

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

¹**Abstract**

2 South Asians are at higher risk of dyslipidaemia–a modifiable risk factor for cardiovascular
3 diseases (CVDs). We aimed to identify protein targets for dyslipidaemia and CVDs in this

3 diseases (CVDs). We aimed to identify protein targets for dyslipidaemia and CVDs in this population. population.

5 We used a two-sample Mendelian randomization (MR) approach, supplemented with MR-
6 Egger, weighted median, colocalization, and generalized MR (GMR), to evaluate the effect 6 Egger, weighted median, colocalization, and generalized MR (GMR), to evaluate the effect
7 of 2.800 plasma proteins on high/low/non-high-density lipoprotein cholesterol (HDL-C/LDL-7 of 2,800 plasma proteins on high/low/non-high-density lipoprotein cholesterol (HDL-C/LDL-
8 C/nonHDL-C), total cholesterol, and triglycerides. Observational analyses were conducted on 8 C/nonHDL-C), total cholesterol, and triglycerides. Observational analyses were conducted on
9 MR findings with strong colocalization (posterior probability> 80%) and GMR findings. 9 MR findings with strong colocalization (posterior probability≥ 80%) and GMR findings.
10 Univariate MR assessed lipid-associated proteins' effect on CVDs. Finally, we compared the 10 Univariate MR assessed lipid-associated proteins' effect on CVDs. Finally, we compared the
11 potential causal effects of plasma proteins on lipids in South Asians with those in Europeans 11 potential causal effects of plasma proteins on lipids in South Asians with those in Europeans
12 to study heterogeneity in the MR effects. to study heterogeneity in the MR effects.

13 We identified 29 genetically proxied proteins potentially causal to at least one lipid measure,
14 12 of which showed strong colocalization and GMR evidence, including ANGPTL3 and 14 12 of which showed strong colocalization and GMR evidence, including ANGPTL3 and
15 PCSK9. Notably. PCSK9 demonstrated a stronger association with LDL-C in European 15 PCSK9. Notably, PCSK9 demonstrated a stronger association with LDL-C in European
16 compared to South Asian ($\beta_{\text{European}} = 0.37$: 95% Confidence Interval (CI)= (0.36, 0.38). β_{South} 16 compared to South Asian (β_{European}= 0.37; 95% Confidence Interval (CI)= (0.36, 0.38), β_{South}
17 $\epsilon_{\text{sim}} = 0.16$: 95% CI= (0.11, 0.21)). Observational analysis suggested significant interaction 17 $\text{Asian} = 0.16$; 95% CI= (0.11, 0.21)). Observational analysis suggested significant interaction
18 between PCSK9 levels with LDL-C levels in South Asians with South Asians having a 18 between PCSK9 levels with LDL-C levels in South Asians with South Asians having a
19 significantly lower effect compared to other ethnicities (PCSK9*South Asian: 6= -0.14: 95% 19 significantly lower effect compared to other ethnicities (PCSK9*South Asian; β= -0.14; 95%
20 CI= (-0.174, -0.107)). Additionally, we showed that CELSR2 is also linked with CAD in 20 CI= (-0.174, -0.107)). Additionally, we showed that CELSR2 is also linked with CAD in
21 South Asians. 21 South Asians. 22 Our study highlighted potential causal links between plasma proteins, dyslipidaemia, and
23 CVD in South Asians with significant heterogeneity across genetic ancestry groups. Larger

23 CVD in South Asians, with significant heterogeneity across genetic ancestry groups. Larger
24 studies in South Asians are needed to validate these findings. studies in South Asians are needed to validate these findings.

25

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

1 **Introduction**
2 Routinely measure

2 Routinely measured lipid parameters including high/low/non-high-density lipoprotein
3 cholesterol (HDL-C/LDL-C/nonHDL-C), total cholesterol (TC), and triglycerides (TG), are 3 cholesterol (HDL-C/LDL-C/nonHDL-C), total cholesterol (TC), and triglycerides (TG), are
4 well-established risk factors for cardiovascular diseases (CVDs), including coronary artery well-established risk factors for cardiovascular diseases (CVDs), including coronary artery
5 disease (CAD), stroke, and heart failure, (1-4) all of which are leading causes of mortality 5 disease (CAD), stroke, and heart failure, (1-4) all of which are leading causes of mortality
6 and morbidity worldwide (5). HDL-C and TC have been included in several risk equations 6 and morbidity worldwide (5). HDL-C and TC have been included in several risk equations
7 like SCORE and PCE as important predictors for CVDs (6, 7). The role of HDL-C and TG as 7 like SCORE and PCE as important predictors for CVDs (6, 7). The role of HDL-C and TG as
8 the target for CVD intervention is inconclusive but under investigation (8, 9). LDL-C and 8 the target for CVD intervention is inconclusive but under investigation (8, 9). LDL-C and
9 nonHDL-C are not only prognostic risk markers but also established therapeutic targets for 9 nonHDL-C are not only prognostic risk markers but also established therapeutic targets for
10 CVDs (10, 11). Building on this, evidence suggests that individuals of South Asian ancestry. 10 CVDs (10, 11). Building on this, evidence suggests that individuals of South Asian ancestry,
11 compared with non-Hispanic white population have a higher prevalence of dyslinidaemia 11 compared with non-Hispanic white population, have a higher prevalence of dyslipidaemia
12 (12) and are more susceptible to cardiometabolic diseases closely related to dyslipidaemia 12 (12) and are more susceptible to cardiometabolic diseases closely related to dyslipidaemia
13 including CAD (13, 14), stroke (15), and type 2 diabetes (16). Therefore, understanding the 13 including CAD (13, 14), stroke (15), and type 2 diabetes (16). Therefore, understanding the
14 genomic and proteomic makeup of plasma lipids and identifying causal factors in South 14 genomic and proteomic makeup of plasma lipids and identifying causal factors in South
15 Asians is crucial for intervening in dyslipidaemia and preventing lipid-related cardiovascular 15 Asians is crucial for intervening in dyslipidaemia and preventing lipid-related cardiovascular
16 conditions although the residual risk of lipid modifying medications remains. conditions although the residual risk of lipid modifying medications remains.

17 Circulating plasma proteins are key to disease mechanisms and are promising drug targets.
18 Previous studies have identified proteins associated with linids some of which are targeted for 18 Previous studies have identified proteins associated with lipids some of which are targeted for
19 dyslipidaemia treatment (17). The most prominent example is Apolipoprotein B (APOB), 19 dyslipidaemia treatment (17). The most prominent example is Apolipoprotein B (APOB),
20 which is one of the most prognostic and best therapeutic targets for CVDs (18). Despite 20 which is one of the most prognostic and best therapeutic targets for CVDs (18). Despite
21 recent advances, most plasma proteins linked to lipids have been discovered in the European 21 recent advances, most plasma proteins linked to lipids have been discovered in the European
22 repondation (19) The proteomic findings in non-European individuals are limited especially 22 population (19). The proteomic findings in non-European individuals are limited, especially
23 in South Asians, which is the most rapidly growing but neglected population with higher 23 in South Asians, which is the most rapidly growing but neglected population with higher
24 susceptibility to dyslipidaemia (12) We recently reported that certain plasma proteins may 24 susceptibility to dyslipidaemia (12). We recently reported that certain plasma proteins may
25 exert ancestry-specific causal effects on certain CVDs (20). Therefore, understanding the 25 exert ancestry-specific causal effects on certain CVDs (20). Therefore, understanding the
26 proteomic features of dyslipidaemia in South Asians is essential for developing more 26 proteomic features of dyslipidaemia in South Asians is essential for developing more
27 effective disease prevention strategies and drug discovery approaches for this high-risk 27 effective disease prevention strategies and drug discovery approaches for this high-risk
28 population. population.

29 Recent genome wide association studies (GWAS) involving South Asian populations have
20 made resources on circulating plasma proteins and lipid traits available (21, 22). Furthermore, 30 made resources on circulating plasma proteins and lipid traits available (21, 22). Furthermore,
31 advances in methodologies for causal inference in epidemiology, including the Mendelian 31 advances in methodologies for causal inference in epidemiology, including the Mendelian
32 randomization (MR) framework, now allow us to investigate the potential causal relationship 32 randomization (MR) framework, now allow us to investigate the potential causal relationship
33 between an exposure (e.g. plasma proteins) and an outcome (e.g. lipid traits) for which 3 between an exposure (e.g. plasma proteins) and an outcome (e.g. lipid traits) for which

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

1 GWAS summary statistics are available (23). MR utilises genetic variants as instrument
2 variables (IVs), and is less susceptible to confounders and reverse causation bias than other 2 variables (IVs), and is less susceptible to confounders and reverse causation bias than other
3 study designs, and can be applied to for causal inference (24). Such approach was applied to 3 study designs, and can be applied to for causal inference (24). Such approach was applied to
4 systematically evaluate plasma proteins' effect on lipid traits in EUR, but not yet in SAS (25). systematically evaluate plasma proteins' effect on lipid traits in EUR, but not yet in SAS (25). 5 In this study, we aimed to (1) systematically evaluate bi-directional causal effects of plasma
6 proteins on five lipid traits in South Asians using a two-sample proteome-wide MR approach 6 proteins on five lipid traits in South Asians using a two-sample proteome-wide MR approach

2 (2) Investigate whether lipid-associated plasma proteins affect CAD and stroke risk in South 7 (2) Investigate whether lipid-associated plasma proteins affect CAD and stroke risk in South
8 Asians (3) Compare the effects of plasma proteins on lipid fractions in South Asians and 8 Asians (3) Compare the effects of plasma proteins on lipid fractions in South Asians and
9 Europeans. Europeans.

10

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

¹**Method**

²**Data source**

³**Genetic associations for genetically predicted plasma proteins**

In the UK Biobank Pharma Proteomics Project (UKBPPP), 2,940 probes capturing 2,922
5 unique proteins were made available (22). We defined each probe as a protein and extracted 5 unique proteins were made available (22). We defined each probe as a protein and extracted
6 eenetic associations of the 2.940 plasma proteins in individuals of Central/South Asian 6 genetic associations of the 2,940 plasma proteins in individuals of Central/South Asian
7 ancestry (CSA, N=920) and European ancestry (EUR, N=34.557) from the UKBPPP (22). 7 ancestry (CSA, N=920) and European ancestry (EUR, N=34,557) from the UKBPPP (22).
8 The ancestries were defined with the pan-UK Biobank (UKBB) definitions of genetic 8 The ancestries were defined with the pan-UK Biobank (UKBB) definitions of genetic
9 ancestry (available in UKBB return dataset 2442). We included the 2.800 plasma proteins as 9 ancestry (available in UKBB return dataset 2442). We included the 2,800 plasma proteins as
10 the primary exposures after excluding the ones whose cognate gene is ambiguous ($N = 15$) 10 the primary exposures after excluding the ones whose cognate gene is ambiguous ($N = 15$)
11 and those encoded by genes lying on the X chromosome (due to unavailability of X 11 and those encoded by genes lying on the X chromosome (due to unavailability of X
12 chromosome data on outcome: $N = 88$) or the MHC region (due to complex LD in this 12 chromosome data on outcome; $N = 88$) or the MHC region (due to complex LD in this region: $N = 37$). region; $N = 37$).

¹⁴**Genetic associations for genetically predicted lipid traits and CVDs**

15 Five lipid traits including HDL-C, LDL-C, nonHDL-C, TC, and TG were included as the
16 primary outcome in MR analysis Ancestry-specific $(N_{SAS} \Box \equiv \Box$ to 40.963 16 primary outcome in MR analysis. Ancestry-specific (N_{SAS} \Box = \Box up to 40,963,
17 N_{EUP} \Box = \Box 1 320 016) genetic associations of the 5 traits were publicly available from the 17 $N_{EUR} \square = \square 1,320,016$ genetic associations of the 5 traits were publicly available from the Global Lipids Genetics Consortium (21). The GWAS statistics for CAD were obtained from 18 Global Lipids Genetics Consortium (21). The GWAS statistics for CAD were obtained from
19 the East London Genes & Health (ELGH) study for SAS (26) and from the 19 the East London Genes & Health (ELGH) study for SAS (26) and from the
20 CARDIoGRAMplusC4D Consortium for EUR (27). The GWAS summary data on stroke for 20 CARDIoGRAMplusC4D Consortium for EUR (27). The GWAS summary data on stroke for
21 EUR and SAS were sourced from the GIGASTROKE consortium (28) EUR and SAS were sourced from the GIGASTROKE consortium (28).

²²**Proteome-wide MR and colocalization analysis on lipid fractions in SAS**

²³**Instrument selection and MR:**

24 To obtain genetic instruments for plasma proteins, biallelic single nucleotide polymorphisms
25 (SNP) lying within $+/-500$ kilobase (KB) from the coding gene (defined as cis-acting SNPs). 25 (SNP) lying within $+/-$ 500 kilobase (KB) from the coding gene (defined as cis-acting SNPs),
26 with minor allele frequency (MAF) > 0.05 and reaching genome wide significant level (P < 26 with minor allele frequency (MAF) > 0.05, and reaching genome wide significant level (P < $\sim 5 \times 10^{-8}$) were extracted. The SNPs were further harmonized to the 5 linid fractions and 5×10^{-8}) were extracted. The SNPs were further harmonized to the 5 lipid fractions and
28 clumped to an $r^2 < 0.001$ to ensure independence across the instruments. Subsequently a F-28 clumped to an r^2 < 0.001 to ensure independence across the instruments. Subsequently, a F-
29 statistic was calculated for each SNP and SNPs with F-statistic < 10 were excluded to avoid 29 statistic was calculated for each SNP and SNPs with F-statistic < 10 were excluded to avoid
20 weak instrument bias (29). We also applied the Steiger filtering and excluded the SNPs with 30 weak instrument bias (29). We also applied the Steiger filtering and excluded the SNPs with
31 potential reverse causality (30). After applying these filters, we ended up with 708 proteins potential reverse causality (30). After applying these filters, we ended up with 708 proteins

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

 $\ddot{}$ $\frac{1}{2}$ 1 with at least 1 instrument available which were carried forward for the downstream analysis
2 (Figure S1). The Wald ratio method was applied to proteins with 1 single SNP as the
1 instrument while the inverse-variance we 2 (Figure S1). The Wald ratio method was applied to proteins with 1 single SNP as the instrument while the inverse-variance weighted (IVW) model was applied to proteins which were instrumented by 2 or more SNPs (23). To te 3 instrument while the inverse-variance weighted (IVW) model was applied to proteins which
4 were instrumented by 2 or more SNPs (23). To test for horizontal pleiotropy and to ensure
5 robustness of the proteome-wide MR fi were instrumented by 2 or more SNPs (23). To test for horizontal pleiotropy and to ensure
5 robustness of the proteome-wide MR findings, we applied MR-Egger and Weighted median
6 to the significant associations where at le 5 robustness of the proteome-wide MR findings, we applied MR-Egger and Weighted median
6 to the significant associations where at least 3 instruments were available (31, 32). For
7 multiple testing correction, a false disc to the significant associations where at least 3 instruments were available (31, 32). For
multiple testing correction, a false discovery rate (FDR) was calculated by applying the
Benjamini-Hochberg (BH) adjustment to each 7 multiple testing correction, a false discovery rate (FDR) was calculated by applying the Benjamini-Hochberg (BH) adjustment to each lipid fraction (33) and an FDR < 0.05 was defined to be significant. 8 Benjamini-Hochberg (BH) adjustment to each lipid fraction (33) and an FDR < 0.05 was
defined to be significant.
0 To avoid bias due to sample overlap between UKBPPP and GLGC, we also performed MR

9 defined to be significant.
0 To avoid bias due to sam
1 using GLGC data exclude 10 To avoid bias due to sample overlap between UKBPPP and GLGC, we also performed MR
11 using GLGC data excluding UKBB participants. A correlation analysis was performed to
12 compare the beta estimates derived from GLGC d 11 using GLGC data excluding UKBB participants. A correlation analysis was performed to
12 compare the beta estimates derived from GLGC data with or without UKBB individuals.
13 **Bayesian colocalization**: A Bayesian coloca

12 compare the beta estimates derived from GLGC data with or without UKBB individuals.
 Bayesian colocalization: A Bayesian colocalization analysis was conducted on all pro

lipid associations with FDR-corrected $P < 0.0$ 13 **Bayesian colocalization**: A Bayesian colocalization analysis was conducted on all protein-
14 lipid associations with FDR-corrected P < 0.05 to determine if they shared the same causal
15 variant (34). The colocalizati 14 lipid associations with FDR-corrected $P < 0.05$ to determine if they shared the same causal
15 variant (34). The colocalization enabled us to minimise horizontal pleiotropy caused by
16 linkage disequilibrium (LD) wher 15 variant (34). The colocalization enabled us to minimise horizontal pleiotropy caused by
16 linkage disequilibrium (LD) where the plasma protein levels and lipid traits were influenced
17 by 2 distinct variants in LD wit 16 linkage disequilibrium (LD) where the plasma protein levels and lipid traits were influenced
17 by 2 distinct variants in LD with each other (35). Colocalization was performed on the same
18 window as the previous prote 17 by 2 distinct variants in LD with each other (35). Colocalization was performed on the same
18 window as the previous proteome-wide MR (within $+/-$ 500KB of the cognate gene), with
19 rare variants (MAF < 0.05) residin 18 window as the previous proteome-wide MR (within $+/-$ 500KB of the cognate gene), with
19 rare variants (MAF < 0.05) residing in the window dropped. Default priors as described in the
20 original paper were applied (34) 19 19 19 19 20 rare variants (MAF 21 original paper were applied (34). A posterior probability of colocalization (PPH4) \geq 80% variants indicated strong colocalization. while 60% \leq PPH4 \lt 80% was considered sugg 20 original paper were applied (34). A posterior probability of colocalization (PPH4) ≥ 80% indicated strong colocalization, while $60\% \leq PPH4 < 80\%$ was considered suggestive 21 indicated strong colocalization, while $60\% \le PPH4 < 80\%$ was considered suggestive
22 evidence for colocalization.
23 **MR** generalized to correlated instruments (GMR): Due to stringent genetic instrument

evidence for colocalization.

22 evidence for colocalization.

23 **MR generalized to correlated instruments (GMR)**: Due to stringent genetic instrument

24 selection, most proteins had few pQTLs, which could bias the MR es 23 **MR generalized to correlated instruments (GMR):** Due to stringent genetic instrument selection, most proteins had few pQTLs, which could bias the MR estimates due to unknown pleiotropy and make MR-Egger and WM unapplic 24 selection, most proteins had few pQTLs, which could bias the MR estimates due to unknown
25 pleiotropy and make MR-Egger and WM unapplicable. Therefore, to improve robustness, we
26 included additional instruments at 25 pleiotropy and make MR-Egger and WM unapplicable. Therefore, to improve robustness, we included additional instruments at $p < 1 \times 10$ and clumped them with $r^2 = 0.4$ (36). To account for correlation between instrument 26 included additional instruments at $p < 1 \times 10$ and clumped them with $r^2 = 0.4$ (36). To
27 account for correlation between instruments, a generalized inverse variance weighted
28 regression (gIVW) was applied as the p 27 account for correlation between instruments, a generalized inverse variance weighted
28 regression (gIVW) was applied as the primary method (37). Where applicable, we also
29 performed the MR-Egger generalized to correl 28 regression (gIVW) was applied as the primary method (37). Where applicable, we also performed the MR-Egger generalized to correlated variants (gEgger) (38) and weighted median (32). 29 performed the MR-Egger generalized to correlated variants (gEgger) (38) and weighted
30 median (32).
Causal effects of lipid-associated proteins on CAD and stroke

30 median (32).
31 **Causal effect** 31 **Causal effects of lipid-associated proteins on CAD and stroke**

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

 \mathbf{I} $\frac{1}{2}$ 1 A *cis*-MR analysis was conducted with lipid-associated plasma proteins (P-FDR< 0.05) as
2 exposures and CAD and stroke (any stroke, any ischemic stroke, large artery stroke,
3 cardioembolic stroke, and small vessel stro exposures and CAD and stroke (any stroke, any ischemic stroke, large artery stroke, cardioembolic stroke, and small vessel stroke) as outcomes. An FDR correction was applied separately for each outcome while colocalization 3 cardioembolic stroke, and small vessel stroke) as outcomes. An FDR correction was applied
4 separately for each outcome while colocalization and supplemental MR methods (gIVW,
5 gEgger, and weighted median) were performe 4 separately for each outcome while colocalization and supplemental MR methods (gIVW,
5 gEgger, and weighted median) were performed to validate the genetically proxied
6 associations surviving the FDR correction (32-34, 37 5 gEgger, and weighted median) were performed to validate the genetically proxied
6 associations surviving the FDR correction (32-34, 37, 38). The instrument selection criteria
7 and parameters for MR and colocalization in 6 associations surviving the FDR correction (32-34, 37, 38). The instrument selection criteria
3 and parameters for MR and colocalization in this step were set as described above.
8 Additionally, univariate MR was performe

9 similar causal effects on CVDs in South Asians compare to Europeans. Genetic instruments 8 Additionally, univariate MR was performed to assess whether the lipids measures show
9 similar causal effects on CVDs in South Asians compare to Europeans. Genetic instruments
0 for lipid fractions were extracted from th 9 similar causal effects on CVDs in South Asians compare to Europeans. Genetic instruments
0 for lipid fractions were extracted from the whole genome (autosomes only), and other MR
1 criteria and methods were the same as t 10 for lipid fractions were extracted from the whole genome (autosomes only), and other MR
11 criteria and methods were the same as those applied in *cis-MR* (29, 30). Where applicable,
12 MR-Egger and weighted median were 11 criteria and methods were the same as those applied in *cis-MR* (29, 30). Where applicable,
12 MR-Egger and weighted median were applied as well (31, 32).
13 Furthermore, for plasma proteins associated with both lipids

12 MR-Egger and weighted median were applied as well (31, 32).
13 Furthermore, for plasma proteins associated with both lipids a
14 multi-trait colocalization whether the 3 traits have a sh 13 Furthermore, for plasma proteins associated with both lipids and CVDs, we evaluated with
14 multi-trait colocalization whether the 3 traits have a shared causal variant in the
15 corresponding gene region (39). The colo 14 multi-trait colocalization whether the 3 traits have a shared causal variant in the corresponding gene region (39). The colocalization analysis was applied on plasma protein, lipid fractions, and the CVD on the genomic 15 corresponding gene region (39). The colocalization analysis was applied on plasma protein,
16 lipid fractions, and the CVD on the genomic region $+/-$ 500KB extended from the cognate
17 gene (39). Rare variants with MAF 16 lipid fractions, and the CVD on the genomic region +/- 500KB extended from the cognate
17 gene (39). Rare variants with MAF < 0.05 were dropped. Priors for multi-trait colocalization
18 were set to default values: the p 17 gene (39). Rare variants with MAF < 0.05 were dropped. Priors for multi-trait colocalization
18 were set to default values: the probability of any SNP within the colocalization window being
19 exclusively associated wi 18 were set to default values: the probability of any SNP within the colocalization window being
20 exclusively associated with one of the three traits was 1×100 , with two traits was 1×100 , and with all three trai 19 exclusively associated with one of the three traits was $1 \times 10\Box$, with two traits was $1 \times 10\Box$, and with all three traits was $1 \times 10\Box$ (39).
21 **Reverse MR with lipid fractions as exposures and plasma proteins as**

21 Reverse MR with lipid fractions as exposures and plasma proteins as outcomes in SAS
22 Furthermore to understand the potential causal effects of dyslipidaemia on plasma proteing

22 Furthermore, to understand the potential causal effects of dyslipidaemia on plasma protein
23 abundance, we conducted reverse MR using lipid fractions (exposure) and plasma proteins 22 Furthermore, to understand the potential causal effects of dyslipidaemia on plasma protein
23 abundance, we conducted reverse MR using lipid fractions (exposure) and plasma proteins
24 (outcomes). Genetic instruments we 23 abundance, we conducted reverse MR using lipid fractions (exposure) and plasma proteins
24 (outcomes). Genetic instruments were selected from the 22 autosomes using the filtering
25 criteria as previously described. 24 (outcomes). Genetic instruments were selected from the 22 autosomes using the filtering
25 criteria as previously described.
26 **Comparison with European population** 25 criteria as previously described.
26 **Comparison with European population**
27 For plasma proteins with potential effects

27 For plasma proteins with potential effects on lipid fractions in SAS, we estimate their effects
28 in EUR using a two-sample MR approach, applying the same genetic instrument criteria used 28 in EUR using a two-sample MR approach, applying the same genetic instrument criteria used
29 in SAS. We checked the consistency of the MR estimates and applied a correlation analysis
30 between EUR and SAS population es in SAS. We checked the consistency of the MR estimates and applied a correlation analysis 29 in SAS. We checked the consistency of the MR estimates and applied a correlation analysis
30 between EUR and SAS population estimates for proteins that were consistent and significant 30 between EUR and SAS population estimates for proteins that were consistent and significant
 $\frac{7}{100}$

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

- $\frac{1}{2}$ 1 in both groups. We also compared the 95% confidence intervals, defining a significant
2 difference in the genetically proxied MR estimates when the intervals did not overlap.
3 **Observational associations of protein leve**
- $\frac{1}{\sqrt{2}}$ 2 difference in the genetically proxied MR estimates when the intervals did not overlap.

2 **Observational associations of protein levels with lipid fractions**

4 For plasma proteins associated to LDL-C and HDL-C with all

3 **Observations** For plasma proteins associated to LDL-C and HDL-C with all the MR, strong colocalization,
3 and GMR evidence, we conducted observational analysis using linear regression for each 4 and GMR evidence, we conducted observational analysis using linear regression for each
6 protein and its associated lipid fraction. Each regression model was adjusted for age, sex,
7 Townsend deprivation, BMI, HbA1c, cho 5 and GMR evidence, we conducted observational analysis using linear regression for each
6 protein and its associated lipid fraction. Each regression model was adjusted for age, sex,
7 Townsend deprivation, BMI, HbA1c, cho 6 protein and its associated lipid fraction. Each regression model was adjusted for age, sex,
7 Townsend deprivation, BMI, HbA1c, cholesterol medication, smoking, systolic blood
8 pressure, blood pressure medication, and e 7 Townsend deprivation, BMI, HbA1c, cholesterol medication, smoking, systolic blood
8 pressure, blood pressure medication, and ethnicity. To evaluate potential effect modification
9 by ancestry, we included an interaction 8 pressure, blood pressure medication, and ethnicity. To evaluate potential effect modification
9 by ancestry, we included an interaction term between a binary variable for South Asian
0 ancestry and protein level. All con 9 by ancestry, we included an interaction term between a binary variable for South Asian
0 ancestry and protein level. All continuous variables and lipid fractions were standardised
1 before modelling and p values were adj 10 ancestry and protein level. All continuous variables and lipid fractions were standardised
11 before modelling and p values were adjusted for multiple testing using FDR at 5%.
12 11 before modelling and p values were adjusted for multiple testing using FDR at 5%.
12

12
13

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

1 **Result**
2 An overv

²An overview of the study design is shown in **Figure 1**.

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

 \overline{a} $\frac{1}{1}$ 1 **Figure 1**. Overview of the study design. MR; Mendelian randomization, FDR; false discovery rate, GMR; generalized MR, PPH4; posterior probability of hypothesis 4, HDL-C and LDL-C; high- and low-density lipoprotein chole

2 discovery rate, GMR; generalized MR, PPH4; posterior probability of hypothesis 4, HDL-C
and LDL-C; high- and low-density lipoprotein cholesterol
Proteome-wide MR identified 29 plasma proteins associated with lipid fract

3 and LDL-C; high- and low-density lipoprotein cholesterol
 4 Proteome-wide MR identified 29 plasma proteins assoc

5 Excluding proteins with ambiguous cognate genes and tho

4 Excluding proteins with ambiguous cognate genes and those encoded by genes located on the
4 Ex-chromosome or within the *MHC* region (**Table S1**), 2,800 proteins were included in our 5 Excluding proteins with ambiguous cognate genes and those encoded by genes located on the X-chromosome or within the *MHC* region (**Table S1**), 2,800 proteins were included in our study. Of these, 708 had at least one ge 8 X-chromosome or within the *MHC* region (**Table S1**), 2,800 proteins were included in our study. Of these, 708 had at least one genetic instrument variable available (**Figure S1**) and were carried forward for the proteom 7 study. Of these, 708 had at least one genetic instrument variable available (**Figure S1**) and
8 were carried forward for the proteome-wide MR study. Using the Wald ratio (IV=1) or IVW
9 (IVs≥2) approach as the primary M were carried forward for the proteome-wide MR study. Using the Wald ratio (IV=1) or IVW (IVs \geq 2) approach as the primary MR method, 186 plasma proteins showed a potential causal effect on at least one of the lipid frac 9 (IVs≥2) approach as the primary MR method, 186 plasma proteins showed a potential causal
0 effect on at least one of the lipid fractions (309 associations in total; $P < 0.05$; Table S2).
4 After adjusting for multiple t 10 effect on at least one of the lipid fractions (309 associations in total; P < 0.05; **Table S2**).
11 After adjusting for multiple testing, a total of 29 genetically proxied plasma proteins showed
12 potential causal effe 11 After adjusting for multiple testing, a total of 29 genetically proxied plasma proteins showed
12 potential causal effect on at least one of the 5 lipid fractions (53 associations in total; FDR < 0.05, Figure 2, Table 12 potential causal effect on at least one of the 5 lipid fractions (53 associations in total; FDR < 0.05, **Figure 2, Table S2**).
14 To avoid bias due to sample overlap between UKBPPP and GLGC, we performed MR using

13 0.05, **Figure 2, Table S2**).
14 To avoid bias due to sampl
15 GLGC data without UKBB 14 To avoid bias due to sample overlap between UKBPPP and GLGC, we performed MR using
15 GLGC data without UKBB participants and the MR estimates were largely consistent across
16 GLGC with and without UKBB (Pearson $r^2 =$

GLGC with and without UKBB (Pearson r^2

15 GLGC data without UKBB participants and the MR estimates were largely consistent across
16 GLGC with and without UKBB (Pearson $r^2 = 0.93$, P< 0.001; **Figure S2A, Figure S2B**).
17 Among these 29 lipids-associated prote 16 GLGC with and without UKBB (Pearson $r^2 = 0.93$, P< 0.001; **Figure S2A, Figure S2B**).
17 Among these 29 lipids-associated proteins (53 associations), 12 (25 associations) supported by strong colocalization evidence (PP 17 Among these 29 lipids-associated proteins (53 associations), 12 (25 associations) were
18 supported by strong colocalization evidence (PPH4 > 80%) with one of the lipids fractions
19 including ANGPTL3, APOE, CELSR2, EPP 18 supported by strong colocalization evidence (PPH4 > 80%) with one of the lipids fractions
19 including ANGPTL3, APOE, CELSR2, EPPK1, GAS6, GSTA1, GSTA3, HGFAC, LPA,
20 PCSK9, PLA2G15, and PLTP (**Table S3**). Particularly including ANGPTL3, APOE, CELSR2, EPPK1, GAS6, GSTA1, GSTA3, HGFAC, LPA,
PCSK9, PLA2G15, and PLTP (**Table S3**). Particularly, we identified strong colocalization of
ANGPTL3 with all lipid fractions except HDL, CELSR2 with a 20 PCSK9, PLA2G15, and PLTP (**Table S3**). Particularly, we identified strong colocalization of
21 ANGPTL3 with all lipid fractions except HDL, CELSR2 with all lipid fractions except TG,
22 LPA with LDL-C and TC, and PCSK9 21 ANGPTL3 with all lipid fractions except HDL, CELSR2 with all lipid fractions except TG,
22 LPA with LDL-C and TC, and PCSK9 with HDL, TC, and non-HDL. Additionally, 7 plasma
23 proteins (7 associations in total) showed 22 LPA with LDL-C and TC, and PCSK9 with HDL, TC, and non-HDL. Additionally, 7 plasma
23 proteins (7 associations in total) showed suggestive colocalization evidence (80% > PPH4 \ge 60%) with the tested lipid fractions (23 proteins (7 associations in total) showed suggestive colocalization evidence (80% > PPH4 ≥ 60%) with the tested lipid fractions (**Figure 2, Figure S3, Table S3**).
25 To further validate our findings in the proteome-

24 60%) with the tested lipid fractions (**Figure 2, Figure S3, Table S3**).
25 To further validate our findings in the proteome-wide cis-MR, we correlated SNPs $(r^2 < 0.4)$ that are adequately associated with the plas 25 To further validate our findings in the proteome-wide cis-MR, we included moderately
26 correlated SNPs (r^2 < 0.4) that are adequately associated with the plasma proteins (p < 1 × 10⁻
27⁴) and applied gIVW (36 correlated SNPs (r^2 < 0.4) that are adequately associated with the plasma proteins ($p < 1 \times 10^{-1}$) 27
28 ⁴) and applied gIVW (36). Out of the 53 associations of lipid fractions with plasma proteins,
gIVW produced consistent estimates with FDR corrected *P* value < 0.05 for 43 associations.
Subsequently, gEgger and Weighted 28 gIVW produced consistent estimates with FDR corrected *P* value < 0.05 for 43 associations.
29 Subsequently, gEgger and Weighted median were applied to the 42 associations with≥ 3 SNP instruments. Weighted median produ 29 Subsequently, gEgger and Weighted median were applied to the 42 associations with \geq 3 SNP instruments. Weighted median produced highly consistent estimates for all 42 associations.
21 The gEgger detected no horizont 30 instruments. Weighted median produced highly consistent estimates for all 42 associations.
31 The gEgger detected no horizontal pleiotropy (FDR corrected P for intercept < 0.05) but 31 The gEgger detected no horizontal pleiotropy (FDR corrected *P* for intercept < 0.05) but

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

- .
.
.
- $\frac{1}{2}$
- 1 derived inconsistent estimates for the association of TNFSF13 with TC and nonHDL-C,
2 THOP1 with HDL-C, and EPPK1 with TC (**Figure 2, Table S4**).
4 Altogether, 30 associations of 14 genetically proxied proteins with lipi THOP1 with HDL-C, and EPPK1 with TC (**Figure 2, Table S4**).

Altogether, 30 associations of 14 genetically proxied proteins

identified by proteome-wide MR, further validated by coloc 3 Altogether, 30 associations of 14 genetically proxied proteins with lipid fractions were
4 identified by proteome-wide MR, further validated by colocalization (either strong or
5 suggestive) and the subsequent MR analysi 4 identified by proteome-wide MR, further validated by colocalization (either strong or suggestive) and the subsequent MR analysis generalized to correlated instruments (GMR).
The top findings include CELSR2 associated wit 5 suggestive) and the subsequent MR analysis generalized to correlated instruments (GMR).

The top findings include CELSR2 associated with all lipid fractions except TG, PCSK9 and

HGFAC with TC, LDL-C, and non-HDL-C, LPA
- 6 The top findings include CELSR2 associated with all lipid fractions except TG, PCSK9 and HGFAC with TC, LDL-C, and non-HDL-C, LPA with LDL-C and TC, and ANGPTL3 with all lipid fractions (**Figure 2**).
-

 $\begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array}$ 10 Figure 2. Volcano plots showing the causal effect of circulating plasma protein on the 5 lipid traits A)
11 LDL-C; B) HDL-C; C) TC; D) TG; E) non-HDL-C. Each dot indicates a plasma protein with the x-
axis showing the W 11 LDL-C; B) HDL-C; C) TC; D) TG; E) non-HDL-C. Each dot indicates a plasma protein with the x-
axis showing the Wald ratio or IVW while the y-axis showing -log₁₀ FDR-corrected P from the MR
analysis. F) Circular plot sh 21 axis showing the Wald ratio or IVW while the y-axis showing -log₁₀ FDR-corrected P from the MR
13 analysis. F) Circular plot showing the overlap of plasma proteins with the 5 lipid traits tested.
14 **Potential causal** analysis. F) Circular plot showing the overlap of plasma proteins with the 5 lipid traits tested.
Potential causal effects of lipid-associated proteins on CVD outcomes

$\frac{15}{16}$

16 Subsequently, we investigated whether lipid-associated proteins (N=29) identified by
17 proteome-wide MR have potential causal effects on risk of CAD and stroke in SAS. After 17 proteome-wide MR have potential causal effects on risk of CAD and stroke in SAS. After
18 correcting for multiple testing, only genetically predicted CELSR2 had a causal association
19 with CAD (Odds ratio (OR) = 0.64 18 correcting for multiple testing, only genetically predicted CELSR2 had a causal association
19 with CAD (Odds ratio (OR) = 0.64, 95% Confidence Interval (CI) = (0.50, 0.81), FDR =
20 0.003; **Figure 3A, Table S5**) which 18 correcting for multiple testing, only genetically predicted CELSR2 had a causal association
19 with CAD (Odds ratio (OR) = 0.64, 95% Confidence Interval (CI) = (0.50, 0.81), FDR =
10.003; **Figure 3A, Table S5**) which w 19 with CAD (Odds ratio (OR) = 0.64, 95% Confidence Interval (CI) = (0.50, 0.81), FDR = 0.003; **Figure 3A, Table S5**) which was also supported by strong colocalization evidence (PPH4 = 93.6%, **Figure 3B, Table S6**). Addit 20 0.003; **Figure 3A, Table S5**) which was also supported by strong colocalization evidence
21 (PPH4 = 93.6%, **Figure 3B, Table S6**). Additionally, gIVW, gEgger, and weighted median 21 (PPH4 = 93.6%, **Figure 3B**, **Table S6**). Additionally, gIVW, gEgger, and weighted median

- $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$
- $\frac{1}{1}$ produced consistent estimates and no pleiotropy was detected (**Figure 3C, Table S7**).
2 Notably, ANGPTL3 and LPA showed suggestive associations with CVDs (P< 0.05) but did not pass the 5% FDR threshold (**Figure 3A**).
-

5
5
6
7 **Figure 3**. Effect of lipid-associated plasma proteins on CVDs A) Forest plot showing all plasma
for proteins that were nominally associated with CVDs (P< 0.05). B) Stacked genomic plot showing
evidence for colocalization evidence for colocalization between CELSR2 and CAD; and C) Forest plot showing the effect of 27 evidence for colocalization between CELSR2 and CAD; and C) Forest plot showing the effect of

28 CELSR2 on CAD estimated by gIVW, gEgger, and weighted median.

29 Since the efficacy of lipid-associated proteins may be p

8 CELSR2 on CAD estimated by gIVW, gEgger, and weighted median.
9
0 Since the efficacy of lipid-associated proteins may be pr -
0
1 10 Since the efficacy of lipid-associated proteins may be prioritized depending on the
11 association of lipid fractions with CVDs, we assessed the causal effects of five lipid fractions
12 on CVDs to indicate proteins mor 11 association of lipid fractions with CVDs, we assessed the causal effects of five lipid fractions
12 on CVDs to indicate proteins more promising to CVD treatment. Our univariable MR
13 identified four associations reach 12 on CVDs to indicate proteins more promising to CVD treatment. Our univariable MR identified four associations reaching nominal significance (P < 0.05), including LDL-C with CAD (β_{IVW} = 1.64; 95% CI = (1.03, 2.62) an 13 identified four associations reaching nominal significance (P < 0.05), including LDL-C with

14 CAD ($\beta_{\text{IVW}} = 1.64$; 95% CI = (1.03, 2.62) and cardioembolic stroke ($\beta_{\text{IVW}} = 1.78$; 95% CI = (1.01, 3.15); **Table S** CAD (β_{IVW} = 1.64; 95% CI = (1.03, 2.62) and cardioembolic stroke (β_{IVW} = 1.78; 95% CI = (1.01, 3.15); **Table S8**). The MR-Egger and weighted median methods produced estimates consistent in direction with the inverse 15 (1.01, 3.15); **Table S8**). The MR-Egger and weighted median methods produced estimates
16 consistent in direction with the inverse variance weighted method (P< 0.05).
17 Since CELSR2 showed potential causal effects on b

16 consistent in direction with the inverse variance weighted method (P< 0.05).
17 Since CELSR2 showed potential causal effects on both LDL-C and CAD v
18 proxied LDL-C was also causally associated with CAD risk, a multi-t 17 Since CELSR2 showed potential causal effects on both LDL-C and CAD while genetically
18 proxied LDL-C was also causally associated with CAD risk, a multi-trait colocalization was
19 performed on the 3 traits in the geno 18 proxied LDL-C was also causally associated with CAD risk, a multi-trait colocalization was
19 performed on the 3 traits in the genomic region +/- 500KB extended from the CELSR2 gene.
10 The multi-trait colocalization pr 19 performed on the 3 traits in the genomic region +/- 500KB extended from the CELSR2 gene.

20 The multi-trait colocalization produced a posterior probability of 70.0% that CELSR2, LDL-

21 C, and CAD colocalized in this 20 The multi-trait colocalization produced a posterior probability of 70.0% that CELSR2, LDL-
21 C, and CAD colocalized in this region (**Figure 4, Table S9**). The posterior probabilities for
22 all 15 scenarios of multi-tr 21 C, and CAD colocalized in this region (**Figure 4, Table S9**). The posterior probabilities for all 15 scenarios of multi-trait colocalization were presented in **Table S9**. 22 all 15 scenarios of multi-trait colocalization were presented in **Table S9**.

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

 $\mathbf{1}$

 $\frac{1}{2}$ **Figure 4**. Stacked regional genomic plot from multi-trait colocalization showing the colocalized genetic variant rs660240 across LDL-C, CAD at the CELSR2 locus in SAS.

5 Reverse MR with lipid fractions as exposures and plasma proteins as outcomes 5
6

5 To understand the plasma proteins modified by lipid fractions, a reverse MR was performed
7 using lipid fractions as exposures and plasma proteins as outcomes. This analysis identified 6 Sepectically proxied associations between TG and LDLR, and HDL-C with 3 proteins:
9 APOA1, MENT, and FGFBP2 ($\beta_{\text{range}} = 0.55$ to 0.61, FDR-corrected P< 0.05, Table S10, This analysis identified

1772 using lipid fractions as exposures and plasma proteins as outcomes. This analysis identified

1783 genetically proxied associations between TG and LDLR, and HDL-C with 3 proteins:

1893 APOA 8 genetically proxied associations between TG and LDLR, and HDL-C with 3 proteins:
9 APOA1, MENT, and FGFBP2 ($\beta_{\text{range}} = 0.55$ to 0.61, FDR-corrected P< 0.05, **Table S10,**
6 **Figure S4**). Subsequent MR-Egger and weighted 9 APOA1, MENT, and FGFBP2 ($β_{range} = 0.55$ to 0.61, FDR-corrected P< 0.05, **Table S10,**
 Figure S4). Subsequent MR-Egger and weighted median produced consistent estimates and detected no horizontal pleiotropy (**Table S11,** 10 **Figure S4**). Subsequent MR-Egger and weighted median produced consistent estimates and detected no horizontal pleiotropy (**Table S11, Figure S5**).
12 11 detected no horizontal pleiotropy (**Table S11, Figure S5**).

12

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

¹**Comparison with EUR population**

2 Among the 53 proteome-wide MR identified associations in SAS, 29 of them had consistent
3 beta estimates in EUR (FDR-corrected $P < 0.05$). The correlation analysis on the 30 beta 3 beta estimates in EUR (FDR-corrected $P < 0.05$). The correlation analysis on the 30 beta
4 estimates in SAS against those in EUR produced a correlation coefficient of 0.85 ($P = 6.5$) 4 estimates in SAS against those in EUR produced a correlation coefficient of 0.85 ($P = 6.5$
5 $\times 10^{-9}$. Figure S6). Out of the 30 associations in SAS with MR, colocalization, and GMR $\times 10^{-9}$, **Figure S6).** Out of the 30 associations in SAS with MR, colocalization, and GMR evidence. 22 were verified in EUR population (**Table S12** and **Figure 5**). Proteome-wide MR. ⁶evidence, 22 were verified in EUR population (**Table S12** and **Figure 5**). Proteome-wide MR, 7 colocalization, and GMR identified 6 proteins associated with HDL-C in SAS, but only
8 ANGPTL3 and APOE were also significant in EUR (Figure 5). Of the 6 proteins linked to 8 ANGPTL3 and APOE were also significant in EUR (**Figure 5**). Of the 6 proteins linked to
9 LDL-C in SAS, ANGPTL3, CELSR2, LPA, and PCSK9 were verified in EUR, with PCSK9 ⁹LDL-C in SAS, ANGPTL3, CELSR2, LPA, and PCSK9 were verified in EUR, with PCSK9 10 showing a stronger effect in EUR ($\beta_{\text{EUR}} = 0.37$; 95% CI = (0.36, 0.38), $\beta_{\text{SAS}} = 0.16$; 95% CI = 11 (0.11, 0.21); **Figure 5**). Among the remaining 18 associations with nonHDL-C, TC, or TG, 11 (0.11, 0.21); **Figure 5**). Among the remaining 18 associations with nonHDL-C, TC, or TG,
12 only the association of EPPK1 with nonHDL-C and HGFAC with TC were not significant in 12 only the association of EPPK1 with nonHDL-C and HGFAC with TC were not significant in
13 EUR. EUR.

15 Figure 5. Effect of plasma proteins on lipid fractions using cis-MR in SAS and EUR on A)
16 HDL-C: B) LDL-C: and C) Scatter plot for comparison of causal effect estimates from MR

16 HDL-C; B) LDL-C; and C) Scatter plot for comparison of causal effect estimates from MR
17 between EUR and SAS for LDL-C.

18 ¹⁹**Observational associations of protein levels with lipid fractions**

¹⁷ between EUR and SAS for LDL-C.
18

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

 $\ddot{}$ $\frac{1}{2}$ 1 There were 3 proteins with strong evidence (MR, strong colocalization, and GMR) of
association HDL-C (APOE, CELSR2, and PLA2G15) and 6 proteins with LDL-C
(ANGPTL3, CELSR2, EPPK1, HGFAC, LPA, and PCSK9) that were tested association HDL-C (APOE, CELSR2, and PLA2G15) and 6 proteins with LDL-C
(ANGPTL3, CELSR2, EPPK1, HGFAC, LPA, and PCSK9) that were tested in observational
analysis. After adjusting for multiple testing, all 9 proteins had a 3 (ANGPTL3, CELSR2, EPPK1, HGFAC, LPA, and PCSK9) that were tested in observational
4 analysis. After adjusting for multiple testing, all 9 proteins had a significant association with
5 their genetically associated lipid 4 5 beir genetically associated lipid fraction (FDR adjusted $P < 0.05$) but only 4 had a consistent 4 direction of effect with MR estimates (Table S13); APOE ($\beta = 0.078$; 95% CI = (0.075, their genetically associated lipid fraction (FDR adjusted *P* < 0.05) but only 4 had a consistent
direction of effect with MR estimates (**Table S13**); APOE (β = 0.078; 95% CI = (0.075,
0.082)) and PLA2G15 (β = -0.06 direction of effect with MR estimates (**Table S13**); APOE (β = 0.078; 95% CI = (0.075,
0.082)) and PLA2G15 (β = -0.067; 95% CI = (-0.071, -0.063)) with HDL, and ANGPTL3 (β
= 0.180; 95% CI = (0.175, 0.184)) and PC 7 0.082)) and PLA2G15 (β = -0.067; 95% CI = (-0.071, -0.063)) with HDL, and ANGPTL3 (β = 0.180; 95% CI = (0.175, 0.184)) and PCSK9 (β = 0.196; 95% CI = (0.192, 0.200)) with LDL. Out of these four proteins, there was a si 8 = 0.180; 95% CI = (0.175, 0.184)) and PCSK9 (β = 0.196; 95% CI = (0.192, 0.200)) with

19 LDL. Out of these four proteins, there was a significant interaction between ANGPTL3 and

10 PCSK9 levels with LDL-Cin South Asi 9 LDL. Out of these four proteins, there was a significant interaction between ANGPTL3 and
PCSK9 levels with LDL-Cin South Asians with South Asians having a significantly lower
effect compared to other ancestries (ANGPLT3 10 PCSK9 levels with LDL-Cin South Asians with South Asians having a significantly lower
11 effect compared to other ancestries (ANGPLT3*South Asian; $\beta = -0.072$; 95% CI = (-0.103, -
12 0.040), PCSK9*South Asian; $\beta = -0.1$ 11 effect compared to other ancestries (ANGPLT3*South Asian; β = -0.072; 95% CI = (-0.103, -
0.040), PCSK9*South Asian; β = -0.140; 95% CI = (-0.174, -0.107)).

12 0.040), PCSK9*South Asian; β = -0.140; 95% CI = (-0.174, -0.107)).

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

$\begin{array}{c} \n\end{array}$

1 **Discussion**
2 **Key findings in this study**
3 Here, we performed a bidi

1 Here, we performed a bidirectional proteome-wide MR on five lipid fractions, and linked
1 lipid-related proteins to cardiovascular outcomes in SAS. Our study confirmed key proteins 3 Here, Here, Western and Same in Same and the proteins (PCSK9, ANGPTL3, LPA), identified novel targets (GSTA1, GSTA3, EPPK1, PECR, and PLA2G12), and strengthened evidence for CELSR2 and GAS6 in dyslipidaemia. Notably, 4 (PCSK9, ANGPTL3, LPA), identified novel targets (GSTA1, GSTA3, EPPK1, PECR, and
6 PLA2G12), and strengthened evidence for CELSR2 and GAS6 in dyslipidaemia. Notably,
7 our results highlight significant heterogeneity in MR 5 PLA2G12), and strengthened evidence for CELSR2 and GAS6 in dyslipidaemia. Notably,
7 our results highlight significant heterogeneity in MR estimates across genetic ancestry
8 groups, particularly for the effect of PCSK9 FLA2G12), and strengthened evidence for CELSR2 and GAS6 in dyslipidaemia. Notably,
7 our results highlight significant heterogeneity in MR estimates across genetic ancestry
8 groups, particularly for the effect of PCSK9 on 7 our results highlight significant heterogeneity in MR estimates across genetic ancestry
8 groups, particularly for the effect of PCSK9 on LDL-C. We also report CELSR2 with
9 evidence for its effect on LDL-C and CAD risk. 8 groups, particularly for the effect of PCSK9 on LDL-C. We also report CELSR2 with
9 evidence for its effect on LDL-C and CAD risk. Reverse MR identified LDLR as modifiable
0 by TG, and APOA1, MENT, and FGFBP2 by HDL-C. 9 evidence for its effect on LDL-C and CAD risk. Reverse MR identified LDLR as modifiable
0 by TG, and APOA1, MENT, and FGFBP2 by HDL-C.
1 **Enhanced role of CELSR2 and GAS6 in lipid metabolism and cardiovascular outcomes**

10 by TG, and APOA1, MENT, and FGFBP2 by HDL-C.
11 **Enhanced role of CELSR2 and GAS6 in lipid metab**
12 We found an inverse association of genetically proxiec

11 Enhanced role of CELSR2 and GAS6 in lipid metabolism and cardiovascular outcomes
12 We found an inverse association of genetically proxied CELSR2 with LDL-C and CAD risk
13 in SAS. CELSR2 is a transmembrane protein be 12 We found an inverse association of genetically proxied CELSR2 with LDL-C and CAD risk
13 in SAS. CELSR2 is a transmembrane protein belonging to the flamingo family of cadherin
14 superfamily (40). Although the biologica in SAS. CELSR2 is a transmembrane protein belonging to the flamingo family of cadherin
14 superfamily (40). Although the biological function of CELSR2 is not well understood, the
15 role of CELSR2 in lipid metabolism was i 14 superfamily (40). Although the biological function of CELSR2 is not well understood, the role of CELSR2 in lipid metabolism was indicated by some previous studies. A locus in the vicinity of CELSR2, rs599839 (in LD with 15 role of CELSR2 in lipid metabolism was indicated by some previous studies. A locus in the vicinity of CELSR2, rs599839 (in LD with rs660240, the instrument of CELSR2 in this study $(r^2 = 0.989 \text{ in SAS}, r^2 = 0.871 \text{ in EUR})$, wa 16 vicinity of CELSR2, rs599839 (in LD with rs660240, the instrument of CELSR2 in this study
17 $(r^2 = 0.989$ in SAS, $r^2 = 0.871$ in EUR)), was first reportedly associated with CAD, LDL-C
18 and TC by 2 European ancestral $(r^2 = 0.989 \text{ in SAS, } r^2)$ $(r^2 = 0.989$ in SAS, $r^2 = 0.871$ in EUR)), was first reportedly associated with CAD, LDL-C
and TC by 2 European ancestral GWAS (41-43). rs660240, a 3' UTR variant, is an eQTL for
CELSR2, PSRC1, and *SORT1* in liver tis 18 and TC by 2 European ancestral GWAS (41-43). rs660240, a 3' UTR variant, is an eQTL for
19 CELSR2, PSRC1, and SORT1 in liver tissue (Open Target Genetics). It shows slight
20 variations in allele frequencies between Sou 19 *CELSR2*, *PSRC1*, and *SORT1* in liver tissue (Open Target Genetics). It shows slight variations in allele frequencies between South Asians, Europeans, and East Asians, which could have implications for studies related 20 variations in allele frequencies between South Asians, Europeans, and East Asians, which
21 could have implications for studies related to disease susceptibility and treatment response.
22 Furthermore, a transcriptomic 21 could have implications for studies related to disease susceptibility and treatment response.
22 Furthermore, a transcriptomic study revealed the risk allele of rs599839 to CAD and high
23 LDL-C also suppressed the expr 22 Furthermore, a transcriptomic study revealed the risk allele of rs599839 to CAD and high
23 LDL-C also suppressed the expression of *CELSR2* gene in liver (44). Extending to non-
24 European populations, the association 23 LDL-C also suppressed the expression of *CELSR2* gene in liver (44). Extending to non-
24 European populations, the association of *CELSR2* variants with lipid fractions and CAD risk
25 was also verified in the South As 24 European populations, the association of *CELSR2* variants with lipid fractions and CAD risk
25 was also verified in the South Asian population (45, 46). However, although the effect of
26 CELRS2 on lipid metabolism was 25 was also verified in the South Asian population (45, 46). However, although the effect of CELRS2 on lipid metabolism was indicated by genetic and transcriptomic studies, the underlying mechanism is less clear. One study 26 CELRS2 on lipid metabolism was indicated by genetic and transcriptomic studies, the underlying mechanism is less clear. One study demonstrated that CELSR2 deficiency can elevate reactive oxygen species of hepatocytes, w 27 underlying mechanism is less clear. One study demonstrated that CELSR2 deficiency can
28 elevate reactive oxygen species of hepatocytes, which impairs lipid homeostasis and
29 physiological unfolded protein response(47) 28 elevate reactive oxygen species of hepatocytes, which impairs lipid homeostasis and
29 physiological unfolded protein response(47). In conclusion, our result is consistent with the
30 significant role of CELSR2 in CAD a 29 physiological unfolded protein response(47). In conclusion, our result is consistent with the significant role of CELSR2 in CAD and lipid metabolism suggested in earlier studies. Our study also extends the generalizabil 30 significant role of CELSR2 in CAD and lipid metabolism suggested in earlier studies. Our study also extends the generalizability of CELSR2's function to the South Asian population. 31 study also extends the generalizability of CELSR2's function to the South Asian population.

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

 $\frac{1}{2}$ $\frac{1}{2}$ 1 We identified an inverse association of genetically proxied GAS6 with TC and TG. GAS6 is
2 a ligand for TAM receptor protein tyrosine kinases including AXL, TYRO3 and MER. The
3 GAS6 - TAM pathway was found implicated in 2 a ligand for TAM receptor protein tyrosine kinases including AXL, TYRO3 and MER. The
3 GAS6 - TAM pathway was found implicated in carcinoma, inflammation, and haemostasis
4 and has been targeted for the treatment of carc 3 GAS6 - TAM pathway was found implicated in carcinoma, inflammation, and haemostasis
4 and has been targeted for the treatment of carcinoma (48-50). In recent years, there is also
5 evidence that GAS6 plays a role in regu and has been targeted for the treatment of carcinoma (48-50). In recent years, there is also
so evidence that GAS6 plays a role in regulating obesity and lipid metabolism (51, 52) as
plasma gamma-glutamyl carboxylated GAS6 5 evidence that GAS6 plays a role in regulating obesity and lipid metabolism (51, 52) as
6 plasma gamma-glutamyl carboxylated GAS6 (Gla-GAS6) was found significantly lower in
7 hyperlipidaemic individuals compared with hea 6 plasma gamma-glutamyl carboxylated GAS6 (Gla-GAS6) was found significantly lower in
hyperlipidaemic individuals compared with healthy controls (51). The subsequent experiment
showed that higher Gla-GAS6 expression induce 7 hyperlipidaemic individuals compared with healthy controls (51). The subsequent experiment
8 showed that higher Gla-GAS6 expression induced by vitamin K in plasma and hepatocyte
9 could reduce the plasma lipid level in h 8 showed that higher Gla-GAS6 expression induced by vitamin K in plasma and hepatocyte
could reduce the plasma lipid level in hyperlipidaemic mice (51). The Gla-GAS6 takes effect
by regulating the AMPK/SREBP1/PPARα signal 9 could reduce the plasma lipid level in hyperlipidaemic mice (51). The Gla-GAS6 takes effect
0 by regulating the AMPK/SREBP1/PPARα signalling pathways of hepatic lipid metabolism
1 (51). Our study supports this hypothesi 10 by regulating the AMPK/SREBP1/PPARα signalling pathways of hepatic lipid metabolism (51). Our study supports this hypothesis by providing genetic evidence for GAS6 as a potential regulator of lipid metabolism. 11 (51). Our study supports this hypothesis by providing genetic evidence for GAS6 as a potential regulator of lipid metabolism.
13 **Known targets and novel findings in lipid metabolism**

12 potential regulator of lipid metabolism.
13 **Known targets and novel findings in lipid metabolism**
14 In this study, we replicated previously established protein

14 In this study, we replicated previously established protein associations with lipid metabolism,
15 including ANGPTL3, APOE, LPA, PCSK9, and PLTP, some of which are drug targets for 14 In this study, we replicated previously established protein associations with lipid metabolism,
15 including ANGPTL3, APOE, LPA, PCSK9, and PLTP, some of which are drug targets for
16 dyslipidemia treatment (53). Howeve including ANGPTL3, APOE, LPA, PCSK9, and PLTP, some of which are drug targets for
16 dyslipidemia treatment (53). However, despite the concordance in direction and significance
17 for the association of PCSK9 on LDL-C in b 16 dyslipidemia treatment (53). However, despite the concordance in direction and significance
17 for the association of PCSK9 on LDL-C in both SAS and EUR, we identified a significantly
18 reduced effect in SAS as support 17 for the association of PCSK9 on LDL-C in both SAS and EUR, we identified a significantly
18 reduced effect in SAS as supported by both MR and observational analysis. Since drug
19 responses to lipid-lowering therapy can 18 reduced effect in SAS as supported by both MR and observational analysis. Since drug
19 responses to lipid-lowering therapy can vary across ethnicities, this finding may have
20 important clinical implications. Previous 19 responses to lipid-lowering therapy can vary across ethnicities, this finding may have

20 important clinical implications. Previous studies suggest that atorvastatin and simvastatin

21 have similar lipid-lowering effe 20 important clinical implications. Previous studies suggest that atorvastatin and simvastatin
21 have similar lipid-lowering effects in SAS patients compared to those in EUR (54). Therefore,
22 PCSK9, as a novel target fo 21 have similar lipid-lowering effects in SAS patients compared to those in EUR (54). Therefore,
22 PCSK9, as a novel target for lipid-lowering medication, warrants further investigation for
23 their effect in more diverse PCSK9, as a novel target for lipid-lowering medication, warrants further investigation for
their effect in more diverse ethnicity. Mechanisms accounting for the ancestral heterogeneity
in PCSK9's effect are limited. Howeve 23 their effect in more diverse ethnicity. Mechanisms accounting for the ancestral heterogeneity
24 in PCSK9's effect are limited. However, a recent study sequencing PCSK9 gene in Indians
25 indicated difference in prevale 24 in PCSK9's effect are limited. However, a recent study sequencing PCSK9 gene in Indians
25 indicated difference in prevalence of mutation in SAS. The by-ancestry heterogeneous
26 mutation pattern can result in heterogen 25 indicated difference in prevalence of mutation in SAS. The by-ancestry heterogeneous
26 mutation pattern can result in heterogeneity of PCSK9 structure and activity, which may
27 modify the effect of PCSK9 abundance (55 26 mutation pattern can result in heterogeneity of PCSK9 structure and activity, which may
27 modify the effect of PCSK9 abundance (55).
28 Additionally, our study identified novel associations, including HGFAC, GIPC3, EPP

27 modify the effect of PCSK9 abundance (55).
28 Additionally, our study identified novel associations, including HGFAC, GIPC3, EPPK1,
29 GSTA1 GSTA3 PECR and PLA2G12 HGFAC activates henatocyte growth factor (HGF) 29 GSTA1, GSTA3, PECR, and PLA2G12. HGFAC activates hepatocyte growth factor (HGF)
20 by converting it to a heterodimer, which then binds to the MET receptor to activate
31 downstream signalling. A previous study linked a 29 GSTA1, GSTA3, PECR, and PLA2G12. HGFAC activates hepatocyte growth factor (HGF)
20 by converting it to a heterodimer, which then binds to the MET receptor to activate
21 downstream signalling. A previous study linked a 30 by converting it to a heterodimer, which then binds to the MET receptor to activate
31 downstream signalling. A previous study linked a putative HGFAC loss-of-function variant
32 to elevated serum TG and LDL-C (56). In 32 to elevated serum TG and LDL-C (56). In another animal experiment, an increase of-
33 circulating TG was present in both male and female HGFAC-KO mice while higher level of-32 to elevated serum TG and LDL-C (56). In another animal experiment, an increase of circulating TG was present in both male and female HGFAC-KO mice while higher level of 17 33 circulating TG was present in both male and female HGFAC-KO mice while higher level of 17

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

 $\frac{1}{\sqrt{2}}$ 1 circulating TC was present in male HGFAC-KO mice (57). The association of GIPC3 with

2 HDL-C was firstly identified in our study. However, a previous study reported that GIPC1,

3 another subtype of GIPC family, is invo 2 HDL-C was firstly identified in our study. However, a previous study reported that GIPC1,
3 another subtype of GIPC family, is involved in lipid metabolism by regulating the SR-B1
4 expression (58). Lastly, we also ident

3 another subtype of GIPC family, is involved in lipid metabolism by regulating the SR-B1 expression (58). Lastly, we also identified GSTA1, GSTA3, EPPK1, PECR, and PLA2G12 associated with lipid traits, all of which are no expression (58). Lastly, we also identified GSTA1, GSTA3, EPPK1, PECR, and PLA2G12
s associated with lipid traits, all of which are novel findings.
Reverse effect of lipid fractions on plasma proteins

5 associated with lipid traits, all of which are novel findings.
 6 Reverse effect of lipid fractions on plasma proteins
 7 Converselv, we applied MR to identify plasma proteins mo

6 **Reversely, we applied MR** to identify plasma proteins modified by lipid fractions. We found
8 **Reverse that genetically proxied TG** levels were associated with increased plasma LDLR. LDLR is a 7 Conversely, we applied MR to identify plasma proteins modified by lipid fractions. We found
that genetically proxied TG levels were associated with increased plasma LDLR. LDLR is a
cell membrane glycoprotein that regulat 8 that genetically proxied TG levels were associated with increased plasma LDLR. LDLR is a
9 cell membrane glycoprotein that regulates lipid homeostasis by binding and internalizing
0 circulating cholesterol-containing lip 9 cell membrane glycoprotein that regulates lipid homeostasis by binding and internalizing
0 circulating cholesterol-containing lipoprotein particles, including LDL-C, VLDL-C, and
1 chylomicron remnants (59). Deficiency in 10 circulating cholesterol-containing lipoprotein particles, including LDL-C, VLDL-C, and
11 chylomicron remnants (59). Deficiency in LDLR can result in dyslipidaemia. Therefore, we
12 interpreted that the LDLR upregulatio 11 chylomicron remnants (59). Deficiency in LDLR can result in dyslipidaemia. Therefore, we
12 interpreted that the LDLR upregulation triggered by genetically proxied TG is likely to
13 reverse the hyperlipidaemia. In addi 12 interpreted that the LDLR upregulation triggered by genetically proxied TG is likely to
13 reverse the hyperlipidaemia. In addition, we observed genetically proxied HDL-C level
14 associated with increased plasma APOA1, associated with increased plasma APOA1, MENT, and FGFBP2 level. All 3 proteins play
important roles in various disorders and mechanisms, including cholesterol transport, 14 associated with increased plasma APOA1, MENT, and FGFBP2 level. All 3 proteins play
15 important roles in various disorders and mechanisms, including cholesterol transport,
16 angiogenesis, tissue repair, and cellular m 15 important roles in various disorders and mechanisms, including cholesterol transport,
16 angiogenesis, tissue repair, and cellular metabolism. Therefore, further investigation is
17 necessary for detailed biological int 16 angiogenesis, tissue repair, and cellular metabolism. Therefore, further investigation is
17 necessary for detailed biological interpretation of these findings.
18 **Strength and Limitations** 17 necessary for detailed biological interpretation of these findings.

18 **Strength and Limitations**

19 Our study has several strengths. First, to the best of our knowledge, this is the first MR study

18 **Strength and Limitations**
19 Our study has several strengths. First, to the best of our knowledge, this is the first MR study
19 to systematically evaluate the potential causal association between plasma proteins and l 19 Our study has several strengths. First, to the best of our knowledge, this is the first MR study
10 to systematically evaluate the potential causal association between plasma proteins and lipid
11 traits in South Asians 20 to systematically evaluate the potential causal association between plasma proteins and lipid
21 traits in South Asians, replicating known protein-stroke associations and discovering novel
22 targets. Secondly, we combi 21 traits in South Asians, replicating known protein-stroke associations and discovering novel
22 targets. Secondly, we combined MR with colocalization, to reduce bias from LD and reverse
23 causation — a potential limitat 22 targets. Secondly, we combined MR with colocalization, to reduce bias from LD and reverse
23 causation — a potential limitation of conventional observational studies. Finally, by
24 incorporating GWAS on CVDs, we linked 23 causation — a potential limitation of conventional observational studies. Finally, by
24 incorporating GWAS on CVDs, we linked lipid associated proteins to cardiovascular
25 outcomes and identified CELSR2 as a promising 24 incorporating GWAS on CVDs, we linked lipid associated proteins to cardiovascular
25 outcomes and identified CELSR2 as a promising target for both LDL-C and CAD.
26 There are also limitations in our study. First, the ge

25 outcomes and identified CELSR2 as a promising target for both LDL-C and CAD.
26 There are also limitations in our study. First, the genetic associations for plasma proteins in
27 SAS were based on a relatively small sam 26 There are also limitations in our study. First, the genetic associations for plasma proteins in
27 SAS were based on a relatively small sample size, compared with current GWAS standards,
28 so our findings should be int 27 SAS were based on a relatively small sample size, compared with current GWAS standards,
28 so our findings should be interpreted cautiously, and larger GWAS are needed for validation.
29 Second, despite using various MR 28 so our findings should be interpreted cautiously, and larger GWAS are needed for validation.
29 Second, despite using various MR methods and sensitivity analyses, unaccounted pleiotropy
20 may still bias our results. Th 29 Second, despite using various MR methods and sensitivity analyses, unaccounted pleiotropy
20 may still bias our results. Third, our analysis assumes the absence of SNP-SNP and 30 may still bias our results. Third, our analysis assumes the absence of SNP-SNP and

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

- 1 SNP \Box environment interaction. Lastly, our findings may not be generalizable to ancestry
2 groups other than the ones studied here.
3 **Conclusion**
- $\sum_{k=1}^{n}$ 2 groups other than the ones studied here.
3 **Conclusion**
4 Our comprehensive study triangulated e

- 4 Our comprehensive study triangulated evidence from MR, colocalization, and observational
3 analyses, highlighting several novel proteins associated with lipid fractions in South Asians.
- 2004 Our comprehensive study triangulated evidence from MR, colocalization, and observational
5 analyses, highlighting several novel proteins associated with lipid fractions in South Asians.
5 Notably, our analysis suggest
-
- 5 analyses, highlighting several novel proteins associated with lipid fractions in South Asians.
6 Notably, our analysis suggests that the causal effect of PCSK9 on LDL-C may be ancestry-
5 specific. Future studies with la 6 Notably, our analysis suggests that the causal effect of PCSK9 on LDL-C may be ancestry-
specific. Future studies with larger sample sizes are needed to validate our findings, along
with further mechanistic and clinical
- 17 specific. Future studies with larger sample sizes are needed to validate our findings, along
18 with further mechanistic and clinical studies to confirm the role of PCSK9 on LDL-C in
19 South Asians and Europeans. 8 with further mechanistic and clinical studies to confirm the role of PCSK9 on LDL-C in
9 South Asians and Europeans.
0
- 9 South Asians and Europeans.
10

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

 $\frac{1}{2}$

 $\frac{1}{2}$ 1 **Contributors**: DM, IT and AD contributed to the conception and design of the study. SW
2 and AS contributed to acquisition and statistical analysis of data. DM, IT, AD supervised the
3 project. All authors contributed t

2 and AS contributed to acquisition and statistical analysis of data. DM, IT, AD supervised the
2 project. All authors contributed to the drafting and critical revision of the manuscript.
2 **Acknowledgments**: The authors w

3 project. All authors contributed to the drafting and critical revision of the manuscript.
4 **Acknowledgments**: The authors would like to thank the staff and participan
5 contributed to the LIK Biobank LIK RPDP GLGC the E Acknowledgments: The authors would like to thank the staff and participants who contributed to the UK Biobank, UKBPPP, GLGC, the East London Genes & Health (ELGH) study, the CARDIoGRAMplusC4D consortium, GIGASTROKE consort 5 contributed to the UK Biobank, UKBPPP, GLGC, the East London Genes & Health (ELGH)
6 study, the CARDIoGRAMplusC4D consortium, GIGASTROKE consortium for making the
7 resources available. 6 study, the CARDIoGRAMplusC4D consortium, GIGASTROKE consortium for making the

7 resources available.

8 Funding: This work is supported by the British Heart Equadation Besearch Excellence

7 resources available.
8 **Funding**: This work 8 **Funding**: This work is supported by the British Heart Foundation Research Excellence
9 Award (4) (RE/24/130023). IT and AR are supported by NIH R01 HL162300⁻¹02.

9 Award (4) (RE/24/130023). IT and AR are supported by NIH R01 HL162300 \square 02.
0 **Conflict of interest**: A.R. reported consultancies for Raydel, Johnson & Johnson 10 **Conflict of interest**: A.R. reported consultancies for Raydel, Johnson & Johnson, and JP
11 Morgan, all with modest compensation, and received a significant research grant from CSL
12 Limited but with no salary support 11 Morgan, all with modest compensation, and received a significant research grant from CSL
12 Limited but with no salary support included. The remaining authors have no disclosures to
13 report. 12 Limited but with no salary support included. The remaining authors have no disclosures to
13 report.
14 **Dete Aveilability:** The UKBPP data can be developeded from http://ukb.ppp.gwg.eu

13 report.
 14 Data Availability: The UKBPPP data can be downloaded from http://ukb-ppp.gwas.eu.

15 GLGC dataset can be downloaded from https://cs*o* sph.umich.edu/willer/public/gloc-14 **Data Availability**: The UKBPPP data can be downloaded from http://ukb-ppp.gwas.eu.
15 GLGC dataset can be downloaded from https://csg.sph.umich.edu/willer/public/glgc-
16 lipids2021/results/ancestry_specific/. ELGH dat 15 GLGC dataset can be downloaded from https://csg.sph.umich.edu/willer/public/glgc-
16 lipids2021/results/ancestry_specific/. ELGH data can be downloaded from
17 https://www.genesandhealth.org/research/scientific-data-dow 16 lipids2021/results/ancestry_specific/. ELGH data can be downloaded from
17 https://www.genesandhealth.org/research/scientific-data-downloads/gwas-data-genes-health-
18 feb-2020-datafreeze. GWAS summary statistics for st 17 https://www.genesandhealth.org/research/scientific-data-downloads/gwas-data-genes-health-
18 feb-2020-datafreeze. GWAS summary statistics for stroke and its subtypes for South Asian
19 and Europeans can be obtained from 18 feb-2020-datafreeze. GWAS summary statistics for stroke and its subtypes for South Asians
19 and Europeans can be obtained from the GWAS catalogue with accession GCST90104559-
10 GCST90104563, and GCST90104539-GCST90104 19 and Europeans can be obtained from the GWAS catalogue with accession GCST90104559-

10 GCST90104563, and GCST90104539-GCST90104543, respectively.

21 GCST90104563, and GCST90104539-GCST90104543, respectively.

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

$\frac{1}{2}$

References
1. Reiner Ž. Hypertriglyceridaemia and risk of coronary artery disease. Nature Reviews
3. Cardiology. 2017;14(7):401-11.
2. Yaghi S. Elkind MSV. Lipids and Cerebrovascular Disease. Stroke. 2015;46(11):3322-8 2. Yaghi S, Elkind MSV. Lipids and Cerebrovascular Disease. Stroke. 2015;46(11):3322-8.
5 3. Holmes MV, Millwood IY, Kartsonaki C, Hill MR, Bennett DA, Boxall R, et al. Lipids,
6 Lipoproteins. and Metabolites and Risk of M 3 Cardiology. 2017;14(17): 4 2. Yaghi S, Andrew March March March March March March 2017;44, 2017
1997 - Holmes MV, Millwood IY, Kartsonaki C, Hill MR, Bennett DA, Boxall R, et al. Lipids,
1997 - American College of Cardiology. 2018;71(6):620-32. 5 Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. Journal of the
7 American College of Cardiology. 2018;71(6):620-32.
8 4. Velagaleti RS, Massaro J, Vasan RS, Robins SJ, Kannel WB, Levy D. Relat 6 American College of Cardiology. 2018;71(6):620-32.
6 Lipoprotes A. Strokens, and Metabolites and Metabolites and Metabolites Andres Andres Andres Andres Stroke.
6 Concentrations to Heart Failure Incidence. Circulation. 2 7 Concentrations to Heart Failure Incidence. Circulation. 2009;120(23):2345-51.
10 5. Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF,
11 regional, and national age-sex specific mortality for 264 causes 8 5. Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, et al. Global,
11 regional, and national age-sex specific mortality for 264 causes of death,
12 1980–:2016: a systematic analysis for the Glob 9 S. Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, (
1 regional, and national age-sex specific mortality for 264 causes of death,
2 1980–2016: a systematic analysis for 11 regional, and national age-sex specific mortality for 264 causes of death,
12 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. The
13 Lancet. 2017;390(10100):1151-210. 12 1980–2016: a systematic analysis for the Global Burden of Diseas
13 Lancet. 2017;390(10100):1151-210.
14 6. Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer 13 Lancet. 2017;390(10100):1151-210.
14 6. Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer G, et al.
15 Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. 14 6. Conroy RM, Pyörälä K, Fitzge
15 Estimation of ten-year risk of fatal ca
16 European Heart Journal. 2003;24(11 15 Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE pro
16 European Heart Journal. 2003;24(11):987-1003.
17 T. Lloyd-Jones DM, Braun LT, Ndumele CE, Smith SC, Sperling LS, Virani SS, et 16 European Heart Journal. 2003;24(11):987-1003.
17 1. – Lloyd-Jones DM, Braun LT, Ndumele CE, Smith SC, Sperling LS, Virani SS, et al. Use of
18 Risk Assessment Tools to Guide Decision-Making in the Primary Prevention of
 18 Risk Assessment Tools to Guide Decision-Making in the Primary Prevention of
19 Atherosclerotic Cardiovascular Disease: A Special Report From the American H
20 Association and American College of Cardiology. Journal of t 18 Risk Assessment Tools to Guide Decision-Making in the Primary Prevention of
19 Atherosclerotic Cardiovascular Disease: A Special Report From the American Heart
20 Association and American College of Cardiology. Journal 19 Atherosclerotic Cardiovascular Disease: A Special Report From the American Heart
18 Association and American College of Cardiology. Journal of the American College of
18 Cardiology. 2019;73(24):3153-67.
18 Brewer HB. Hi 20 Association and American College of Cardiology. Journal of the American College of
21 Cardiology. 2019;73(24):3153-67.
22 8. Brewer HB. High-Density Lipoproteins: A New Potential Therapeutic Target 21 Association and American College of Cardiology. 2019;73(24):3153-67.
22 8. Brewer HB. High-Density Lipoproteins: A New Potential Therapeutic Target for Prevention of Cardiovascular Disease. Arteriosclerosis, Thrombosis, 22 8. Brewer HB. High-Density Li
23 Prevention of Cardiovascular Disea
24 2004;24(3):387-91. 22 8. Brevention of Cardiovascular Disease. Arteriosclerosis, Thrombosis, and Vascular Biology.
24 2004;24(3):387-91.
25 9. Drexel H, Tamargo J, Kaski JC, Lewis BS, Saely CH, Fraunberger P, et al. Triglycerides 24 2004;24(3):387-91.
25 9. Drexel H, Tamargo J, Kaski JC, Lewis BS, Saely CH, Fraunberger P, et al. Triglyceride
26 revisited: is hypertriglyceridaemia a necessary therapeutic target in cardiovascular disease 25 9. Drexel H, Tar
26 revisited: is hypertri
27 European Heart Jou 26 10. Previsited: is hypertriglyceridaemia a necessary therapeutic target in cardiovascular disease?
27 European Heart Journal - Cardiovascular Pharmacotherapy. 2023;9(6):570-82.
28 10. Boekholdt SM, Arsenault BJ, Mora S, 27 European Heart Journal - Cardiovascular Pharmacotherapy. 2023;9(6):570-82.
28 10. Boekholdt SM, Arsenault BJ, Mora S, Pedersen TR, LaRosa JC, Nestel PJ, et al.
29 Association of LDL Cholesterol, Non—HDL Cholesterol, and 28 10. Boekholdt SM, Arsenault BJ, Mora S, Pedersen TR, LaRosa JC, Nestel PJ,
29 Association of LDL Cholesterol, Non–HDL Cholesterol, and Apolipoprotein B Le
30 of Cardiovascular Events Among Patients Treated With Statins: 29 Association of LDL Cholesterol, Non-HDL Cholesterol, and Apolipoprotein B Levels W
30 of Cardiovascular Events Among Patients Treated With Statins: A Meta-analysis. JAM
31 2012;307(12):1302-9. 2012;307(12):1302-9. 30 Association of Cardiovascular Events Among Patients Treated With Statins: A Meta-analysis. JAMA.
2012;307(12):1302-9. 32 11. Pischon T, Girman CJ, Sacks FM, Rifai N, Stampfer MJ, Rimm EB. Non–High-D 31 of Cardiovascular Events Among Patients Among Patients Among Patients Among 2012;307(12):1302-9.
32 11. Pischon T, Girman CJ, Sacks FM, Rifai N, Stampfer MJ, Rimm EB. Non–High-Dens
33 Lipoprotein Cholesterol and Apolipo 32 11. Pischon T, Girn
33 Lipoprotein Cholester
34 Men. Circulation. 2005 33 Lipoprotein Cholesterol and Apolipoprotein B in the Prediction of Coronary Heart Disease i
34 Men. Circulation. 2005;112(22):3375-83.
35 12. Frank ATH, Zhao B, Jose PO, Azar KMJ, Fortmann SP, Palaniappan LP. Racial/Ethn 34 Men. Circulation. 2005;112(22):3375-83.
35 12. Frank ATH, Zhao B, Jose PO, Azar KMJ, Fortmann SP, Palaniappan LP. Racial/Ethnic
36 Differences in Dyslipidemia Patterns. Circulation. 2014;129(5):570-9.
37 13. Anand SS, Y 36 Differences in Dyslipidemia Patterns. Circulation. 2014;129(5):570-9.
37 13. Anand SS, Yusuf S, Vuksan V, Devanesen S, Teo KK, Montague
38 in risk factors, atherosclerosis, and cardiovascular disease between e 36 Differences in Dyslipidemia Patterns. Circulation. 2014;129(5):570-9.
37 13. Anand SS, Yusuf S, Vuksan V, Devanesen S, Teo KK, Montague PA, et al. Differences
38 in risk factors, atherosclerosis, and cardiovascular dise 39 the Study of Health Assessment and Risk in Ethnic groups (SHARE). Lancet.
40 2000;356(9226):279-84. 38 in risk factors, atherosclerosis, and cardiovascular disease between ethnic groups in Canada
39 the Study of Health Assessment and Risk in Ethnic groups (SHARE). Lancet.
40 2000;356(9226):279-84. 39 the Study of Health Assessment and Risk in Ethnic groups (SHARE). Lancet.
40 2000;356(9226):279-84.
41 14. Bhopal R, Unwin N, White M, Yallop J, Walker L, Alberti KGMM, et al. Heterogeneity 39 19 the Study of Health Assessment and Richard Study of 2000;356(9226):279-84.
39 the Shopal R, Unwin N, White M, Yallop J, Walker L, Alberti KGMM, et a
32 of coronary heart disease risk factors in Indian, Pakistani, Ban 11 14. Bhopal R, Unwin N, White M, Yallop J, Walker L, Alberti K
42 of coronary heart disease risk factors in Indian, Pakistani, Bangla
43 populations: cross sectional study. BMJ. 1999;319(7204):215-20.
44 15. Wasay M, Kha 42 of coronary heart disease risk factors in Indian, Pakistani, Bangladeshi, and European origin
43 oppulations: cross sectional study. BMJ. 1999;319(7204):215-20.
44 15. Wasay M, Khatri IA, Kaul S. Stroke in South Asian c 43 of populations: cross sectional study. BMJ. 1999;319(7204):215-20.
44 15. Wasay M, Khatri IA, Kaul S. Stroke in South Asian countries. Nat Rev Neurol.
45 2014;10(3):135-43. 44 15. Wasay M, Khatri IA, Kaul S. Stroke in South Asian countrie
2014;10(3):135-43. $45. 2014; 10(3): 135-43.$ $\frac{1}{2}$

It is made available under a [CC-BY-ND 4.0 International license](http://creativecommons.org/licenses/by-nd/4.0/) .

 $\frac{1}{\epsilon}$ $\begin{array}{c} \begin{array}{c} \begin{array}{c} \end{array} \\ \begin{array}{c} \end{array} \end{array} \end{array}$ 1 16. Kanaya AM, Herrington D, Vittinghoff E, Ewing SK, Liu K, Blaha MJ, et al.
2 Understanding the High Prevalence of Diabetes in U.S. South Asians Compared With Four
13 Racial/Ethnic Groups: The MASALA and MESA Studies. $\mathbf{1}$ 16. 2 Racial/Ethnic Groups: The MASALA and MESA Studies. Diabetes Care. 2014;37(6):1621-8.
2 17. Hegele RA, Tsimikas S. Lipid-Lowering Agents. Circulation Research. 2019;124(3):3
2 18. De Oliveira-Gomes D, Joshi PH, Peterson E 3 17. Hegele RA, Tsimikas S. Lipid-Lowering Agents. Circulation Research. 2019;124(3):386-
5 404. De Oliveira-Gomes D, Joshi PH, Peterson ED, Rohatgi A, Khera A, Navar AM.
7 Apolipoprotein B: Bridging the Gap Between Evide 4
5 18.
7 Apoli
8 2024 7 Apolipoprotein B: Bridging the Gap Between Evidence and Clinical Practice. Circulation.
8 2024:150(1):62-79. 2024;150(1):62-79.
7 Apolison B: Bridging the Gap Between European European Evidence and Clinical Practice. Circulation.
7 Associations of Circulating Protein Levels With Lipid Fractions in the General Population. 8 2022
8 19. Figarska SM,
8 Associations of Circu
1 Arteriosclerosis, Thr 19. Figarska SM, Gustafsson S, Sundström J, Ärnlöv J, Mälarstig A, Elmståhl S, et al.
10 Associations of Circulating Protein Levels With Lipid Fractions in the General Populatio 9 11. Associations of Circulating Protein Levels With Lipid Fractions in the General Population
1 1. Arteriosclerosis, Thrombosis, and Vascular Biology. 2018;38(10):2505-18.
2 20. 11 Wu S, Meena D, Yarmolinsky J, Gill D, S 11 Arteriosclerosis, Thrombosis, and Vascular Biology. 2018;38(10):2505-18.
12 20. Wu S, Meena D, Yarmolinsky J, Gill D, Smith A, Dib MJ, et al. Mendelian
13 Randomization and Bayesian Colocalization Analysis Implicate Gly 12 20. Wu S, Meena D, Yarmolinsky J, Gill D, Smith A, Dib MJ, et al. Mend
13 Randomization and Bayesian Colocalization Analysis Implicate Glycoprote
14 Drug Target for Cardioembolic Stroke in South Asian Populations. Journ 13 Randomization and Bayesian Colocalization Analysis Implicate Glycoprotein VI a
14 Drug Target for Cardioembolic Stroke in South Asian Populations. Journal of the
15 Heart Association. 2024;13(16):e035008. 13 Randomization and Bayesian Colocalization Analysis Implicate Glycoprotein VI as a Potenti
14 Drug Target for Cardioembolic Stroke in South Asian Populations. Journal of the American
15 Heart Association. 2024;13(16):e03 15 Drug Target for 2024;13(16):e035008.
16 21. Graham SE, Clarke SL, Wu KH, Kanoni S, Zajac GJM, Ramdas S, et al. The power of genetic diversity in genome-wide association studies of lipids. Nature. 2021;600(7890):675 16 21. Graham SE, Clarke SL, Wu KH, Kan
17 genetic diversity in genome-wide associat
18 9.9. PR SLL LLT LL LL R 17 genetic diversity in genome-wide association studies of lipids. Nature. 2021;600(7890):67
18 9.
19 22. Sun BB, Chiou J, Traylor M, Benner C, Hsu Y-H, Richardson TG, et al. Plasma 18 9.
19 22. Sun BB, Chiou J, Traylor M, Benner C, Hsu Y-H, Richardson TG, et al. Plasma
20 proteomic associations with genetics and health in the UK Biobank. Nature.
21 2023;622(7982):329-38. 19 22
20 pro
21 20 19 20 proteomic associations with genetics and health in the UK Biobank. Nature.
19 2023;622(7982):329-38.
19 23 23. Burgess S, Butterworth A, Thompson SG. Mendelian Randomization Analysis With
19 23 Multiple Genetic Varia 21 protes 2023;622(7982):329-38.
22 23. Burgess S, Butterworth A, Thompson SG. Mendelian Randomization.
23 Multiple Genetic Variants Using Summarized Data. Genetic Epidemiology. 20 22 23. Burgess S, Butter
23 Multiple Genetic Variant:
24 65. 22 23 Multiple Genetic Variants Using Summarized Data. Genetic Epidemiology. 2013;37(7):658.
24 65. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian 24 65.
25 24. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian
26 mandomization: Using genes as instruments for making causal inferences in epidemiology. 25 24.
26 rand
27 Stat 26 25 25 25. Lawlow Daniel Ramsen M, Franch Principle Principle Principle Principle Principle Statistics in Medicine. 2008;27(8):1133-63.
25 25. Kim MS, Song M, Kim B, Shim I, Kim DS, Natarajan P, et al. Prioritization of 27 Statistics in Medicine. 2008;27(8):1133-63.
28 25. Kim MS, Song M, Kim B, Shim I, Kim DS, Natarajan P, et al. Prioritization of
29 therapeutic targets for dyslipidemia using integrative multi-omics and multi-trait analy 28 25. Kim MS, Song M, Kim B, Shim I, Kim
29 therapeutic targets for dyslipidemia using in
30 Cell Reports Medicine. 2023;4(9). 29 therapeutic targets for dyslipidemia using integrative multi-omics and multi-trait a
20 Cell Reports Medicine. 2023;4(9).
31 26. Finer S, Martin HC, Khan A, Hunt KA, MacLaughlin B, Ahmed Z, et al. Cohor 20 thell Reports Medicine. 2023;4(9).
29 theraph MacLaughlin B, Ahman A Shmed Z, et al. Cohort Profile:
22 therapen Samp; Health (ELGH), a community-based population genomics and
23 thealth study in British Bangladeshi and 31 26. Finer S, Martin HC, Khan A,
32 East London Genes & amp; Health
33 health study in British Bangladesh 32 East London Genes & Health (ELGH), a community-based population genomics and
33 health study in British Bangladeshi and British Pakistani people. International Journal of
34 Epidemiology. 2019;49(1):20-1i. 33 Bealth study in British Bangladeshi and British Pakistani people. International Journal of
34 Epidemiology. 2019;49(1):20-1i.
35 27. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensi 34 Feidemiology. 2019;49(1):20-1i.
35 health Study II:20-1i.
36 health Study M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehens
36 health Study Journal of Art and Art association meta-analysis of coron Epidemiology. 2019;49(1):20-1i.
35 27. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive
36 1,000 Genomes-based genome-wide association meta-analysis of coronary artery dise 36 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease.
37 Nat Genet. 2015;47(10):1121-30.
38 28. Mishra A, Malik R, Hachiya T, Jürgenson T, Namba S, Posner DC, et al. Stroke genetic 37 Nat Genet. 2015;47(10):1121-30.
38 28. Mishra A, Malik R, Hachiya T, Jürgenson T, Namba S, Posner DC, et al. Stroke genetion
39 informs drug discovery and risk prediction across ancestries. Nature. 2022;611(7934):115-38 28. Mishra A, Malik R, Hachiya
39 informs drug discovery and risk pr
40 23. 39 informs drug discovery and risk prediction across ancestries. Nature. 2022;611(7934):115-
40 23. Pierce BL, Ahsan H, VanderWeele TJ. Power and instrument strength requirements
41 29. Pierce BL, Ahsan H, VanderWeele TJ. 39 information generation, and risk prediction across antisometric across and across 23.
39 april 29. Pierce BL, Ahsan H, VanderWeele TJ. Power and instrument strength requirements
42 for Mendelian randomization studies us 41 29.
 42 for l
 43 of E 42 for Mendelian randomization studies using multiple genetic variants. International Journal
43 of Epidemiology. 2011;40(3):740-52.
44 30. Li J, Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship betwee 43 of Epidemiology. 2011;40(3):740-52.
44 30. Li J, Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between
45 imprecisely measured traits using GWAS summary data. PLOS Genetics. 2017;13(11). 44 30. Li J, Hemani G, Tilling K, Davey
45 imprecisely measured traits using GW 14 30. Line Engineering C, Timing C, Davey Smith G. Orienting the candidate relationship between
45 imprecisely measured traits using GWAS summary data. PLOS Genetics. 2017;13(11). $\frac{1}{2}$ imprecisely measured traits using GWAS summary data. PLOS Generics. $P = \frac{1}{2}$.

 $\frac{1}{2}$ ֞֝֝֝֝֝֝֝֝֝֝֝֝֝֝֝֝֝֝֝֝֝
֧֝
֧֪֝֬֝֝֬֝<u>֚</u>֪֪֪֦֦֪֖֚֚֘֝֘֝֘֝֘֝֓ 1 31. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid
2 instruments: effect estimation and bias detection through Egger regression. International
3 Journal of Epidemiology. 2015;44(2):512-25.
4 32. $\mathbf{1}$ 31. 2 instruments: effect entity and bias detection through Egger regression in
2 incurred bias detection and bias detection and bias detection of the Mendelian Randomization with Some Invalid Instruments Using a Weighted Medi 3 Journal of Epidemiology. 2015; Alaycock PC
3 Mendelian Randomization with Some Invalid
6 Estimator. Genetic Epidemiology. 2016;40(4) 4 14 32. Boundary Service Press, 1994.
1994 - Mendelian Randomization with Some Invalid Instruments Using a Weighted Media
1994 - Estimator. Genetic Epidemiology. 2016;40(4):304-14.
1994 - Benjamini Y, Hochberg Y. Controll 5 Mendelian Randomization Randomization Pendelian Randomization Estimator. Genetic Epidemiology. 2016;40(4):304-14.
1986 - Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and
1986 - Powerful Appr 8 Table 1 Benjamini Y, Hochberg Y. Controlling the False
8 Powerful Approach to Multiple Testing. Journal of the
9 (Methodological). 1995;57(1):289-300. 7 23. Benjamini Y, Hochberg Y, Homology Waren Michael Approach to Multiple Testing. Journal of the Royal Statistical Society: Series
19 (Methodological). 1995;57(1):289-300.
24. Williams SM, Giambartolomei C, Vukcevic D, S 8 Powerful Approach to Multiple Testing. J. 289-300.
8 Powerful Approach to Multiple Testing. Journal of the Royal Statistical Statistical Statistical Society: Ser
8 Bayesian Test for Colocalisation between Pairs of Geneti e (Methodological). 1994. Milliams SM, Giambartolomei C, 1995.
1996: 1997: 1995: 1995: 1996. 1997. 1998. 1998. 1998. 1998. 1998. 1999. 1999. 1999. 1999. 1999. 1999. 1999. 1
1997: 1999. 1999. 1999. 1999. 1999. 1999. 1999. 1 10 34. Williams SM, Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, et al.
11 Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using
12 Summary Statistics. PLoS Genetics. 2014; 12 Summary Statistics. PLoS Genetics. 2014;10(5).
13 35. Zuber V, Grinberg NF, Gill D, Manipur I, Slob EAW, Patel A, et al. Combining e
14 from Mendelian randomization and colocalization: Review and comparison of appr 14 from Mendelian randomization and colocalization: Review and comparison of approaches.
15 The American Journal of Human Genetics. 2022;109(5):767-82.
16 36. Gordillo-Marañón M, Zwierzyna M, Charoen P, Drenos F, Chopade S 15 From Mendelian Journal of Human Genetics. 2022;109(5):767-82.
16 36. Gordillo-Marañón M, Zwierzyna M, Charoen P, Drenos F, Chopade S, Shah T, et al.
17 Validation of lipid-related therapeutic targets for coronary heart 15 The American Journal of Human Genetics. 2022;109(5):767-82.
16 36. Gordillo-Marañón M, Zwierzyna M, Charoen P, Drenos F, Chopade S, Shah T, et al. 16 36. Gordillo-Marañón M, Zwierzyna M, Charoen P, Drenos F
17 Validation of lipid-related therapeutic targets for coronary hear
18 human genetics. Nature Communications. 2021;12(1). 17 Validation of lipid-related therapeutic targets for coronary heart disease prevention using
18 human genetics. Nature Communications. 2021;12(1).
19 37. Burgess S, Dudbridge F, Thompson SG. Combining information on mult 18 International communications. 2021;12(1).
19 37. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple
20 Instrumental variables in Mendelian randomization: comparison of allele score and 19 37. Burgess S, Dudbridge F, Thompson SG. Combini
20 instrumental variables in Mendelian randomization: co
21 summarized data methods. Statistics in Medicine. 2015 19 11 20 11 instrumental variables in Mendelian randomization: comparison of allele score and
19 11 19 summarized data methods. Statistics in Medicine. 2015;35(11):1880-906.
19 12 13 14 15 15 16 17 18 19 19 19 19 19 19 19 21 summarized data methods. Statistics in Medicine. 2015;35(11):1880-906.
22 38. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization:
23 the MR-Egger method. European Journal of Epidemiology. 2017;32 22 38. Burgess S, Thompson SG. Interpreting findings from Mendelian rar
23 the MR-Egger method. European Journal of Epidemiology. 2017;32(5):377
24 39. Giambartolomei C, Zhenli Liu J, Zhang W, Hauberg M, Shi H, Booco 23 the MR-Egger method. European Journal of Epidemiology. 2017;32(5):377-89.
24 39. Giambartolomei C, Zhenli Liu J, Zhang W, Hauberg M, Shi H, Boocock J, et al. A
25 Bayesian framework for multiple trait colocalization fro 24 39. Giambartolomei C, Zhenli Liu J, Zhang W, Hauberg M, Shi H, Boocock J,
25 Bayesian framework for multiple trait colocalization from summary association
2017: 2018:34(15):2538-45. 25 Bayesian framework for multiple trait colocalization from summary association statisti
26 Bioinformatics. 2018;34(15):2538-45.
27 40. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et a 26 Bioinformatics. 2018;34(15):2538-45.
27 40. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al.
28 Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419 27 40. Uhlén M, Fagerberg L, Hallströ
28 Proteomics. Tissue-based map of the l
29 41. Kathiresan S, Melander O, Guid 28 Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):12604
29 41. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, et al. Six new
30 associated with blood low-density lipoprotei Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419.

29 41. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, et al. Six new loci

30 associated with blood low-density lip 31 cholesterol or triglycerides in humans. Nature Genetics. 2008;40(2):189-97.
32 42. Samani NJ, Braund PS, Erdmann J, Götz A, Tomaszewski M, Linsel-Nitschke F
33 The novel genetic variant predisposing to coronary artery d 32 42. Samani NJ, Braund PS, Erdmann J, Götz A, Tomaszewski M, Linsel-Nit
33 The novel genetic variant predisposing to coronary artery disease in the regi
34 and CELSR2 genes on chromosome 1 associates with serum cholester 22 42. Samani NJ, Braund PS, Erdmann J, Götz A, Tomaszewski M, Linsel-Nitschke P, et al.
23 The novel genetic variant predisposing to coronary artery disease in the region of the PSR 33 The novel genetic variant predisposing to coronary artery disease in the region of the PSRC
34 and CELSR2 genes on chromosome 1 associates with serum cholesterol. J Mol Med (Berl).
35 2008;86(11):1233-41. 34 and CELSR2 genes on chromosome 1 associates with serum cholesterol. J Mol Med (Berl).
35 2008;86(11):1233-41.
36 43. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al.
37 Genomewide association an 35 2008;86(11):1233-41.
36 43. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al.
37 Genomewide association analysis of coronary artery disease. N Engl J Med.
28. 2003 257(5): 113 53 36 43. Samani NJ, Erd
37 Genomewide associat
38 2007;357(5):443-53. 37 Genomewide association analysis of coronary artery disease. N Engl J Med.
38 2007;357(5):443-53.
39 44. Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, et al. Mapping the ge 37 Genomewide association analysis of coronary artery disease. N Engl J Med. 39 44. Schadt EE, Mo
40 architecture of gene
41 45. Saleheen D, S 39 39 40. Schwarz Expression in human liver. PLoS Biol. 2008;6(5):e107.
39 41 45. Saleheen D, Soranzo N, Rasheed A, Scharnagl H, Gwilliam R, Alexander M, et al.
32 Genetic determinants of major blood lipids in Pakistanis c 41 45. Saleheen D, Soranzo N, Rasheed A, Scharnagl H, Gwilliam R, Alexan
42 Genetic determinants of major blood lipids in Pakistanis compared with Eu
43 Cardiovasc Genet. 2010;3(4):348-57. 41 45. Saleheen D, Soranzo N, Rasheed A, Scharnagl H, Gwilliam R, Alexander M, et al.
42 Genetic determinants of major blood lipids in Pakistanis compared with Europeans. Circ
43 Cardiovasc Genet. 2010;3(4):348-57.
46. Wal 43 Cardiovasc Genet. 2010;3(4):348-57.
44 46. Walia GK, Gupta V, Aggarwal A, Asghar M, Dudbridge F, Timpson N, et al. Associa
45 of Common Genetic Variants with Lipid Traits in the Indian Population. PLOS ONE. 44 46. Walia GK, Gupta V, Aggarwal *I*
45 of Common Genetic Variants with Lip
46 2014;9(7):e101688. 45. Walia Gammaria, Gupta V, Aggarwal A, Agaia M, Dudbridge F, Timpson A, Dudbridge F, Timpson N, 2014;9(7):et
46. 2014;9(7):e101688. 46 $2014;9(7):e101688.$

 $\frac{4}{3}$

 $\overline{}$ $\frac{2}{3}$ 47. Tan J, Che Y, Liu Y, Hu J, Wang W, Hu L, et al. CELSR2 deficiency suppresses lipid
2 accumulation in hepatocyte by impairing the UPR and elevating ROS level. Faseb j.
2021;35(10):e21908.
4 48. Linger RM, Keating AK, Ea $\mathbf{1}$ 47. 2 2021;35(10):e21908.
2 accumulation in hepatocyte by impairing the UPR and electron in hepatocyte by impairing RAS linger RM, Keating AK, Earp HS, Graham DK. TAM receptor tyrosine kinases
5 bunctions, signaling, and poten 2 2021;
3 2 2008;100:35-83.
5 2008;100:35-83. 49. Linger RM, Arring AM, Earp HS, Fraham Dimitriculp Linguan minister at the gradient functions, signaling, and potential therapeutic targeting in human cancer. Adv Cancer Res.
2008;100:35-83.
49. Linguan der Meer JH, van 5 2008;100:35-83.
7 49. van der Meer JH, van der Poll T, van 't Veer C. TAM receptors, Gas6, and protein S:
8 roles in inflammation and hemostasis. Blood. 2014;123(16):2460-9.
9 50. Wu G. Ma Z. Hu W. Wang D. Go 8 roles in inflammation and hemostasis. Blood. 2014;123(16):2460-9.
50. Wu G, Ma Z, Hu W, Wang D, Gong B, Fan C, et al. Molecular
10 cancer development and therapy. Cell Death Dis. 2017;8(3):e2700. 7 49. van der Meer JH, van der Poll T, van 't Veer C. TAM receptors, Gas6, and protein S: erton Cancer development and therapy. Cell Death Dis. 2017;8(3):e2700.
1991. Bordoloi J, Ozah D, Bora T, Kalita J, Manna P. Gamma-glutamyl carboxylated Gas6
29. Mediates the beneficial effect of vitamin K on lowering hyper 8 10 cancer development and therapy. Cell Death Dis. 2017;8(3):e2700.
11 51. Bordoloi J, Ozah D, Bora T, Kalita J, Manna P. Gamma-glutar
12 mediates the beneficial effect of vitamin K on lowering hyperlipide 11 51. Bordoloi J, Ozah D, Bora T, Kalita J, Manna P. Gamma-glutan
12 mediates the beneficial effect of vitamin K on lowering hyperlipider
13 AMPK/SREBP1/PPARα signaling cascade of lipid metabolism. J Nutr 12 mediates the beneficial effect of vitamin K on lowering hyperlipidemia via regulating the
13 AMPK/SREBP1/PPAR α signaling cascade of lipid metabolism. J Nutr Biochem. 2019;70:174
14 84. 13 AMPK/SREBP1/PPARα signaling cascade of lipid metabolism. J Nutr Biochem. 2019;70:17
14 84. Wu KS, Hung YJ, Lee CH, Hsiao FC, Hsieh PS. The Involvement of GAS6 Signaling in 13 AMPK/SREBP1/PPARα signaling cascade of lipid metabolism. J Nutr Biochem. 2019;70:17
14 84. Wu KS, Hung YJ, Lee CH, Hsiao FC, Hsieh PS. The Involvement of GAS6 Signaling in
16 the Development of Obesity and Associated I 15 52.
16 the
17 201 17 2015;2015:202513.
18 53. Ochoa D, Hercules A, Carmona M, Suveges D, Baker J, Malangone C,
19 generation Open Targets Platform: reimagined, redesigned, rebuilt. Nucleic 18 53. Ochoa D, He
19 generation Open Ta
20 2023;51(D1):D1353. 18 53. Ochoa D, Hercules A, Carmona M, Suveges D, Baker J, Malangone C, et al. The next-
19 eeneration Open Targets Platform: reimagined, redesigned, rebuilt. Nucleic Acids Res. 20 2023;51(D1):D1353-d9.
19 general Computan Open Targets Platform: reimagined, redesigned, redesigned, redesigned, redesigned, redesig
19 generation Open Targets Platform: 1 Clin Pharmacol. 2009;49(7):831-7. 21 54. Gupta M, Braga I
22 cholesterol in South Asia
23 55. Reddy LL, Shah S 18 54. Gupta M, Braga MF, Teoh H, Tsigoulis M, Verma S. Statin effects on LDL and HDL
18 ocholesterol in South Asian and white populations. J Clin Pharmacol. 2009:49(7):831-7. 22 cholesterol in South Asian and white populations. J Clin Pharmacol. 2009;49(7):831-7.
23 55. Reddy LL, Shah SAV, Ponde CK, Dalal JJ, Jatale RG, Dalal RJ, et al. Screening of PCS
24 and LDLR genetic variants in Familial 22 cholester and LDLR genetic variants in Familial Hypercholesterolemia (FH) patients in India. Journal of P
24 and LDLR genetic variants in Familial Hypercholesterolemia (FH) patients in India. Journal Senetics. 2021;66(1 24 and LDLR genetic variants in Familial Hypercholesterolemia (FH) patients in India. Journal of
25 Human Genetics. 2021;66(10):983-93.
26 56. (AMP) AMP. Common Metabolic Diseases Knowledge Portal: rs3748034 [Available 24 and LDLR genetic variants in Familial Hypercholesterolemia (FH) patients in India. Journal c
25 Human Genetics. 2021;66(10):983-93.
26 56. (AMP) AMP. Common Metabolic Diseases Knowledge Portal: rs3748034 [Available
27 f 27 from: https://hugeamp.org/variant.html?variant=4%3A3446091%3AG%3AT.
28 57. Sargsyan A, Doridot L, Hannou SA, Tong W, Srinivasan H, Ivison R, et al. HGFAC is a
29 ChREBP-regulated hepatokine that enhances glucose and lip 27 from: https://hugeamp.org/variant.html?variant=4%3A3446091%3AG%3AT.
28 57. Sargsyan A, Doridot L, Hannou SA, Tong W, Srinivasan H, Ivison R, et a
29 ChREBP-regulated hepatokine that enhances glucose and lipid homeostasi 28 57. Sargsyan A, Doridot L, Hannou SA, Tong W, Srinivasan H, Ivison R, et a
29 ChREBP-regulated hepatokine that enhances glucose and lipid homeostasis. J
2023;8(1). 29 ChREBP-regulated hepatokine that enhances glucose and lipid homeostasis. JCI Insight.
30 2023;8(1).
31 58. Zhang Z, Zhou Q, Liu R, Liu L, Shen WJ, Azhar S, et al. The adaptor protein GIPC1 2023;8(1).
29 Chram-regulation in that enhances glucose and lipid homeostasis.
29 S. – Zhang Z, Zhou Q, Liu R, Liu L, Shen WJ, Azhar S, et al. The adaptor protein GIPC1
22 Stabilizes the scavenger receptor SR-B1 and increa 31 58. Zha
32 stabilizes t
33 2021;296:1
31 53 32 stabilizes the scavenger receptor SR-B1 and increases its cholesterol uptake. J Biol Chem
33 2021;296:100616.
59. Go GW, Mani A. Low-density lipoprotein receptor (LDLR) family orchestrates 33 2021;296:100616.
34 59. Go GW, Mani A. Low-density lipoprotein receptor (LDLR) family orchestrates
35 cholesterol homeostasis. Yale J Biol Med. 2012;85(1):19-28.
36 34 59. Go GW, Ma
35 cholesterol homeo
36 35 cholesterol homeostasis. Yale J Biol Med. 2012;85(1):19-28.
36 35 cholesterol homeostasis. Yale J Biol Med. 2012;85(1):19-28.
36