



Phenotypic and Genetic Characterization of Lower LDL Cholesterol and Increased Type 2 Diabetes Risk in the UK Biobank

Yann C. Klimentidis,^{1,2} Amit Arora,¹ Michelle Newell,¹ Jin Zhou,¹ Jose M. Ordovas,^{3,4} Benjamin J. Renquist,⁵ and Alexis C. Wood⁶

Diabetes 2020;69:2194–2205 | <https://doi.org/10.2337/db19-1134>

Although hyperlipidemia is traditionally considered a risk factor for type 2 diabetes (T2D), evidence has emerged from statin trials and candidate gene investigations suggesting that lower LDL cholesterol (LDL-C) increases T2D risk. We thus sought to more comprehensively examine the phenotypic and genotypic relationships of LDL-C with T2D. Using data from the UK Biobank, we found that levels of circulating LDL-C were negatively associated with T2D prevalence (odds ratio 0.41 [95% CI 0.39, 0.43] per mmol/L unit of LDL-C), despite positive associations of circulating LDL-C with HbA_{1c} and BMI. We then performed the first genome-wide exploration of variants simultaneously associated with lower circulating LDL-C and increased T2D risk, using data on LDL-C from the UK Biobank ($n = 431,167$) and the Global Lipids Genetics Consortium ($n = 188,577$), and data on T2D from the Diabetes Genetics Replication and Meta-Analysis consortium ($n = 898,130$). We identified 31 loci associated with lower circulating LDL-C and increased T2D, capturing several potential mechanisms. Seven of these loci have previously been identified for this dual phenotype, and nine have previously been implicated in nonalcoholic fatty liver disease. These findings extend our current understanding of the higher T2D risk among individuals with low circulating LDL-C and of the underlying mechanisms, including those responsible for the diabetogenic effect of LDL-C-lowering medications.

Rates of cardiovascular disease (CVD) and type 2 diabetes (T2D) are among the most pressing health concerns worldwide. These two diseases share many risk factors and tend to co-occur, because there is an excess of CVD among individuals with T2D (1,2). Yet, controversy remains over whether all risk factors exert similar effects on the development of these two conditions. LDL cholesterol (LDL-C) is a class of highly atherogenic particles, and circulating levels of LDL-C are a causal risk factor for CVD across the life span (3). However, several lines of evidence suggest that decreased levels of circulating LDL-C are associated with an increased T2D risk.

Lipid-lowering medications, in particular from the statin drug class, are effective at lowering levels of circulating LDL-C and rates of adverse cardiovascular events (4) but convey an increased T2D risk (odds ratio [OR] 1.09) (5,6) in a dose-dependent manner (7). This increased risk, however, is outweighed at a population level by the cardiovascular event rate reduction. An increased T2D risk has also been reported in observational studies. Individuals with low levels of circulating LDL-C (e.g., <60 mg/dL) exhibit a higher risk of prevalent and incident T2D (8,9), and among individuals with coronary disease, LDL-C and T2D are inversely related (10). In addition, individuals with familial hypercholesterolemia exhibit a decreased risk of T2D as well as lower BMI and triglyceride (TG) levels (11).

¹Department of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ

²BIO5 Institute, University of Arizona, Tucson, AZ

³Nutrition and Genomics Laboratory, Jean Mayer U.S. Department of Agriculture Human Nutrition Research Center on Aging, Tufts University, Boston, MA

⁴Instituto Madrileño de Estudios Avanzados (IMDEA) Food Institute, Campus de Excelencia Internacional Universidad Autónoma de Madrid + Consejo Superior de Investigaciones Científicas, Madrid, Spain

⁵School of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson, AZ

⁶U.S. Department of Agriculture/Agricultural Research Service Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX

Corresponding author: Yann C. Klimentidis, yann@email.arizona.edu

Received 12 November 2019 and accepted 29 May 2020

This article contains supplementary material online at <https://doi.org/10.2337/figshare.12389273>.

© 2020 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/content/license>.

See accompanying article, p. 2058.

Genetic studies have lent further support to inverse phenotypic associations between LDL-C and T2D, with recent studies pointing to genetic loci that harbor variants exerting opposing effects on LDL-C and T2D. These include loci containing the *HMGCR* (12), *APOE* (13,14), *PCSK9* (12,14,15), *NPC1L1* (12,14), *PNPLA3* (14), *TM6SF2* (14), *GCKR* (14), and *HNF4A* (14) genes. Furthermore, Fall et al. (16) and White et al. (17) have both found that genetically predicted higher circulating LDL-C was associated with a lower risk of T2D. Yet, genetic findings show that not all variants have opposing effects on circulating LDL-C levels and T2D risk. LDL-C-lowering variants in *ABCG5/G8* and *LDLR* genes were not shown to alter T2D risk (12), and subsets of LDL-C-lowering alleles pose a stronger risk for T2D than the full gamut (16). Circulating LDL-C levels, like T2D, are reflective of a number of physiological processes. The findings outlined above suggest that there is heterogeneity in T2D outcomes, depending on which pathways are the primary LDL-C-lowering mechanisms, and that genetic studies may give us insights into these pathways. For example, it is not clear whether these associations are driven by changes in circulating levels of LDL-C or by changes in intracellular levels of cholesterol. A better understanding of which genetic loci lower circulating levels of LDL-C and increase T2D risk may yield mechanistic insights that could help develop therapeutic options that lower lipid levels without raising the risk of T2D and help identify individuals at greater risk for T2D with statin use.

Here, we first examined the relationship of directly measured circulating LDL-C levels with prevalent T2D, HbA_{1c}, and BMI. We then sought to identify, for the first time on a genome-wide scale, loci simultaneously associated with lower LDL-C and increased T2D (and vice versa). Upon identifying variants, we sought to generate additional mechanistic insights by testing of associations with seven other traits in the UK Biobank related to T2D, LDL-C, and nonalcoholic fatty liver disease (NAFLD).

RESEARCH DESIGN AND METHODS

UK Biobank

Data from the UK Biobank were used for 1) phenotypic data analysis, which examined the associations of circulating levels of LDL-C and TG with T2D, HbA_{1c}, and BMI, and 2) discovery genome-wide association study (GWAS) for variants that are associated with lower circulating levels of LDL-C and higher T2D. The UK Biobank is a prospective cohort study of ~500,000 individuals between the ages of 39 and 72 years living throughout the U.K. Participants attended 1 of 21 assessment centers in the U.K. and had their blood drawn for biomarker and genetic analysis and weight and height measured to derive BMI (kg/m²). Directly measured circulating LDL-C, HbA_{1c}, HDL-C, TG, alanine aminotransferase (ALT), and AST were obtained from all UK Biobank participants at the baseline visit between 2006 and 2010 in a nonfasting

state. LDL-C was assessed by enzymatic protective selection analysis on a Beckman Coulter AU5800.

To define prevalent T2D case and control subjects, we used criteria previously used by Yaghootkar et al. (18) and Eastwood et al. (19). We first excluded individuals with a missing age of T2D diagnosis, reporting a T2D diagnosis within 1 year of the baseline examination, those self-reporting type 1 diabetes in the verbal interview, and women reporting only gestational diabetes on the touchscreen or verbal interview. Prevalent T2D was defined using the following criteria: 1) self-reported diabetes diagnosed by a doctor during the touchscreen, or self-reported T2D or generic diabetes in verbal interviews; 2) having a nonmissing age of diagnosis and an age of diagnosis >35 years of age (>30 years of age for participants reporting an ethnicity of South Asian or African Caribbean); and 3) not using insulin within 1 year of diagnosis to exclude possible type 1 diabetes case subjects. Control subjects were participants with no self-reported diabetes of any type from the touchscreen or verbal interview, no self-reported insulin use in the touchscreen or verbal interview, those not excluded according to the aforementioned criteria, and those not reporting nonmetformin T2D medication (see the list in Supplementary Table 2).

UK Biobank Genotypes

Genotypes in the UK Biobank were obtained with the Affymetrix UK Biobank Axiom Array (Santa Clara, CA), whereas 10% of participants were genotyped with the Affymetrix UK BiLEVE Axiom Array. Details regarding imputation, principal components analysis, and quality control procedures are described elsewhere (20). The analysis excluded individuals with unusually high heterozygosity, with a high (>5%) missing rate, or with a mismatch between self-reported and genetically inferred sex. Single nucleotide polymorphisms (SNPs) out of Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$), with a high missing rate (>1.5%), with a low minor allele frequency (<0.1%), or with a low imputation accuracy (info <0.4) were excluded from analyses. This resulted in the availability of ~15 million SNPs for analysis.

Diabetes Genetics Replication and Meta-Analysis and Global Lipids Genetics Consortium GWAS Meta-Analysis Summary Statistics

The latest GWAS meta-analysis summary statistics for T2D (unadjusted for BMI) were obtained from the Diabetes Genetics Replication and Meta-Analysis consortium (DIAGRAM), which includes data on up to 898,130 individuals (74,124 case and 824,006 control subjects), including UK Biobank individuals (21). We used the results of our GWAS of circulating LDL-C in UK Biobank, along with the aforementioned DIAGRAM-T2D results, for the discovery of inverse association signals. We then replicated LDL-C associations of our top hits with an independent GWAS meta-analysis of LDL-C from the Global Lipids Genetics Consortium (GLGC) (22) ($n = 188,577$);

this meta-analysis does not include the UK Biobank study. Across the UK Biobank, DIAGRAM, and GLGC summary statistics, we aligned all SNP alleles and their corresponding effects by using the `harmonize` function in the `TwoSampleMR` package in R software (23).

Statistical Analyses

To evaluate and plot the prevalence of T2D and of BMI and HbA_{1c} by decile of circulating LDL-C and TG in the UK Biobank, we excluded all participants who self-reported (at baseline) use of cholesterol-lowering medications during the touchscreen survey, or cholesterol-lowering medication during the verbal interview (see Supplementary Table 1 for list of medications). Approximately 91% of individuals taking cholesterol-lowering medications were taking statins. To examine levels of HbA_{1c} by decile of circulating LDL-C, we excluded participants defined as T2D cases (see above). We further excluded individuals with outlier values of HbA_{1c} >4 SDs from the mean. Deciles were calculated using the “`quantcut`” function in the “`gtools v3.5.0`” library in R software. Once decile were established, T2D prevalence by LDL-C/TG decile was calculated and plotted with CIs determined by the Clopper-Pearson interval (24). Mean HbA_{1c} and BMI and their distributions are shown in boxplots for each decile of circulating LDL-C. We further examined T2D prevalence by circulating LDL-C decile separately in men and women, and in different age-groups (40–49 years, 50–59 years, and 60–69 years).

To statistically evaluate these phenotypic associations, we performed logistic regression with T2D prevalence as the outcome and linear regression with HbA_{1c} and BMI as outcomes. As mentioned above, all individuals on cholesterol-lowering medication were excluded. To normalize residuals, we transformed circulating LDL-C, TG, HbA_{1c}, and BMI by inverse normalization for all linear regression analyses. For each analysis, we used the same exclusion criteria as those mentioned above and adjusted for “last eating” time (excluding individuals reporting extreme values, >16 h), age, sex, and center. We considered an expanded model with additional covariates: education (college/university degree or not), Townsend Deprivation Index, BMI, hypertension status (self-reported status or hypertension medication), ethnicity (white/European or not), family history of T2D (at least one first-degree family member), smoking status (never, past, current), and alcohol consumption (never or only special occasions, one to three times per month, one to two times per week, three to four times per week, daily/almost daily). In analyzing the association of circulating LDL-C with T2D, we also tested for interactions with sex and age and provided stratified analyses accordingly. Finally, we also examined the association of LDL-C with T2D among only the individuals taking cholesterol-lowering medication.

To address possible ascertainment bias of prevalent T2D case subjects due to exclusion of people taking cholesterol-lowering medication, we performed a sensitivity analysis

using propensity score matching to remove bias between the two groups due to observed covariates (25). We used the propensity score to match on the probability of taking cholesterol-lowering medication given the set of baseline characteristics listed in Supplementary Table 3. All covariates were selected based on previous literature or if considered potential significant confounders for cholesterol-lowering medication use or T2D (26–28). Matching analyses were performed using R software and the package `MatchIt v3.0.2` with 1:1 nearest-neighbor matching and a caliper width equal to 0.1 to achieve balanced covariates between the two groups (29). Standardized mean differences were used to assess covariate balance before and after matching. Standardized mean differences <0.1 were considered adequately balanced to reduce significant differences between the two groups (30). Subsequent analyses were conducted among all individuals, regardless of cholesterol-lowering medication use, using logistic regression for T2D as the outcome and linear regression for HbA_{1c} as the outcome. Model 1 was adjusted for time since eating, age, sex, and center. Model 2 was additionally adjusted for BMI and use of cholesterol-lowering medication. BMI was included as a covariate because the addition of BMI to the unadjusted model in the matched sample changed the regression coefficient for LDL-C by >10%. Although cholesterol medication use did not result in this magnitude of change, we adjusted for cholesterol medication use in model 2 to address any potential confounding. The only covariates considered in these regression analyses were those used in the main analyses stratified by cholesterol-lowering medication (see above and model 2 in Supplementary Tables 6 and 7). To meet the assumptions of linear regression, HbA_{1c}, LDL-C, and BMI were inverse normalized in all linear regression models of HbA_{1c} regressed on LDL-C. We used complete case analysis to address missing data. As a result of greater missingness for HDL-C ($n = 35,382$), analyses were repeated excluding HDL-C from the matching covariates. However, the regression results were not attenuated, and only results including HDL-C were reported.

For the GWAS of LDL-C in the UK Biobank, the circulating LDL-C level of individuals on cholesterol-lowering medication was corrected by dividing it by a correction factor of 0.63 (31). We also ran a GWAS only on individuals not taking cholesterol-lowering medication. We transformed LDL-C by inverse normalization. We used BOLT-LMM software (32) to perform GWAS on individuals of European descent ($n = 431,167$) and included “last eating” time (see above), sex, age, age², center, genotyping chip, and the first 10 principal components as covariates. BOLT-LMM performs a linear mixed model regression that includes a random effect of all SNP genotypes other than the one being tested. We aligned effect sizes across the GWAS summary statistics of each trait to the same effect allele using the `harmonize` function, as mentioned above. We used `metaCCA v1.12.0` (33) to perform a multivariate GWAS with the LDL-C and T2D GWAS summary

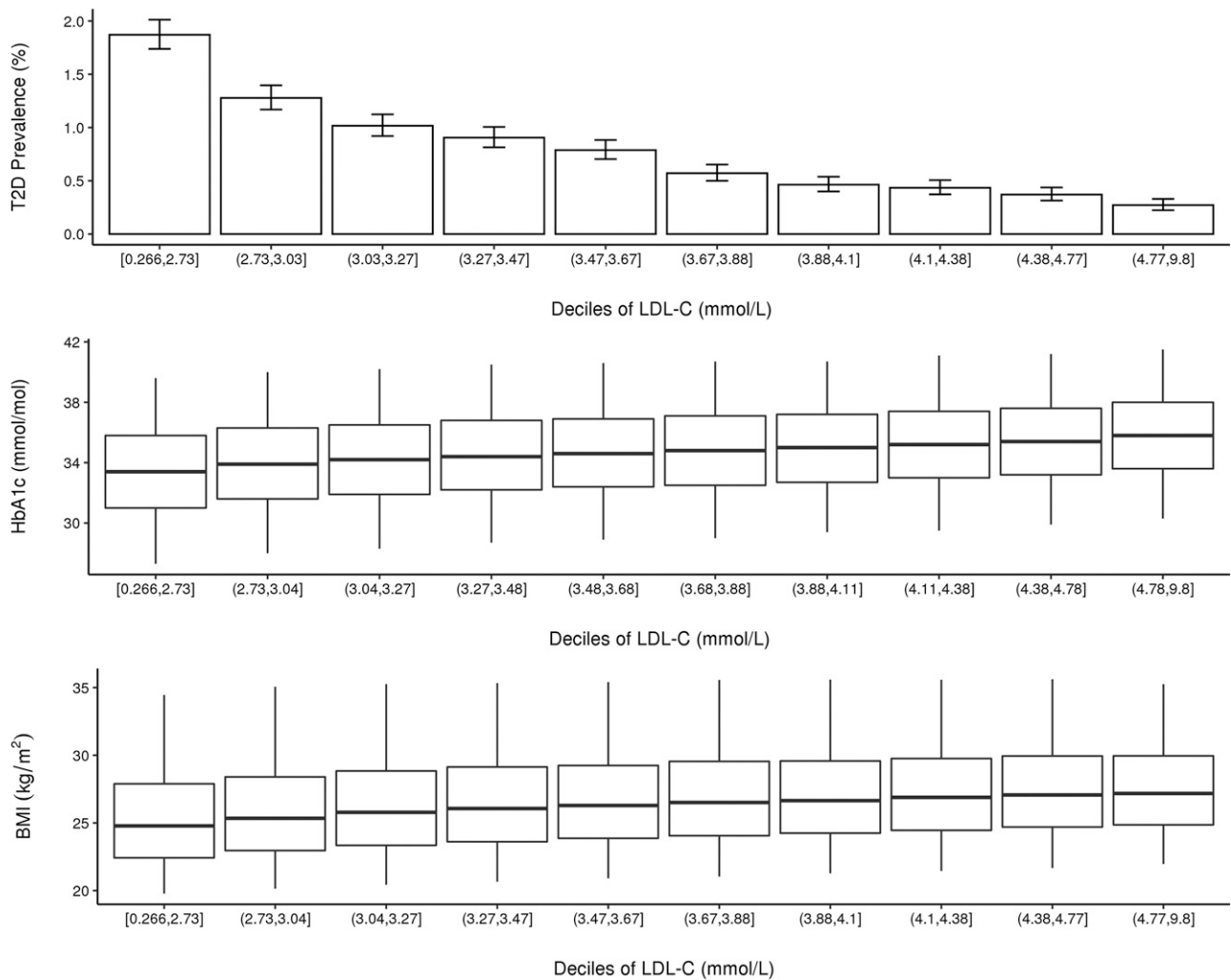


Figure 1—T2D prevalence, HbA_{1c}, and BMI by circulating LDL-C deciles in the UK Biobank. T2D prevalence is shown as a percentage, with error bars corresponding to the Clopper-Pearson CI. Whisker plots show the median value (horizontal line in box), the 25th and 75th percentile delimited by the box, and the vertical lines extending to the 5th and 95th percentile.

statistics. Briefly, metaCCA implements a canonical correlation analysis on GWAS summary statistic data in which the phenotype correlation structure was estimated from the univariate GWAS summary statistics. We first selected only those SNPs that exhibited opposite directions of univariate effects for LDL-C and T2D and having a metaCCA $P < 5 \times 10^{-8}$. To further minimize the potential of selecting false-positive loci, we selected among these SNPs only those with a univariate association $P < 5 \times 10^{-5}$ for each of LDL-C and T2D. At this univariate P -value threshold, the prior probability of a given SNP associated with two traits and with discordant direction of effect under the null hypothesis corresponds to $0.00005 \times 0.000025 = 1.25 \times 10^{-9}$ (34). SNPs within <500 kb of each other or in linkage disequilibrium of $r^2 > 0.05$ were clumped together, and the SNP with the lowest metaCCA P value was reported.

For the replication of the 44 discovered loci, we considered both the univariate results for LDL-C from GLGC

and multivariate results from metaCCA using the GLGC LDL-C and DIAGRAM T2D. Because of incomplete overlap of SNPs in GLGC with those in the UK Biobank and DIAGRAM and differences in population composition, we examined all SNPs within each locus identified in the discovery stage (i.e., the base pair range at a given locus for which all SNPs satisfied the above univariate and multivariate criteria in the discovery analysis). After further restricting to only variants for which the effect size for (GLGC) LDL-C and T2D exhibited opposite directions of effect, we chose the SNP with the lowest metaCCA P value. A locus was considered to be successfully replicated if this top SNP had a univariate LDL-C $P < 5 \times 10^{-3}$ and a metaCCA $P < 5 \times 10^{-5}$. Among the replicated loci, we tested for colocalization using the DIAGRAM T2D and UK Biobank LDL-C results to determine whether, at a given locus, the two traits are likely to be affected by the same causal variant. Specifically at each of the replicated loci, we used the coloc v3.2-1 package in R software (35) to test for colocalization

Table 1—List of replicated genetic loci in which all SNPs have $P < 5E-5$ for both circulating LDL-C and T2D and with opposite directions of effect, for UK Biobank LDL-C and DIAGRAM T2D

Chr	bp_min	bp_max	Nearest_genes	UK Biobank LDL-C and DIAGRAM T2D										GLGC		
				Top SNP-metaCCA	metaCCA-min P	T2D β	T2D P	LDL β	LDL P	Top SNP-metaCCA	LDL P	metaCCA P				
1	109,285,825	109,285,825	<i>FNDC7-STXBP3</i>	rs12144658	6.7E-10	-0.095	8.3E-06	0.028	7.0E-06	rs12144658	1.7E-03	4.2E-07				
1	109,843,746	110,078,255	<i>SORT1-PSMA5</i>	rs56218586	2.0E-178	-0.059	8.1E-06	0.095	2.6E-118	rs17647543	6.3E-27	2.5E-32				
2	27,161,476	28,246,841	<i>GCKR</i>	rs6547692	1.4E-194	-0.055	1.6E-17	0.043	8.1E-113	rs1260326	8.4E-10	1.3E-32				
2	43,452,183	44,020,441	<i>THADA</i>	rs116097055	2.2E-15	0.18	1.5E-10	-0.034	1.7E-06	rs55877805	1.1E-03	2.4E-29				
2	121,306,440	121,350,667	<i>LINC01101-GLI2</i>	rs17050272	2.0E-51	-0.031	3.5E-06	0.026	6.4E-42	rs896499	8.9E-09	4.7E-12				
2	234,508,963	234,565,917	<i>UGT1A8</i>	rs2741038	8.5E-09	-0.03	1.5E-05	0.009	3.8E-05	rs2741019	3.5E-06	2.5E-09				
3	12,058,246	12,456,689	<i>PPARG</i>	rs17036170	1.8E-19	-0.13	2.7E-06	0.073	9.9E-15	rs17036326	1.5E-08	1.2E-24				
3	124,912,207	129,301,874	<i>SLC12A8</i>	rs9873519	1.8E-16	-0.039	1.4E-09	0.011	1.9E-08	rs4679370	3.2E-04	2.7E-10				
3	170,631,635	170,743,747	<i>SLC2A2</i>	rs12496734	8.0E-18	0.05	2.5E-07	-0.020	4.5E-12	rs7430672	4.2E-05	1.1E-15				
5	74,352,180	75,042,913	<i>HMGCR-POC5</i>	rs34361	6.9E-322	-0.05	8.7E-14	0.054	3.3E-156	rs7703051	3.0E-86	7.5E-118				
5	78,350,390	78,385,845	<i>BHMT2</i>	rs592052	1.7E-10	0.031	1.1E-06	-0.008	1.3E-05	rs557302	3.0E-03	2.3E-07				
6	31,145,920	31,372,569	<i>HLA-C/B</i>	rs36057735	1.4E-15	-0.037	6.7E-06	0.016	2.3E-11	rs7743761	2.5E-11	2.5E-10				
6	32,127,477	33,514,567	<i>HLA-DQA1/B1</i>	rs3130276	1.8E-10	0.031	3.8E-06	-0.009	4.4E-06	rs9275221	3.6E-15	1.0E-23				
6	126,629,425	127,167,072	<i>CENPW</i>	rs9388490	3.2E-29	-0.049	4.5E-14	0.015	3.2E-15	rs9388490	3.8E-03	8.9E-15				
7	44,585,052	44,709,372	<i>NPC1L1-DDX56</i>	rs2008428	2.9E-78	0.031	4.8E-05	-0.038	4.5E-63	rs217386	6.3E-22	4.3E-23				
7	100,266,081	100,414,386	<i>GIGYF1-POP7-EPO</i>	rs221793	2.5E-24	-0.047	1.9E-05	0.029	6.7E-20	rs506597	4.8E-04	3.5E-08				
8	9,168,146	9,740,317	<i>PPP1R3B</i>	rs9987289	1.6E-88	0.066	3.9E-09	-0.057	1.1E-65	rs9987289	1.4E-27	1.4E-36				
10	64,988,931	65,340,897	<i>REEP3</i>	rs3847326	8.4E-15	-0.028	1.3E-05	0.012	9.2E-11	rs12775231	1.1E-03	6.7E-06				
10	101,900,579	102,039,463	<i>ERLIN1-CWF19L1</i>	rs35614792	2.8E-15	0.031	1.9E-06	-0.012	2.2E-10	rs369136533	1.8E-03	8.5E-08				
11	65,422,853	65,604,195	<i>OVOL1-SNX32</i>	rs34381009	1.2E-09	0.037	2.3E-05	-0.012	2.8E-06	rs3814707	4.3E-04	6.5E-08				
12	50,881,148	50,987,378	<i>DIP2B</i>	rs7316864	3.3E-11	-0.029	4.6E-05	0.011	6.7E-08	rs10876041	1.1E-06	1.2E-06				
12	121,435,778	123,891,209	<i>HNF1A-OASL</i>	rs12231806	8.5E-17	-0.05	8.0E-08	0.018	1.4E-10	rs7131696	4.4E-06	1.4E-15				
15	62,383,155	63,314,519	<i>C2CD4A/B</i>	rs1881415	2.8E-17	0.047	4.7E-13	-0.008	1.8E-05	rs4502156	1.1E-04	5.0E-14				
17	29,628,549	29,749,444	<i>NF1-EV12A</i>	rs7502556	9.7E-19	0.032	5.3E-07	-0.014	2.8E-13	rs1129506	5.8E-06	9.2E-08				
17	46,044,031	46,114,620	<i>COP22</i>	rs12935890	3.0E-17	-0.032	6.6E-06	0.015	6.2E-13	rs11651275	3.6E-03	2.4E-07				
19	19,320,825	19,793,545	<i>TM6SF2</i>	rs59148799	9.9E-324	-0.054	3.4E-08	0.077	1.4E-162	rs10401969	4.6E-61	1.6E-86				

Continued on p. 2199

Table 1—Continued

Chr	bp_min	bp_max	Nearest_genes	UK Biobank LDL-C and DIAGRAM T2D					GLGC			
				Top SNP-metaCCA	metaCCA-min P	T2D β	T2D P	LDL β	LDL P	Top SNP-metaCCA	LDL P	metaCCA P
19	45,324,756	45,428,234	APOE	rs101119	4.9E-42	0.043	3.5E-09	-0.024	3.1E-31	rs20756650	0.0E+00	0.0E+00
19	46,123,929	46,314,504	GIPR	rs2238689	2.0E-33	-0.05	1.3E-14	0.017	9.4E-19	rs10422601	1.9E-09	3.4E-22
20	42,958,768	43,042,364	HNF4A	rs1800961	5.3E-72	-0.16	3.2E-20	0.076	1.8E-42	rs1800961	5.2E-11	7.1E-31
22	18,415,158	18,486,514	MICAL3	rs5992928	3.6E-11	0.035	3.1E-06	-0.011	7.9E-07	rs7286607	3.4E-05	2.4E-08
22	44,324,730	44,394,402	PNPLA3	rs738408	4.6E-19	-0.049	1.8E-10	0.014	7.1E-10	rs738408	2.9E-05	1.9E-11

β-Coefficients refer to the SNP with the lowest P value for each trait in the respective region. The minimum univariate P value for LDL-C in GLGC summary statistics as well as the top SNP and P value of the top metaCCA SNP for the region discovered in UK Biobank is also shown. bp_min and bp_max are the minimum and maximum base pair position for which all SNPs have P < 5E-5 for LDL-C and T2D, with opposite directions of effect, and with metaCCA P < 5E-8. Human build GRCh37/hg19; β-coefficients correspond to log-ORs for T2D and SDs for LDL-C. P values from GLGC LDL-C associations are for SNPs showing opposite direction of association with T2D (DIAGRAM). Chr, chromosome.

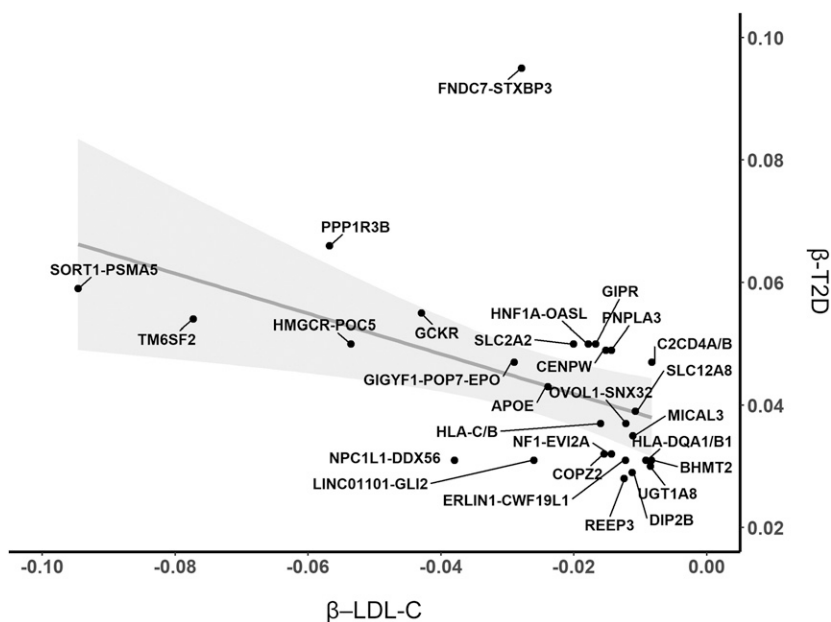


Figure 2—Plot of β -coefficients for circulating LDL-C vs. T2D for SNPs with opposite directions of effect on these two traits. β -Coefficients correspond to log-ORs for T2D, and SDs for circulating LDL-C. Shaded area corresponds to the 95% CI for the best fit regression line.

using all SNPs within 250 kb of the SNP with the lowest metaCCA P value. We used default parameters and priors. We considered that there was evidence for colocalization if the posterior probability for a shared causal variant hypothesis 4 (PP.H4) was $>80\%$.

To test the association of the 31 SNPs (T2D increasing allele) that we identified with a range of other cardiometabolic traits that are known to be related to LDL-C and T2D and are available in the UK Biobank, we used similar methods described above for the circulating LDL-C GWAS. For TG and HDL-C, we excluded individuals reporting cholesterol-lowering medication. For ALT and AST, we excluded 15,138 individuals with medical conditions, other than NAFLD, that could affect liver enzyme levels (36). For HbA_{1c}, we excluded individuals with prevalent T2D (see above). For the waist-to-hip ratio, we additionally adjusted for BMI before inverse normalization and subsequent GWAS. We inverse normalized all traits before the GWAS. We tested the association of each of the 31 SNPs with each of these seven additional phenotypes. We then normalized the effect sizes by dividing the β -coefficients by the corresponding SEs and dividing by the square root of the respective sample size. We used hierarchical clustering to group the identified variants according to their pattern of association with all nine traits, including T2D and circulating LDL-C. Clustering was performed with the `hclust` v3.6.2 function in R, with the Euclidian metric to calculate distances, and the Ward clustering method (37). Cluster stability was assessed by using the `cValid` v0.6-6 package in R software, evaluating the hierarchical, k -means, and partitioning around medoids methods, and evaluating 2–10 clusters (38). Finally, using UK Biobank individual-level data, we used a multivariate approach, MultiPhen v2.0.3 (39),

which uses ordinal regression to model each SNP as the outcome and includes all traits as covariates (except for T2D) in addition to age and sex. We present only β -coefficients from these models because the P values are nearly all >0.05 , possibly due to the inclusion of many correlated phenotypes into each model.

Data and Resource Availability

The data that support the findings of this study are available to researchers, upon application, from the UK Biobank, but restrictions apply to the availability of these data, which were used under license for the current study. Data from the DIAGRAM and GLGC consortia are publically available at their respective websites: <https://www.diagram-consortium.org/> and <http://lipidgenetics.org/>.

RESULTS

Participant Characteristics

In a sample size of 375,783 individuals after exclusion of individuals on lipid-lowering medication, T2D prevalence was 0.8% and was higher in men (1.15%) than in women (0.54%). Individuals with prevalent T2D had lower circulating LDL-C and HDL-C, higher circulating TG, higher HbA_{1c}, and higher BMI (Supplementary Table 4). Among individuals on cholesterol-lowering medication ($n = 78,626$), T2D prevalence was 18.2%, and individuals with T2D had lower circulating LDL-C and HDL-C, and higher circulating TG, HbA_{1c}, and BMI (Supplementary Table 5).

Association of Circulating LDL-C With T2D

We observed an inverse relationship between circulating LDL-C and T2D prevalence (OR 0.41 [95% CI 0.39, 0.43]

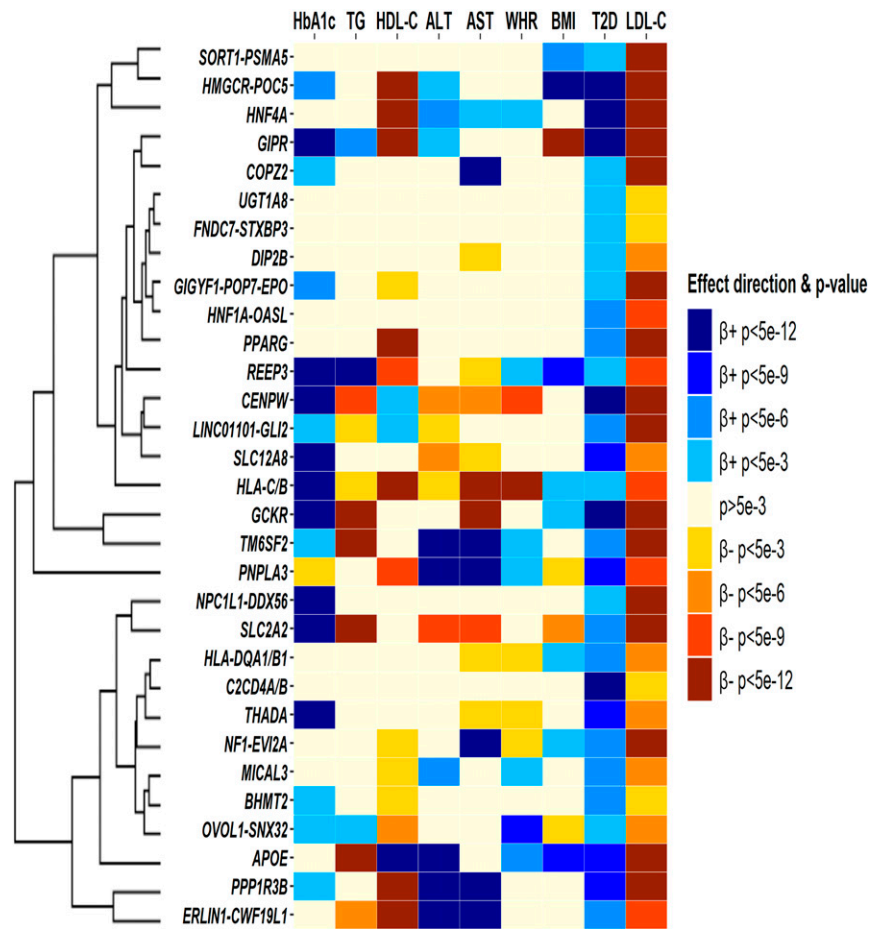


Figure 3—Association of T2D-increasing allele at 31 identified SNPs with nine cardiometabolic traits, based on univariate analyses, and hierarchical clustering dendrogram based on corresponding standardized effect sizes. WHR, waist-to-hip ratio.

per mmol/L unit of LDL-C, $P = 1.26 \times 10^{-263}$). Individuals in the lowest decile of circulating LDL-C exhibited the highest prevalence of T2D, and a consistent decrease in T2D prevalence was observed with increasing circulating LDL-C (Fig. 1). We found a very similar negative association of circulating LDL-C with T2D among only the individuals reporting the use of cholesterol-lowering medication. We found a significant interaction of circulating LDL-C with sex ($P = 1.52 \times 10^{-13}$), whereby the association of circulating LDL-C with T2D prevalence was stronger among men (OR 0.35 [95% CI 0.32, 0.37] per mmol/L unit of LDL-C, $P = 7.37 \times 10^{-215}$) than among women (OR 0.51 [0.47, 0.55] per mmol/L unit of LDL-C, $P = 3.63 \times 10^{-62}$) (Supplementary Table 6 and Supplementary Fig. 1). We also observed a stronger inverse association between circulating LDL-C and T2D prevalence among older individuals ($P_{\text{interaction}} = 3.54 \times 10^{-13}$) (Supplementary Table 6 and Supplementary Fig. 3). Positive associations were found between circulating LDL-C and both HbA_{1c} (after exclusion of individuals with T2D; $\beta = 0.14$, SE = 0.0017, $P < 5.0 \times 10^{-300}$) and BMI ($\beta = 0.16$, SE = 0.0016, $P < 5.0 \times 10^{-300}$) (Fig. 1 and Supplementary Table 6). We also observed a positive association between circulating TG

and T2D prevalence (OR 1.34 [95% CI 1.31, 1.38], $P = 8.03 \times 10^{-109}$) (Supplementary Table 6 and Supplementary Fig. 4). Among individuals on cholesterol-lowering medications, we found a nearly identical negative association of circulating LDL-C and T2D prevalence but a much weaker positive association with HbA_{1c} and a negative association with BMI (Supplementary Table 7 and Supplementary Fig. 1). In models including additional covariates, the results remained very similar (Supplementary Tables 6 and 7). Results were also very similar in propensity score-matching analyses. In a total sample size of ~70,000 individuals, the T2D prevalence was 6.69%, and these analyses showed similar negative associations of circulating LDL-C with T2D (OR 0.51 [95% CI 0.49, 0.54] in model 2) and positive associations of circulating LDL-C with HbA_{1c} (Supplementary Tables 3 and 8).

Loci Associated Inversely With LDL-C and T2D

We identified 44 loci associated in opposite directions with circulating LDL-C and T2D using the UK Biobank LDL-C and the DIAGRAM T2D results (Supplementary Table 9). In an analysis in which we used a GWAS of circulating LDL-C excluding individuals on cholesterol-lowering

medication, we observed nearly identical results (Supplementary Table 10). Among these 44 loci, 31 replicated with respect to LDL-C association when using the GLGC LDL-C GWAS results instead of UK Biobank (Table 1). Several loci are previously known or suspected to be inversely associated with circulating LDL-C and T2D (*HMGCR*, *APOE*, *NPC1L1*, *PNPLA3*, *TM6SF2*, *GCKR*, and *HNF4A*). However, most of the loci are novel for this LDL-C–T2D trait. Of these novel loci, 12 have previously been identified for LDL-C in the GLGC GWAS, 14 were previously identified in T2D GWAS, and 14 have not been identified previously with either trait. The loci with the strongest degree of opposing effects include *FNDCC7-STXBP3*, *SORT1-PSMA5*, *HMGCR-POC5*, *PPP1R3B*, and *GCKR* (Fig. 2). Colocalization analyses suggest that of the 31 loci, *GCKR*, *PPP1R3B*, *TM6SF2*, *HNF4A*, *MICAL3*, and *PNPLA3* have the same causal variant(s) influencing circulating LDL-C and T2D (PP.H4 > 0.8). Although most SNPs showed colocalization at shared or distinct causal variants, a few loci showed no evidence of colocalization (Supplementary Table 11).

The variants that we have identified can be linked with genes that affect de novo fatty acid synthesis, hepatic lipid uptake, hepatic lipid export, peripheral tissue lipid balance, fatty liver of unknown origin, insulin secretion, and insulin action (Supplementary Table 12). They are associated in distinct patterns across a range of cardiometabolic traits (Fig. 3). At these loci, the T2D-increasing alleles are generally associated with higher HbA_{1c} levels and lower HDL-C levels, although this pattern is not entirely consistent across all 31 SNPs. According to cluster stability evaluation, two clusters were optimally identified by hierarchical clustering (Supplementary Table 13). However, it is difficult to discern any consistent trait association patterns that differentiate the two sets of loci. In Supplementary Fig. 6 we present the trait-specific β -coefficients based on MultiPhen, some of which are substantially different from the univariate results.

DISCUSSION

We used the largest sample to date to examine the association of circulating LDL-C with T2D prevalence and found that individuals with low circulating LDL-C exhibit a higher prevalence of T2D. Then, in the first genome-wide analysis aimed at identifying variants associated with both lower circulating LDL-C and higher T2D risk, we identified 24 novel loci exerting opposite-direction effects on these traits. Our analyses lend weight to the notion that the association between lower circulating LDL-C and increased T2D risk is driven, at least in part, by a specific group of genetic variants that may be implicated via diverse mechanisms, including hepatic lipid synthesis, export, and uptake, as well as insulin secretion and action. These variants provide insight into the heterogeneous outcomes for different lipid and glucose metabolism pathways.

We found that low circulating LDL-C is associated with greater T2D prevalence, which is consistent with two previous studies examining T2D prevalence (8) and

incidence (9). In addition, we found that lower levels of circulating LDL-C are associated with lower HbA_{1c} (among individuals without T2D) and lower BMI. Our finding that lower circulating LDL-C is associated with increased T2D prevalence but lower HbA_{1c} appears counterintuitive. It is important to note that the latter association was performed in a slightly different subset than the first association (i.e., excluding those with T2D). We may be observing a threshold effect, whereby the etiology of “normal” HbA_{1c} variation is somewhat distinct from the etiology of crossing into overt T2D (e.g., 37). It is also possible that our results could be affected by collider bias because individuals on cholesterol-lowering medications are excluded from our main analysis. However, we observed a similar inverse relationship of circulating LDL-C with T2D in the set of people on cholesterol-lowering medication and in a propensity score-matching analysis. We also find that unlike LDL-C, TG levels are positively associated with T2D prevalence. This opposing relationship of circulating LDL-C and TG with T2D prevalence may suggest that LDL particles are being overfilled in individuals with T2D.

Previous research into loci that jointly alter the risk for circulating LDL-C and T2D has focused on the genomic targets of lipid-lowering medications in the hope that these analyses will give specific insights into associated T2D risk. On one hand, our analyses confirmed that variants in *HMGCR* (41) and *NPC1L1* (14) are associated with lower circulating LDL-C and increased T2D risk. On the other hand, our analyses did not identify variants at *PCSK9*. The lowest T2D *P* value was 0.003 in this region for a SNP with opposite direction coefficient. However, our analyses identified a fourth target of lipid-lowering medications: variants in the peroxisome proliferator-activated receptor (*PPARG*) gene, the target of fibrates and thiazolidinediones.

We observed nine variants previously identified as being associated with NAFLD: *PNPLA3*, *GCKR*, *TM6SF2*, *PPP1R3B*, *ERLIN1-CWF19L1*, *REEP3*, *HNF1A*, *SLC2A2*, and *MICAL3* (42–44). Furthermore, five of the seven colocalizing loci are among these nine. This enrichment for NAFLD-related genes may reflect increased synthesis and storage of TG and reduced export/secretion of VLDL, leading to reduced circulating LDL-C. Indeed, the LDL-C-decreasing alleles at most of these loci are associated with increased liver enzymes, indicative of hepatic steatosis, with the exception of *GCKR* and *SLC2A2*, consistent with a previous finding (45). In turn, lower levels of circulating LDL-C along with increased liver enzymes would be expected to indicate increased NAFLD and T2D. A recent bidirectional Mendelian randomization study provides support for this hypothesized causal effect of NAFLD on T2D (46). Our findings that liver fat may be an important mediator of the effect of cholesterol lowering on T2D is consistent with a report showing that liver fat may help identify statin-taking individuals at risk for T2D (47). Finally, it is noteworthy that the *HMGCR* variant that lowers circulating LDL-C is not associated with any

significant change in liver enzymes, potentially reflecting the lack of an increase in NAFLD incidence seen with statin medications (48).

Our analyses identified a number of variants previously implicated in lipid and glucose metabolism. Sortilin 1 (*SORT1*) is highly expressed in adipocytes, and the sortilin gene product facilitates the formation and export of VLDL from the liver (49,50). The role of *SORT1* in T2D risk is not well understood. Sortilin 1 is required for insulin-dependent glucose uptake (51–53), yet *Sort1*-knockout mice may show reduced glucose and glycolic intermediates in the fasted state (54). This highlights again the potential for heterogeneous paths in T2D risk and the dependence on multiple pathways of lipid and glucose metabolism to explain our findings.

Several loci were also identified that are known to be related to T2D, without known associations with circulating LDL-C. These include *THADA*, *C2CD4A*, *CENPW*, and *SLC12A8* (21). In addition, we identified several variants associated with lower circulating LDL-C but increased T2D risk with no known biological pathways linking these loci to either trait. *SLC2A2*, which encodes GLUT2, has not previously been associated with circulating LDL-C or T2D in the large respective GWAS consortia. However, GLUT2 is key to hepatic glucose uptake after a meal and the associated hepatic de novo lipogenesis (55). In fact, liver-specific GLUT2 knockout decreases liver TG concentrations. Importantly, GLUT2 expression in the β -cell is required for the glucose-stimulated insulin response (56). In turn, a locus that decreases GLUT2 expression would be expected to limit serum insulin, increase HbA_{1c}, and decrease circulating LDL-C.

Our approach is subject to several limitations. We used prevalent T2D in the UK Biobank, which limits inferences related to the direction of causality. As incident T2D cases develop in the UK Biobank, it will be important to examine the association of circulating LDL-C at baseline with incident T2D. The risk of a false-positive finding (i.e., a SNP that is associated with two traits in opposite directions, each with $P < 5 \times 10^{-5}$, and with a genome-wide significant metaCCA P value) is extremely low. However, a limitation of our study is that the replication is limited to a replication of the LDL-C effect estimates of these SNPs. This could lead to an increased risk of false-positive signals with respect to the associations of the SNPs with T2D. It is also difficult to identify the causal gene at identified loci. Although we annotated these loci according to nearby genes and/or previous annotation, the listed and mentioned genes may not necessarily be directly implicated, if at all. Our identification of loci associated with both LDL-C and T2D does not necessarily imply that in each case, the effects of the genetic variant on each trait are linked by a common pathway or mechanism. In other words, it is possible that the way in which a variant causes a lowering of LDL-C could be distinct (different tissue and/or pathway) from the way in which it increases T2D. It is thus

possible that some of the variants identified are not having effects on a common pathway. Furthermore, because we are not necessarily identifying a single causal variant, it is possible that within a locus, multiple different variants affect each trait. Results from colocalization do suggest that in many cases, there are different causal variants. However, differences in patterns of linkage disequilibrium between the UK Biobank and DIAGRAM consortium studies could reduce our ability to colocalize causal variants. If there was not a shared pathway, we might expect that the effect size of LDL-C would be directly proportional to the effect on T2D across all LDL-C-lowering alleles. However, there are many variants that are known to be strongly associated with LDL-C that are not identified in this analysis (e.g., *LDLR*, *APOB*, *ABCG5/8*).

In conclusion, our results suggest that low circulating LDL-C may be a risk factor for T2D, although further study is warranted. We have identified a collection of genetic variants that may provide insight into the mechanisms underlying the diabetogenic risk of low circulating LDL-C and of lipid-lowering medications, and the decreased T2D risk among individuals with familial hypercholesterolemia.

Acknowledgments. The authors acknowledge the vital contributions of the GLGC and DIAGRAM as well as all organizers and participants of individual participating studies. This research was conducted using the UK Biobank resource under application number 15678. The authors thank the participants and organizers of the UK Biobank.

Funding. The study received support from the National Heart, Lung, and Blood Institute (R01-HL-136528). A.C.W. was funded, in part, by U.S. Department of Agriculture/Agricultural Research Service cooperative agreement no. 58-3092-5-001.

The contents of this publication do not necessarily reflect the views or policies of the U.S. Department of Agriculture, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Author Contributions. Y.C.K. conceived and designed the study. Y.C.K., A.A., M.N., and J.Z. performed data analyses. Y.C.K., M.N., B.J.R., and A.C.W. wrote the manuscript. J.M.O., B.J.R., and A.C.W. contributed to writing of the introduction and discussion. All authors read and edited the full manuscript. Y.C.K. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the American Society of Human Genetics 2019 Annual Meeting, Houston, Texas, 15–19 October 2020.

References

1. Rawshani A, Rawshani A, Franzén S, et al. Mortality and cardiovascular disease in type 1 and type 2 diabetes. *N Engl J Med* 2017;376:1407–1418
2. Gregg EW, Li Y, Wang J, et al. Changes in diabetes-related complications in the United States, 1990–2010. *N Engl J Med* 2014;370:1514–1523
3. Postmus I, Deelen J, Sedaghat S, et al. LDL cholesterol still a problem in old age? A Mendelian randomization study. *Int J Epidemiol* 2015;44:604–612
4. Boekholdt SM, Arsenault BJ, Mora S, et al. Association of LDL cholesterol, non-HDL cholesterol, and apolipoprotein B levels with risk of cardiovascular events among patients treated with statins: a meta-analysis. *JAMA* 2012;307:1302–1309

5. Sattar N, Preiss D, Murray HM, et al. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. *Lancet* 2010;375:735–742
6. Crandall JP, Mather K, Rajpathak SN, et al. Statin use and risk of developing diabetes: results from the Diabetes Prevention Program. *BMJ Open Diabetes Res Care* 2017;5:e000438
7. Preiss D, Seshasai SRK, Welsh P, et al. Risk of incident diabetes with intensive-dose compared with moderate-dose statin therapy: a meta-analysis. *JAMA* 2011;305:2556–2564
8. Feng Q, Wei WQ, Chung CP, et al. Relationship between very low low-density lipoprotein cholesterol concentrations not due to statin therapy and risk of type 2 diabetes: a US-based cross-sectional observational study using electronic health records. *PLoS Med* 2018;15:e1002642
9. Andersson C, Lyass A, Larson MG, Robins SJ, Vasan RS. Low-density-lipoprotein cholesterol concentrations and risk of incident diabetes: epidemiological and genetic insights from the Framingham Heart Study. *Diabetologia* 2015;58:2774–2780
10. Sacks FM, Tonkin AM, Craven T, et al. Coronary heart disease in patients with low LDL-cholesterol: benefit of pravastatin in diabetics and enhanced role for HDL-cholesterol and triglycerides as risk factors. *Circulation* 2002;105:1424–1428
11. Besseling J, Kastelein JJP, Defesche JC, Hutten BA, Hovingh GK. Association between familial hypercholesterolemia and prevalence of type 2 diabetes mellitus. *JAMA* 2015;313:1029–1036
12. Lotta LA, Sharp SJ, Burgess S, et al. Association between low-density lipoprotein cholesterol-lowering genetic variants and risk of type 2 diabetes: a meta-analysis. *JAMA* 2016;316:1383–1391
13. Zhao W, Rasheed A, Tikkanen E, et al.; CHD Exome+ Consortium; EPIC-CVD Consortium; EPIC-Interact Consortium; Michigan Biobank. Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nat Genet* 2017;49:1450–1457
14. Liu DJ, Peloso GM, Yu H, et al.; Charge Diabetes Working Group; EPIC-InterAct Consortium; EPIC-CVD Consortium; GOLD Consortium; VA Million Veteran Program. Exome-wide association study of plasma lipids in >300,000 individuals. *Nat Genet* 2017;49:1758–1766
15. Schmidt AF, Swerdlow DI, Holmes MV, et al.; LifeLines Cohort study group; UCLEB consortium. PCSK9 genetic variants and risk of type 2 diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol* 2017;5:97–105
16. Fall T, Xie W, Poon W, et al.; GENESIS Consortium. Using genetic variants to assess the relationship between circulating lipids and type 2 diabetes. *Diabetes* 2015;64:2676–2684
17. White J, Swerdlow DI, Preiss D, et al. Association of lipid fractions with risks for coronary artery disease and diabetes. *JAMA Cardiol* 2016;1:692–699
18. Yaghoobkar H, Lotta LA, Tyrrell J, et al. Genetic evidence for a link between favorable adiposity and lower risk of type 2 diabetes, hypertension, and heart disease. *Diabetes* 2016;65:2448–2460
19. Eastwood SV, Mathur R, Atkinson M, et al. Algorithms for the capture and adjudication of prevalent and incident diabetes in UK Biobank. *PLoS One* 2016;11:e0162388
20. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203–209
21. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* 2018;50:1505–1513
22. Willer CJ, Schmidt EM, Sengupta S, et al.; Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013;45:1274–1283
23. Hemani G, Zheng J, Elsworth B, et al. The MR-base platform supports systematic causal inference across the human genome. *Elife* 2018;7:e34408
24. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 1934;26:404–413
25. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41–55
26. Ko MJ, Jo AJ, Kim YJ, et al. Time- and dose-dependent association of statin use with risk of clinically relevant new-onset diabetes mellitus in primary prevention: a nationwide observational cohort study. *J Am Heart Assoc* 2019;8:e011320
27. Hyun MH, Jang JW, Choi BG, et al. Risk of insulin resistance with statin therapy in individuals without dyslipidemia: a propensity-matched analysis in a registry population. *Clin Exp Pharmacol Physiol* 2020;47:947–954
28. Yamazaki K, Takahashi Y, Teduka K, Nakayama T, Nishida Y, Asai S. Assessment of effect modification of statins on new-onset diabetes based on various medical backgrounds: a retrospective cohort study. *BMC Pharmacol Toxicol* 2019;20:34
29. Ho DE, Imai K, King G, Stuart EA. MatchIt: nonparametric preprocessing for parametric causal inference. *J Stat Softw* 2011;42:1–28
30. Zhang Z, Kim HJ, Lonjon G, Zhu Y; written on behalf of AME Big-Data Clinical Trial Collaborative Group. Balance diagnostics after propensity score matching. *Ann Transl Med* 2019;7:16
31. Nissen SE, Tuzcu EM, Schoenhagen P, et al.; Reversal of Atherosclerosis with Aggressive Lipid Lowering (REVERSAL) Investigators. Statin therapy, LDL cholesterol, C-reactive protein, and coronary artery disease. *N Engl J Med* 2005;352:29–38
32. Loh P-R, Kichaev G, Gazal S, Schoech AP, Price AL. Mixed-model association for biobank-scale datasets. *Nat Gen* 2018;50:906–908
33. Cichonska A, Rousu J, Martinen P, et al. metaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis. *Bioinformatics* 2016;32:1981–1989
34. Lotta LA, Gulati P, Day FR, et al.; EPIC-InterAct Consortium; Cambridge FPLD1 Consortium. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nat Genet* 2017;49:17–26
35. Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* 2014;10:e1004383
36. Pratt DS, Kaplan MM. Evaluation of abnormal liver-enzyme results in asymptomatic patients. *N Engl J Med* 2000;342:1266–1271
37. Murtagh F. *Lectures in Computational Statistics: Multidimensional Clustering Algorithms. COMPSTAT Lectures 4*, Vienna, Physika Verlag, 1985
38. Brock G, Pihur V, Datta S, Datta S. cValid: an R package for cluster validation [article online]. *J Stat Software* 2008;25:1–22. Available from <https://www.jstatsoft.org/v025/i04>. Accessed 23 April 2020
39. O'Reilly PF, Hoggart CJ, Pomyen Y, et al. MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. *PLoS One* 2012;7:e34861.
40. Dupuis J, Langenberg C, Prokopenko I, et al.; DIAGRAM Consortium; GIANT Consortium; Global BPgen Consortium; Anders Hamsten on behalf of Procardis Consortium; MAGIC Investigators. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 2010;42:105–116
41. Higuchi S, Izquierdo MC, Haeusler RA. Unexplained reciprocal regulation of diabetes and lipoproteins. *Curr Opin Lipidol* 2018;29:186–193
42. Speliotes EK, Yerges-Armstrong LM, Wu J, et al.; NASH CRN; GIANT Consortium; MAGIC Investigators; GOLD Consortium. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet* 2011;7:e1001324
43. Yuan X, Waterworth D, Perry JRB, et al. Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. *Am J Hum Genet* 2008;83:520–528.
44. Chambers JC, Zhang W, Sehmi J, et al.; Alcohol Genome-wide Association (AlcGen) Consortium; Diabetes Genetics Replication and Meta-analyses (DIAGRAM+) Study; Genetic Investigation of Anthropometric Traits (GIANT) Consortium; Global Lipids Genetics Consortium; Genetics of Liver Disease (GOLD) Consortium; International Consortium for Blood Pressure (ICBP-GWAS); Meta-analyses of Glucose and Insulin-Related Traits Consortium (MAGIC). Genome-wide association study

- identifies loci influencing concentrations of liver enzymes in plasma. *Nat Genet* 2011;43:1131–1138
45. Sliz E, Sebert S, Würtz P, et al. NAFLD risk alleles in PNPLA3, TM6SF2, GCKR and LYPLAL1 show divergent metabolic effects. *Hum Mol Genet* 2018;27:2214–2223
46. De Silva NMG, Borges MC, Hingorani AD, et al.; UCLEB consortium. Liver function and risk of type 2 diabetes: bidirectional Mendelian randomization study. *Diabetes* 2019;68:1681–1691
47. Shah RV, Allison MA, Lima JAC, et al. Liver fat, statin use, and incident diabetes: the Multi-Ethnic Study of Atherosclerosis. *Atherosclerosis* 2015;242:211–217
48. Sigler MA, Congdon L, Edwards KL. An evidence-based review of statin use in patients with nonalcoholic fatty liver disease. *Clin Med Insights Gastroenterol* 2018;11:1179552218787502
49. Chen C, Li J, Matye DJ, Wang Y, Li T. Hepatocyte sortilin 1 knockout and treatment with a sortilin 1 inhibitor reduced plasma cholesterol in Western diet-fed mice. *J Lipid Res* 2019;60:539–549
50. Brown MS, Goldstein JL. Sterol regulatory element binding proteins (SREBPs): controllers of lipid synthesis and cellular uptake. *Nutr Rev* 1998;56:S1–S3; discussion S54–S75
51. Shi J, Kandror KV. Sortilin is essential and sufficient for the formation of Glut4 storage vesicles in 3T3-L1 adipocytes. *Dev Cell* 2005;9:99–108
52. Huang G, Buckler-Pena D, Nauta T, et al. Insulin responsiveness of glucose transporter 4 in 3T3-L1 cells depends on the presence of sortilin. *Mol Biol Cell* 2013;24:3115–3122
53. Tsuchiya Y, Hatakeyama H, Emoto N, Wagatsuma F, Matsushita S, Kanzaki M. Palmitate-induced down-regulation of sortilin and impaired GLUT4 trafficking in C2C12 myotubes. *J Biol Chem* 2010;285:34371–34381
54. Li J, Matye DJ, Wang Y, Li T. Sortilin 1 knockout alters basal adipose glucose metabolism but not diet-induced obesity in mice. *FEBS Lett* 2017;591:1018–1028
55. Seyer P, Vallois D, Poitry-Yamate C, et al. Hepatic glucose sensing is required to preserve β cell glucose competence. *J Clin Invest* 2013;123:1662–1676
56. Guillam MT, Dupraz P, Thorens B. Glucose uptake, utilization, and signaling in GLUT2-null islets. *Diabetes* 2000;49:1485–1491