	0.0015 (0.0001)	0.0052	8.60×10^{-131}	0.005 (0.0001)	0.0256	$<10^{-307}$	0.006 (0.0001)	0.0483	$<10^{-307}$
Current smoking (%)	--	--	--	−0.0028 (0.0009)	0.0001	0.0018	0.0029 (0.0009)	0.0001	0.0012
rs3846662 (GG vs. GA vs. AA)	−0.0034 (0.0004)	0.0020	4.25×10^{-18}	−0.0051 (0.0006)	0.0022	3.78×10^{-18}	−0.0048 (0.0005)	0.0022	5.40×10^{-20}
rs13354746 (CC vs. CT vs. TT)	0.003 (0.0004)	0.0004	3.17×10^{-11}	0.005 (0.0007)	0.0005	5.58×10^{-14}	0.0043 (0.0006)	0.0004	9.62×10^{-13}
rs191835914 (AA vs. AC vs. CC)	−0.0113 (0.0017)	0.0004	1.24×10^{-11}	−0.0186 (0.0025)	0.0005	8.99×10^{-14}	−0.0161 (0.0022)	0.0004	7.11×10^{-13}

BMI: Body mass index, Beta: β coefficient, SE: Standard error, TC: Total cholesterol, LDL-C: Low-density lipoprotein cholesterol, HDL-C: High-density lipoprotein cholesterol, *HMGCR*: 3-Hydroxy-3-methylglutaryl-CoA reductase

Table 3: Summary of coefficients used for standard Mendelian randomization analysis: Low-density lipoprotein cholesterol levels and diabetes mellitus

T_A	T_B	G_A	T_A–T_B		G_A–T_A		G_A–T_B		IV_A–T_B		IV_A–T_B–adjT_A
			Beta (SE)	*P*[a]	Beta (SE)	*P*[a]	Beta (SE)	*P*[a]	Beta (SE)	*P*	*P*[b]
LDL-C	DM	rs3846662	−2.3131 (0.0988)	$3.31\times10_{-121}$	−0.0078 (0.0005)	4.92×10^{-57}	0.0301 (0.0167)	0.0715	−3.7669 (2.0899)	0.0715[a] (0.0176[c])	0.3367
		rs191835914	−2.3131 (0.0988)	$3.31\times10_{-121}$	−0.0227 (0.0025)	4.16×10^{-20}	0.196 (0.0786)	0.0127	−8.5200 (3.4187)	0.0127[a] (0.0050[c])	0.0678
		WGRS_*HMGCR*_2SNPs	−2.3131 (0.0988)	$3.31\times10_{-121}$	0.9049 (0.0522)	3.31×10^{-67}	−4.6167 (1.7563)	0.0086	−5.1014 (1.9406)	0.0086[a] (0.0011[c])	0.1028
		rs13354746	−2.3131 (0.0988)	$3.31\times10_{-121}$	0.0084 (0.0006)	4.08×10^{-49}	−0.0211 (0.0196)	0.2805	−2.6420 (2.4482)	0.2805[a] (0.1608[c])	0.7114
		WGRS_*HMGCR*_3SNPs	−2.3131 (0.0988)	$3.31\times10_{-121}$	0.6497 (0.0345)	8.76×10^{-79}	−2.6535 (1.1782)	0.0243	−4.0823 (1.8126)	0.0243[a] (0.0044[c])	0.2112
		WGRS_LDL-C_47SNPs	−2.3131 (0.0988)	$3.31\times10_{-121}$	0.9818 (0.0175)	$<10^{-307}$	−2.6235 (0.6070)	1.70×10^{-5}	−2.6716 (0.6211)	1.70×10−5[a] (1.66×10−8[c])	0.9236

[a]Adjusted for sex, age, smoking, and BMI, [b]After further correction for LDL-C level, [c]Adjusted for gender, age, smoking, BMI, and other possible confounders (e.g., estimated glomerular filtration rate, platelet count, hemoglobin, triglyceride, aspartate aminotransferase, and RC). All *P* values were calculated using the enter method with logistic regression, except for G_A–T_A, which employed linear regression. WGRS_*HMGCR*_2SNPs: WGRSs of 2 *HMGCR* variants *HMGCR* rs3846662, *HMGCR* rs191835914; WGRS_*HMGCR*_3SNPs: WGRSs of 3 *HMGCR* variants *HMGCR* rs3846662, *HMGCR* rs191835914, *HMGCR* rs13354746, T_A and T_B: phenotypes A (LDL-C level) and B (DM), respectively. G_A: Genotypes determining T_A, IVA: Instrumental variables for GA, adjTA: Adjusting LDL-C level, DM: Diabetes mellitus, LDL-C: Low-density lipoprotein cholesterol, *HMGCR*: 3-Hydroxy-3-methylglutaryl-CoA reductase, RC: Remnant cholesterol, BMI: Body mass index

results indicated that the genotypes of rs3846662, rs13354746, and rs191835914 individually contributed to 0.20%, 0.04%, and 0.04% of the variance in TC levels; 0.22%, 0.05%, and 0.05% of the variance in LDL-C levels; and 0.22%, 0.04%, and 0.04% of the variance in non-HDL-C levels [Table 2].

Cause-and-effect relationship between low-density lipoprotein cholesterol and DM susceptibility: Mendelian randomization analysis results for weighted genetic risk scores of 3-Hydroxy-3-methylglutaryl-CoA reductase variants and of lead variants from genome-wide association study

We employed a 2SLS IV study to investigate the direction and causation of the connection between LDL-C and the occurrence of DM [Table 3]. The analysis encompassed three *HMGCR* variants independently associated with LDL-C levels to derive WGRS, displaying a significant association with DM prevalence ($P = 0.0243$). While the association of *HMGCR* WGRSs with DM stayed significant after accounting for possible variables linked to LDL-C levels ($P = 0.0044$), it lost significance upon further correction for LDL-C levels ($P = 0.2112$). Notably, the analysis of WGRSs for the two potentially functional *HMGCR* variants, rs3846662 and rs191835914, showed significant associations with DM prevalence ($P = 0.0086$). These associations gained significance with additional adjustments for multiple variables ($P = 0.0011$) but decreased in significance after further correction for LDL-C levels ($P = 0.1028$). Importantly, all *F*-statistic values for the LDL-C-level determining genotypes derived from the instruments were >10, indicating a low probability of weak instrument bias [Supplementary Table 4]. From the GWAS for LDL-C levels [Supplementary Figure 3 and Supplementary Table 5], 54 SNPs associated with LDL-C levels at the genome-wide significance level were identified after adjusting for sex, age, current smoking, and BMI. After further adjustment for LDL-C, the significant association between 7 of these variants and DM was found to be independent of LDL-C and thus excluded as IVs [Supplementary Table 6]. We then investigated the associations between the WGRSs derived from 47 LDL-C-determining variants and DM status [Table 3]. The WGRSs of the 47 gene variants and DM status were found to be significantly associated in the

2SLS IV analysis ($P = 1.70 \times 10^{-5}$). These associations persisted even after adjusting for several parameters related to LDL-C levels ($P = 1.66 \times 10^{-8}$); however, they became insignificant when LDL-C levels were adjusted ($P = 0.9236$). We further compared the odds ratios for the WGRSs derived from LDL-C-level–determining genotypes with observational LDL levels for DM risk [Figure 2]. Upon reanalyzing the 108,880 TWB participants, the results revealed an opposite association between DM prevalence and genetically predisposed LDL-C levels. The corresponding effect size for *HMGCR* variants was 0.502–0.574 (per 1 mmol/L increase in LDL-C level). As more LDL-determining SNPs included as IVs, the effect size of DM risk approaches the observational odds from plasma LDL levels.

Standard Mendelian randomization sensitivity analyses for causal inference on low-density lipoprotein cholesterol and diabetes mellitus using multiple genetic variants

Further sensitivity analyses confirmed the causal effects determined using the 2SLS method, with consistent effect size estimates across various approaches, such as IVW, simple and weighted median, and MR-Egger regression (slope) methods [Supplementary Table 7]. The symmetrical funnel plots suggested no directional pleiotropy effects across all IV estimates and IV strengths [Supplementary Figure 4]. Scatter plots indicated that the relationships between alleles determining LDL-C levels and DM are dependent on LDL-C levels [Supplementary Figure 5]. Cochran's Q did not identify significant heterogeneity [Supplementary Table 8].

Discussion

In this study, we performed a thorough analysis of the association between *HMGCR* variants and the serum lipid profile in participants from the TWB. Three *HMGCR* variants, including the Asian-specific nonsynonymous mutation rs191835914, were identified as autonomously linked with TC, LDL-C, and non-HDL-C levels. Collectively, these variants accounted for 0.28%, 0.32%, and 0.30% of the variations in TC, LDL-C, and non-HDL-C levels, respectively. In addition, employing MR analyses with 2SLS regression using *HMGCR* WGRSs as genetic instruments, we observed a robust inverse relationship between heritably predisposed LDL-C levels and the occurrence of DM. These findings contribute to the growing body of evidence supporting the crucial role of *HMGCR* variants in influencing the cardiovascular outcomes within the Taiwanese population. Consequently, they could potentially inform future approaches to the prevention and treatment of cardiovascular diseases.

3-Hydroxy-3-methylglutaryl-CoA reductase rs3846662 variant: A splicing-site mutation

A functional splicing variant, rs3846662 is identified associated with the exclusion of 53 amino acids in exon 13, resulting in a HMGCR isoform with catalytically lower or null activity [21,23]. An *in vitro* pharmacogenetic study revealed that statin treatment modulates alternative splicing, influenced by rs3846662 genotypes, and is linked to an inverse *in vivo* statin response [23]. Positioned in the binding motif of competing splicing factors, the rs3846662 variant's minor A allele disrupts the motif, promoting exon 13 retention. This disruption leads to preferential binding of nuclear ribonucleoprotein A1, resulting in exon 13 skipping [29]. Thus, our findings reinforce the notion that rs3846662 is the most potent known functional *HMGCR* variant, exerting a profound impact on LDL-C levels in the individuals of Taiwanese descent.

3-Hydroxy-3-methylglutaryl-CoA reductase rs191835914 variant: An Asian-specific nonsynonymous mutation

The association between *HMGCR* nonsynonymous mutations and LDL-C levels has been a topic of uncertainty. In a study by Das *et al*. [30], where 375 missense mutations in *HMGCR* were analyzed, 7 rare high-risk mutations specific to certain ethnic populations were identified. Among these, rs147043821 and rs193026499 were highlighted as potentially damaging mutations in the *HMGCR* PolyPhen program. However, conclusive links between these mutations and circulating total TC or LDL-C levels have yet to be established. In contrast, previous studies by Lu *et al*. [31,32] indicated that the rs191835914 genotypes are linked to both LDL-C levels and coronary artery disease risk. This current study provides novel insights by confirming that rs191835914 (p.Y311S), an Asian-specific nonsynonymous mutation exhibiting robust LD with the lead variant rs150454634, is strongly associated with TC, LDL-C, and non-HDL-C levels among individuals in Taiwan. With a MAF of 0.0098 in the analyzed participants from the TWB, rs191835914 demonstrates a higher MAF in East Asian populations (0.0103–0.0277) compared to South Asian (0.007), European (0.0001–0.0002), and African (0–0.00006) populations, according to PubMed data (PubMed.gov). While no other *HMGCR* missense mutation has been reported to be linked to LDL-C levels in the NHGRI-EBI GWAS Catalog database (www.ebi.ac.uk/gwas; [33]), Yogev *et al*.[34] reported a family with autosomal recessive limb-girdle muscular disease due to a loss-of-function *HMGCR* mutation, resulting in statin myopathy. In this family, affected members exhibited elevated fasting blood sugar

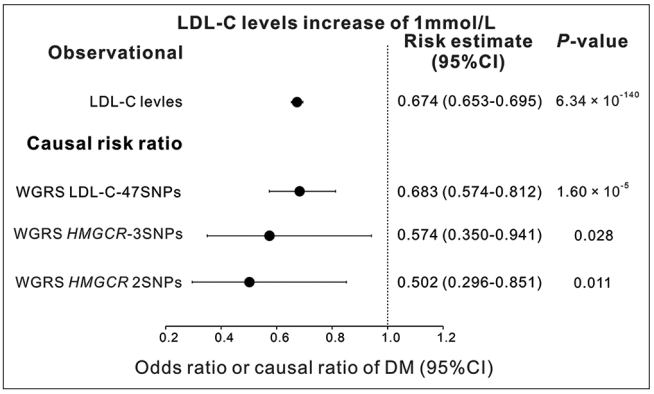


Figure 2: Forest plot demonstrating observational and associations of elevated Low-density lipoprotein cholesterol (LDL-C) levels with diabetes mellitus (DM) risk. For causal risk ratios, each weighted genetic risk score was oriented to mimic the putative estimate effects corresponding to the relative odds for DM per 1 mmol/L in elevated LDL-C levels. *P*: Adjusted for age, sex, current smoking status, BMI, and other possible confounders (e.g. estimated glomerular filtration rate, platelet count, hemo-globin, triglyceride, aspartate aminotransferase, and remnant cholesterol). LDL-C: Low-density lipoprotein cholesterol, DM: Diabetes mellitus, WGRS: Weighted genetic risk score

levels (>126 mg/dL) alongside low, yet within the normal range and blood cholesterol levels. However, the pathogenicity of *HMGCR* mutations and their association with clinical diseases are needed to further determine through *in vitro* or *in vivo* studies.

3-Hydroxy-3-methylglutaryl-CoA reductase rs13354746 variant

In our investigation, we observed that the rs13354746 genotypes are independently linked to LDL-C levels. This variant, situated in the intergenic region between *ANKRD31* and *HMGCR*, has previously been associated with coronary artery disease and myocardial infarction [35-37]. In a metabolome-wide association study utilizing 249 metabolomic measures, Davyson *et al.* [38] found that the rs13354746 genotypes are significantly correlated with cholesterol to total lipids in small LDL percentage. Our current study adds further evidence suggesting that the association between rs13354746 genotypes and ischemic heart disease phenotypes may be attributed to the effect of these genotypes on LDL-C levels.

Mendelian randomization analysis

The observed causal connection between LDL-C and diabetes may exhibit specificity related to genes or mechanisms [11,18]. Our previous analysis of 7 common and rare *PCSK9* variants, 3 nonsynonymous *APOB* mutations, and 41 lead SNPs identified in a GWAS for LDL-C levels in TWB participants provided evidence strengthening a causal connection between LDL-C levels and the occurrence of DM [9,25]. In our prior GWAS with a relatively small population (75,441 TWB participants), the lead *HMGCR* SNP variant rs3064191 showed no significant association with DM (P = 0.5325). In the present study, with an expanded sample size of over 100,000 participants, we conducted a more robust investigation of the potential association by combining variants with WGRSs from *HMGCR* and the whole-genome GWAS [Figure 2]. Our findings on *HMGCR* variants and WGRSs offer evidence supporting an opposite correlation between LDL-C levels and susceptibility to DM. Exploring additional candidate genes linked to LDL-C levels may contribute to an improved understanding of the complex genetic pathways influencing the susceptibility to DM.

Limitations

This study's cross-sectional design introduces the potential for survival bias and inherently constrains the exploration of baseline LDL-C effects on DM incidence. Furthermore, due to the genetic diversity across various ethnicities, the generalizability of our findings to other ethnic groups may be limited. Although we did not incorporate a second cohort for result replication within our Taiwanese cohort, the consistency of causal effects between LDL-C levels and DM observed in several studies involving individuals from the diverse ethnic backgrounds supports the robustness of our conclusions [8,10,13,15,16].

Conclusions

Our study reveals the independent associations of both coding and noncoding *HMGCR* variants with LDL-C levels among individuals in Taiwan. Utilizing genetic instruments in MR analyses, including WGRSs from *HMGCR* and other candidate gene variants, we identified an inverse causal relationship between LDL-C levels and DM. These findings underscore the significant impact of *HMGCR* variants on cardiometabolic outcomes and offer valuable insights for the prevention and treatment of cardiovascular diseases.

Acknowledgments

We express our sincere gratitude for the technical support provided by the Core Laboratory of the Taipei Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, and for the expert statistical analysis assistance from Tsung-Han Hsieh.

Data availability statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Financial support and sponsorship

Financial support for this study was provided through grants from the National Science and Technology Council (NSTC 112-2314-B-303-023-MY3), Taipei Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation (TCRD-TPE-111-12), and Buddhist Tzu Chi Medical Foundation (TCMF-A 112-02, TCMF-EP 111-02).

Conflicts of interest

Dr. Yu-Lin Ko, an editorial board member at *Tzu Chi Medical Journal*, had no role in the peer review process of or decision to publish this article. The other authors declared no conflicts of interest in writing this paper.

References

1. Johannesen CD, Langsted A, Mortensen MB, Nordestgaard BG. Association between low density lipoprotein and all cause and cause specific mortality in Denmark: Prospective cohort study. BMJ 2020;371:m4266.
2. Prospective Studies Collaboration, Lewington S, Whitlock G, Clarke R, Sherliker P, Emberson J, et al. Blood cholesterol and vascular mortality by age, sex, and blood pressure: A meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. Lancet 2007;370:1829-39.
3. Cholesterol Treatment Trialists' (CTT) Collaboration, Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet 2010;376:1670-81.
4. Cholesterol Treatment Trialists' (CTT) Collaboration, Fulcher J, O'Connell R, Voysey M, Emberson J, Blackwell L, et al. Efficacy and safety of LDL-lowering therapy among men and women: Meta-analysis of individual data from 174,000 participants in 27 randomised trials. Lancet 2015;385:1397-405.
5. Luirink IK, Wiegman A, Kusters DM, Hof MH, Groothoff JW, de Groot E, et al. 20-year follow-up of statins in children with familial hypercholesterolemia. N Engl J Med 2019;381:1547-56.
6. Silverman MG, Ference BA, Im K, Wiviott SD, Giugliano RP, Grundy SM, et al. Association between lowering LDL-C and cardiovascular risk reduction among different therapeutic interventions: A systematic review and meta-analysis. JAMA 2016;316:1289-97.

7. Yusuf S, Lonn E, Pais P, Bosch J, López-Jaramillo P, Zhu J, et al. Blood-pressure and cholesterol lowering in persons without cardiovascular disease. N Engl J Med 2016;374:2032-43.
8. Ference BA, Robinson JG, Brook RD, Catapano AL, Chapman MJ, Neff DR, et al. Variation in PCSK9 and HMGCR and risk of cardiovascular disease and diabetes. N Engl J Med 2016;375:2144-53.
9. Hsu LA, Teng MS, Wu S, Chou HH, Ko YL. Common and rare PCSK9 variants associated with low-density lipoprotein cholesterol levels and the risk of diabetes mellitus: A Mendelian randomization study. Int J Mol Sci 2022;23:10418.
10. Schmidt AF, Swerdlow DI, Holmes MV, Patel RS, Fairhurst-Hunter Z, Lyall DM, et al. PCSK9 genetic variants and risk of type 2 diabetes: A Mendelian randomisation study. Lancet Diabetes Endocrinol 2017;5:97-105.
11. Tragante V, Asselbergs FW, Swerdlow DI, Palmer TM, Moore JH, de Bakker PI, et al. Harnessing publicly available genetic data to prioritize lipid modifying therapeutic targets for prevention of coronary heart disease based on dysglycemic risk. Hum Genet 2016;135:453-67.
12. White J, Swerdlow DI, Preiss D, Fairhurst-Hunter Z, Keating BJ, Asselbergs FW, et al. Association of lipid fractions with risks for coronary artery disease and diabetes. JAMA Cardiol 2016;1:692-9.
13. Filippatos TD, Panagiotopoulou T, Tzavella E, Elisaf MS. Hypolipidemic drugs and diabetes mellitus-mechanisms and data from genetic trials. J Cardiovasc Pharmacol Ther 2018;23:187-91.
14. Schmidt AF, Holmes MV, Preiss D, Swerdlow DI, Denaxas S, Fatemifar G, et al. Phenome-wide association analysis of LDL-cholesterol lowering genetic variants in PCSK9. BMC Cardiovasc Disord 2019;19:240.
15. Swerdlow DI, Preiss D, Kuchenbaecker KB, Holmes MV, Engmann JE, Shah T, et al. HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: Evidence from genetic analysis and randomised trials. Lancet 2015;385:351-61.
16. Lotta LA, Sharp SJ, Burgess S, Perry JR, Stewart ID, Willems SM, et al. Association between low-density lipoprotein cholesterol-lowering genetic variants and risk of type 2 diabetes: A meta-analysis. JAMA 2016;316:1383-91.
17. Reith C, Preiss D, Blackwell L, Emberson J, Spata E, Davies K, *et al*. Effects of statin therapy on diagnoses of new-onset diabetes and worsening glycaemia in large-scale randomised blinded statin trials: An individual participant data meta-analysis. Lancet Diabetes Endocrinol 2024;12:306-19.
18. Pan W, Sun W, Yang S, Zhuang H, Jiang H, Ju H, et al. LDL-C plays a causal role on T2DM: A Mendelian randomization analysis. Aging (Albany NY) 2020;12:2584-94.
19. Schachter M. Chemical, pharmacokinetic and pharmacodynamic properties of statins: An update. Fundam Clin Pharmacol 2005;19:117-25.
20. Endo A. The discovery and development of HMG-CoA reductase inhibitors. J Lipid Res 1992;33:1569-82.
21. Burkhardt R, Kenny EE, Lowe JK, Birkeland A, Josowitz R, Noel M, et al. Common SNPs in HMGCR in micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon13. Arterioscler Thromb Vasc Biol 2008;28:2078-84.
22. Leduc V, Bourque L, Poirier J, Dufour R. Role of rs3846662 and HMGCR alternative splicing in statin efficacy and baseline lipid levels in familial hypercholesterolemia. Pharmacogenet Genomics 2016;26:1-11.
23. Medina MW, Gao F, Ruan W, Rotter JI, Krauss RM. Alternative splicing of 3-hydroxy-3-methylglutaryl coenzyme A reductase is associated with plasma low-density lipoprotein cholesterol response to simvastatin. Circulation 2008;118:355-62.
24. Ference BA, Ray KK, Catapano AL, Ference TB, Burgess S, Neff DR, et al. Mendelian randomization study of ACLY and cardiovascular disease. N Engl J Med 2019;380:1033-42.
25. Jang SJ, Tuan WL, Hsu LA, Er LK, Teng MS, Wu S, et al. Pleiotropic effects of APOB variants on lipid profiles, metabolic syndrome, and the risk of diabetes mellitus. Int J Mol Sci 2022;23:14963.
26. Teng MS, Yeh KH, Hsu LA, Chou HH, Er LK, Wu S, et al. Differential effects of ABCG5/G8 gene region variants on lipid profile, blood pressure status, and gallstone disease history in Taiwan. Genes (Basel) 2023;14:754.
27. Yeh KH, Wan HL, Teng MS, Chou HH, Hsu LA, Ko YL. Genetic variants at the APOE locus predict cardiometabolic traits and metabolic syndrome: A Taiwan biobank study. Genes (Basel) 2022;13:1366.
28. Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. Stat Methods Med Res 2012;21:223-42.
29. Yu CY, Theusch E, Lo K, Mangravite LM, Naidoo D, Kutilova M, et al. HNRNPA1 regulates HMGCR alternative splicing and modulates cellular cholesterol metabolism. Hum Mol Genet 2014;23:319-32.
30. Das KC, Hossain MU, Moniruzzaman M, Salimullah M, Akhteruzzaman S. High-risk polymorphisms associated with the molecular function of human HMGCR gene infer the inhibition of cholesterol biosynthesis. Biomed Res Int 2022;2022:4558867.
31. Lu X, Li J, Li H, Chen Y, Wang L, He M, et al. Coding-sequence variants are associated with blood lipid levels in 14,473 Chinese. Hum Mol Genet 2016;25:4107-16.
32. Lu X, Peloso GM, Liu DJ, Wu Y, Zhang H, Zhou W, et al. Exome chip meta-analysis identifies novel loci and East Asian-specific coding variants that contribute to lipid levels and coronary artery disease. Nat Genet 2017;49:1722-30.
33. Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, et al. The NHGRI-EBI GWAS catalog: Knowledgebase and deposition resource. Nucleic Acids Res 2023;51:D977-85.
34. Yogev Y, Shorer Z, Koifman A, Wormser O, Drabkin M, Halperin D, et al. Limb girdle muscular disease caused by HMGCR mutation and statin myopathy treatable with mevalonolactone. Proc Natl Acad Sci U S A 2023;120:e2217831120.
35. Aragam KG, Jiang T, Goel A, Kanoni S, Wolford BN, Atri DS, et al. Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. Nat Genet 2022;54:1803-15.
36. Koyama S, Ito K, Terao C, Akiyama M, Horikoshi M, Momozawa Y, et al. Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. Nat Genet 2020;52:1169-77.
37. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, et al. A cross-population atlas of genetic associations for 220 human phenotypes. Nat Genet 2021;53:1415-24.
38. Davyson E, Shen X, Gadd DA, Bernabeu E, Hillary RF, Mccartney DL, et al. Metabolomic investigation of major depressive disorder identifies a potentially causal association with polyunsaturated fatty acids. Biol Psychiatry 2023;94:630-9.

Supplementary material

Supplementary Method 1

Sensitivity analysis for determining causal relationships using standard Mendelian randomization (MR) with multiple genetic variants.

To explore the causal relationships between dependent and independent variables simultaneously, we first conducted multivariate analysis to minimize the influence of measurable confounding factors. We employed funnel plots, a straightforward method for detecting directional pleiotropy, to display the causal precision of instrumental variables (IVs) on the vertical axis against the IV estimates for each genetic variant on the horizontal axis. The ratios of the IV estimates obtained from individual genetic variants were aggregated using an inverse-variance weighted (IVW) meta-analysis formula (Burgess *et al*., 2013; Johnson *et al*., 2012 [S1,S2] to derive an overall estimate. The summary of the IVW analysis yielded similar mean and median estimates to those obtained using the two-stage least squares (2SLS) method commonly applied to individual-level data. The IVW method remains unbiased as the number of single nucleotide variants increases. In the IVW model, a random effects model was used to obtain the corrected standard error (Bowden *et al*., 2017) [S3].

The IVW estimate proved to be an efficient analysis method when all genetic variants were considered valid IVs. Conversely, a simple median estimator offered a reliable estimate of the causal effect when fewer than 50% of the examined genetic variants met the validity criteria (Bowden *et al*., 2016) [S4]. The IVW regression analysis was conducted using Meta-Essentials_1.4, an analytical software available at www.erim.eur.nl/research-support/meta-essentials. When there was considerable variability in individual estimates, relying solely on the simple median estimator proved inadequate. In such cases, the weighted median approach emerged as a more robust method. This method uses the median of the weighted ratios, employing standardized weights to ensure that the sum of the weights equals 1. Notably, the weighted median provides a consistent estimate as long as at least 50% of the weight comes from valid IVs. Both forms of median regression were performed using SPSS 22 statistical software (SPSS Inc., Chicago, IL, USA).

The Mendelian randomization (MR)-Egger regression approach, which originated in the field of meta-analysis (Egger *et al*., 1997) [S5], was used to evaluate small-study biases and identify overall directional pleiotropy across different genetic variants when the intercept term is not equal to zero (Bowden *et al*., 2016) [S4]. This method offers a reliable estimate of the true causal effect, maintaining consistency even when all genetic variants do not meet the IV assumption criteria, as described by Bowden *et al* [S4]. We implemented this methodology using Meta-Essentials_1.4 and SPSS analysis software. To simultaneously investigate the causal relationships between dependent and independent variables, we initially conducted multivariate analysis to mitigate the influence of measurable confounding factors. Funnel plots, a simple method for detecting directional pleiotropy, were used to illustrate the causal precisions of IVs on the vertical axis against the IV estimates for each genetic variant on the horizontal axis. The ratios of the IV estimates obtained from individual genetic variants were combined using a formula based on IVW meta-analysis (Burgess *et al*., 2013; Johnson *et al*., 2012) [S1,S2] to derive an overall estimate. The summary of IVW analysis give similar mean and median estimates to the 2SLS that commonly used with individual-level data. The IVW method remains unbiased as the number of single nucleotide variation increases. In IVW model, fitting a random effect was employed to receive the corrected standard error (Bowden *et al*., 2017) [S3].

The IVW estimate proved to be an efficient analysis method when all genetic variants were deemed valid IVs. Conversely, a simple median estimator offered a reliable estimate of the causal effect when fewer than 50% of the examined genetic variants meet the validity criteria (Bowden *et al*., 2016) [S4]. The IVW regression analysis was conducted using Meta-Essentials_1.4, an analytical software available at www.erim.eur.nl/research-support/meta-essentials. When there was considerable variability in individual estimates, relying solely on the simple median estimator proves inadequate. In such cases, the weighted median emerges as a more robust method. This method utilizes the median of the weighted ratio, employing standardized weights to ensure that the sum of the weights equals 1. Notably, the weighted median furnishes a consistent estimate provided that at least 50% of the weight stems from valid IVs. SPSS 22 statistical software (SPSS Inc., Chicago, IL, USA) was used to perform both forms of median regression. The MR-Egger regression approach, originating from the field of meta-analysis (Egger *et al*., 1997) [S5], served the purpose of evaluating small-study biases and identifying overall directional pleiotropy across distinct genetic variants when the intercept term is not equal to zero (Bowden *et al*., 2016) [S4]. It offers a reliable estimation of the genuine causal effect, maintaining consistency even when all genetic variants fail to meet the IV assumption criteria, as described by Bowden *et al* [S4]. We implemented this methodology utilizing Meta-Essentials_1.4 and SPSS analysis software.

According to the methodology described by Dai *et al*. [S6], scatter plots were employed to evaluate the consistency of IV effects. Consistent associations of independent genetic variants from different genomic regions with the outcome of interest suggest a robust estimated causal relationship. To examine the heterogeneity of causal estimates for each genetic variant associated with the outcome, we visually inspected the scatter plots and applied Cochran's or Rücker's Q test. These causal estimates were then compared with the genetic associations related to the exposure of interest. The procedures for these assessments are comprehensively detailed in the studies by Burgess *et al* [S7]. and Bowden *et al* [S8].

References

S1. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013;37:658-65.

S2. Johnson T, Uk S, editors. Efficient calculation for multi-SNP genetic risk scores. American Society of Human Genetics Annual Meeting; 2012: San Francisco.

S3. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med 2017;36:1783-802.

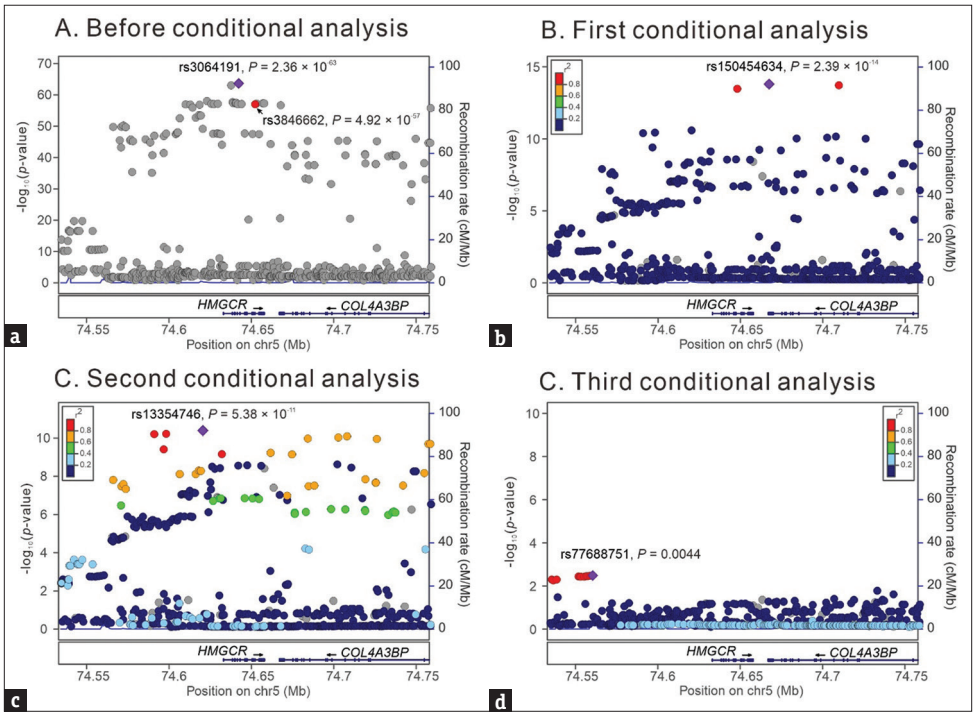
S4. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol 2016;40:304-14.

S5. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-34.

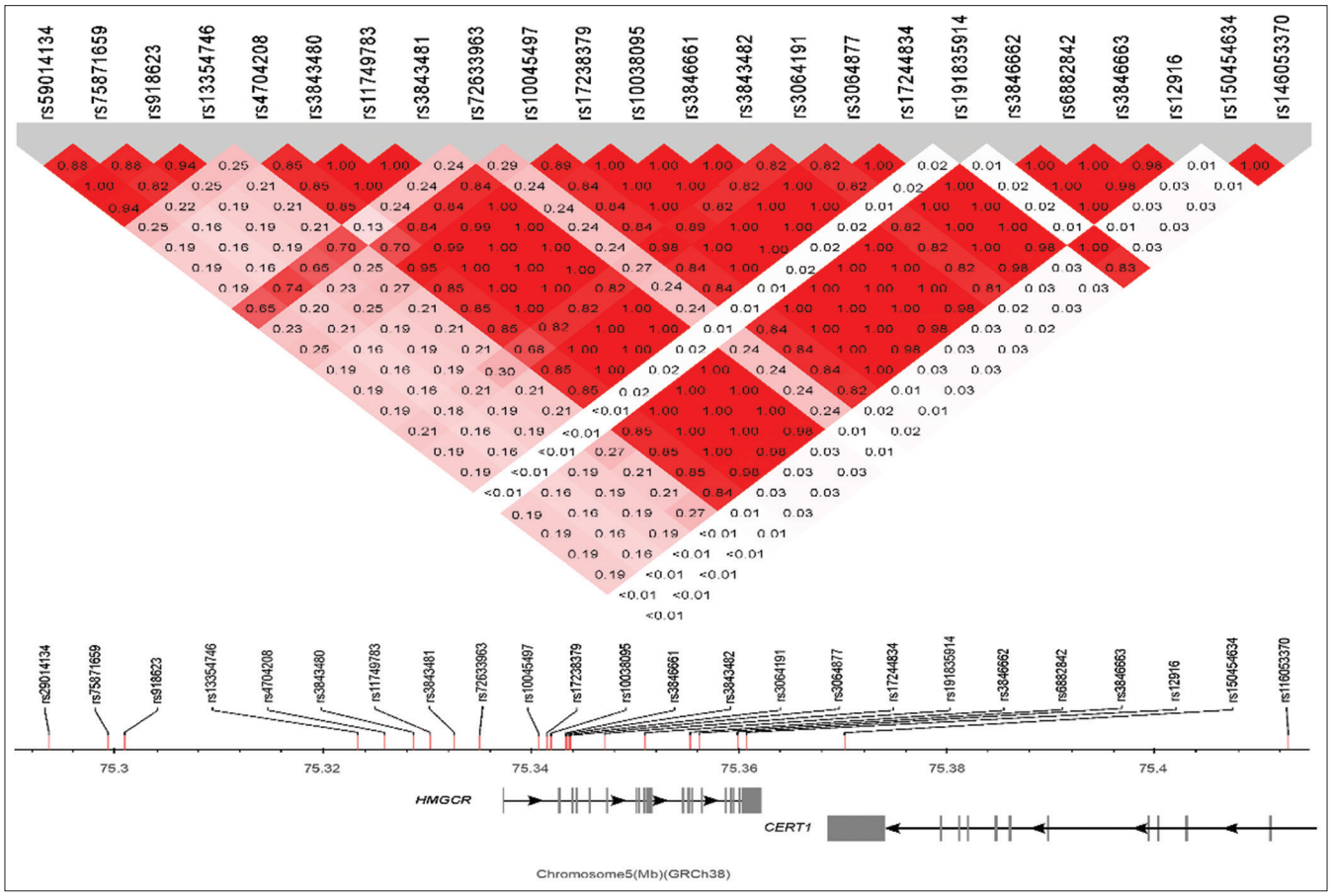
S6. Dai JY, Chan KC, Hsu L. Testing concordance of instrumental variable effects in generalized linear models with application to Mendelian randomization. Stat Med 2014;33:3986-4007.

S7. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. Epidemiology 2017;28:30-42.

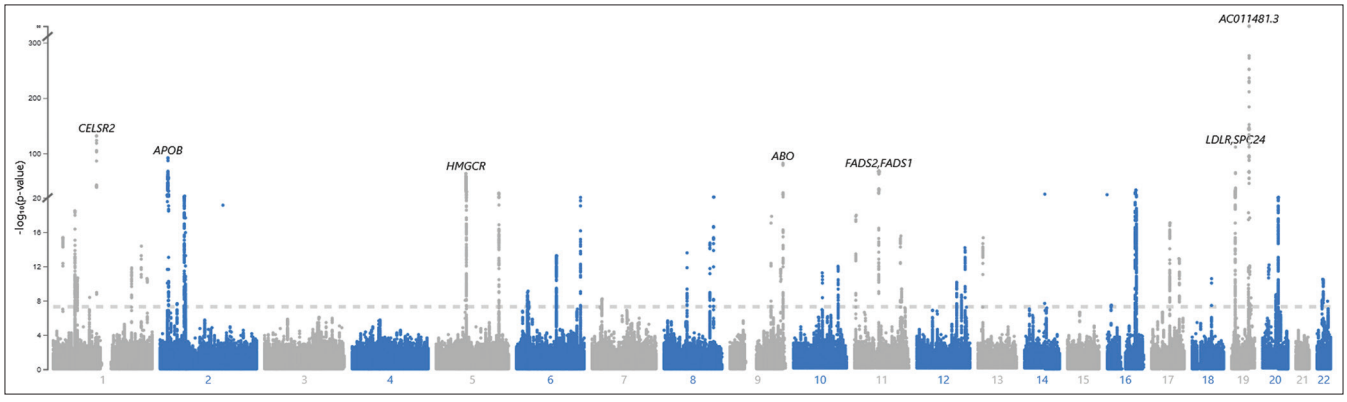
S8. Bowden J, Spiller W, Del Greco MF, Sheehan N, Thompson J, Minelli C, et al. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. Int J Epidemiol 2018;47:2100.



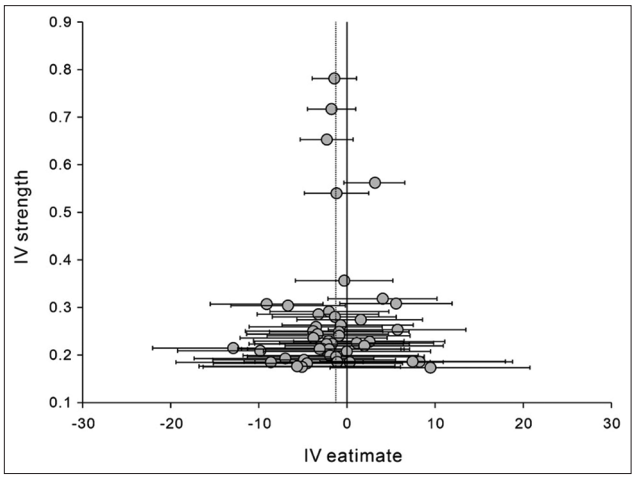
Supplementary Figure 1: Regional association analysis for correlation between 3-Hydroxy-3-methylglutaryl-CoA reductase (HMGCR) genetic variants and low-density lipoprotein cholesterol (LDL-C) levels. Regional association analysis using linear regression determined the genetic variants in a 230-kb region surrounding the *HMGCR* locus and their relationship with LDL-C levels. Without (a) and with sequential conditional adjustments for the genotypes of rs3064191 (b), rs150454634 (c), and rs13354746 (d). *P*: Adjusted for sex, age, body mass index, and current smoking status. HMGCR: 3-Hydroxy-3-methylglutaryl-CoA reductase



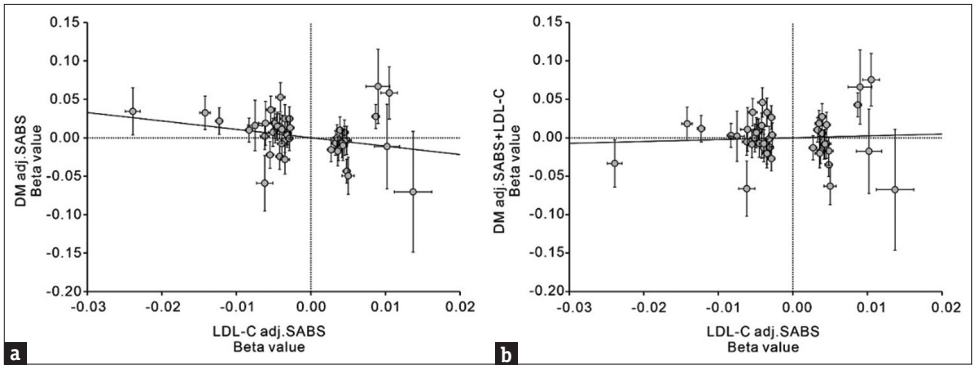
Supplementary Figure 2: Location and linkage disequilibrium (LD) map for 24 3-Hydroxy-3-methylglutaryl-CoA reductase (HMGCR) region variants in strong LD with any of the three lead HMGCR variants in regional association analysis. Red and gray shading denotes the strength of the pairwise LD, as indicated by r^2, and numbers indicate r^2 values. HMGCR: 3-Hydroxy-3-methylglutaryl-CoA reductase



Supplementary Figure 3: Genome-wide association study analysis for low-density lipoprotein cholesterol (LDL-C) levels in Taiwanese population. Manhattan plot revealed the genome wide significant gene loci for LDL-C levels of participants from Taiwan biobank. *P* value was adjusted for age, sex, current smoking status and body mass index. HMGCR: 3-Hydroxy-3-methylglutaryl-CoA reductase



Supplementary Figure 4: Funnel plots of the instrumental variable (IV) strength against IV estimates for each genetic variant separately for standard Mendelian randomization for the low-density lipoprotein cholesterol on diabetes mellitus. Horizontal lines represent 95% confident intervals for the IV estimates. Solid vertical lines are at the null, and dashed vertical lines are the (fixed-effect) inverse-variance weighted estimates. Instrumental variable strength: ($\frac{\hat{\beta}x \mid G}{SE(\hat{\beta}Y \mid G)}$) and instrumental variable estimates: ($\frac{\beta Y \mid G}{\beta X \mid G}$). IV: Instrumental variable



Supplementary Figure 5: Scatter plots for putative causal relationships between low-density lipoprotein cholesterol (LDL-C)-determining alleles on diabetes mellitus (DM) without (a) and with (b) adjustment of LDL-C. We exhibit the effect sizes of LDL-C-determining alleles on LDL-C and DM. Lines represent one standard error. LDL-C: Low-density lipoprotein cholesterol, DM: Diabetes mellitus

Supplementary Table 1: Selected 3-Hydroxy-3-methylglutaryl-CoA reductase variants with strong linkage disequilibrium (r^2>0.8) with the lead variant rs3064191 in regional association analysis before condition analysis

Chr	SNP	BP (GRCh37)	Ref/Alt	Func. refGene	Gene.refGene	ExonicFunc. refGene	AAChange. refGene	MAF	HWE	BETA	SE	*P*
5	rs4704208	74621671	A/G	Intergenic	*ANKRD31; HMGCR*	NA	NA	0.4683	0.9645	−0.0079	0.0005	5.475×10^{-58}
5	rs3843480	74624482	T/C	Intergenic	*ANKRD31; HMGCR*	NA	NA	0.4625	0.8284	−0.0078	0.0005	3.571×10^{-57}
5	rs11749783	74626082	T/C	Intergenic	*ANKRD31; HMGCR*	NA	NA	0.4626	0.8371	−0.0078	0.0005	4.158×10^{-57}
5	rs3843481	74628375	A/T	Intergenic	*ANKRD31; HMGCR*	NA	NA	0.4630	0.9115	−0.0078	0.0005	3.642×10^{-57}
5	rs10045497	74636484	C/A	Intronic	*HMGCR*	NA	NA	0.4917	0.8682	−0.0082	0.0005	4.853×10^{-63}
5	rs17238379	74637352	-/CTTA	Intronic	*HMGCR*	NA	NA	0.4638	0.9247	−0.0079	0.0005	1.107×10^{-57}
5	rs10038095	74637711	A/T	Intronic	*HMGCR*	NA	NA	0.4629	0.9247	−0.0078	0.0005	2.117×10^{-57}
5	rs3846661	74639178	A/G	Intronic	*HMGCR*	NA	NA	0.4625	0.9512	−0.0078	0.0005	2.661×10^{-57}
5	rs3843482	74639259	T/G	Intronic	*HMGCR*	NA	NA	0.4628	0.9468	−0.0078	0.0005	1.845×10^{-57}
5	rs3064191	74639549	TTGTA/-	Intronic	*HMGCR*	NA	NA	0.4923	0.8420	−0.0082	0.0005	2.36×10^{-63}
5	rs3064877	74639576	-/AA	Intronic	*HMGCR*	NA	NA	0.4628	0.9468	−0.0078	0.0005	1.845×10^{-57}
5	rs17244834	74642848	A/T	Intronic	*HMGCR*	NA	NA	0.4628	0.9867	−0.0078	0.0005	4.18×10^{-57}
5	rs3846662	74651084	A/G	Intronic	*HMGCR*	NA	NA	0.4625	0.9956	−0.0078	0.0005	4.918×10^{-57}
5	rs6882842	74651909	G/A	Intronic	*HMGCR*	NA	NA	0.4626	0.9911	−0.0078	0.0005	2.859×10^{-57}
5	rs3846663	74655726	C/T	Intronic	*HMGCR*	NA	NA	0.4631	1.0000	−0.0078	0.0005	2.471×10^{-57}
5	rs12916	74656539	T/C	UTR3	*HMGCR*	NA	NA	0.4628	0.9956	−0.0078	0.0005	2.495×10^{-57}

P: Adjusted for sex, age, BMI, and current smoking status and determined using the enter method with linear regression analysis.
BP: Position, MAF: Minor allele frequency, HWE: Hardy–Weinberg equilibrium, SE: Standard error, BMI: Body mass index, NA: Not available, *HMGCR*: 3-Hydroxy-3-methylglutaryl-CoA reductase, SNP: Single-nucleotide polymorphism

Supplementary Table 2: Selected 3-Hydroxy-3-methylglutaryl-CoA reductase variants with strong linkage disequilibrium (r^2>0.8) with the lead variant rs191835914 in regional association analysis during first conditional analysis

Chr	SNP	BP (GRCh37)	Ref/ Alt	Func. refGene	Gene. refGene	ExonicFunc.refGene	AAChange.refGene	MAF	HWE	BETA	SE	*P*
5	rs191835914	74646765	A/C	Exonic	*HMGCR*	Nonsynonymous SNV	exon9:c.A932C: p.Y311S	0.0098	0.566	−0.0187	0.0025	5.07×10^{-14}
5	rs150454634	74665994	C/A	Downstream	*COL4A3BP*	NA	NA	0.0098	0.566	−0.0190	0.0025	2.39×10^{-14}
5	rs146053370	74708473	A/G	Intronic	*COL4A3BP*	NA	NA	0.0098	0.567	−0.0189	0.0025	2.89×10^{-14}

P: adjusted for sex, age, BMI, current smoking status, and rs3064191 as well as determined by the enter method with linear regression analysis.
BP: Position, MAF: Minor allele frequency, HWE: Hardy–Weinberg equilibrium, SE: Standard error, BMI: Body mass index, NA: Not available, *HMGCR*: 3-Hydroxy-3-methylglutaryl-CoA reductase, SNP: Single-nucleotide polymorphism

Supplementary Table 3: Selected 3-Hydroxy-3-methylglutaryl-CoA reductase variants with strong linkage disequilibrium (r^2>0.8) with the lead variant rs13354746 in regional association analysis during the second conditional analysis

Chr	SNP	BP (GRCh37)	Ref/ Alt	Func. refGene	Gene.refGene	ExonicFunc. refGene	AAChange. refGene	MAF	HWE	BETA	SE	*P*
5	rs59014134	74589385	C/T	Intergenic	*ANKRD31; HMGCR*	NA	NA	0.237	0.1977	0.0045	0.0007	8.78×10^{-11}
5	rs75871659	74595090	C/G	Intergenic	*ANKRD31; HMGCR*	NA	NA	0.2078	0.0428	0.0043	0.0007	5.53×10^{-10}
5	rs918623	74596682	T/G	Intergenic	*ANKRD31; HMGCR*	NA	NA	0.2371	0.1959	0.0045	0.0007	8.35×10^{-11}
5	rs13354746	74619132	C/T	Intergenic	*ANKRD31; HMGCR*	NA	NA	0.2456	0.2256	0.0045	0.0007	5.38×10^{-11}
5	rs72633963	74630829	G/A	Intergenic	*ANKRD31; HMGCR*	NA	NA	0.2203	0.0688	0.0043	0.0007	9.10×10^{-10}

P: adjusted for sex, age, BMI, current smoking status, rs3064191 and rs150454634, as well as calculated using the enter method with linear regression analysis. BP: Position, MAF: Minor allele frequency, HWE: Hardy–Weinberg equilibrium, SE: Standard error, BMI: Body mass index, NA: Not available, *HMGCR*: 3-Hydroxy-3-methylglutaryl-CoA reductase, SNP: Single-nucleotide polymorphism

Supplementary Table 4: *F*-statistic values for 3-Hydroxy-3-methylglutaryl-CoA reductase variants

Chr	SNP	BP (GRCh37)	Ref/ Alt	Func. refGene	Gene.refGene	Exonic Func. refGene	AA Change. refGene	MAF	HWE	BETA	SE	*P*	*F*-statistics
5	rs13354746	74619132	C/T	Intergenic	*ANKRD31;HMGCR*	NA	NA	0.2456	0.2256	0.0045	0.0007	5.38×10^{-11}	2040.7
5	rs191835914	74646765	A/C	Exonic	*HMGCR*	Nonsynonymous SNV	exon9:c.A932C: p.Y311S	0.0098	0.5660	−0.0187	0.0025	5.07×10^{-14}	2000.7
5	rs3846662	74651084	A/G	Intronic	*HMGCR*	NA	NA	0.4625	0.9956	−0.0078	0.0005	4.918×10^{-57}	2066.4

P: Adjusted for sex, age, BMI, and current smoking status. MAF: Minor allele frequency, HWE: Hardy–Weinberg equilibrium, SE: Standard error, BMI: Body mass index, NA: Not available, *HMGCR*: 3-Hydroxy-3-methylglutaryl-CoA reductase, SNP: Single-nucleotide polymorphism

Supplementary Table 5: The association between variants and low-density lipoprotein cholesterol by genome-wide association study analysis

Chr	BP	SNP	Ref/Alt	Candidate gene	MAF	HWE	LogLDL-C_SABS			*F*-statistic	R^2
							BETA	SE	*P*		
19	44912383	rs445925	G/A	*AC011481.3*	0.0839	0.9410	−0.0586	0.0009	0.00E+00	1702.5647	0.0725
1	109274968	rs12740374	G/T	*CELSR2*	0.0667	0.2012	−0.0239	0.0010	1.08E-132	866.6680	0.0383
19	11131631	rs2738464	C/G	*LDLR, SPC24*	0.2757	0.5216	−0.0123	0.0005	2.51E-112	845.4500	0.0374
2	21024193	rs57825321	T/A	*APOB*	0.1476	0.0247	−0.0142	0.0007	5.12E-93	824.3578	0.0367
9	133266456	rs2519093	C/T	*ABO*	0.1817	0.5035	0.0123	0.0006	3.80E-83	818.3697	0.0362
11	61824164	rs174565	G/C	*FADS2, FADS1*	0.4143	0.0024	0.0087	0.0005	3.16E-69	800.2466	0.0356
5	75343719	rs3064191	ATTGT/-	*HMGCR*	0.4928	0.6405	−0.0083	0.0005	2.41E-64	800.9567	0.0356
16	72194607	rs3852789	A/C	*PMFBP1*	0.3167	0.6839	−0.0065	0.0005	6.89E-35	772.3752	0.0342
5	156970662	rs6882345	A/G	*TIMD4*	0.2700	0.2491	−0.0062	0.0006	4.83E-29	755.6535	0.0339
14	69778476	rs2296651	G/A	*SLC10A1*	0.0964	0.0002	−0.0089	0.0008	2.69E-27	765.7491	0.0340
16	377772	rs375498857	C/A	*TMEM8A*	0.0162	3.78E-07	−0.0208	0.0020	2.46E-26	757.9851	0.0341
2	62660749	rs6727888	C/T	*AC092155.1*	0.2757	0.9772	−0.0055	0.0005	1.06E-23	761.7250	0.0339
8	125467120	rs6982502	T/C	*AC091114.1*	0.4369	0.7683	0.0048	0.0005	5.78E-22	760.2358	0.0337
20	41325964	rs6124341	A/G	*ADI1P1*	0.2420	0.0707	−0.0054	0.0006	2.04E-21	758.9977	0.0338
6	160596331	rs73596816	G/A	*LPA*	0.0520	0.5259	0.0105	0.0011	2.87E-21	759.8713	0.0337
2	157625480	rs10164853	A/G	*ACVR1C*	0.3149	0.3634	−0.0048	0.0005	7.42E-20	758.1013	0.0336
1	55033483	rs7525649	T/C	*PCSK9*	0.3470	0.3157	−0.0046	0.0005	3.42E-19	758.2521	0.0336
11	5663172	rs1976339	C/G	*AC104389.5*	0.4549	0.9260	−0.0044	0.0005	1.09E-18	737.6861	0.0336
9	104903458	rs2575876	G/A	*ABCA1*	0.2309	0.0246	−0.0051	0.0006	1.46E-18	756.7675	0.0336
17	47685639	rs4794047	A/T	*KPNB1*	0.3444	0.6852	0.0045	0.0005	8.82E-18	742.3376	0.0334
11	116782580	rs2075290	T/C	*ZPR1*	0.2233	0.6172	0.0048	0.0006	2.87E-16	755.0261	0.0335
1	25449191	rs12027110	A/T	*MACO1*	0.2649	0.7257	0.0045	0.0006	4.41E-16	752.2152	0.0335
13	32402221	rs10492397	G/A	*N4BP2L1*	0.3548	0.1720	−0.0042	0.0005	4.63E-16	735.8722	0.0336
8	115658120	rs3808477	C/T	*TRPS1*	0.3071	0.4865	−0.0042	0.0005	1.92E-15	752.7866	0.0335
1	220799001	rs867772	G/A	*MARC1, AL445423.3*	0.1935	0.2748	−0.0049	0.0006	4.37E-15	744.8037	0.0336
12	120977112	rs2255531	G/A	*HNF1A-AS1*	0.3959	0.3027	0.0039	0.0005	6.77E-15	747.1827	0.0333
8	58493931	rs112784971	C/T	*CYP7A1*	0.2304	0.8087	−0.0045	0.0006	2.68E-14	737.1911	0.0332
6	100989473	rs873649	C/T	*AL133338.2*	0.2115	0.1680	0.0045	0.0006	5.78E-14	745.2236	0.0334
17	69086125	rs12162136	A/G	*ABCA6*	0.4270	1.0000	0.0037	0.0005	1.26E-13	737.1306	0.0332
20	17614241	rs116933131	C/T	*RRBP1*	0.0592	0.1721	−0.0075	0.0010	6.76E-13	740.7874	0.0335
10	112153464	rs2297991	C/T	*GPAM*	0.2844	0.6550	0.0039	0.0005	9.70E-13	751.7283	0.0333
1	196729914	rs10801558	T/G	*CFH*	0.4426	0.2688	−0.0035	0.0005	1.50E-12	751.8510	0.0334
9	127966535	rs74987020	AG/-	*FAM102A*	0.2175	0.5307	0.0042	0.0006	2.20E-12	747.1967	0.0332
20	17865277	rs2328223	A/C	*AL035045.1*	0.1976	0.3376	0.0043	0.0006	2.23E-12	751.3954	0.0333
10	72932888	rs41280378	T/G	*OIT3*	0.2539	0.9273	−0.0039	0.0006	5.72E-12	733.9323	0.0335
1	62586017	rs12030293	C/T	*DOCK7*	0.2010	0.1133	−0.0041	0.0006	2.33E-11	750.5629	0.0333
1	234712862	rs486142	A/G	*AL160408.6*	0.2252	0.5629	−0.0039	0.0006	2.59E-11	749.8484	0.0333
18	49594230	rs9953437	G/A	*LIPG*	0.4039	0.9574	0.0033	0.0005	2.71E-11	740.9514	0.0332
22	30036855	rs5844888	C/-	*HORMAD2-AS1*	0.1205	0.9893	0.0050	0.0008	3.20E-11	742.6637	0.0333
9	128704210	rs13289095	G/T	*PKN3*	0.0788	0.4550	−0.0061	0.0009	3.21E-11	738.5525	0.0331
12	100528401	rs2373355	G/A	*NR1H4*	0.2929	0.2215	0.0035	0.0005	7.20E-11	740.9944	0.0333
11	118558115	rs10677211	-/GATAA	*IFT46*	0.4815	0.5977	−0.0031	0.0005	4.48E-10	729.9370	0.0329
6	29939440	rs9260028	G/A	*HLA-A*	0.3294	0.1336	−0.0032	0.0005	7.76E-10	746.1832	0.0332
20	35555321	rs224419	A/G	*ERGIC3*	0.2230	0.0863	−0.0035	0.0006	1.95E-09	747.9106	0.0332
12	112399801	rs75919415	G/T	*RPL6*	0.2146	0.6719	0.0036	0.0006	2.22E-09	741.8088	0.0332
1	92219224	rs75573831	C/T	*C1orf146*	0.0206	0.0400	0.0102	0.0017	4.17E-09	742.6020	0.0331
19	49496752	rs2280401	G/A	*RPS11*	0.2213	0.9868	−0.0035	0.0006	4.62E-09	748.3824	0.0332
7	26341732	rs12666571	G/C	*SNX10*	0.4366	0.8370	−0.0029	0.0005	6.09E-09	723.8682	0.0333
7	25975772	rs28537499	A/C	*MIR148A*	0.4417	0.0985	−0.0029	0.0005	7.18E-09	742.6429	0.0331
2	21612193	rs146112956	A/-	*AC018742.1, AC009411.2*	0.0260	0.9548	0.0090	0.0016	9.92E-09	731.5452	0.0331
22	42138592	rs2743450	A/G	*AC254562.2, AC254562.3*	0.4514	0.3015	−0.0028	0.0005	1.15E-08	734.1073	0.0333
14	69445082	rs138089855	T/C	*SLC39A9*	0.0508	0.0472	−0.0062	0.0011	2.04E-08	742.2635	0.0332
2	43738159	rs75832441	G/A	*PLEKHH2*	0.0102	0.8880	0.0137	0.0025	2.16E-08	747.0304	0.0332
16	11539890	rs12596176	A/G	*LITAF*	0.4438	0.1985	0.0027	0.0005	3.34E-08	744.9230	0.0332

P value was calculated using linear regression after adjusting sex, age, BMI, and smoking (SABS). MAF: Minor allele frequency, HWE: Hardy–Weinberg equilibrium, SE: Standard error, BMI: Body mass index, SNP: Single-nucleotide polymorphism, *HMGCR*: 3-Hydroxy-3-methylglutaryl-CoA reductase

Supplementary Table 6: Low-density lipoprotein cholesterol and diabetes mellitus according to the low-density lipoprotein cholesterol-determining genotypes with genome-wide significance ($P<5 \times 10^{-8}$) in Taiwan Biobank participants

Chr	BP	SNP	A1	Candidate gene	LogLDL_C_SABS			DM_SABS			DM_SABS+LogLDL		
					BETA	SE	*P*	BETA	SE	*P*	BETA	SE	*P*
19	44912383	rs445925	G/A	*AC011481.3*	−0.0586	0.0009	0.00E+00	−0.0257	0.0280	0.3583	−0.1778	0.0307	6.78E-09
1	109274968	rs12740374	G/T	*CELSR2*	−0.0239	0.0010	1.08E-132	0.0344	0.0306	0.2651	−0.0333	0.0333	0.3176
19	11131631	rs2738464	C/G	*LDLR, SPC24*	−0.0123	0.0005	2.51E-112	0.0218	0.0172	0.2100	0.0119	0.0188	0.5265
2	21024193	rs57825321	T/A	*APOB*	−0.0142	0.0007	5.12E-93	0.0325	0.0217	0.1319	0.0182	0.0235	0.4384
9	133266456	rs2519093	C/T	*ABO*	0.0123	0.0006	3.80E-83	0.0237	0.0198	0.2226	0.0594	0.0217	0.0062
11	61824164	rs174565	G/C	*FADS2, FADS1*	0.0087	0.0005	3.16E-69	0.0276	0.0155	0.0786	0.0427	0.0170	0.0122
5	75343719	rs3064191	ATTGT/-	*HMGCR*	−0.0083	0.0005	2.41E-64	0.0100	0.0154	0.5293	0.0031	0.0169	0.8547
16	72194607	rs3852789	A/C	*PMFBP1*	−0.0065	0.0005	6.89E-35	−0.0131	0.0166	0.4299	−0.0570	0.0182	0.0018
5	156970662	rs6882345	A/G	*TIMD4*	−0.0062	0.0006	4.83E-29	0.0020	0.0174	0.9173	−0.0053	0.0190	0.7821
14	69778476	rs2296651	G/A	*SLC10A1*	−0.0089	0.0008	2.69E-27	−0.0831	0.0267	0.0019	−0.1219	0.0293	3.10E-05
16	377772	rs375498857	C/A	*TMEM8A*	−0.0208	0.0020	2.46E-26	0.2295	0.0574	0.0001	0.2388	0.0618	0.0001
2	62660749	rs6727888	C/T	*AC092155.1*	−0.0055	0.0005	1.06E-23	−0.0223	0.0173	0.1965	−0.0085	0.0189	0.6538
8	125467120	rs6982502	T/C	*AC091114.1*	0.0048	0.0005	5.78E-22	−0.0434	0.0155	0.0051	−0.0351	0.0171	0.0395
20	41325964	rs6124341	A/G	*ADI1P1*	−0.0054	0.0006	2.04E-21	0.0363	0.0178	0.0433	0.0332	0.0194	0.0875
6	160596331	rs73596816	G/A	*LPA*	0.0105	0.0011	2.87E-21	0.0583	0.0340	0.0872	0.0755	0.0374	0.0436
2	157625480	rs10164853	A/G	*ACVR1C*	−0.0048	0.0005	7.42E-20	0.0100	0.0165	0.5648	0.0077	0.0181	0.6705
1	55033483	rs7525649	T/C	*PCSK9*	−0.0046	0.0005	3.42E-19	0.0149	0.0161	0.3494	0.0087	0.0176	0.6216
11	5663172	rs1976339	C/G	*AC104389.5*	−0.0044	0.0005	1.09E-18	−0.0273	0.0157	0.0818	−0.0564	0.0172	0.0010
9	104903458	rs2575876	G/A	*ABCA1*	−0.0051	0.0006	1.46E-18	0.0070	0.0181	0.6930	−0.0002	0.0197	0.9921
17	47685639	rs4794047	A/T	*KPNB1*	0.0045	0.0005	8.82E-18	0.0070	0.0163	0.6874	0.0168	0.0179	0.3461
11	116782580	rs2075290	T/C	*ZPR1*	0.0048	0.0006	2.87E-16	−0.0035	0.0184	0.9800	−0.0173	0.0205	0.3969
1	25449191	rs12027110	A/T	*MACO1*	0.0045	0.0006	4.41E-16	0.0488	0.0173	0.0050	0.0637	0.0190	0.0008
13	32402221	rs10492397	G/A	*N4BP2L1*	−0.0042	0.0005	4.63E-16	0.0149	0.0163	0.3536	0.0158	0.0179	0.3765
8	115658120	rs3808477	C/T	*TRPS1*	−0.0042	0.0005	1.92E-15	−0.0243	0.0167	0.1468	−0.0097	0.0183	0.5953
1	220799001	rs867772	G/A	*MARC1, AL445423.3*	−0.0049	0.0006	4.37E-15	0.0188	0.0196	0.3462	0.0066	0.0213	0.7580
12	120977112	rs2255531	G/A	*HNF1A-AS1*	0.0039	0.0005	6.77E-15	−0.0036	0.0158	0.8199	0.0059	0.0174	0.7334
8	58493931	rs112784971	C/T	*CYP7A1*	−0.0045	0.0006	2.68E-14	0.0149	0.0183	0.4138	−0.0072	0.0200	0.7183
6	100989473	rs873649	C/T	*AL133338.2*	0.0045	0.0006	5.78E-14	−0.0041	0.0189	0.8272	−0.0078	0.0208	0.7058
17	69086125	rs12162136	A/G	*ABCA6*	0.0037	0.0005	1.26E-13	−0.0141	0.0157	0.3678	−0.0177	0.0172	0.3041
20	17614241	rs116933131	C/T	*RRBP1*	−0.0075	0.0010	6.76E-13	0.0159	0.0328	0.6347	0.0021	0.0357	0.9539
10	112153464	rs2297991	C/T	*GPAM*	0.0039	0.0005	9.70E-13	0.0100	0.0170	0.5662	0.0273	0.0187	0.1437
1	196729914	rs10801558	T/G	*CFH*	−0.0035	0.0005	1.50E-12	−0.0038	0.0155	0.8044	−0.0186	0.0170	0.2713
9	127966535	rs74987020	AG/-	*FAM102A*	0.0042	0.0006	2.20E-12	−0.0071	0.0187	0.7044	−0.0089	0.0206	0.6644
20	17865277	rs2328223	A/C	*AL035045.1*	0.0043	0.0006	2.23E-12	−0.0101	0.0194	0.6042	−0.0083	0.0213	0.6958
10	72932888	rs41280378	T/G	*OIT3*	−0.0039	0.0006	5.72E-12	−0.0077	0.0179	0.6650	−0.0132	0.0196	0.4983
1	62586017	rs12030293	C/T	*DOCK7*	−0.0041	0.0006	2.33E-11	0.0526	0.0190	0.0058	0.0460	0.0207	0.0266
1	234712862	rs486142	A/G	*AL160408.6*	−0.0039	0.0006	2.59E-11	0.0109	0.0184	0.5577	−0.0076	0.0202	0.7065
18	49594230	rs9953437	G/A	*LIPG*	0.0033	0.0005	2.71E-11	−0.0068	0.0158	0.6664	0.0103	0.0173	0.5502
22	30036855	rs5844888	C/-	*HORMAD2-AS1*	0.0050	0.0008	3.20E-11	−0.0495	0.0241	0.0398	−0.0630	0.0266	0.0179
9	128704210	rs13289095	G/T	*PKN3*	−0.0061	0.0009	3.21E-11	0.0188	0.0285	0.5081	0.0108	0.0311	0.7281
12	100528401	rs2373355	G/A	*NR1H4*	0.0035	0.0005	7.20E-11	−0.0001	0.0170	0.9930	0.0186	0.0186	0.3170
11	118558115	rs10677211	-/GATAA	*IFT46*	−0.0031	0.0005	4.48E-10	0.0060	0.0155	0.7227	−0.0126	0.0169	0.4581
6	29939440	rs9260028	G/A	*HLA-A*	−0.0032	0.0005	7.76E-10	0.0040	0.0164	0.8116	−0.0014	0.0179	0.9358
20	35555321	rs224419	A/G	*ERGIC3*	−0.0035	0.0006	1.95E-09	0.0247	0.0183	0.1692	0.0330	0.0200	0.0979
12	112399801	rs75919415	G/T	*RPL6*	0.0036	0.0006	2.22E-09	−0.0176	0.0190	0.3551	−0.0202	0.0208	0.3323
1	92219224	rs75573831	C/T	*C1orf146*	0.0102	0.0017	4.17E-09	−0.0114	0.0549	0.8358	−0.0176	0.0610	0.7730
19	49496752	rs2280401	G/A	*RPS11*	−0.0035	0.0006	4.62E-09	−0.0284	0.0186	0.1261	−0.0205	0.0202	0.3113
7	26341732	rs12666571	G/C	*SNX10*	−0.0029	0.0005	6.09E-09	−0.0009	0.0158	0.9568	−0.0271	0.0173	0.1173
7	25975772	rs28537499	A/C	*MIR148A*	−0.0029	0.0005	7.18E-09	0.0247	0.0155	0.1040	0.0263	0.0170	0.1207
2	21612193	rs146112956	A/-	*AC018742.1, AC009411.2*	0.0090	0.0016	9.92E-09	0.0667	0.0482	0.1649	0.0660	0.0538	0.2197
22	42138592	rs2743450	A/G	*AC254562.2, AC254562.3*	−0.0028	0.0005	1.15E-08	0.0129	0.0156	0.4154	0.0036	0.0171	0.8320
14	69445082	rs138089855	T/C	*SLC39A9*	−0.0062	0.0011	2.04E-08	−0.0591	0.0361	0.1008	−0.0662	0.0393	0.0920
2	43738159	rs75832441	G/A	*PLEKHH2*	0.0137	0.0025	2.16E-08	−0.0703	0.0785	0.3705	−0.0675	0.0877	0.4413
16	11539890	rs12596176	A/G	*LITAF*	0.0027	0.0005	3.34E-08	−0.0155	0.0155	0.3169	−0.0130	0.0170	0.4434

SE: Standard error, SNP: Single-nucleotide polymorphism, *HMGCR*: 3-Hydroxy-3-methylglutaryl-CoA reductase

Supplementary Table 7: Estimates of causal effects between the low-density lipoprotein cholesterol and diabetes mellitus when using standard Mendelian randomization sensitivity analysis

Analysis method	LDL-C-determining alleles on DM			
	Coefficient*	SE	95% CI	*P*
IVW				
Fixed-effect	−1.2774	0.4712	−2.2259–−0.3289	0.0067
Random-effects	−1.3165	0.5571	−2.4380–−0.1951	0.0181
Egger regression, intercept				
Fixed-effect	0.0015	0.0030	−0.0046–0.0076	0.6309
Random-effects	0.0016	0.0031	−0.0047–0.0079	0.6163
Egger regression, slope				
Fixed-effect	−1.2141	0.5330	−2.2875–−0.1406	0.0275
Random-effects	−1.1936	0.5349	−2.2709–−0.1163	0.0307
Simple median	−1.7635	0.4129	−2.7958–−0.9189	1.9×10^{-5}
Weighted median	−1.4406	0.1335	−1.7635–−1.4406	<0.00001

SE: Standard error, CI: Confidence interval, DM: Diabetes mellitus, LDL-C: Low-density lipoprotein cholesterol, IVW: Inverse-variance weighted

Supplementary Table 8: Cochran's *Q* and Rücker's *Q*-tests for heterogeneity

Exposure	Outcome	Instrumental variables	Analysis method*	Coefficient*	SE	95% CI	*P*	Cochran's *Q*/Rücker's *Q***	*P1****
LDL-C	DM	LDL-C-determining alleles	IVW method	−1.2774	0.4712	−2.2259–−0.3289	0.0067	53.97	0.1959
			Egger regression, slope	−1.2141	0.5330	−2.2875–−0.1406	0.0275	59.85	0.0825

*All with fixed effect, **For IVW method, Cochran's *Q*-test was performed and for Egger regression method, we used Rücker's *Q*-test, ***$P1$: *P* value for the Cochran's *Q*/Rücker's *Q*-tests. IVW: Inverse-variance weighted, I^2: 14.77%. SE: Standard error, CI: Confidence interval, DM: Diabetes mellitus, LDL-C: Low-density lipoprotein cholesterol