Genetically predicted glucose-dependent insulinotropic polypeptide (GIP) levels and cardiovascular disease risk are driven by distinct causal variants in the *GIPR* region

Short title: Distinct variants drive GIP levels and cardiovascular disease risk

Nicholas Bowker[1], Robert Hansford[1], Stephen Burgess[2,3], Christopher N. Foley[2,3], Victoria P.W. Auyeung[1], A. Mesut Erzurumluoglu[1], Isobel D. Stewart[1], Eleanor Wheeler[1], Maik Pietzner[1], Fiona Gribble[4], Frank Reimann[4], Pallav Bhatnagar[5], Matthew P. Coghlan[5], Nicholas J. Wareham[1], Claudia Langenberg[1]*

[1]MRC Epidemiology Unit, University of Cambridge, Wellcome Trust-MRC Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, CB2 0QQ, UK.

[2]MRC Biostatistics Unit, Cambridge Institute of Public Health Robinson Way, Cambridge, CB2 0SR, UK.

[3]Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, CB1 8RN, UK.

[4]University of Cambridge, Wellcome Trust/MRC Institute of Metabolic Science (IMS) & MRC Metabolic Diseases Unit, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 0QQ, UK.

[5]Diabetes and Complications Therapy Area, Eli Lilly & Company, Lilly Corporate Center Indianapolis, IN 46285 U.S.A.

*Claudia Langenberg (claudia.langenberg@mrc-epid.cam.ac.uk)
MRC Epidemiology Unit,
University of Cambridge,
Institute of Metabolic Science,
Addenbrookes Hospital,
Cambridge,
CB2 0QQ
Tel. +44 (0)1223 330315
Fax. +44 (0)1223 330316

Word count: 4,889
Number of tables: 3
Number of figures: 3

Abstract: There is considerable interest in GIPR agonism to enhance the insulinotropic and extra-pancreatic effects of GIP, thereby improving glycaemic and weight control in type 2 diabetes (T2D) and obesity. Recent genetic epidemiological evidence has implicated higher GIPR-mediated GIP levels in raising coronary artery disease (CAD) risk, a potential safety concern for GIPR agonism. We therefore aimed to quantitatively assess whether the association between higher GIPR-mediated fasting GIP levels and CAD risk is mediated via GIPR or is instead the result of linkage disequilibrium (LD) confounding between variants at the *GIPR* locus. Using Bayesian multi-trait colocalisation, we identified a *GIPR* missense variant rs1800437 (G allele; E354) as the putatively causal variant shared between fasting GIP levels, glycaemic traits and adiposity-related traits (posterior probability for colocalisation, PP_{coloc}>0.97; PP explained by the candidate variant; $PP_{explained}$=1) that was independent from a cluster of CAD and lipid traits driven by a known missense variant in *APOE* (rs7412; distance to E354 ~770Kb; R^2 with E354 = 0.004; PP_{coloc}>0.99; $PP_{explained}$=1). Further, conditioning the association between E354 and CAD on the residual LD with rs7412, we observed slight attenuation in association, but it remained significant (OR per copy of E354 after adjustment 1.03; 95% CI, 1.02, 1.04; P=0.003). Instead, E354's association with CAD was completely attenuated when conditioning on an additional established CAD signal, rs1964272, (R^2 with E354=0.27), an intronic variant in *SNRPD2* (OR for E354 after adjustment for rs1964272: 1.01; 95% CI, 0.99, 1.03; P=0.06). We demonstrate that associations with GIP, anthropometric and glycaemic traits are driven by distinct genetic signals from those driving CAD and lipid traits in the *GIPR* region, and higher E354-mediated fasting GIP levels are not associated with CAD risk. These findings provide evidence that the inclusion of GIPR agonism in dual GIPR/GLP-1R agonists could potentiate the protective effect of GLP-1 agonists on diabetes without undue CAD risk, an aspect which has yet to be assessed in clinical trials.

The incretin hormones glucose-dependent insulinotropic polypeptide (GIP) and glucagon-like peptide-1 (GLP-1) are well known for their insulinotropic activity(1,2), which is diminished in type 2 diabetes (T2D)(3–6). This has prompted significant therapeutic interest in the agonism of their respective receptors, GIPR and GLP1R, to enhance their insulinotropic and extra-pancreatic effects(7,8). Moreover, preclinical and clinical data demonstrate that dual agonism of the GIPR and GLP1R delivers superior glycaemic and weight control efficacy compared to selective GLP1R agonism(9–12). Clinical proof for the superiority of tirzepatide, a dual GIPR/GLP1R agonist, versus GLP1R agonism was established in a 6-month dose range finding Phase 2b trial in subjects with type 2 diabetes(11). Post hoc analysis reported a beneficial effect on cardiovascular risk biomarkers compared to the blinded GLP1R agonist included in the trial(13,14).

There exists little direct preclinical experimental evidence for GIPR agonism contributing to cardiovascular disease (CVD) risk (15,16). GIP exhibits anti-atherogenic effects on vascular endothelial cells(17–20) with the exception that it has been reported to stimulate expression of osteopontin in the vasculature in an endothelin-1 dependent manner(21). Additionally, GIP exerts anti-inflammatory effects on monocytes/macrophages(17,22). These *in vitro* findings are reflected by cardioprotective GIP pharmacology in mouse models of atherosclerosis irrespective of their diabetic condition(17,22,23). Further, GIP infusion or overexpression is protective in mouse models of restenosis and cardiac remodelling(17,24). Whilst germline or cardiomyocyte-selective knock-out of GIPR protected against ischemic injury, GIP itself was not deleterious(25). Further, cardiac selective knock-out of the GIPR was not protective in experimental models of heart failure(25). In contrast with these preclinical experimental findings, recent evidence suggests that fasting GIP levels are associated with increased carotid intimal thickening(26). In addition, evidence from a recent meta-analysis(27) of two large

population-based cohort studies suggests that higher fasting but not post-challenge GIP levels were associated with increased risk of CVD mortality (HR, 1.30; 95% CI, 1.11, 1.52; P=0.001). GLP-1 was not associated with CVD mortality, consistent with clinical trial data(28–31) and genetic evidence(32) highlighting the beneficial effects of GLP-1R agonism.

Genetic evidence from two-sample Mendelian randomisation (2SMR) has reinforced suggestions that higher GIP levels raise CVD risk(27). A missense variant in *GIPR*, rs1800437 (E354Q), encoding a substitution of glutamic acid for glutamine at position 354 of the GIPR protein, was used as an instrumental variable for fasting GIP levels(27). The 354Q allele has been reported to reduce GIPR signalling by increasing receptor desensitisation and down-regulation(33). This variant has previously been associated with higher 2-hour glucose(34), BMI(35) and fasting and 2-hour GIP levels(36). In line with a predicted causal direction from fasting GIP levels to coronary artery disease (CAD) risk, estimates in the reverse direction showed no significant effect of CAD on fasting GIP levels(27). These estimates should be interpreted with caution, however, as (1) they represent the association of a single variant with CAD risk and do not model the effects of other variants in the region which may dampen or modulate this effect, and (2) they do not take into account that the association between E354 and CAD may be entirely synthetic due to linkage disequilibrium (LD) between this variant and the true CAD causal variant.

Considering the pharmacological interest in modulating this pathway as a potential T2D therapeutic, increases in CVD risk would represent a major concern regarding the safety and continued development of these therapies. We aimed to quantitatively assess whether the association between higher GIPR-mediated fasting GIP levels and CAD risk is mediated via GIPR or the result of LD between variants in GIPR and other variants in the region. Using

2SMR, we aimed to quantify the association of higher fasting GIP levels with CAD and other metabolically relevant traits, including ~6000 'omics biomarkers, using E354 as an instrumental variable. Next, using Bayesian colocalisation, we aimed to partition the traits associated with E354 into distinct clusters driven by shared independent variants. Finally, using conditional analysis we aimed to assess whether any of these associations are confounded by LD between E354 and other variants in the GIPR region.

Materials and Methods

Study design

Three sets of genetic analyses were used to investigate the relationship between higher GIPR-mediated fasting GIP levels and CVD risk. Firstly, using univariate 2SMR, we explored the association of higher fasting GIP levels with CAD and 23 different cardiometabolic diseases, along with anthropometric, glycaemic, lipid traits and ~6,000 'omics biomarkers from both in-house and publicly available data, using E354 as a proxy (**Table S1**). Next, Bayesian multi-trait colocalisation was used to partition the traits associated with E354 into distinct clusters driven by shared causal variants. Finally, conditional analyses were used to assess whether any of the associations with E354 are confounded by LD between E354 and other variants in the *GIPR* region, implying that their associations are not mediated via *GIPR* but other genes in the region.

Study participants

EPIC-Norfolk(37) (**Table S2**) is a population-based prospective cohort of individuals aged between 40-79 years and living in Norfolk (a county of the United Kingdom) at the time of recruitment from primary-care outpatient clinics in the city of Norwich and surrounding areas. EPIC-Norfolk(37) consists of two sub-cohorts, a T2D case-cohort and a quasi-random selection of participants from the larger EPIC(38,39) study. The study was approved by the Norfolk Research Ethics Committee (ref. 05/Q0101/191) and all participants gave their written consent before entering the study.

Fenland(40) (**Table S2**) is a population-based cohort study of individuals without diabetes who were born between the years of 1950 and 1975 and recruited through population-based general practice registers in Cambridge, Ely and Wisbech (Cambridgeshire county, United Kingdom).

Ethical approval for the study was given by the Cambridge Local Ethics committee (ref. 04/Q0108/19) and all participants gave their written consent prior to entering the study.

UK Biobank(41) (**Table S2**) is a population-based cohort study of individuals recruited from 22 rural and urban recruitment centres in the United Kingdom. European ancestry participants with available genome-wide genotyping and phenotypic data were included in this study. Ethical approval for the UK Biobank study was given by the North West - Haydock Research Ethics Committee (16/NW/0274). This research was conducted using application 44448. Participants gave their electronic consent to use their anonymised data and samples for health-related research, to be re-contacted for further sub-studies, and for access to their health-related records.

Genotyping and imputation

Genome-wide genotyping in the Fenland cohort was performed in 3 sub-cohorts using either the Affymetrix genome-wide Human variant Array 5.0, the Affymetrix UK Biobank Axiom Array or the Illumina CoreExome-24 v1 chip, with imputation to the Haplotype reference consortium v1.1(42), the 1000 genomes project(43) and the UK10K(44) reference panels. Samples from EPIC-Norfolk and UK Biobank were genotyped using the Affymetrix UK Biobank Axiom Array and imputed to the same reference panels.

Profiling of the plasma proteome

Fasted EDTA plasma samples from 12,084 participants from the Fenland(40) study were subjected to proteomic profiling by SomaLogic Inc. (Boulder, US) using an aptamer-based technology (SOMAscan v4). The relative abundances of 4,775 human proteins were measured using 4,979 SOMAmers(45). To account for within run hybridisation variability, control probes were used to generate a scaling factor for each sample. Differences in total signal

between samples as a result of variation in overall protein concentration or technical variability such as reagent concentration, pipetting or assay timing, were accounted for using the ratio between each SOMAmer's measured value and a reference value. The median of these ratios was computed for each dilution set (40%, 1% and 0.005%) and applied to each dilution set. Samples were removed if they failed SomaLogic QC measures or did not meet the acceptance criteria of between 0.25-4 for all scaling factors. A total of 10,078 samples had available genotype data and were used in this study. Aptamer target annotations and mapping to UniProt accession numbers as well as gene identifiers were provided by SomaLogic.

Plasma metabolomic profiling

Within EPIC-Norfolk (37) (described previously), the levels of up to 1,504 metabolites were measured in three batches using the Metabolon DiscoveryHD4 platform(46) (Metabolon, Inc., Durham, USA), in citrate plasma samples collected at baseline. Measurements were made in approximately 12,000 samples, in two sets of approximately 6000 quasi-randomly selected samples, which were preceded by measurements in an incident T2D case-cohort (N= 1503; 857 in the sub-cohort).

Briefly, raw data were extracted, peaks were identified and assessed for quality by Metabolon. Metabolite identification was done by comparing measures to a curated library containing the retention time, mass to charge ratio and chromatographic data of known metabolites. Each metabolite was then quantified using an area-under-the-curve method and the data were normalised to correct for instrument tuning variations across run-days. Data normalisation for each run-day set the median value for each metabolite to 1, normalising each measurement proportionately. Metabolite annotations and pathway classifications are as reported by Metabolon, Inc.

193

194 ***Statistical analysis***

195 *GWAS of plasma proteins and pairwise colocalisation of GIP levels with cardiometabolic traits*

196 GWAS was performed as described in **Table S3**. Two SOMAmers targeted circulating GIP,
197 namely 16292-288 and 5755-29. SOMAmer 16292-288 was selected against amino acids 1-93
198 of the precursor protein (Uniprot ID: P09681), whereas, 5755-29 targeted amino acids 22-153.
199 SOMAmers are relative measures of GIP abundance, therefore, in order to ascertain whether
200 the underlying genetics at *GIPR* were comparable to previous results(36), we performed
201 pairwise genetic colocalisation analyses between GIP measures and cardiometabolic traits.

202

203 T2D, CHD, BMI, 2-hour glucose adjusted for BMI and LDL were included as cardiometabolic
204 traits of interest (**Table S1)**. Summary statistics from a GWAS of 2-hr glucose adjBMI in
205 Fenland (**Table S3**) were preferred to those from previous efforts(34), due to denser variant
206 coverage. Using GWAS summary statistics for each trait, the 1Mb regions either side of E354
207 (Chr19:45181392-47181392) were extracted. Insertions and deletions as well as any variants
208 with a standard error of 0 were removed. Effect estimates were aligned to the GIP-raising
209 alleles. Pairwise colocalisation was conducted using the COLOC(47) R package. Priors, p1
210 and p2, the prior probabilities that a variant is associated with either trait were set to $1x10^{-4}$ and
211 p12, the probability that a single variant is associated with both traits, was set to $1x10^{-5}$. T2D
212 and CHD were treated as case-control traits and all other traits as quantitative. Posterior
213 probabilities (PP_{coloc}) were considered significant if they met the following criteria: (H4 + H3
214 ≥ 0.9 & $H4/H3 \geq 3$).

215

216 *GWAS of plasma metabolites*

GWAS was performed in 2 sets, for all metabolites present in at least 100 individuals in both sets. The first set consisted of up to 5,841 individuals from both the sub-cohort of the T2D case cohort and the first batch of quasi-randomly selected samples. The second set consisted of up to 5,698 individuals from the second batch of quasi-randomly selected samples. GWAS was performed as described in **Table S3**.

Association between E354, cardiometabolic and molecular traits

This work leveraged regional GWAS summary statistics from in-house studies and data from published studies in the 1Mb regions either side of E354 (Chr19:45181392-47181392). Details on all included phenotypes can be found in **Table S1**. GWAS for phenotypes derived in-house were performed as described in **Table S3**. Only self-reported, white European participants were included for all outcomes except for plasma metabolite measures in EPIC-Norfolk(37), where all participants were included. However, participants in EPIC-Norfolk(37) overwhelmingly self-reported as white European.

We performed univariate 2SMR using the Wald ratio method(48) to estimate the potential causal effect of fasting GIP levels on various traits (**Table S1**). Genetically predicted fasting GIP levels were used as the exposure with E354 as the instrumental variable (HUGO gene: *GIPR*; NCBI transcript NM_000164.4 c.1060G>C; protein change, E354Q; E345 variant is encoded by the G allele). All summary statistics were aligned to the fasting GIP raising allele (G) of E354. Bonferroni corrected significance thresholds were used to ascertain statistical significance of E354 across all outcomes.

Partial correlations between X-12283 and known metabolites

To estimate the metabolite class and putative functional pathway of X-12283, we estimated partial correlations between X-12283 levels and the levels of other metabolites measured in 11,966 participants from EPIC-Norfolk.

First, missing metabolite measures were imputed within each measurement set, using multivariate imputation by chained equations (MICE)(49) with the R package "mice" v3.6.0. To ensure accurate imputation, we only considered the 883 metabolites with less than 50% missingness within both measurement sets. Imputation was repeated a total of 20 times, generating 20 sets of fully imputed results. Following imputation, measures were standardised (mean = 0, SD =1). For each imputation, partial correlations between metabolite pairs were calculated using the R package "GeneNet" v1.2.14. Partial correlation estimates were transformed using Fisher's Z-transformation and the R package "psych" v1.9.12.31, and then pooled across the 20 imputations for each measurement set, using Rubin's rules(50). Estimates for the two measurement sets were then meta-analysed, using a fixed effect, inverse variance weighted method in the R package "meta" v4.12-0, and finally back transformed to correlation estimates. P-values were calculated using the Fisher's transformed partial correlations.

Partial correlation estimates with absolute values of more than 0.1 were then used to draw a gaussian graphical model (GGM) in Cytoscape v3.2.1. Partial correlations were considered significant at a Bonferroni significance threshold of $P \leq 1.28 \times 10^{-7}$, accounting for the 389,403 metabolite pairs tested.

Multi-trait colocalisation across cardiometabolic traits

Multi-trait colocalisation (HyPrColoc)(51) was used at the *GIPR* locus to 1) identify cardiometabolic traits that share a common causal variant, and 2) partition clusters of

cardiometabolic traits driven by distinct causal variants. HyPrColoc was run using the default variant-specific prior configuration, priors 1 and 2 were set at $1x10^{-4}$ and 0.02 respectively, and regional and alignment thresholds of 0.5 were used(51).

Variants were extracted and excluded from GWAS summary statistics for 26 cardiometabolic traits of interest as in the pairwise colocalisations above and all variants in perfect LD ($R^2 = 1$) with E354 were removed. The GIP measures considered were fasting GIP as measured by SOMAmers X16292_288 and 5755-29, as well as fasting and 2-hr GIP measures from the Malmö Diet and Cancer (MDC) sub-cohort of Almgren *et al.* 2017(36). Both the MDC and PPP-botnia cohorts were genotyped using exome-wide arrays, thereby limiting the number of variants included in the analysis when considering variants present across all traits. MDC measures were preferred to those from either the PPP-Botnia sub-cohort or the meta-analysis of the two sub-cohorts due to denser variant coverage, despite PPP-Botnia having a larger sample size. The anthropometric traits adjusted and unadjusted for BMI (where applicable) were BMI, WHR, and hip and waist circumferences. T2D and CAD were included as disease outcomes. Glycaemic measures included non-fasted glucose, HbA1c, 2-hr glucose adjusted for BMI, fasting glucose adjusted for BMI and fasting insulin adjusted for BMI. GWAS summary statistics from Fenland were used for fasting and 2-hour glucose as well as fasting insulin. Finally, lipid traits included were LDL, HDL, total cholesterol, triglycerides, lipoprotein A, apolipoprotein A1 and apolipoprotein B.

To assess sensitivity in the number and size of clusters identified, increasingly stringent prior and threshold configurations were used. Prior 2 values of 0.02, 0.01 and 0.001, and threshold values of 0.5, 0.6, 0.7, 0.8 and 0.9 were considered. T2D and CAD were considered as binary case-control traits and all others were considered quantitative. To estimate the posterior

probability (PP) that the candidate variant is the causal variant (PP_{causal}), we multiplied the PP_{coloc} by the PP explained by the candidate variant ($PP_{explained}$). Trait clusters were reported at the recommended(51) thresholds of prior 2 = 0.02, regional and alignment thresholds = 0.9.

To account for low variant coverage in the MDC cohort, we ran a secondary analysis using the same populations, configuration and sensitivity assessments as above, while excluding the GIP traits measured in MDC.

Finally, heatmaps based on similarity matrices estimating how often trait pairs were clustered together across all algorithm parameter choices were drawn. In addition, regional association plots were drawn for each cluster using the gassocplot R package and. LD data from EPIC-Norfolk. All data analysis was performed using R version 3.6.3.

Conditional analysis at the GIPR locus

To determine whether the association between E354 and CAD was due to LD between E354 and other CAD lead variants in the *GIPR* region, we performed conditional analysis using GCTA(52) v1.93.1. Using full GWAS summary statistics for CAD(53) on chromosome 19, we implemented a step-wise selection to identify independent variants associated with CAD. Selection was performed using a threshold of $P< 1x10^{-5}$, a threshold for collinearity between variants of 0.05 and a minor allele frequency threshold of 1%. An LD reference panel from EPIC-Norfolk was used. The association between E354 and CAD was then conditioned on each independent variant to estimate whether the association was attenuated, implying that the association was due to the residual LD between E354 and an independent variant. This was repeated for all traits associated with E354. If E354 (or a proxy variant in complete LD with E354) was identified as one of the independent variants, conditional analysis was not

316 performed. Following this, regional association plots were generated using LocusZoom v1.2.
317 To determine whether other variants previously found to be associated with fasting GIP
318 levels(36) were associated with CAD, we extracted their estimates from the CAD summary
319 statistics(53).
320
321 *Data and resource availability*
322 The datasets analysed during the current study are publicly available and links are provided in
323 **Table S1**. EPIC-Norfolk or Fenland data are available upon reasonable request via the study
324 websites (https://www.mrc-epid.cam.ac.uk/research/studies/epic-norfolk/ and
325 https://www.mrc-epid.cam.ac.uk/research/studies/fenland/information-for-researchers/). GIP
326 measures from Almgren *et al*. (36) are available from the relevant corresponding author upon
327 reasonable request. All data from UK Biobank are available to approved users upon
328 application. No applicable resources were generated or analysed during the current study.

Results

Characterisation of a missense variant E354 (rs1800437) in GIPR

Among the cardiometabolic disease outcomes examined, higher E354-predicted fasting GIP levels were associated with lower T2D risk (OR per copy of E354, 0.97; 95% CI, 0.96, 0.99; $P=3\times10^{-4}$; **Fig. 1A**), an effect which strengthened following BMI adjustment (0.93; 95% CI, 0.91, 0.95; $P=3\times10^{-14}$). In line with this, lower 2-hour glucose levels were observed (2-hour glucose in mmol/L per copy of E354, -0.09; 95% CI, -0.11, -0.07; $P=2\times10^{-15}$; **Fig. 1B**). Additionally, HbA1c levels were shown to be 0.01 SD units lower per copy of E354. E354 showed a weak positive association with non-fasted glucose levels. As this phenotype captures wide-ranging physiological responses in both the fasted and postprandial state, deconvoluting this association requires further investigation. E354 was associated with higher CAD risk (OR per copy of E354, 1.03; 95% CI, 1.02, 1.05; $P=2x10^{-6}$; **Fig. 1A**and higher levels of several lipid risk factors but lower triglyceride levels (**Fig. 1B**). E354 was not significantly associated with other CVD subtypes in UKBB (**Fig. S1**).

Each copy of E354 was associated with 0.03 SD higher BMI (95% CI, 0.03, 0.04; $P=3\times10^{-59}$; **Fig. 1B**). Similar associations were observed between E354 and higher regional anthropometric measures from bio-impedance data (**Fig. S2**) as well as hip and waist circumferences and waist-to-hip ratio. In addition, significant associations were found with both higher lean and fat mass from a large GWAS based on bio-impedance data (**Fig. S2**).

Of the 19 biomarkers investigated, E354 was significantly associated with lower levels of only two, namely albumin and creatinine (beta in SD units per copy of E354, -0.01; 95% CI, -0.02, -0.01; $P=6\times10^{-6}$; and -0.02; 95% CI, -0.02, -0.01; $P=1\times10^{-11}$, respectively; **Fig. 1B**).

Next, we estimated the association of E354 with the fasting levels of 4,979 human proteins from the SOMAscan® v4 system. Significant associations with the levels of three proteins were found (**Fig. S3**), one of these being 0.08 SD higher fasting GIP levels (95% CI, 0.05, 0.11; $P=4\times10^{-6}$) as measured by SOMAmer 16292-288. Interestingly, our analysis did not find a significant association between the other GIP SOMAmer, 5755-29, and E354. Lower levels of secretoglobin family 3A member 1 (SCGB3A1) and glutaminyl-peptide cyclotransferase-like protein (QPCTL) were also found to be associated with E354. In contrast with a previous report(21), no association between E354 and osteopontin was found.

Lower levels of an unidentified metabolite, X-12283 (beta in SD units per copy of E354, -0.08; 95% CI, -0.12, -0.05; $P=2\times10^{-5}$; **Fig. S4**), analysed in 8,278 participants, were found to be significantly associated with E354. A total of 11 metabolites were significantly correlated with X-12283, of these, six had a partial correlation estimate with X-12283 with absolute values greater than 0.1 (**Fig. S5**). In addition to significant correlations with unknown metabolites, X-12283 was most significantly correlated with indolepropionate (correlation estimate = 0.21; $P=1x10^{-45}$; **Fig. S5**).

Multi-trait colocalisation across cardiometabolic traits at GIPR

A total of 424 variants were included in the main analysis, which was limited due to the inclusion of fasting and 2-hour GIP measures from MDC(36), whereas 5,015 were included in the secondary analysis (**Table 1**). Using the recommended prior and threshold configuration, 5 distinct trait clusters were identified, 3 of which were shared by both analyses (**Table 1**). Cluster similarity across all prior and threshold permutations for the two analyses are summarised in heatmaps (**Fig. 2)**. Results for all permutations for both analyses can be found in **Tables S4 and S5** respectively.

Of the clusters identified, two distinct clusters were of interest. The first, driven by rs7412 a missense variant in the apolipoprotein E gene (*APOE*), contained CAD and lipid traits – many of which are established CVD risk factors. Both PP_{coloc} and PP_{causal} were estimated to be 1 in the two analyses, demonstrating robust evidence for colocalisation (**Table 1 and Fig. S6**). This robustness is further emphasised as the same cluster of traits was identified when using more stringent prior configurations (**Fig. 2, Tables S4 and S5**). A second cluster of GIP, anthropometric and glycaemic traits was driven by rs1800437 (E354) (**Table 1 and Fig. S7**). The PP_{coloc} for both analyses showed robust evidence for colocalisation (Main analysis: PP_{coloc}=0.97; $PP_{explained}$=1; PP_{causal}=0.97; Secondary analysis: PP_{coloc}=0.91; $PP_{explained}$=0.68; PP_{causal}=0.62). A second cluster of BMI and waist circumference driven by E354 was observed in the secondary analysis (**Table 1**). Sensitivity analyses showed that this split was an artefact of the branch and bound clustering algorithm in HyPrColoc and the single causal variant assumption (**Fig. S7**). Removal of the clustering algorithm showed that BMI and waist circumference were part of the larger cluster of GIP, anthropometric and glycaemic traits driven by E354 (PP_{coloc} = 0.95; $PP_{explained}$ = 1; PP_{causal} = 0.95)..

Critically, these results replicate our findings using pairwise-trait colocalisation at this locus, showing that fasting GIP levels and CVD risk are driven by independent variants (R^2 between E354 and rs7412 = 0.004) (**Table 1; Fig. S6-S8; Fig. 2**). Additionally, both colocalisation analyses demonstrate that the underlying genetics at *GIPR* are comparable between GIP levels measured by SOMAmer 16292-288 and the ELISA of previous analyses(36). Together these results robustly demonstrate that the GIP-raising and CVD risk increasing effects at this locus are distinct (**Tables S4 and S5**).

A third cluster including a mixture of glycaemic, anthropometric traits and ApoA1 levels were estimated to colocalise at rs4420638 which was in LD with rs429358 ($R^2 = 0.69$), a missense variant in *APOE* identified as the candidate variant in the secondary analysis (R^2 with E354 = 0.001). . As the secondary analysis included more variants and therefore had greater genomic context, rs429358 is likely to be the candidate variant at which these traits colocalise. The high PP_{coloc} demonstrated robust evidence for colocalisation between these traits at rs429358.

Finally, a cluster between T2D and T2D adj. BMI was identified in the main analysis but was not replicated in the secondary analysis (**Table 1**). Instead, a cluster between triglycerides and hip circumference adj. BMI was identified, driven by an independent variant rs5117 (R^2 with rs8108269 < 0.001) (**Table 1**). This discrepancy is likely to be a result of the number of variants present in the main analysis.

Conditional analysis at the GIPR locus

Our univariate two-sample MR results showed that E354 was associated with a total of 20 traits at a nominal significance threshold (**Fig. 1**). Independent signal selection showed that E354, or proxy variants in high LD ($R^2 > 0.9$) with E354, were identified as independent signals for fasting GIP, 2-hour glucose, total cholesterol levels, BMI and X-12283 levels. A total of 24 variants were independently associated with CAD on chromosome 19, four of which were in the 1Mb regions either side of E354 at the *GIPR* locus (**Table 2**). Conditioning the association between E354 and CAD on the residual LD between E354 and rs7412, the variant estimated to drive the cluster with CAD, resulted in a slight attenuation of this association but remained significant (OR per copy of E354 after adjustment 1.03; 95% CI, 1.02, 1.04; P=0.003). Of the independent variants identified, rs1964272 an intronic variant in small nuclear ribonucleoprotein D2 polypeptide (*SNRPD2*), was estimated to be in the strongest LD with

E354 (R^2=0.27) (**Fig. 3 and Fig. S9**). The association between E354 and CAD risk was attenuated when conditioned on rs1964272 (OR per copy of E354 after adjustment, 1.01; 95% CI, 0.99, 1.03; P = 0.06) (**Table 3**). In line with this, the association between rs1964272 and CAD risk was attenuated but remained significant when conditioning on E354 (beta per copy of rs1964272 after adjustment, 0.02; 95% CI, 0.01, 0.03; P = $7x10^{-4}$; **Table S6**). In addition, the association between E354 and small vessel stroke was also attenuated when conditioned on rs1964272 (**Table 3**). None of the other loci previously shown to be associated with fasting GIP levels were found to be associated with CAD (**Table S7**). Interestingly, rs1964272 was also associated with levels of QPCTL and SCGB3A1, indicating confounding by LD for the proteomics data as well (**Fig. S10**). Conditioning the association between E354 and QPCTL levels on rs1964272 attenuated the association to non-significance (beta QPCTL per copy of E354 after adjustment, 0.01; 95% CI, -0.02, 0.04; P=0.48; **Table 3**).

Conditioning the association of E354 with LDL, ApoB and triglycerides on independent variants for each trait showed that these remained statistically significant despite being attenuated (**Table 3**), suggesting that E354 may have independent effects on lipid metabolism.

Discussion

In this study, we applied Bayesian multi-trait colocalisation and conditional analysis to gain greater understanding of the underlying genetic architecture of CAD and its relation to fasting GIP levels at the *GIPR* locus. Multi-trait colocalisation robustly identified a cluster of CAD and lipid traits at *APOE* that was independent from a cluster of fasting and 2-hour GIP, glycaemic and anthropometric traits driven by E354. Further, conditional analysis robustly attenuated E354's association with CAD, small vessel stroke and QPCTL levels when adjusting for rs1964272 in *SNRPD2*, an established CAD risk locus(53). Together these results show that association signals for CAD at *GIPR* are not mediated by an independent effect of *GIPR* variants on CAD risk but are instead the result of LD confounding between E354 and rs1964272.

Taken together, these findings highlight the specificity of E354's effects on fasting GIP levels and robustly demonstrate that higher E354-mediated fasting GIP levels are not associated with CVD risk. These results contradict recent genetic evidence linking higher fasting GIP levels with increased CVD risk(21,27), which led to concerns that chronic pharmacological GIPR agonism could have detrimental effects on cardiovascular health(27) and represent safety concerns for pharmacological agonism of this pathway(54). We therefore provide evidence that the inclusion of GIPR agonism in dual GIPR/GLP-1R agonists could potentiate the protective effect of GLP-1 agonists on diabetes without undue CVD risk, an aspect which has yet to be assessed in clinical trials. Many studies have shown that GLP1R agonism achieved through chronic pharmacologic therapy, or genetic gain of function, is associated with improved cardiovascular outcomes(28–32). Hence, the available evidence suggests that dual agonism of these receptors may exploit the metabolically favourable combined pharmacology of these incretins without undue CVD risk. However, this proposition requires formal assessment in

clinical trials such as the recently initiated SURPASS cardiovascular outcomes trial of the GIP/GLP1R dual agonist tirzepatide (clinicaltrials.gov: NCT04255433).

This study has potential limitations. Firstly, our analysis focuses on a single locus associated with both fasting GIP levels and CAD. This assumes that the *GIPR* locus is a suitable proxy for fasting GIP levels within which to partition the associations of these two complex traits. Considering that the association at this locus with 2-hour glucose is statistically robust and in line with the established function of GIP, this is a reasonable assumption. In addition, no other locus has been reported to be associated with both fasting GIP and CAD, and examining the association of other variants associated with fasting GIP levels(36) in genes other than *GIPR*, showed no association of any of these variants with CAD. However, this does not preclude the existence of other variants that have not yet been associated with GIP levels may contribute to CVD risk. Patients with T2D are the target of GIPR/GLP-1R agonist treatment. We investigate the genetic association of E354 on CAD using the largest publicly available genome-wide summary statistics(53). Therefore, analyses stratified by T2D status are not possible since such results were not generated and are hence not available. Indeed, pursuing this in individual studies would vastly lower sample sizes and therefore be underpowered to detect whether associations with CAD differ significantly by T2D status. Specifically, to affect our results and conclusions about the E354-CAD association being the result of confounding by LD, the genetic architecture at GIPR would have to differ between European-descent individuals with and without prevalent T2D, such that the residual confounding by LD differs by T2D status. As LD is generally preserved between individuals from the same ethnic group, this is a very unlikely scenario.

Acknowledgments: We thank Emma Ahlqvist for sharing the GWAS summary statistics for fasting and 2-hour GIP measures from Almgren *et al.* 2017. We are grateful to all EPIC Norfolk participants who have been part of the project and to the many members of the study teams at the University of Cambridge who have enabled this research. We are grateful to all Fenland Study volunteers and to the General Practitioners and practice staff for assistance with recruitment. We thank the Fenland Study Investigators, Fenland Study Co-ordination team and the Epidemiology Field, Data and Laboratory teams. This research was conducted using the UK Biobank resource (application: 44448) and data from the EPIC-Norfolk and Fenland studies. Guarantor: N.B. and C.L. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. **Funding**: The EPIC-Norfolk study (https://doi.org/10.22025/2019.10.105.00004) has received funding from the Medical Research Council (MRC) (MR/N003284/1 and MC-UU_12015/1) and Cancer Research UK (C864/A14136). The genetics work in the EPIC-Norfolk study was funded by the MRC (MC_PC_13048). The Fenland Study (10.22025/2017.10.101.00001) is funded by the MRC (MC_UU_12015/1). We further acknowledge support for genomics and metabolomics from the MRC (MC_PC_13046). Fiona Gribble and Frank Reimann acknowledge funding by Wellcome (106262/Z/14/Z and 106263/Z/14/Z) and MRC (MRC_MC_UU_12012/3). NJW is an NIHR Senior Investigator. **Author contributions**: NB designed the study, analysed the data and wrote the first draft of the manuscript. CL, NJW, PB and MPC conceived of the idea under investigation. NB and RH performed the 2SMR analyses and analysed the data. CNF provided statistical support. VAY conducted the partial correlation analysis and created the Gaussian Graphical Model. NB, SB, AME, IDS, EW and MP collated the data and contributed to the data analysis. FR and FG contributed to manuscript revision and result interpretation. All authors contributed to the interpretation of results and writing or revision of the manuscript. **Competing interests**: NB has recently changed positions and is now an employee of GlaxoSmithKline. Throughout the duration of this work NB was a student at the MRC Epidemiology Unit and had nothing to declare. PB and MPC are employees and shareholders of Eli Lilly & Company. FR and FG receive grant funding from Eli Lilly & Company. FR and FG receive grant funding from AstraZeneca and FG is a consultant for Kallyope, but this is unrelated to the presented work. RH, VAY, ME, IDS, EW MP, NJW and CL have no competing interests to declare. **Guarantor statement:** CL is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References:

1. Nauck MA, Homberger E, Siegel EG, Allen RC, Eaton RP, Ebert R, et al. Incretin effects of increasing glucose loads in man calculated from venous insulin and C-peptide responses. J Clin Endocrinol Metab. 1986 Aug;63(2):492–8.
2. Gasbjerg LS, Bergmann NC, Stensen S, Christensen MB, Rosenkilde MM, Holst JJ, et al. Evaluation of the incretin effect in humans using GIP and GLP-1 receptor antagonists. Peptides [Internet]. 2020;125(September 2019):170183. Available from: https://doi.org/10.1016/j.peptides.2019.170183
3. Højberg P V., Vilsbøll T, Rabøl R, Knop FK, Bache M, Krarup T, et al. Four weeks of near-normalisation of blood glucose improves the insulin response to glucagon-like peptide-1 and glucose-dependent insulinotropic polypeptide in patients with type 2 diabetes. Diabetologia. 2009;52(2):199–207.
4. Christensen MB, Calanna S, Holst JJ, Vilsbløll T, Knop FK. Glucose-dependent insulinotropic polypeptide: Blood glucose stabilizing effects in patients with type 2 diabetes. J Clin Endocrinol Metab. 2014;99(3):418–26.
5. Vilsbøll T, Krarup T, Madsbad S, Holst J. Defective amplification of the late phase insulin response to glucose by gip in obese type ii diabetic patients. Diabetologia. 2002;45(8):1111–9.
6. Meier JJ, Hücking K, Holst JJ, Deacon CF, Schmiegel WH, Nauck MA. Reduced Insulinotropic Effect of Gastric Inhibitory Polypeptide in First-Degree Relatives of Patients with Type 2 Diabetes. Diabetes. 2001;50(7–12):2497–504.
7. Bailey CJ. GIP analogues and the treatment of obesity-diabetes. Peptides. 2020;125(November 2019).
8. Knerr PJ, Mowery SA, Finan B, Perez-Tilve D, Tschöp MH, DiMarchi RD. Selection and progression of unimolecular agonists at the GIP, GLP-1, and glucagon receptors as drug candidates. Peptides [Internet]. 2020;125(October 2019):170225. Available from: https://doi.org/10.1016/j.peptides.2019.170225
9. Finan B, Ma T, Ottaway N, Müller TD, Habegger KM, Heppner KM, et al. Unimolecular dual incretins maximize metabolic benefits in rodents, monkeys, and humans. Sci Transl Med. 2013 Oct;5(209).
10. Coskun T, Sloop KW, Loghin C, Alsina-Fernandez J, Urva S, Bokvist KB, et al. LY3298176, a novel dual GIP and GLP-1 receptor agonist for the treatment of type 2 diabetes mellitus: From discovery to clinical proof of concept. Mol Metab [Internet]. 2018;18(October):3–14. Available from: https://doi.org/10.1016/j.molmet.2018.09.009
11. Frias JP, Nauck MA, Van J, Kutner ME, Cui X, Benson C, et al. Efficacy and safety of LY3298176, a novel dual GIP and GLP-1 receptor agonist, in patients with type 2 diabetes: a randomised, placebo-controlled and active comparator-controlled phase 2 trial. Lancet. 2018 Nov;392(10160):2180–93.
12. Samms RJ, Coghlan MP, Sloop KW. How May GIP Enhance the Therapeutic Efficacy of GLP-1? Trends Endocrinol Metab. 2020;31(6):410–21.
13. Wilson JM, Nikooienejad A, Robins DA, Roell WC, Riesmeyer JS, Haupt A, et al. The dual glucose-dependent insulinotropic peptide and glucagon-like peptide-1 receptor agonist, tirzepatide, improves lipoprotein biomarkers associated with insulin resistance and cardiovascular risk in patients with type 2 diabetes. Diabetes, Obes Metab. 2020;
14. Thomas MK, Nikooienejad A, Bray R, Cui X, Wilson J, Duffin K, et al. Dual GIP and GLP-1 Receptor Agonist Tirzepatide Improves Beta-Cell Function and Insulin Sensitivity in Type 2 Diabetes. J Clin Endocrinol Metab [Internet]. 2020 Nov 25;26(6):1–15. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0924977X16300050%5Cnhttp://www.ncbi

.nlm.nih.gov/pubmed/27139079
15. Mori Y, Matsui T, Hirano T, Yamagishi SI. GIP as a potential therapeutic target for atherosclerotic cardiovascular disease– a systematic review. Int J Mol Sci. 2020;21(4):8–10.
16. Greenwell AA, Chahade JJ, Ussher JR. Cardiovascular biology of the GIP receptor. Peptides [Internet]. 2020;125(December 2019):170228. Available from: https://doi.org/10.1016/j.peptides.2019.170228
17. Nagashima M, Watanabe T, Terasaki M, Tomoyasu M, Nohtomi K, Kim-Kaneyama J, et al. Native incretins prevent the development of atherosclerotic lesions in apolipoprotein e knockout mice. Diabetologia. 2011;54(10):2649–59.
18. Lim DM, Park KY, Hwang WM, Kim JY, Kim BJ. Difference in protective effects of GIP and GLP-1 on endothelial cells according to cyclic adenosine monophosphate response. Exp Ther Med. 2017;13(5):2558–64.
19. Ojima A, Matsui T, Maeda S, Takeuchi M, Yamagishi S. Glucose-dependent insulinotropic polypeptide (GIP) inhibits signaling pathways of advanced glycation end products (AGEs) in endothelial cells via its antioxidative properties. Horm Metab Res [Internet]. 2012 Jun 11 [cited 2020 Dec 10];44(7):501–5. Available from: http://www.thieme-connect.de/DOI/DOI?10.1055/s-0032-1312595
20. Ding KH, Zhong Q, Xu J, Isales CM. Glucose-dependent insulinotropic peptide: Differential effects on hepatic artery vs. portal vein endothelial cells. Am J Physiol - Endocrinol Metab. 2004;286(5 49-5):773–9.
21. Berglund LM, Lyssenko V, Ladenvall C, Kotova O, Edsfeldt A, Pilgaard K, et al. Glucose-dependent insulinotropic polypeptide stimulates osteopontin expression in the vasculature via endothelin-1 and CREB. Diabetes. 2016;65(1):239–54.
22. Nogi Y, Nagashima M, Terasaki M, Nohtomi K, Watanabe T, Hirano T. Glucose-dependent insulinotropic polypeptide prevents the progression of macrophage-driven atherosclerosis in diabetic apolipoprotein E-null mice. PLoS One. 2012;7(4):1–8.
23. Kahles F, Liberman A, Halim C, Rau M, Möllmann J, Mertens RW, et al. The incretin hormone GIP is upregulated in patients with atherosclerosis and stabilizes plaques in ApoE−/− mice by blocking monocyte/macrophage activation. Mol Metab [Internet]. 2018;14(May):150–7. Available from: https://doi.org/10.1016/j.molmet.2018.05.014
24. Mori Y, Kushima H, Koshibu M, Saito T, Hiromura M, Kohashi K, et al. Glucose-dependent insulinotropic polypeptide suppresses peripheral arterial remodeling in Male mice. Endocrinology. 2018;159(7):2717–32.
25. Ussher JR, Campbell JE, Mulvihill EE, Baggio LL, Bates HE, McLean BA, et al. Inactivation of the Glucose-Dependent Insulinotropic Polypeptide Receptor Improves Outcomes following Experimental Myocardial Infarction. Cell Metab [Internet]. 2018;27(2):450-460.e6. Available from: https://doi.org/10.1016/j.cmet.2017.11.003
26. Jujić A, Nilsson PM, Atabaki-Pasdar N, Dieden A, Tuomi T, Franks PW, et al. Glucose-Dependent Insulinotropic Peptide in the High-Normal Range Is Associated With Increased Carotid Intima-Media Thickness. Diabetes Care [Internet]. 2020 Nov 18 [cited 2020 Dec 7];dc201318. Available from: http://care.diabetesjournals.org/lookup/doi/10.2337/dc20-1318
27. Jujić A, Atabaki-Pasdar N, Nilsson PM, Almgren P, Hakaste L, Tuomi T, et al. Glucose-dependent insulinotropic peptide and risk of cardiovascular events and mortality: a prospective study. Diabetologia [Internet]. 2020 May 23 [cited 2020 Feb 4];63(5):1043–54. Available from: http://link.springer.com/10.1007/s00125-020-05093-9
28. Gerstein HC, Colhoun HM, Dagenais GR, Diaz R, Lakshmanan M, Pais P, et al. Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-

blind, randomised placebo-controlled trial. Lancet. 2019;394(10193):121–30.
29. Hernandez AF, Green JB, Janmohamed S, D'Agostino RB, Granger CB, Jones NP, et al. Albiglutide and cardiovascular outcomes in patients with type 2 diabetes and cardiovascular disease (Harmony Outcomes): a double-blind, randomised placebo-controlled trial. Lancet. 2018;392(10157):1519–29.
30. Marso SP, Daniels GH, Frandsen KB, Kristensen P, Mann JFE, Nauck MA, et al. Liraglutide and cardiovascular outcomes in type 2 diabetes. N Engl J Med. 2016;375(4):311–22.
31. Marso SP, Bain SC, Consoli A, Eliaschewitz FG, Jodar E, Leiter LA, et al. Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. N Engl J Med. 2016;375(19):1834–44.
32. Scott RA, Amouyel P, Müller-nurasyid M. A genomic approach to therapeutic target validation identifies a glucose-lowering GLP1R variant protective for coronary heart disease. Sci Transl Med. 2016;8(341):1–13.
33. Gabe MBN, van der Velden WJC, Gadgaard S, Smit FX, Hartmann B, Bräuner-Osborne H, et al. Enhanced agonist residence time, internalization rate and signalling of the GIP receptor variant [E354Q] facilitate receptor desensitization and long-term impairment of the GIP system. Basic Clin Pharmacol Toxicol. 2020;126(S6):122–32.
34. Saxena R, Hivert MF, Langenberg C, Tanaka T, Pankow JS, Vollenweider P, et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. Nat Genet [Internet]. 2010 Feb 17 [cited 2019 May 31];42(2):142–8. Available from: http://www.nature.com/articles/ng.521
35. Lyssenko V, Eliasson L, Kotova O, Pilgaard K, Wierup N, Salehi A, et al. Pleiotropic effects of GIP on islet function involve osteopontin. Diabetes. 2011;60(9):2424–33.
36. Almgren P, Lindqvist A, Krus U, Hakaste L, Ottosson-Laakso E, Asplund O, et al. Genetic determinants of circulating GIP and GLP-1 concentrations. JCI Insight [Internet]. 2017 Nov 2 [cited 2019 Oct 11];2(21). Available from: https://insight.jci.org/articles/view/93306
37. Day N, Oakes S, Luben R, Khaw KT, Bingham S, Welch A, et al. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. Br J Cancer [Internet]. 1999 Jul [cited 2018 May 18];80 Suppl 1:95–103. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10466767
38. Riboli E. Nutrition and cancer: Background and rationale of the European prospective investigation into cancer and nutrition (EPIC). Ann Oncol [Internet]. 1992;3(10):783–91. Available from: https://doi.org/10.1093/oxfordjournals.annonc.a058097
39. Riboli E, Hunt K, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr [Internet]. 2002 Dec 2 [cited 2018 May 18];5(6b):1113. Available from: http://www.journals.cambridge.org/abstract_S1368980002001350
40. Lotta LA, Gulati P, Day FR, Payne F, Ongen H, van de Bunt M, et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. Nat Genet [Internet]. 2016;49(1):17–26. Available from: http://www.nature.com/doifinder/10.1038/ng.3714
41. Sudlow C, Gallacher J, Green J, Sprosen T, Pell J, Burton P, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLOS Med [Internet]. 2015 Mar 31 [cited 2018 Sep 17];12(3):e1001779. Available from: http://dx.plos.org/10.1371/journal.pmed.1001779
42. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet [Internet]. 2016 Oct 22 [cited 2018 Aug 31];48(10):1279–83. Available from:

http://www.nature.com/articles/ng.3643
43. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, et al. A global reference for human genetic variation [Internet]. Vol. 526, Nature. 2015 [cited 2018 May 19]. p. 68–74. Available from: https://www.nature.com/articles/nature15393.pdf
44. Walter K, Min JL, Huang J, Crooks L, Memari Y, McCarthy S, et al. The UK10K project identifies rare variants in health and disease. Nature [Internet]. 2015 Oct 14 [cited 2018 Aug 31];526(7571):82–9. Available from: http://www.nature.com/articles/nature14962
45. Williams SA, Kivimaki M, Langenberg C, Hingorani AD, Casas JP, Bouchard C, et al. Plasma protein patterns as comprehensive indicators of health. Nat Med [Internet]. 2019;25(12):1851–7. Available from: http://dx.doi.org/10.1038/s41591-019-0665-2
46. Lotta LA, Pietzner M, Stewart ID, Wittemans LBL, Li C, Bonelli R, et al. Cross-platform genetic discovery of small molecule products of metabolism and application to clinical outcomes. bioRxiv. 2020 Feb 4;2020.02.03.932541.
47. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. Williams SM, editor. PLoS Genet [Internet]. 2014 May 15 [cited 2018 May 21];10(5):e1004383. Available from: http://dx.plos.org/10.1371/journal.pgen.1004383
48. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res. 2017;26(5):2333–55.
49. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. Stat Methods Med Res. 2007;16(3):219–42.
50. Rubin DB. Multiple Imputation for Nonresponse in Surveys [Internet]. Rubin DB, editor. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 1987. (Wiley Series in Probability and Statistics). Available from: http://doi.wiley.com/10.1002/9780470316696
51. Foley CN, Staley JR, Breen PG, Sun BB, Kirk PDW, Burgess S, et al. A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. Nat Commun [Internet]. 2021 Dec 3;12(1):764. Available from: http://dx.doi.org/10.1038/s41467-020-20885-8
52. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88(1):76–82.
53. van der Harst P, Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. Circ Res [Internet]. 2018 Feb 2 [cited 2018 Jul 5];122(3):433–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29212778
54. Buse JB, Ginsberg HN, Bakris GL, Clark NG, Costa F, Eckel R, et al. Primary prevention of cardiovascular diseases in people with diabetes mellitus: a scientific statement from the American Heart Association and the American Diabetes Association. Circulation [Internet]. 2007 Jan 2 [cited 2020 Jul 17];115(1):114–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17192512

Table 1. Clusters of colocalised traits identified by the main and secondary analyses at recommended settings.

		Main Analysis					Secondary Analysis				
Locus	**Candidate LD (R^2) §**	**Colocalised Traits**	**PP Coloc***	**Candidate variant**	**PP explained**	**N variants**	**Colocalised Traits**	**PP Coloc***	**Candidate variant**	**PP explained**	**N variants**
GIPR	1	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	424	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	5,015
GIPR	0.69	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.99	rs4420638	1	424	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.97	rs429358	1	5,015
GIPR	1	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Glucose, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.97	rs1800437	1	424	GIP SOMAmer 16292_288, Hip circumference, 2hr Glucose adjBMI	0.91	rs1800437	0.68	5,015
GIPR	NA						BMI, Waist circumference	1	rs1800437	1	5,015
GIPR	NA	T2D, T2D adjBMI	0.98	rs8108269	0.99	424					
GIPR	NA						Triglycerides, Hip circumference adjBMI	0.98	rs5117	0.93	5,015

Abbreviations: GIPR, Glucose-dependent insulinotropic polypeptide receptor; LD, Linkage disequilibrium; PP, Posterior probability; coloc, Colocalisation; N, Number; variants, Single nucleotide polymorphisms; LDL, Low-density lipoprotein; CAD, Coronary artery disease; HDL, High-density lipoprotein; ApoB, Apolipoprotein B; Glucose, Non-fasted glucose; ApoA1, Apolipoprotein A1; adj., Adjusted for; WHR, Waist-to-hip ratio; BMI, Body mass index; T2D, Type 2 diabetes

* Trait clusters are reported at the recommended thresholds for Hyprcoloc: Prior 2 = 0.02; regional and alignment thresholds = 0.9

‡Blank rows for either analysis indicate a cluster not identified in the respective analysis

§The LD in R^2 between the candidate variants for the main and secondary analyses respectively

Table 2. Independent CAD variants identified using approximate conditional analysis.

Variant *	Chr:pos	Closest gene	EA	EAF	Marginal Beta (SE) †	Marginal P-value †	Conditional Beta (SE) ‡	Conditional P-value ‡	N	R^2 with rs1800437
rs429358	19:45411941	*APOE*	T	0.85	-0.09 (0.008)	2.86×10^{-27}	-0.08 (0.008)	5.87×10^{-23}	286,423	0.001
rs7412	19:45412079	*APOE*	T	0.08	-0.14 (0.011)	1.66×10^{-35}	-0.12 (0.011)	1.58×10^{-28}	275,803	0.004
rs11673093	19:45742094	*EXOC3L2*	A	0.26	0.04 (0.007)	4.11×10^{-11}	0.04 (0.007)	3.09×10^{-10}	300,789	0
rs1964272	19:46190268	*SNRPD2*	A	0.48	-0.03 (0.006)	9.65×10^{-9}	-0.03 (0.006)	1.87×10^{-7}	299,519	0.27

Abbreviations: Chr, Chromosome; pos, Position; EA, Effect allele; EAF, Effect allele frequency; SE, Standard error; N, Number of participants; R^2, Linkage disequilibrium estimate

*The independent CAD variants in the 1Mb region either side of E354 are shown

†Log odds ratios from the original GWAS summary statistics

‡Log odds ratios from the joint model fitted by GCTA

Table 3. Conditioning each of the traits associated with E354 at nominal significance from the 2SMR analysis on independent SNPs for each trait. Estimates of 2-hour glucose, total cholesterol and BMI were not included in this table as the independent signal selection showed that E354 was one of the independent variants.

	2SMR result		Conditional result		Independent variant	
Trait	**Beta (SE)**	**P-value**	**Beta (SE)**	**P-value‡**	**Conditioned on***	**LD with rs1800437†**
T2D	-0.03 (0.007)	$7x10^{-5}$	-0.03 (0.008)	$4x10^{-4}$	rs3810291	0.001
T2DadjBMI	-0.07 (0.009)	$2x10^{-14}$	-0.02 (0.009)	0.04	rs2238689	0.363
CAD	0.03 (0.007)	$2x10^{-6}$	0.01 (0.007)	0.06	rs1964272	0.269
SVS	-0.08 (0.029)	0.009	-0.04 (0.029)	0.12	rs1964272	0.269
Non-fasted plasma glucose	0.02 (0.003)	$3x10^{-8}$	0.01 (0.003)	0.05	rs1964272	0.269
HbA1c	-0.01 (0.003)	$1x10^{-7}$	-0.0003 (0.003)	0.92	rs9676912	0.356
ApoA1	0.01 (0.003)	$3x10^{-6}$	0.002 (0.003)	0.37	rs2238689	0.363
HDL	0.02 (0.003)	$7x10^{-9}$	0.003 (0.003)	0.31	rs2238689	0.363
ApoB	0.02 (0.002)	$5x10^{-13}$	0.01 (0.002)	$2x10^{-5}$	rs7412	0.004
LDL	0.02 (0.003)	$2x10^{-16}$	0.016 (0.003)	$1x10^{-8}$	rs7412	0.004
Triglycerides	-0.01 (0.003)	$2x10^{-5}$	-0.01 (0.003)	$5x10^{-5}$	rs4803936	0.001
CRP	-0.01 (0.002)	0.02	-0.004 (0.002)	0.07	rs7412	0.004
Albumin	-0.01 (0.003)	$6x10^{-6}$	-0.01 (0.003)	0.001	rs35114617	0.061
Creatinine	-0.02 (0.002)	$1x10^{-11}$	-0.02 (0.002)	$3x10^{-11}$	rs7412	0.004
QPCTL	-0.07 (0.016)	$9x10^{-6}$	0.01 (0.016)	0.48	rs1964272	0.269
Secretoglobin family 3A member 1	-0.08 (0.017)	$6x10^{-7}$	-0.04 (0.017)	0.01	rs61703905	0.1

Abbreviations: SE, Standard error; T2D, Type 2 diabetes; adjBMI, Adjusted for BMI; CAD, Coronary artery disease; SVS, Small vessel stroke; HbA1c, Glycated haemoglobin; Apo, Apolipoprotein; HDL, High-density lipoprotein; LDL, Low-density lipoprotein; CRP, C-reactive protein; QPCTL, Glutaminyl-peptide cyclotransferase like* The independent variant showing the greatest attenuation of the E354 association estimate with the respective trait

† LD estimates are in R2 and are quoted from 5 European populations in the LDlink database v4.1.0

‡ A nominal significance threshold of $P \leq 0.05$ was used to ascertain significance for the conditional results

Figure legends:

Fig. 1. Associations between E354 (rs1800437) and cardiometabolic disease endpoints, glycaemic traits, cardiovascular risk factors and lipids, anthropometric traits and biomarkers estimated using 2SMR. **(A)** Associations with cardiometabolic disease endpoints are shown in blue and are represented as odds ratios (95% CI) for each disease per copy of rs1800437. (**B)** Associations with glycaemic traits are shown in orange, cardiovascular and lipid traits in green, anthropometric traits and biomarkers are shown in yellow and purple respectively. Estimates are represented as beta (95% CI) for each outcome per copy of rs1800437. All traits are in SD units aside from fasting and 2-hour glucose which are in mmol/L, fasting insulin in log (pmol/L) and HbA1c in mmol/mol. Fold change insulin represents the fold change in insulin levels between fasting to 2-hour measures. A Bonferroni significance threshold of $P \leq 0.001$ was used, accounting for the number of traits tested.

Abbreviations: 2SMR, Two sample Mendelian randomisation; OR, Odds ratio; CI, Confidence interval; N, Number; BMI, Body mass index; adj., Adjusted; HbA1c, Glycated haemoglobin; ApoA1, Apolipoprotein A1; ApoB, Apolipoprotein B; HDL, High-density lipoprotein; LDL, Low-density lipoprotein; γ, Gamma.

* HbA1c estimates are in SD units per copy of E354. The corresponding clinical units in %(mmol/mol) are as follows: -2.15% (95% CI, -2.15, -2.14) and -0.07mmol/mol (95% CI, -0.07, -0.06).

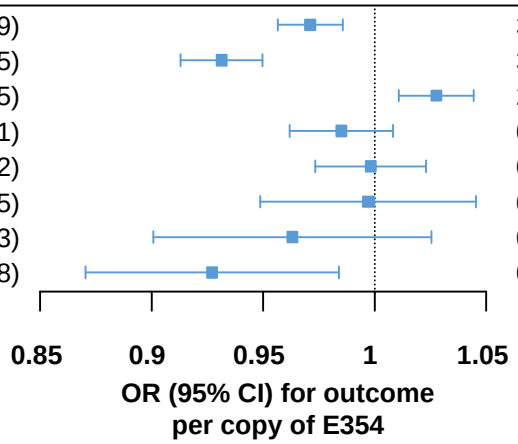
Fig. 2. Similarity heatmap for each cluster at the *GIPR* locus across prior and threshold permutations. Traits that were estimated to colocalise are clustered together. Darker colours represent traits which were estimated to colocalise more often across prior and threshold permutations (prior 2: 0.02, 0.01 and 0.001; thresholds: 0.5, 0.6, 0.7, 0.8 and 0.9). **(A)** Main analysis. (**B)** Secondary analysis

Abbreviations: LDL, Low-density lipoprotein; CAD, Coronary artery disease; HDL, High-density lipoprotein; ApoB, Apolipoprotein B; Glucose, Non-fasted glucose; ApoA1, Apolipoprotein A1; adj., Adjusted for; WHR, Waist-to-hip ratio; BMI, Body mass index; T2D, Type 2 diabetes; WC, Waist circumference; HC, Hip circumference

Fig. 3. Regional association plots depicting CAD lead variants in the *GIPR* region. (**A)** The independent CAD lead variants in the *GIPR* region are labelled and their respective associations with CAD are shown before conditional analysis. The region around rs1800437 (E354) is expanded in the red insert to show the LD and proximity of rs1964272 to rs1800437. (**B).** The associations of variants in the *GIPR* region after conditioning on rs1964272. The region around rs1800437 (E354) is expanded in the red insert to show the attenuation of the E354 signal when conditioned on rs1964272.

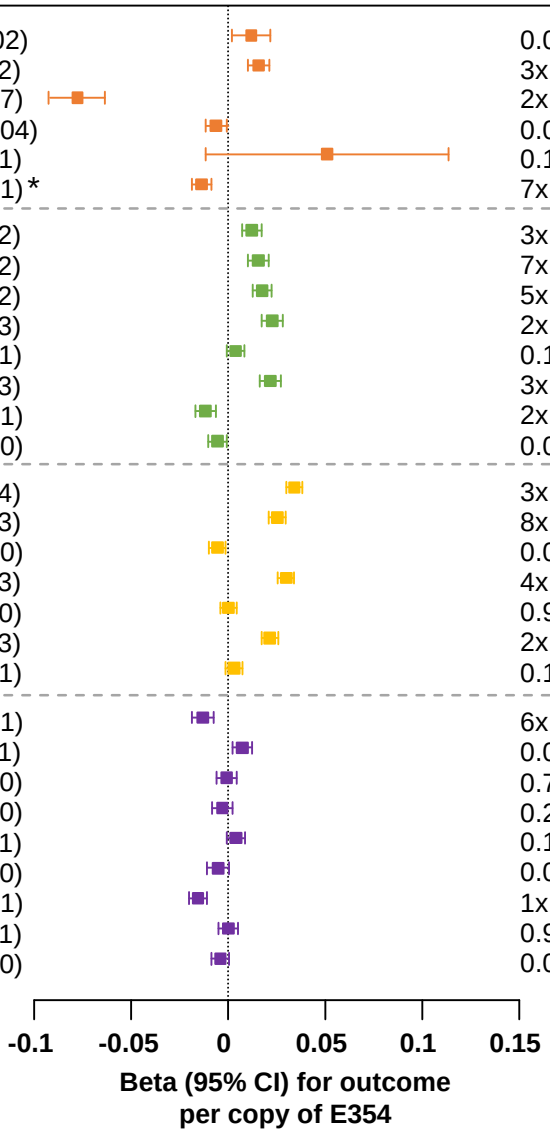
A.

Outcome	Cases	Controls	OR (95% CI)	P-value
Type 2 diabetes	74,124	842,006	0.97 (0.96, 0.99)	$3x10^{-4}$
Type 2 diabetes (BMI adj.)	74,124	842,006	0.93 (0.91, 0.95)	$3x10^{-14}$
Coronary heart disease	85,358	550,908	1.03 (1.02, 1.05)	$2x10^{-6}$
Any Stroke	40,585	406,111	0.99 (0.96, 1.01)	0.2
Any Ischemic Stroke	34,217	406,111	1.00 (0.97, 1.02)	0.89
Cardioembolic Stroke	7,193	406,111	1.00 (0.95, 1.05)	0.91
Large Artery Stroke	4,373	406,111	0.96 (0.90, 1.03)	0.24
Small Vessel Stroke	5,386	406,111	0.93 (0.87, 0.98)	0.009

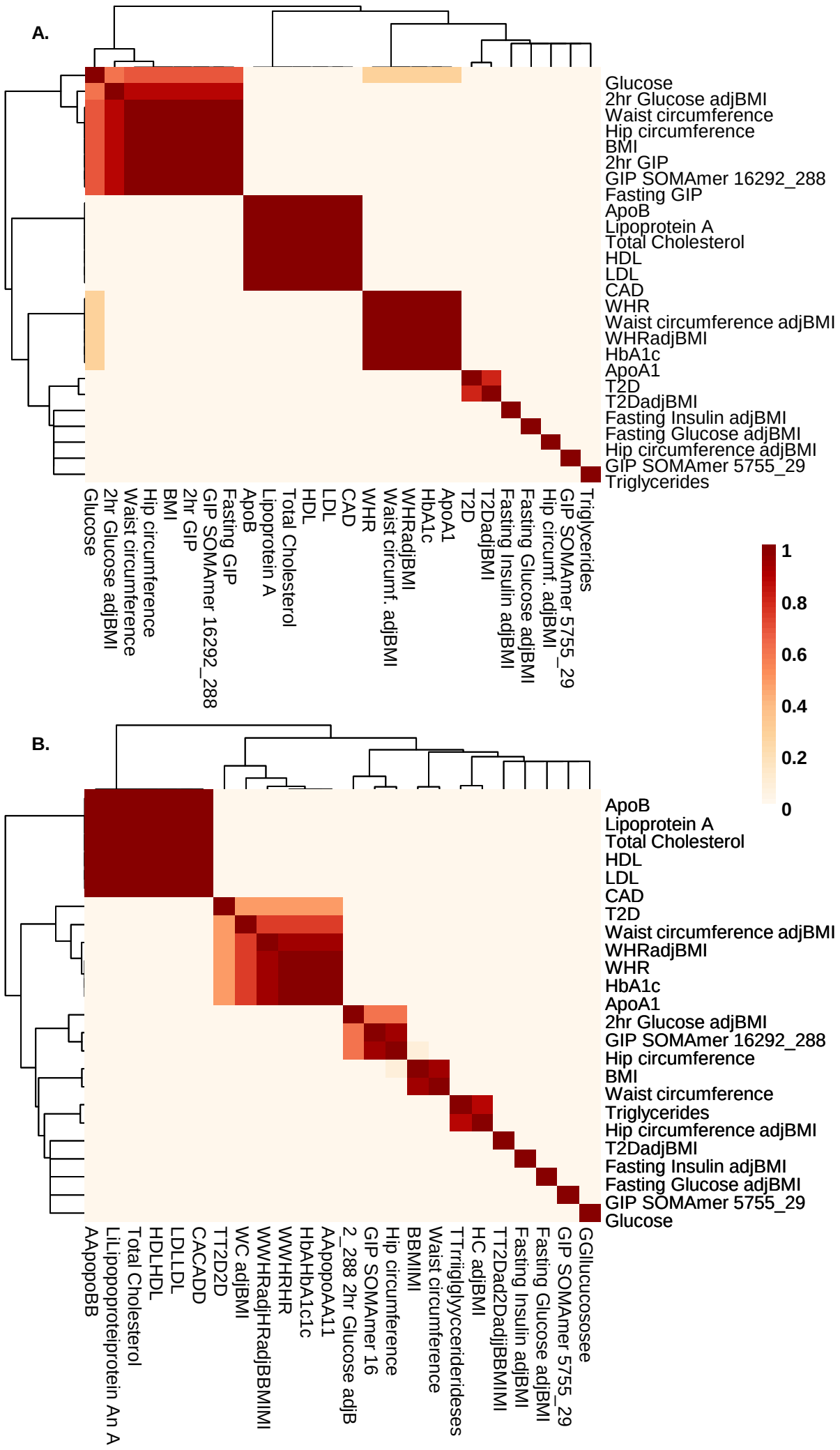


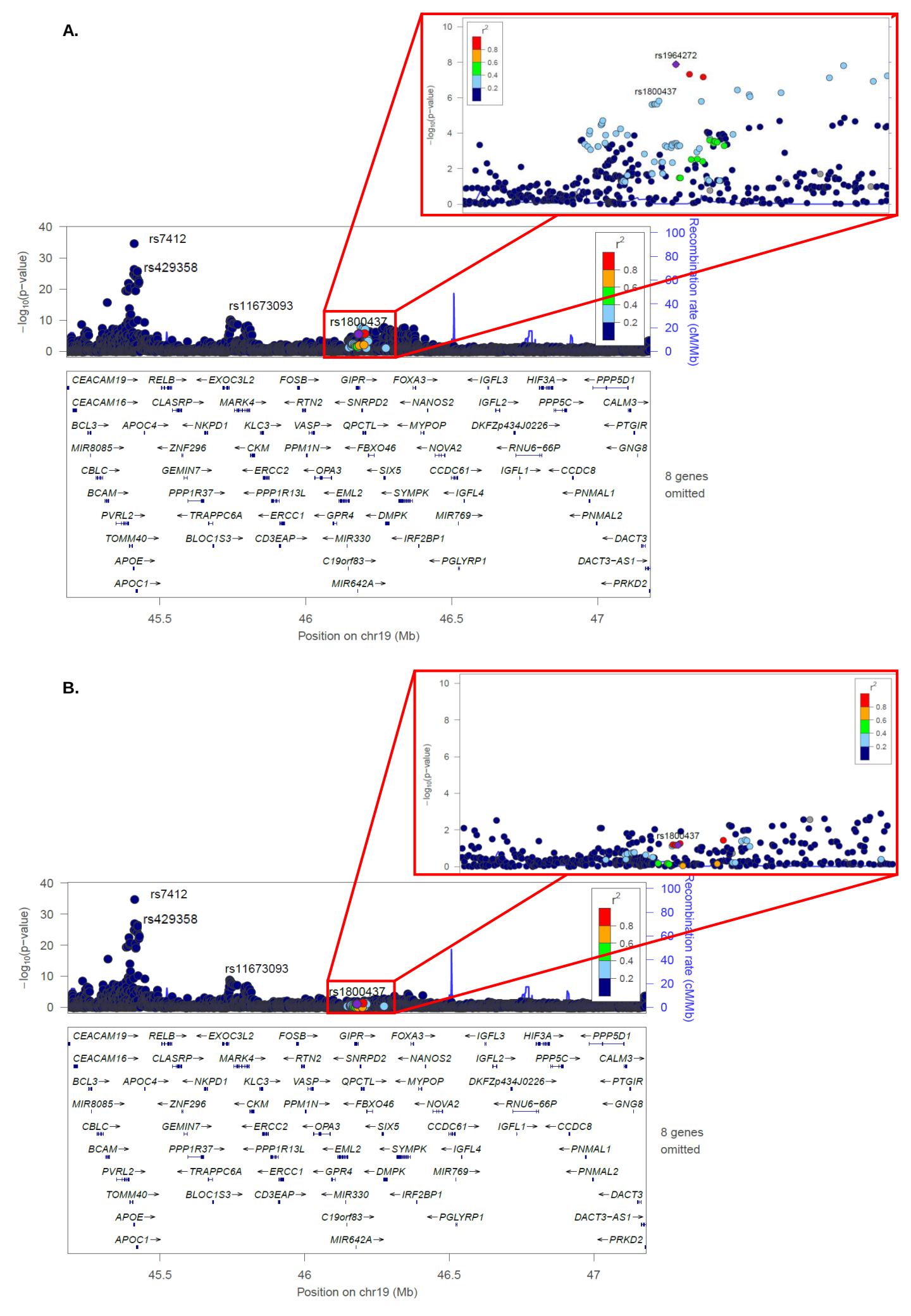
B.

Outcome	N participants	Beta (95% CI)	P-value
Fasting glucose (BMI adj.)	51,750	0.01 ($2x10^{-4}$, 0.02)	0.05
Non-fasted plasma glucose	413,905	0.02 (0.01, 0.02)	$3x10^{-8}$
2-hr glucose (BMI adj.)	41,888	-0.09 (-0.11, -0.07)	$2x10^{-15}$
Fasting insulin (BMI adj.)	51,750	-0.003 (-0.01, -0.004)	0.04
Corrected insulin response	5,318	0.05 (-0.01, 0.11)	0.1
HbA1c	451,782	-0.01 (-0.02, -0.01) *	$7x10^{-8}$
ApoA1	412,328	0.01 (0.01, 0.02)	$3x10^{-6}$
HDL	450,957	0.02 (0.01, 0.02)	$7x10^{-9}$
ApoB	448,859	0.02 (0.01, 0.02)	$5x10^{-13}$
LDL	375,774	0.02 (0.02, 0.03)	$2x10^{-16}$
Lipoprotein A	406,825	0.00 (-0.00, 0.01)	0.12
Total cholesterol	377,031	0.02 (0.02, 0.03)	$3x10^{-15}$
Triglycerides	450,625	-0.01 (-0.02, -0.01)	$2x10^{-5}$
C-reactive protein	465,067	-0.01 (-0.01, -0.00)	0.02
BMI	738,628	0.03 (0.03, 0.04)	$3x10^{-59}$
Hip circumference	568,765	0.03 (0.02, 0.03)	$8x10^{-28}$
Hip circumference (BMI adj.)	633,860	-0.01 (-0.01, -0.00)	0.01
Waist circumference	654,577	0.03 (0.03, 0.03)	$4x10^{-42}$
Waist circumference (BMI adj.)	654,253	0.00 (-0.00, 0.00)	0.96
Waist-to-hip ratio	636,672	0.02 (0.02, 0.03)	$2x10^{-22}$
Waist-to-hip ratio (BMI adj.)	636,282	0.00 (-0.00, 0.01)	0.17
Albumin	415,714	-0.01 (-0.02, -0.01)	$6x10^{-6}$
Alkaline phosphatase	450,743	0.01 (0.00, 0.01)	0.006
Alanine aminotransferase	452,291	-0.00 (-0.01, 0.00)	0.77
Aspartate transaminase	450,594	-0.00 (-0.01, 0.00)	0.24
Bilirubin	448,652	0.00 (-0.00, 0.01)	0.1
Calcium	414,173	-0.01 (-0.01, 0.00)	0.07
Creatinine	451,942	-0.02 (-0.02, -0.01)	$1x10^{-11}$
γ-glutamyl transpeptidase	450,745	0.00 (-0.00, 0.01)	0.98
Urate	451,665	-0.00 (-0.01, 0.00)	0.08



Disease endpoints
Glycaemic traits
Cardiovascular risk factors and lipids
Anthropometric traits
Biomarkers


A.
Glucose
2hr Glucose adjBMI
Waist circumference
Hip circumference
BMI
2hr GIP
GIP SOMAmer 16292_288
Fasting GIP
ApoB
Lipoprotein A
Total Cholesterol
HDL
LDL
CAD
WHR
Waist circumference adjBMI
WHRadjBMI
HbA1c
ApoA1
T2D
T2DadjBMI
Fasting Insulin adjBMI
Fasting Glucose adjBMI
Hip circumference adjBMI
GIP SOMAmer 5755_29
Triglycerides
1
0.8
0.6
0.4
0.2
0
B.
ApoB
Lipoprotein A
Total Cholesterol
HDL
LDL
CAD
T2D
Waist circumference adjBMI
WHRadjBMI
WHR
HbA1c
ApoA1
2hr Glucose adjBMI
GIP SOMAmer 16292_288
Hip circumference
BMI
Waist circumference
Triglycerides
Hip circumference adjBMI
T2DadjBMI
Fasting Insulin adjBMI
Fasting Glucose adjBMI
GIP SOMAmer 5755_29
Glucose


A.
rs1964272
rs1800437
rs7412
rs429358
rs11673093
rs1800437
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
CEACAM19→ RELB→ ←EXOC3L2 FOSB→ GIPR→ FOXA3→ ←IGFL3 HIF3A→ ←PPP5D1
CEACAM16→ CLASRP→ MARK4→ ←RTN2 ←SNRPD2 ←NANOS2 IGFL2→ PPP5C→ CALM3→
BCL3→ APOC4→ ←NKPD1 KLC3→ VASP→ QPCTL→ ←MYPOP DKFZp434J0226→ ←PTGIR
MIR8085→ ←ZNF296 ←CKM PPM1N→ ←FBXO46 ←NOVA2 ←RNU6−66P ←GNG8
CBLC→ GEMIN7→ ←ERCC2 ←OPA3 ←SIX5 CCDC61→ IGFL1→ ←CCDC8
BCAM→ PPP1R37→ ←PPP1R13L ←EML2 ←SYMPK ←IGFL4 ←PNMAL1
PVRL2→ ←TRAPPC6A ←ERCC1 ←GPR4 ←DMPK MIR769→ ←PNMAL2
TOMM40→ BLOC1S3→ CD3EAP→ ←MIR330 ←IRF2BP1 ←DACT3
APOE→ C19orf83→ ←PGLYRP1 DACT3−AS1→
APOC1→ MIR642A→ ←PRKD2
8 genes omitted
45.5
46
46.5
47
Position on chr19 (Mb)
B.
rs1800437
rs7412
rs429358
rs11673093
rs1800437
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
CEACAM19→ RELB→ ←EXOC3L2 FOSB→ GIPR→ FOXA3→ ←IGFL3 HIF3A→ ←PPP5D1
CEACAM16→ CLASRP→ MARK4→ ←RTN2 ←SNRPD2 ←NANOS2 IGFL2→ PPP5C→ CALM3→
BCL3→ APOC4→ ←NKPD1 KLC3→ VASP→ QPCTL→ ←MYPOP DKFZp434J0226→ ←PTGIR
MIR8085→ ←ZNF296 ←CKM PPM1N→ ←FBXO46 ←NOVA2 ←RNU6−66P ←GNG8
CBLC→ GEMIN7→ ←ERCC2 ←OPA3 ←SIX5 CCDC61→ IGFL1→ ←CCDC8
BCAM→ PPP1R37→ ←PPP1R13L ←EML2 ←SYMPK ←IGFL4 ←PNMAL1
PVRL2→ ←TRAPPC6A ←ERCC1 ←GPR4 ←DMPK MIR769→ ←PNMAL2
TOMM40→ BLOC1S3→ CD3EAP→ ←MIR330 ←IRF2BP1 ←DACT3
APOE→ C19orf83→ ←PGLYRP1 DACT3−AS1→
APOC1→ MIR642A→ ←PRKD2
8 genes omitted
45.5
46
46.5
47
Position on chr19 (Mb)

Supplemental material

Genetically predicted glucose-dependent insulinotropic polypeptide (GIP) levels and cardiovascular disease risk are driven by distinct causal variants in the *GIPR* region

Index

Table S1. Summary of the participating studies.

Outcome type	Outcome	Cases overall, N	Non-cases (for case-control studies) or participants (for continuous trait studies) overall, N	Participating study	PubMed ID for cohort description
Disease outcomes	Type 2 diabetes *	74,124	842,006	DIAMANTE	30297969
	Type 2 diabetes (BMI adjusted) *	74,124	842,006	DIAMANTE	30297969
	Coronary artery disease *	34,541	261,984	CARDIoGRAMplusC4D, UK Biobank	29212778
	Any Stroke *	40,585	406,111	MEGASTROKE	29531354
	Any Ischemic Stroke *	34,217			
	Cardioembolic Stroke *	7,193			
	Large Artery Stroke *	4,373			
	Small Vessel Stroke *	5,386			
	Abdominal Aortic Aneurysm *	1,094	366,492	UK Biobank	31756303
	Atrial Fibrillation *	16,945	350,641		
	Aortic Valve Stenosis *	2,244	365,342		
	Coronary Artery Disease *	29,278	338,308		
	Deep Vein Thrombosis *	9,454	358,132		
	Haemorrhagic Stroke (all) *	1,981	365,605		
	Heart Failure *	6,712	360,874		
	Ischaemic Cerebrovascular Disease (all) *	8,084	359,502		
	Pulmonary Embolism *	6,148	361,438		
	Peripheral Vascular Disease *	3,415	364,171		
	Thoracic Aortic Aneurysm *	347	367,239		
	Transient Ischaemic Attack *	3,962	363,624		
	Intracerebral Haemorrhage *	1,064	366,522		
	Subarachnoid Haemorrhage *	1,084	366,502		
	Ischaemic Stroke *	4,602	362,984		
	Ischaemic stroke plus haemorrhagic stroke plus unknown stroke (but not TIA) *	9,652	357,934		
	Venous Thromboembolism (all) *	14,097	353,489		
Glycaemic outcomes	Fasting glucose (BMI adjusted) *		51,750	MAGIC	22581228
	Non-fasted plasma glucose †		413,905	UK Biobank; InterAct	25826379
	2-hr glucose (BMI adjusted) *		41,888	MAGIC	20081857
	Fasting insulin (BMI adjusted) *		51,750	MAGIC	22581228
	Corrected insulin response *		5,318	MAGIC	24699409
	HbA1C †		451,782	UK Biobank; InterAct	25826379
Cardiovascular and lipid-related outcomes	Apolipoprotein A1 †		412,328	UK Biobank; InterAct	25826379
	High-density lipoprotein †		450,957		
	Apolipoprotein B †		448,859		
	Low-density lipoprotein †		375,774		
	Lipoprotein A †		406,825		
	Total cholesterol †		377,031		
	Triglycerides †		450,625		

Outcome type	Outcome	Cases overall, N	Non-cases (for case-control studies) or participants (for continuous trait studies) overall, N	Participating study	PubMed ID for cohort description
	C-reactive protein †		465,067		
Anthropometric outcomes	BMI †		738,628	GIANT, UK Biobank	25673413; 25826379
	Hip circumference †		568,765		
	Hip circumference (BMI adjusted) †		633,860		
	Waist circumference †		654,577		
	Waist circumference (BMI adjusted) †		654,253		
	Waist-to-hip ratio †		636,672		
	Waist-to-hip ratio (BMI adjusted) †		636,282		
Additional biomarker outcomes	Albumin †		415,714	UK Biobank; InterAct	25826379
	Alkaline phosphatase †		450,743		
	Alanine aminotransferase †		452,291		
	Aspartate transaminase †		450,594		
	Bilirubin †		448,652		
	Calcium †		414,173		
	Creatinine †		451,942		
	Gamma-glutamyl transpeptidase †		450,745		
	Urate †		451,665		
Regional adiposity outcomes (measured by bio-impedance)	Android fat mass †		435,387	UK Biobank	25826379
	Arms fat mass †				
	Gynoid fat mass †				
	Legs fat mass †				
	Peripheral fat mass †				
	Subcutaneous fat mass †				
	Total fat mass †				
	Trunk fat mass †				
	Visceral fat mass †				
	Appendicular lean mass †				
	Android lean mass †				
	Arms lean mass †				
	Gynoid lean mass †				
	Legs lean mass †				
	Total lean mass †				
	Trunk lean mass †				
Plasma proteins	4,979 proteins †		10,708	Fenland	27841877
Metabolites	1,008 metabolites †		11,539	EPIC-Norfolk	10466767
GIP measures	Fasting and 2hr GIP *		7,828	MDC and PPP-Botnia	29093273

*Publicly available datasets, the phenotype definitions of which can be found in the original studies (PMID provided)
†In-house datasets, the phenotype definitions of which can be found in Table S3

Table S2. Study participants.

Study	Fenland *	EPIC-Norfolk *	UK Biobank *
Participants, N	10,708	11,539 ‡	452,197
Age at baseline, mean years (SD)	49 (7)	60 (9)	57 (8)
Women, N (%)	5,714 (53)	6,198 (54)	245,277 (54)
Men, N (%)	4,994 (47)	5,341 (46)	206,883 (46)
BMI in kg/m2, mean (SD)	26.9 (4.9)	26.2 (3.7)	27.4 (4.8)
Waist-to-hip ratio, mean (SD)	0.74 (0.08)	0.86 (0.09)	0.87 (0.09)
Systolic blood pressure in mmHg, mean (SD)	123 (15)	136 (18)	138 (19)
Diastolic blood pressure in mmHg, mean (SD)	74 (10)	82 (11)	82 (10)
Fasting glucose in log-pg/mL, median (IQR) †	1.57 (1.50, 1.63)	N/A	N/A
2-hr glucose in log-pg/mL, median (IQR) †	1.63 (1.44, 1.79)	N/A	N/A
Fasting insulin in log-pg/mL, median (IQR) †	3.66 (3.29, 4.06)	N/A	N/A
Study stage	2SMR, colocalisation, conditional analyses	2SMR	2SMR §
Participants with prevalent T2D, N	N/A ‖	N/A ¶	22,610

Abbreviations: N/A, not available; N, number of participants; SD, standard deviation; BMI, body mass index; mmHg; Millimetres Mercury; pg; Picograms; mL, Millilitres; IQR, Interquartile range

*The relevant outcomes that make use of data from each study are described in **Table S1**

†Glycaemic measures from Epic-Norfolk and UK Biobank were not used in this study

‡Participants used in the plasma metabolite GWAS sample

§The publicly available GWAS dataset[1] included UK Biobank samples, however, this table only describes samples used for in-house GWAS analyses.

‖Participants with prevalent T2D were excluded from the study cohort as part of the exclusion criteria

¶Only participants from the quasi-randomly selected samples were used, excluding participants with prevalent T2D

Table S3: Description of the GWAS analyses for in-house datasets and the quality control procedures applied.

Cohort	Trait	Measurement	Transformations applied	Covariates	Variant-level QC ‡‡
Fenland*	Plasma proteins†	Described in main text	Rank-based inverse normal within each genotyping subset	Age, sex, sample collection site and 10 genetic principal components	MAF < 0.001, Imputation quality < 0.4, HWE P-value < $1x10^{-7}$ ||
Fenland	Fasting insulin, fasting glucose, 2hr glucose‡	Fasting glucose and insulin were measured in whole blood after overnight fast. 2hr glucose was measured in plasma two-hours after a 75-gram oral glucose challenge. Glucose levels were quantified using the Dimension RxL Integrated Chemistry System (Siemens, Germany). Insulin levels were quantified using the 1235 AutoDELFIA automatic immunoassay system using a two-step time resolved fluorometric assay (Kit No. B080-101, Perkin Elmer, USA). Individuals were excluded if they had prevalent type 1 or type 2 diabetes (defined by physician diagnosis); reported use of diabetes medication(s); or had fasting glucose levels >=7 mmol/L, 2-hr glucose levels >=11.1mmol/L, or HbA1c >= 6.5%.	Fasting and 2hr glucose: untransformed; fasting insulin: natural log	Age, sex, BMI and the first 10 principal components§	Call rate (< 95%), HWE P<$1x10^{-6}$, imputation quality < 0.4, MAF < 1%, tri-allelic, MAC<3, SE<0, SE>10, missing beta or SE or imputation quality estimate ¶
EPIC-Norfolk*	Plasma metabolites	Described in main text	Natural log-transformed and winsorised to 5 SD	Age, sex and measurement batch	Imputation quality < 0.4, MAC < 10, HWE P<$1x10^{-6}$, abs(beta) > 10, SE<0, SE>10, MAF < 0.0001 #
UK Biobank and InterAct*	ApoA1, HbA1c, HDL, ApoB, LDL, LpA, Total cholesterol, Triglycerides, CRP, Albumin, ALP, ALT, AST, Bilirubin, Calcium, Creatinine, γ-GGT, Urate	All biomarkers in InterAct, except HbA1c, were measured using a Cobas® (Roche Diagnostics, Mannheim, Germany) assay on a Roche Hitachi Modular P analyser. HbA1c was measured on erythrocyte samples using a Tosoh (HLC-723G8) assay on a Tosoh G8 analyser.	Raw measures regressed on age, age2, sex, centre and 10 genetic principal components to generate residuals which were then rank-based inverse normal transformed within each study **	Age, age2, sex, aliquot, genotyping chip, lipid lowering medication and the top 40 principal components	Imputation quality < 0.4, MAC < 10, HWE P<$1x10^{-6}$, abs(beta) > 10, SE<0, SE>10, MAF < 0.0001

UK Biobank and GIANT*	BMI, Hip circumference, Hip circumference adj. BMI, Waist circumference, Waist circumference adj. BMI, WHR, WHRadjBMI	In UK Biobank, weight was measured using a calibrated electronic scale (TANITA model BC-418 MA; Tanita, Tokyo, Japan). Height was measured with a wall-mounted stadiometer (SECA 202; Seca, Birmingham, United Kingdom). BMI (in kg/m2) was calculated as weight divided by height squared. Waist and hip circumferences were measured with a non-stretchable sprung tape measure (Wessex tape, London, United Kingdom). WHR was the ratio between the waist and hip circumferences.	Residuals were generated for each sex independently by regressing each outcome against age, age^2, study-specific covariates and BMI (if applicable) then rank-based inverse normal transformed	Age, sex, genotyping chip, and the top 40 principal components.	Imputation quality < 0.4, MAC < 10, HWE P<1x10^{-6}, abs(beta) > 10, SE<0, SE>10, MAF < 0.0001 ††
UK Biobank	Bio-impedance	Tanita BC418MA body composition analyser (Amsterdam, The Netherlands)	Natural log-transformed and regressed on age (and total fat mass or height^2 – if adjusted) in each sex separately to generate residuals. Residuals were then rank-based inverse normal transformed	Age, sex, genotyping chip, and the top 40 principal components.	Imputation quality < 0.4, MAC < 10, HWE P<1x10^{-6}, abs(beta) > 10, SE<0, SE>10, MAF < 0.0001

Abbreviations: MAF, Minor allele frequency; MAC, Minor allele count; HWE, Hardy-Weinberg equilibrium; SE, Standard error; BMI, Body mass index; adj., Adjusted for; WHR, Waist-to-hip ratio; ApoA1, Apolipoprotein A1; HbA1c, Glycated haemoglobin; HDL, High density lipoprotein cholesterol; ApoB, Apolipoprotein B; LDL, Low density lipoprotein cholesterol; LpA, Lipoprotein A; CRP, C-reactive protein; ALP, Alkaline phosphatase; ALT, Alanine aminotransferase; AST, Aspartate transaminase; γ-GGT, Gamma-glutamyl transpeptidase

*Studies or genotyping subsets were meta-analysed using inverse variance weighted fixed effect meta-analysis in METAL

†GWAS conducted using BGENIE v1.3

‡GWAS conducted using SNPTEST v2.4.1

§Fasting insulin and fasting glucose were also adjusted for age^2

‖Only variants present in the largest genotyping subset were taken forward

¶Only samples genotyped using the Affymetrix UK Biobank Axiom Array were included

#If BOLT-LMM failed, related individuals were excluded (IBD > 0.185) and linear regression models were run using SNPTEST v2.4.1, while also adjusting for the top 4 principal components

**Traits measured in UK Biobank were also rank-based inverse normal transformed within each respective aliquot.

††Variant-level QC only applies to UK Biobank, as GIANT data was publicly available

‡‡Variants were excluded if they were outside of the thresholds listed

Table S4: Clusters of colocalised traits identified by the main analysis across the permutations of prior 2 and the regional and alignment thresholds (prior 2: 0.02, 0.01 and 0.001; thresholds: 0.5, 0.6, 0.7, 0.8 and 0.9). A total of 424 variants were included.

Locus	Colocalised traits	PP Coloc	Candidate variant	PP explained	Prior 2	Regional and alignment threshold
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.02	0.5
GIPR	Glucose, HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.7248	rs4420638	1	0.02	0.5
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.9782	rs1800437	1	0.02	0.5
GIPR	T2D, T2DadjBMI	0.979	rs8108269	0.9967	0.02	0.5
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.02	0.6
GIPR	Glucose, HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.7248	rs4420638	1	0.02	0.6
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.9782	rs1800437	1	0.02	0.6
GIPR	T2D, T2DadjBMI	0.979	rs8108269	0.9967	0.02	0.6
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.02	0.7
GIPR	Glucose, HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.7248	rs4420638	1	0.02	0.7
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.9782	rs1800437	1	0.02	0.7
GIPR	T2D, T2DadjBMI	0.979	rs8108269	0.9967	0.02	0.7
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.02	0.8
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9994	rs4420638	1	0.02	0.8
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Glucose, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.9737	rs1800437	1	0.02	0.8
GIPR	T2D, T2DadjBMI	0.979	rs8108269	0.9967	0.02	0.8
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.02	0.9
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9994	rs4420638	1	0.02	0.9
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Glucose, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.9737	rs1800437	1	0.02	0.9
GIPR	T2D, T2DadjBMI	0.979	rs8108269	0.9967	0.02	0.9
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.01	0.5
GIPR	Glucose, HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.5725	rs4420638	1	0.01	0.5
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.9584	rs1800437	1	0.01	0.5
GIPR	T2D, T2DadjBMI	0.9589	rs8108269	0.9967	0.01	0.5
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.01	0.6
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9989	rs4420638	1	0.01	0.6
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Glucose, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.9499	rs1800437	1	0.01	0.6
GIPR	T2D, T2DadjBMI	0.9589	rs8108269	0.9967	0.01	0.6

GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.01	0.7
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9989	rs4420638	1	0.01	0.7
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Glucose, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.9499	rs1800437	1	0.01	0.7
GIPR	T2D, T2DadjBMI	0.9589	rs8108269	0.9967	0.01	0.7
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.01	0.8
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9989	rs4420638	1	0.01	0.8
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Glucose, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.9499	rs1800437	1	0.01	0.8
GIPR	T2D, T2DadjBMI	0.9589	rs8108269	0.9967	0.01	0.8
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.01	0.9
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9989	rs4420638	1	0.01	0.9
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Glucose, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.9499	rs1800437	1	0.01	0.9
GIPR	T2D, T2DadjBMI	0.9589	rs8108269	0.9967	0.01	0.9
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.001	0.5
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9889	rs4420638	1	0.001	0.5
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Glucose, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.6397	rs1800437	1	0.001	0.5
GIPR	T2D, T2DadjBMI	0.6999	rs8108269	0.9967	0.001	0.5
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.001	0.6
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9889	rs4420638	1	0.001	0.6
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Glucose, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.6397	rs1800437	1	0.001	0.6
GIPR	T2D, T2DadjBMI	0.6999	rs8108269	0.9967	0.001	0.6
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.001	0.7
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9889	rs4420638	1	0.001	0.7
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Glucose, Hip circumference, Waist circumference, 2hr Glucose adjBMI	0.6397	rs1800437	1	0.001	0.7
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.001	0.8
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9889	rs4420638	1	0.001	0.8
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Glucose, Hip circumference, Waist circumference	0.8098	rs1800437	1	0.001	0.8
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.001	0.9
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9889	rs4420638	1	0.001	0.9
GIPR	GIP SOMAmer 16292_288, Fasting GIP, 2hr GIP, BMI, Hip circumference, Waist circumference	0.934	rs1800437	1	0.001	0.9

Abbreviations: PP, Posterior probability; N, Number; LDL, Low-density lipoprotein; CAD, Coronary artery disease; HDL, High-density lipoprotein; ApoB, Apolipoprotein B; ApoA1, Apolipoprotein A1; HbA1c, Glycated haemoglobin; WHR, Waist-to-hip ratio; adjBMI, Adjusted for BMI; BMI, Body mass index; GIP, Gastric inhibitory polypeptide

Table S5: Clusters of colocalised traits identified by the secondary analysis across the permutations of prior 2 and the regional and alignment thresholds (prior 2: 0.02, 0.01 and 0.001; thresholds: 0.5, 0.6, 0.7, 0.8 and 0.9). A total of 5,015 variants were included.

Locus	Colocalised traits	PP coloc	Candidate variant	PP explained	Prior 2	Regional and alignment threshold
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.02	0.5
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR, T2D	0.8471	rs429358	1	0.02	0.5
GIPR	BMI, Waist circumference	1	rs1800437	1	0.02	0.5
GIPR	Triglycerides, Hip circumference adjBMI	0.983	rs5117	0.9328	0.02	0.5
GIPR	GIP SOMAmer 16292_288, Hip circumference, 2hr Glucose adjBMI	0.9079	rs1800437	0.6768	0.02	0.5
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.02	0.6
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR, T2D	0.8471	rs429358	1	0.02	0.6
GIPR	BMI, Waist circumference	1	rs1800437	1	0.02	0.6
GIPR	Triglycerides, Hip circumference adjBMI	0.983	rs5117	0.9328	0.02	0.6
GIPR	GIP SOMAmer 16292_288, Hip circumference, 2hr Glucose adjBMI	0.9079	rs1800437	0.6768	0.02	0.6
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.02	0.7
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR, T2D	0.8471	rs429358	1	0.02	0.7
GIPR	BMI, Waist circumference	1	rs1800437	1	0.02	0.7
GIPR	Triglycerides, Hip circumference adjBMI	0.983	rs5117	0.9328	0.02	0.7
GIPR	GIP SOMAmer 16292_288, Hip circumference, 2hr Glucose adjBMI	0.9079	rs1800437	0.6768	0.02	0.7
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.02	0.8
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR, T2D	0.8471	rs429358	1	0.02	0.8
GIPR	BMI, Waist circumference	1	rs1800437	1	0.02	0.8
GIPR	Triglycerides, Hip circumference adjBMI	0.983	rs5117	0.9328	0.02	0.8
GIPR	GIP SOMAmer 16292_288, Hip circumference, 2hr Glucose adjBMI	0.9079	rs1800437	0.6768	0.02	0.8
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.02	0.9
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9651	rs429358	1	0.02	0.9
GIPR	BMI, Waist circumference	1	rs1800437	1	0.02	0.9
GIPR	Triglycerides, Hip circumference adjBMI	0.983	rs5117	0.9328	0.02	0.9
GIPR	GIP SOMAmer 16292_288, Hip circumference, 2hr Glucose adjBMI	0.9079	rs1800437	0.6768	0.02	0.9
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.01	0.5
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR, T2D	0.7384	rs429358	1	0.01	0.5
GIPR	BMI, Waist circumference	1	rs1800437	1	0.01	0.5
GIPR	Triglycerides, Hip circumference adjBMI	0.9665	rs5117	0.9328	0.01	0.5
GIPR	GIP SOMAmer 16292_288, Hip circumference, 2hr Glucose adjBMI	0.8288	rs1800437	0.6768	0.01	0.5
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.01	0.6
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR, T2D	0.7384	rs429358	1	0.01	0.6
GIPR	BMI, Waist circumference	1	rs1800437	1	0.01	0.6
GIPR	Triglycerides, Hip circumference adjBMI	0.9665	rs5117	0.9328	0.01	0.6
GIPR	GIP SOMAmer 16292_288, Hip circumference, 2hr Glucose adjBMI	0.8288	rs1800437	0.6768	0.01	0.6
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.01	0.7

GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR, T2D	0.7384	rs429358	1	0.01	0.7
GIPR	BMI, Waist circumference	1	rs1800437	1	0.01	0.7
GIPR	Triglycerides, Hip circumference adjBMI	0.9665	rs5117	0.9328	0.01	0.7
GIPR	GIP SOMAmer 16292_288, Hip circumference, 2hr Glucose adjBMI	0.8288	rs1800437	0.6768	0.01	0.7
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.01	0.8
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9334	rs429358	1	0.01	0.8
GIPR	BMI, Waist circumference	1	rs1800437	1	0.01	0.8
GIPR	Triglycerides, Hip circumference adjBMI	0.9665	rs5117	0.9328	0.01	0.8
GIPR	GIP SOMAmer 16292_288, Hip circumference, 2hr Glucose adjBMI	0.8288	rs1800437	0.6768	0.01	0.8
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.01	0.9
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.9334	rs429358	1	0.01	0.9
GIPR	BMI, Waist circumference	1	rs1800437	1	0.01	0.9
GIPR	Triglycerides, Hip circumference adjBMI	0.9665	rs5117	0.9328	0.01	0.9
GIPR	GIP SOMAmer 16292_288, Hip circumference	0.9614	rs1800437	0.681	0.01	0.9
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.001	0.5
GIPR	HbA1c, ApoA1, WHRadjBMI, Waist circumference adjBMI, WHR	0.5827	rs429358	1	0.001	0.5
GIPR	BMI, Waist circumference	1	rs1800437	1	0.001	0.5
GIPR	Triglycerides, Hip circumference adjBMI	0.7428	rs5117	0.9328	0.001	0.5
GIPR	GIP SOMAmer 16292_288, Hip circumference	0.6959	rs1800437	0.681	0.001	0.5
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.001	0.6
GIPR	HbA1c, ApoA1, WHRadjBMI, WHR	0.8888	rs429358	1	0.001	0.6
GIPR	BMI, Waist circumference	1	rs1800437	1	0.001	0.6
GIPR	Triglycerides, Hip circumference adjBMI	0.7428	rs5117	0.9328	0.001	0.6
GIPR	GIP SOMAmer 16292_288, Hip circumference	0.6959	rs1800437	0.681	0.001	0.6
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.001	0.7
GIPR	HbA1c, ApoA1, WHRadjBMI, WHR	0.8888	rs429358	1	0.001	0.7
GIPR	BMI, Waist circumference	1	rs1800437	1	0.001	0.7
GIPR	Triglycerides, Hip circumference adjBMI	0.7428	rs5117	0.9328	0.001	0.7
GIPR	GIP SOMAmer 16292_288, Hip circumference	0.6959	rs1800437	0.681	0.001	0.7
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.001	0.8
GIPR	HbA1c, ApoA1, WHRadjBMI, WHR	0.8888	rs429358	1	0.001	0.8
GIPR	BMI, Waist circumference	1	rs1800437	1	0.001	0.8
GIPR	GIP SOMAmer 16292_288, Hip circumference	0.6959	rs1800437	0.681	0.001	0.8
GIPR	LDL, CAD, HDL, Total Cholesterol, Lipoprotein A, ApoB	1	rs7412	1	0.001	0.9
GIPR	HbA1c, ApoA1, WHR	0.9999	rs429358	1	0.001	0.9
GIPR	BMI, Hip circumference	0.9806	rs1800437	1	0.001	0.9

Abbreviations: PP, Posterior probability; N, Number; LDL, Low-density lipoprotein; CAD, Coronary artery disease; HDL, High-density lipoprotein; ApoB, Apolipoprotein B; ApoA1, Apolipoprotein A1; HbA1c, Glycated haemoglobin; WHR, Waist-to-hip ratio; adjBMI, Adjusted for BMI; BMI, Body mass index; GIP, Gastric inhibitory polypeptide

Table S6. Association of rs1964272 with CAD after conditioning on E354.

Variant	Chr:pos	EA	EAF	Beta (SE)	P-value	Beta (SE)	P-value	N
rs1964272	19:46190268	G	0.5193	0.03 (0.006)	9.65×10^{-9}	0.02 (0.006)	7.18×10^{-4}	299519

Abbreviations: Chr, Chromosome; pos, Position; EA, Effect allele; EAF, Effect allele frequency; SE, Standard error; N, Number of participants

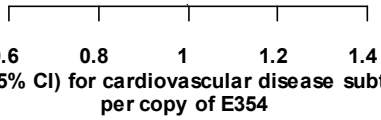
Table S7. Association of other previously identified fasting GIP variants with CAD. The association of rs2287019 was not considered due to its high LD with E354.

Variant	Chr:pos	EA	EAF	Beta	SE	P-value	Cases	Controls
rs17681684	17:9792768	A	0.3082	-0.0074	0.0057	0.1925	34,541	261,984

Abbreviations: Chr, Chromosome; pos, Position; EA, Effect allele; EAF, Effect allele frequency; SE, Standard error

Fig. S1. Association of E354 and cardiovascular disease sub-types in UK Biobank. Cardiovascular disease sub-types were defined in UK Biobank and tested for association with E354 using multivariable logistic regression adjusting for age, sex and 10 principal components[2]. Estimates for each disease are expressed per copy of E354. A Bonferroni corrected significance threshold of P<0.0029 was used.

Cardiovascular disease subtype	Cases	Controls	OR (95% CI)	P-value
Ischaemic, haemorrhagic and unknown stroke (but not TIA)	9,652	357,934	1.00 [0.97, 1.04]	0.87
Ischaemic Cerebrovascular Disease (all)	8,084	359,502	1.00 [0.96, 1.04]	0.91
Ischaemic Stroke	4,602	362,984	1.00 [0.95, 1.05]	0.97
Transient Ischaemic Attack	3,962	363,624	1.00 [0.94, 1.05]	0.91
Haemorrhagic Stroke (all)	1,981	365,605	0.97 [0.89, 1.05]	0.43
Intracerebral Haemorrhage	1,064	366,522	0.98 [0.88, 1.09]	0.75
Subarachnoid Haemorrhage	1,084	366,502	0.98 [0.88, 1.09]	0.73
Abdominal Aortic Aneurysm	1,094	366,492	1.19 [1.07, 1.30]	0.003
Thoracic Aortic Aneurysm	347	367,239	0.86 [0.68, 1.04]	0.1
Venous Thromboembolism (all)	14,097	353,489	1.03 [1.00, 1.06]	0.03
Deep Vein Thrombosis	9,454	358,132	1.05 [1.01, 1.08]	0.02
Pulmonary Embolism	6,148	361,438	1.04 [0.99, 1.08]	0.13
Peripheral Vascular Disease	3,415	364,171	1.05 [0.99, 1.11]	0.13
Aortic Valve Stenosis	2,244	365,342	0.93 [0.86, 1.00]	0.05
Atrial Fibrillation	16,945	350,641	1.00 [0.97, 1.03]	0.89
Heart Failure	6,712	360,874	1.01 [0.97, 1.05]	0.65



Abbreviations: TIA, Transient ischaemic attack; PC, principal component; OR, Odds ratio; CI, Confidence interval

Fig. S2. Associations between E354 and regional adiposity compartments in 435,387 participants measured by bio-impedance. Fat mass in each compartment is shown in orange and lean mass in blue. Estimates for each compartment are in SD per copy of E354 (rs1800437). All estimates are adjusted for age, sex, genotyping chip, and the top 40 principal components. A Bonferroni significance threshold of $P \leq 0.003$ was used to ascertain significance.
Abbreviations: CI, Confidence interval

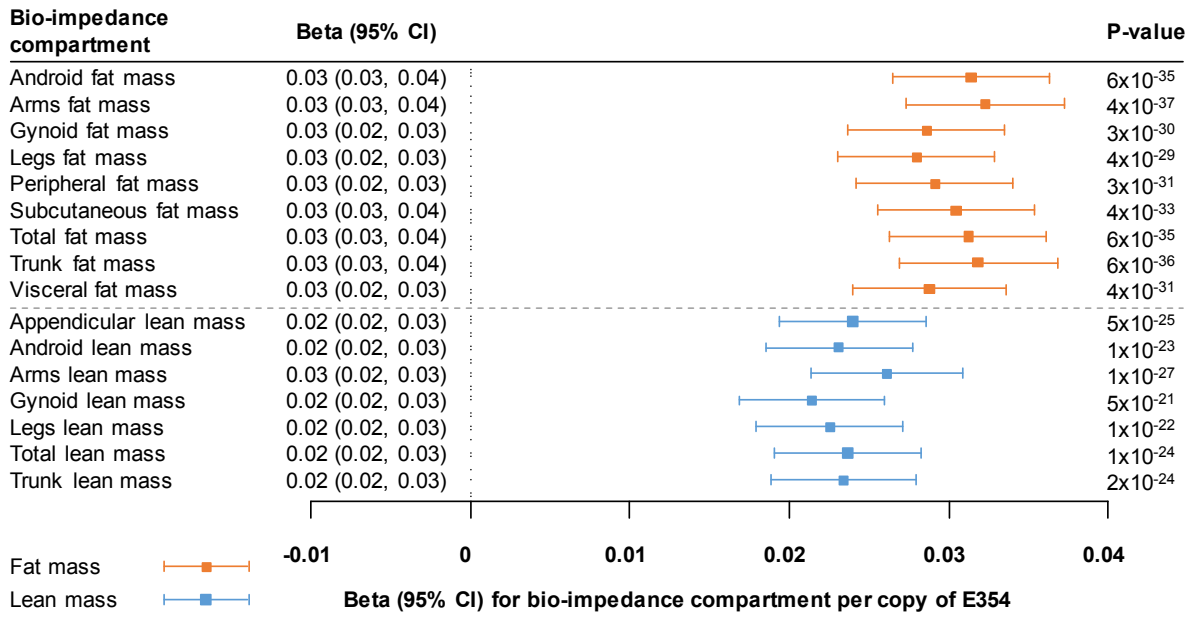


Fig. S3. Associations between E354 and human protein levels. All estimates are adjusted for age, sex, sample collection site and 10 genetic principal components. **Panel A.** Volcano plot showing the associations between E354 and 4,979 human protein levels. The dashed line indicates the Bonferroni significance threshold $P \leq 1x10^{-5}$. The point size for each protein is proportional to its effect size. Significant protein associations with E354 are shown in blue, non-significant proteins are shown in yellow. Associations with significant proteins and proteins of interest are labelled. Two SOMAmers from the SOMAscan® 4k assay target GIP levels, both are labelled. **Panels B & C.** Regional association plots depicting the E354 (rs1800437) association with both GIP SOMAmers, X16292_288 and X5755_29 respectively.

Abbreviations: QPCTL, Glutaminyl-peptide cyclotransferase like; GIP, Gastric inhibitory polypeptide.

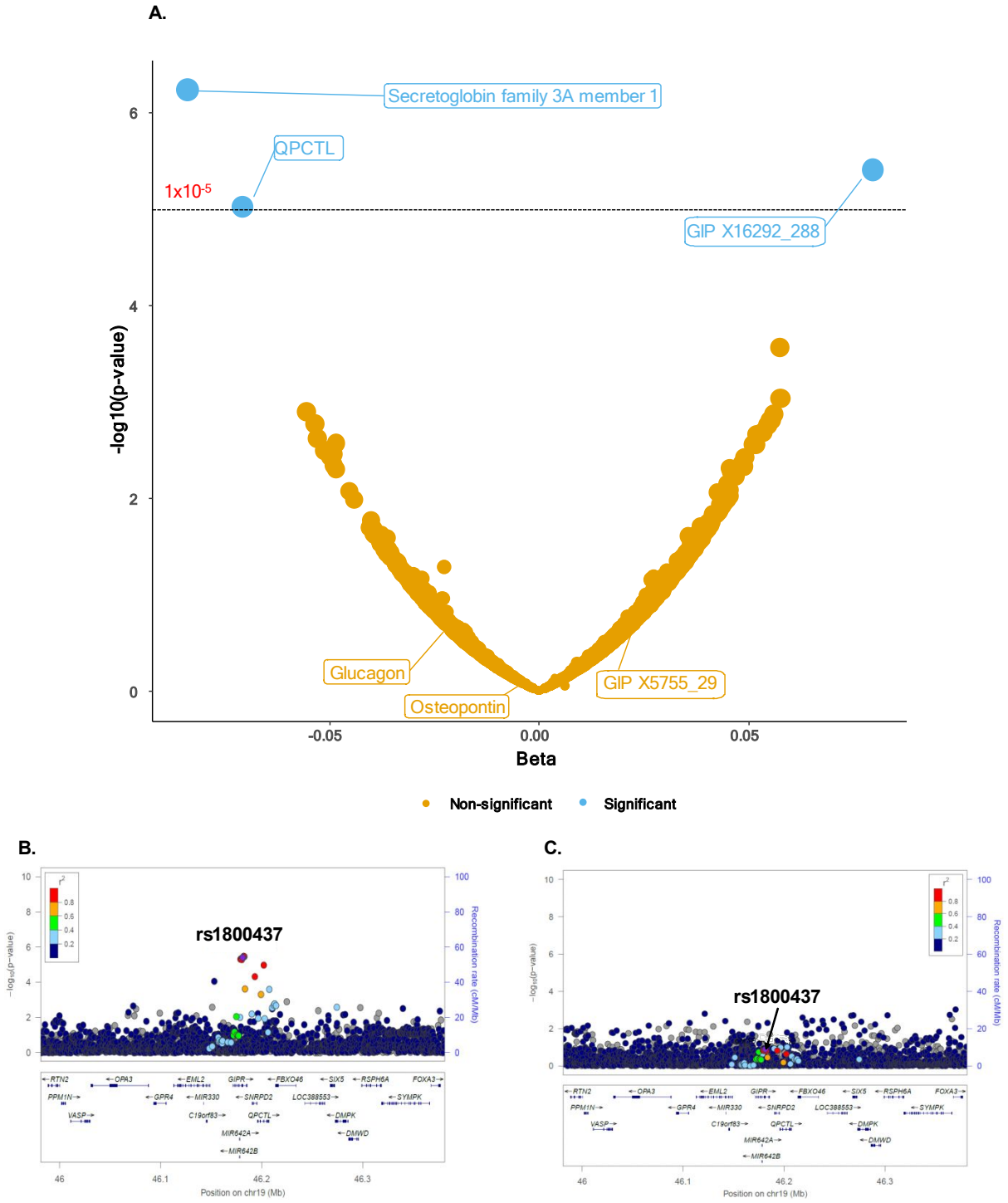


Fig. S4. Associations between E354 and human metabolite levels. Volcano plot showing the associations between E354 and the levels of 1,008 human plasma metabolites. All estimates are adjusted for age, sex and measurement batch. The dashed line indicates the Bonferroni significance threshold P ≤ $5x10^{-5}$. The point size for each protein is proportional to its effect size. Metabolites are coloured according to their metabolite class. Significant metabolite associations with E354 are labelled in orange.

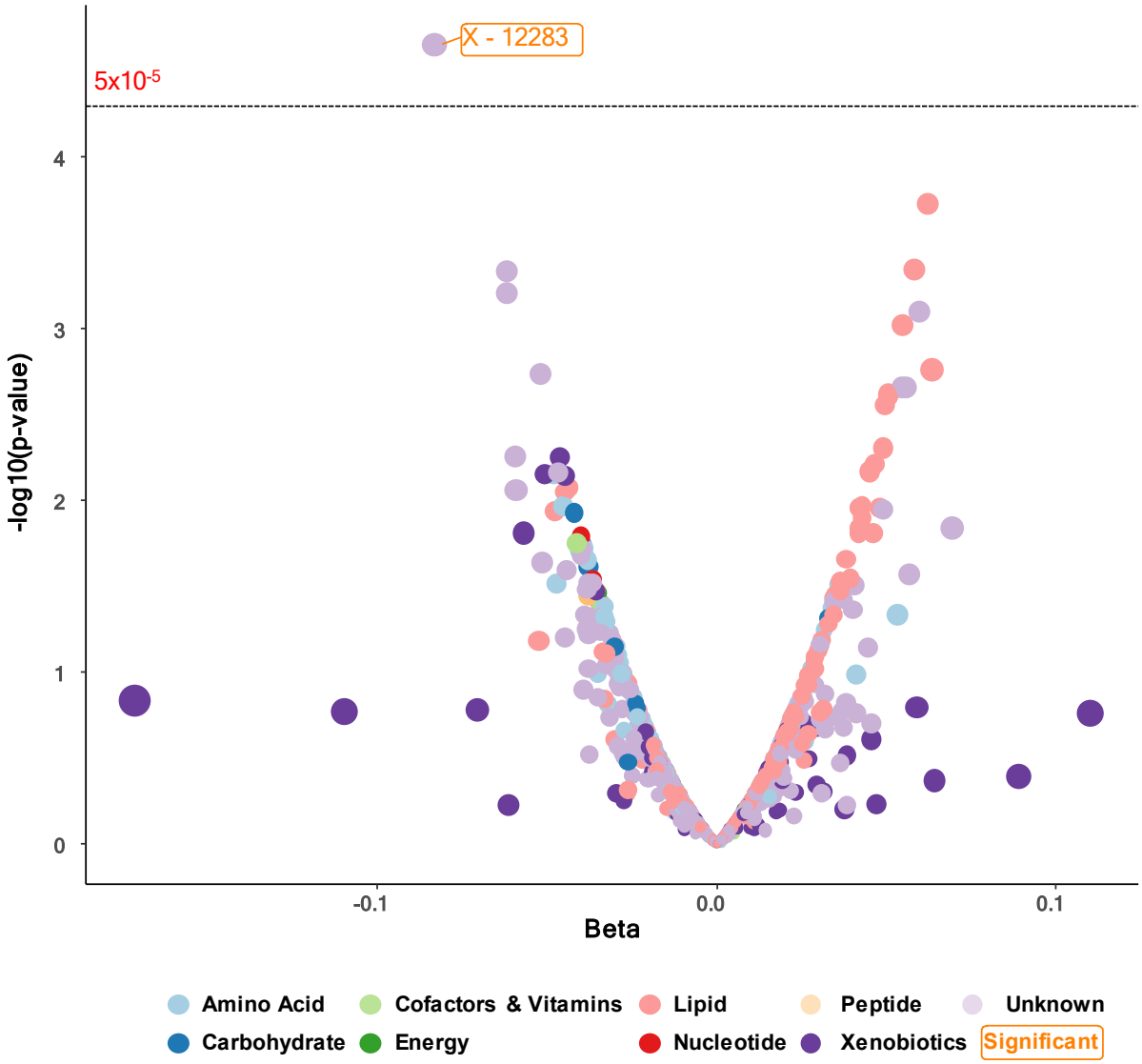
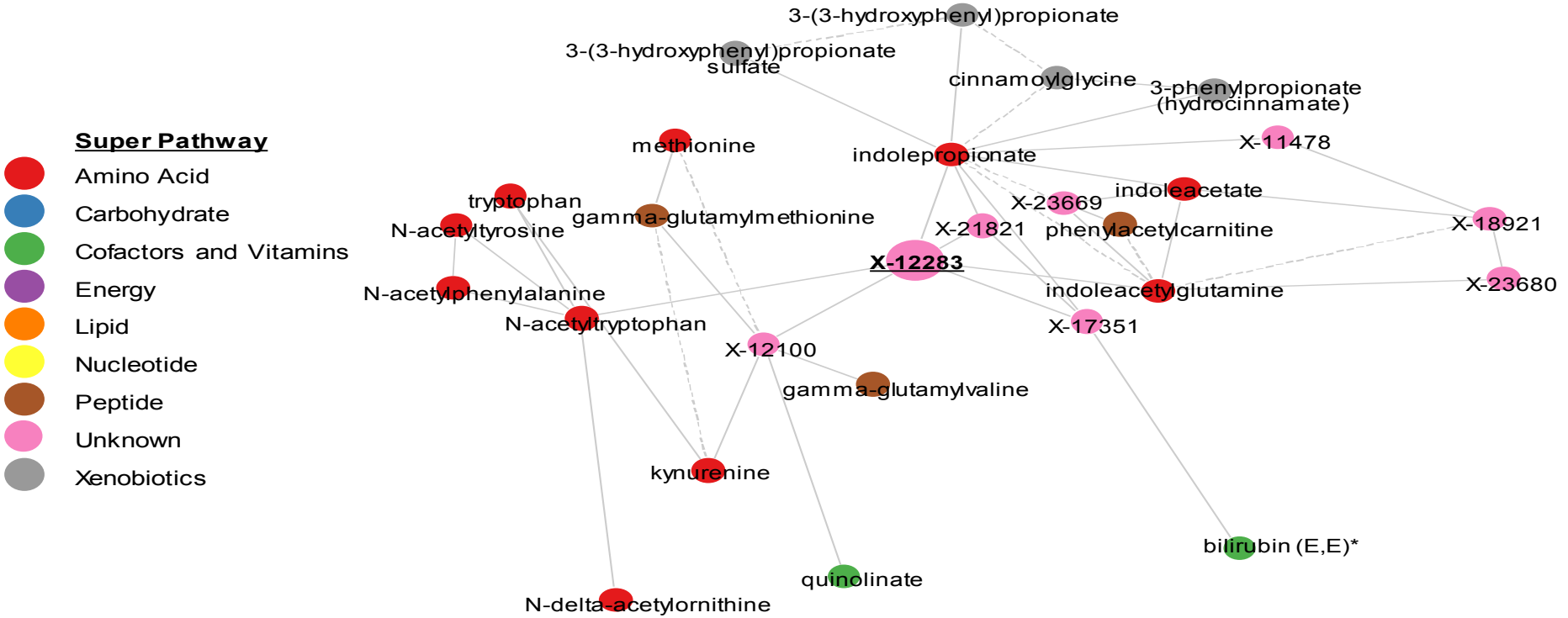


Fig. S5. Gaussian graphical model illustrating the partial correlation network in 11,966 participants between X-12283 and first and second order connections most correlated with X-12283. Positive partial correlation estimates between metabolites are denoted with solid lines whereas negative estimates are shown with dashed lines. Metabolites directly connected with X-12283 represent first order connections, others are second order connections. Metabolites clustered closest to X-12283 are more strongly correlated. Metabolite nodes are coloured by their super pathway. The table outlines the 6 metabolites with a first order connection to X-12283 and shows their partial correlation coefficients and related P-values.



Metabolite	Partial correlation with X-12283	P-value
X-21821	0.41	1.98×10^{-252}
X-17351	0.28	1.57×10^{-117}
Indolepropionate	0.21	1.31×10^{-45}
N-acetyltryptophan	0.16	3.81×10^{-38}
X-12100	0.15	9.00×10^{-39}
Indoleacetylglutamine	0.10	2.78×10^{-14}

Fig. S6. Stacked regional association plot showing the cluster of cardiovascular-related traits which colocalise near the *GIPR* locus. The purple diamond represents the rs7412 variant, a missense variant in *APOE*. variant markers are coloured by their LD with rs7412, with red indicating LD ($R^2 > 0.8$).

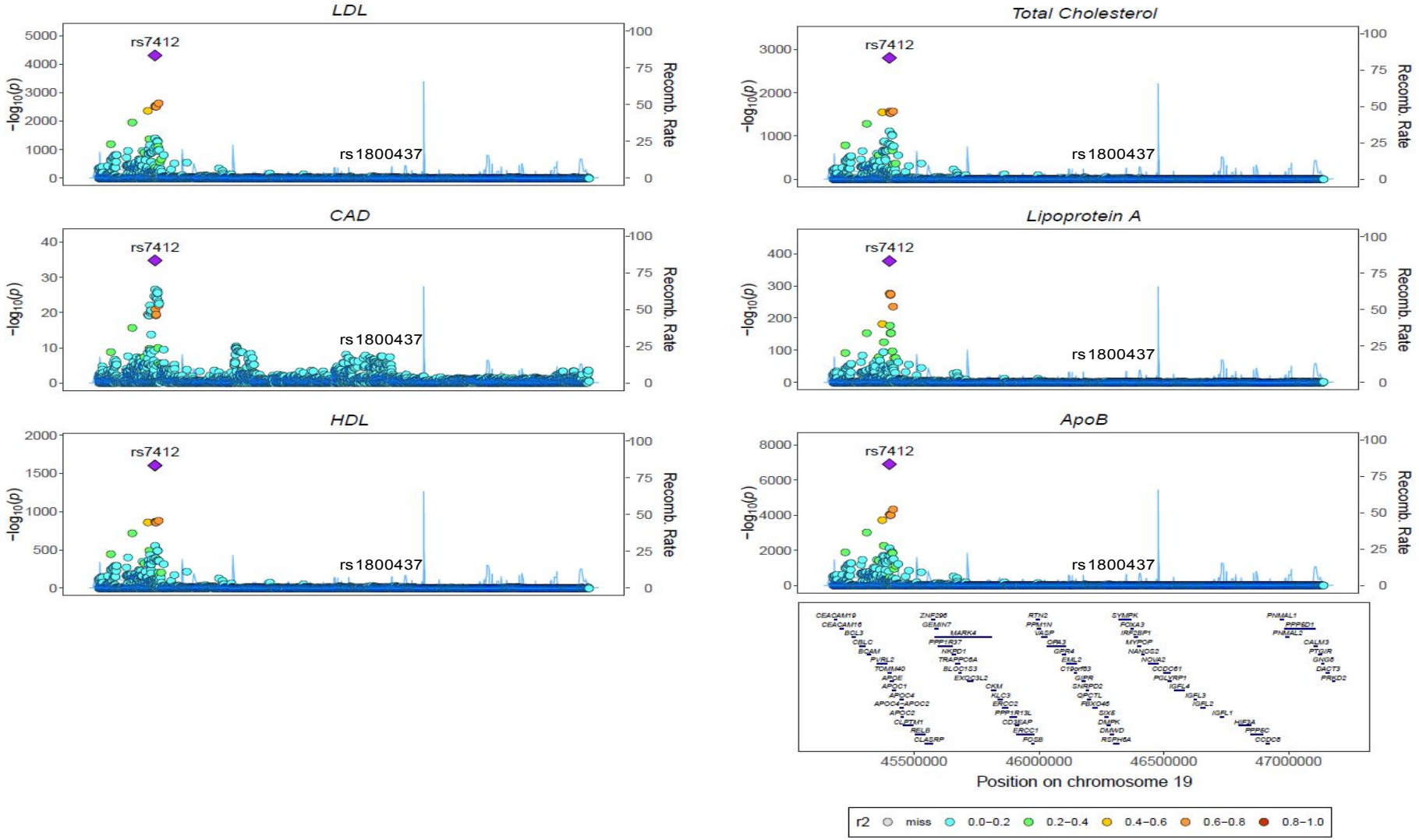


Fig. S7. Regional association plot illustrating the cluster of traits which colocalise with the GIP measures at the *GIPR* locus. The purple diamond represents the rs1800437 variant (E354). Variant markers are coloured by their LD with rs1800437, with red indicating LD ($R^2 > 0.8$). Fasting and 2-hour GIP levels are from the MDC cohort of Almgren *et al.* 2017[3].

Abbreviations: GIPR, Gastric inhibitory polypeptide receptor; LD, Linkage disequilibrium; adj, Adjusted for; BMI, Body mass index; HbA1c, Glycated haemoglobin

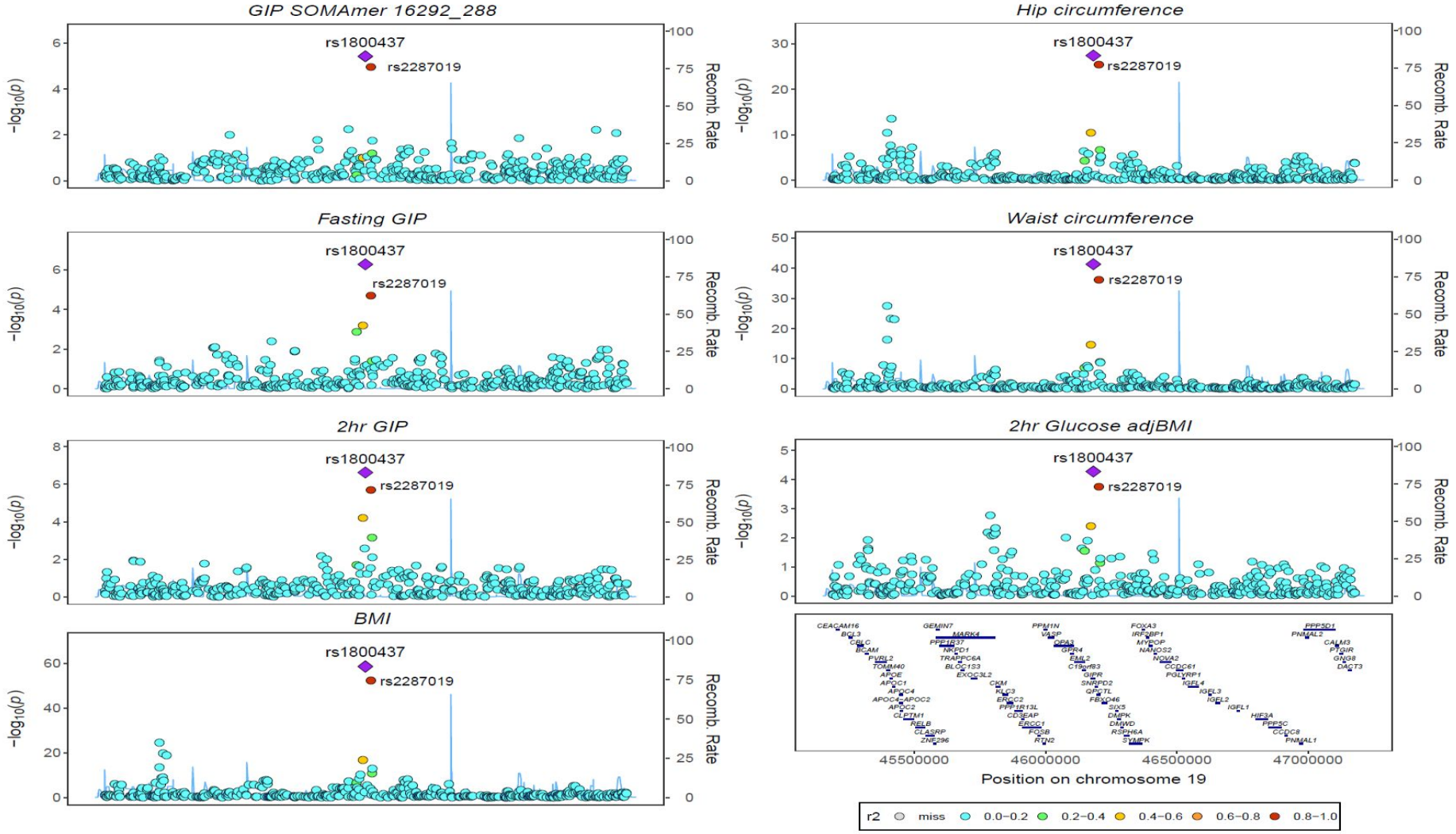


Fig. S8. Heatmap matrix depicting the largest pairwise colocalisation estimate between fasting GIP measures from SomaLogic, fasting and 2-hour GIP measures from Almgren *et al*. 2017[3], 2hr glucose adjusted for BMI ,BMI, LDL, CAD and T2D. Each colocalisation hypothesis is coloured differently with the colour saturation referring to the evidential strength. Posterior probabilities (PP_{coloc}) were considered significant if they met the following criteria: (H4 + H3 ≥ 0.9 & H4/H3 ≥ 3). Trait-pairs with significant posterior probability estimates of colocalisation were outlined in black. To discriminate between H1 and H2 hypotheses, traits along the X-axis were used as "Trait 1" in the analysis and traits listed on the Y-axis were used as "Trait 2".

Abbreviations: H, Hypothesis; BMI, Body mass index; CHD, Coronary heart disease; GIP, Gastric inhibitory polypeptide; 2hr, 2-hour; LDL, Low-density lipoprotein; T2D, Type 2 diabetes

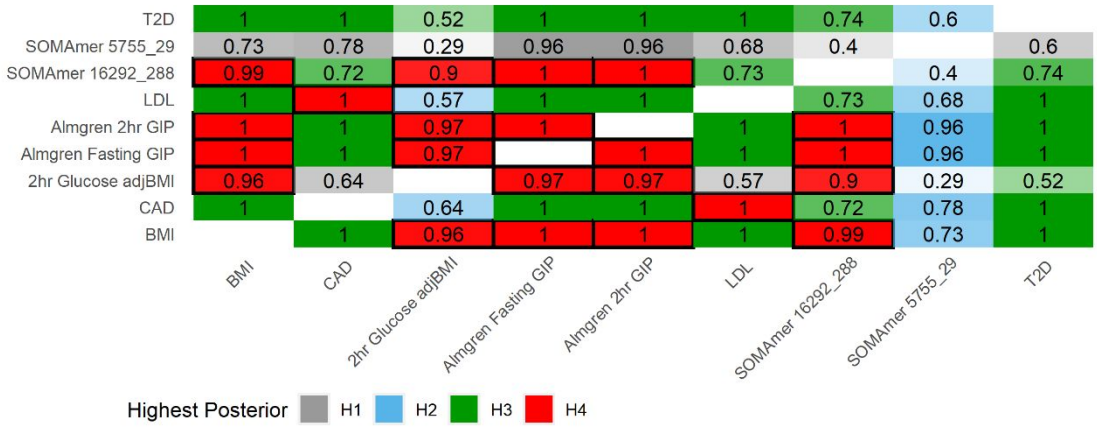


Fig. S9. Matrix illustrating the LD between each of the independent CAD variants and rs1800437 (E354) estimated using 5 European populations in LDlink[4]. Pairwise R^2 values between variants are shown in red in the lower triangle, whereas D' values are shown in blue in the top triangle. Colour saturation represents the strength of the LD estimate between two variants. The LD between rs1800437 and rs1964272 ($R^2 = 0.27$) is depicted in light pink, whereas the very low LD between rs1800437 and the other CAD variants are shown as blank spaces.

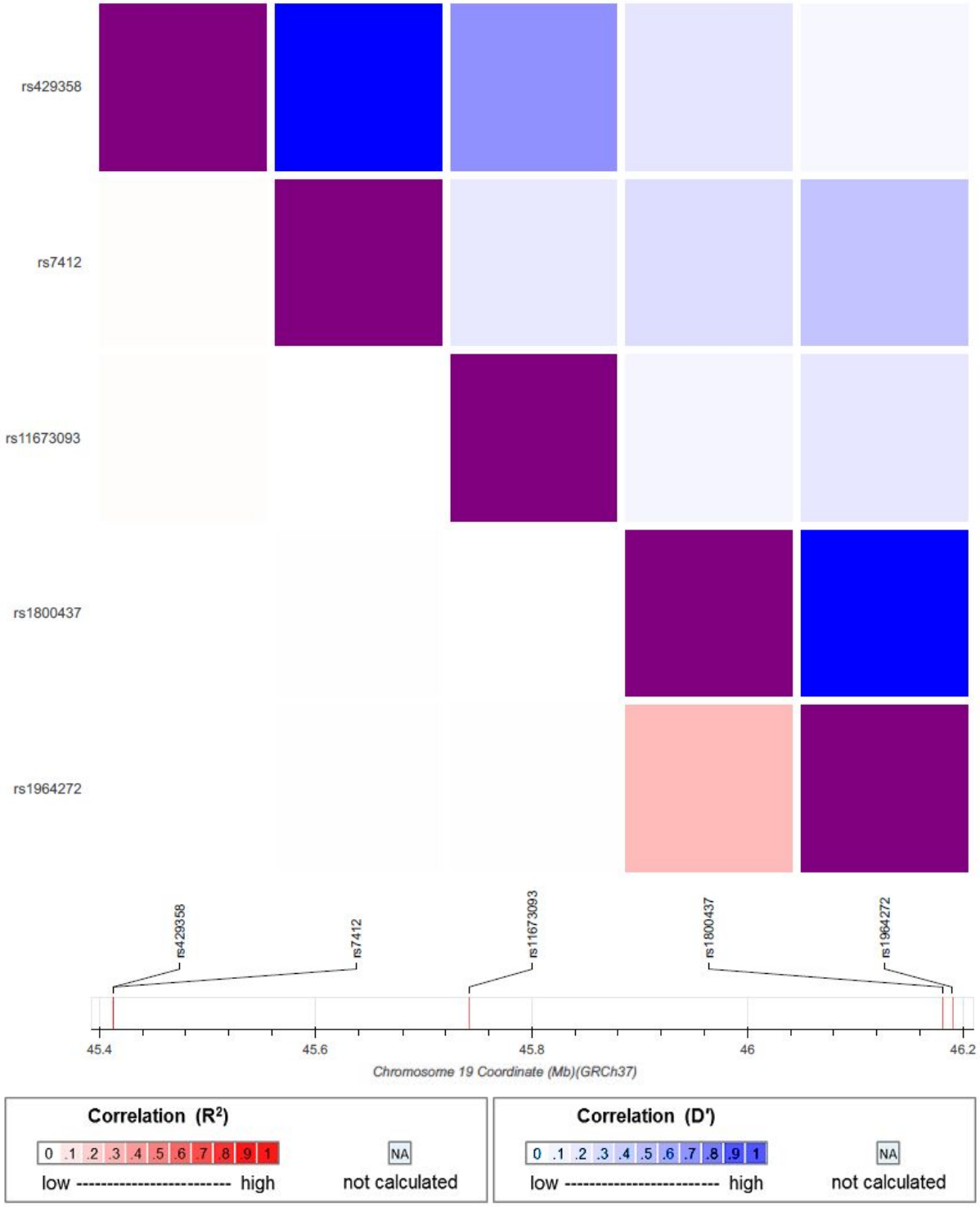
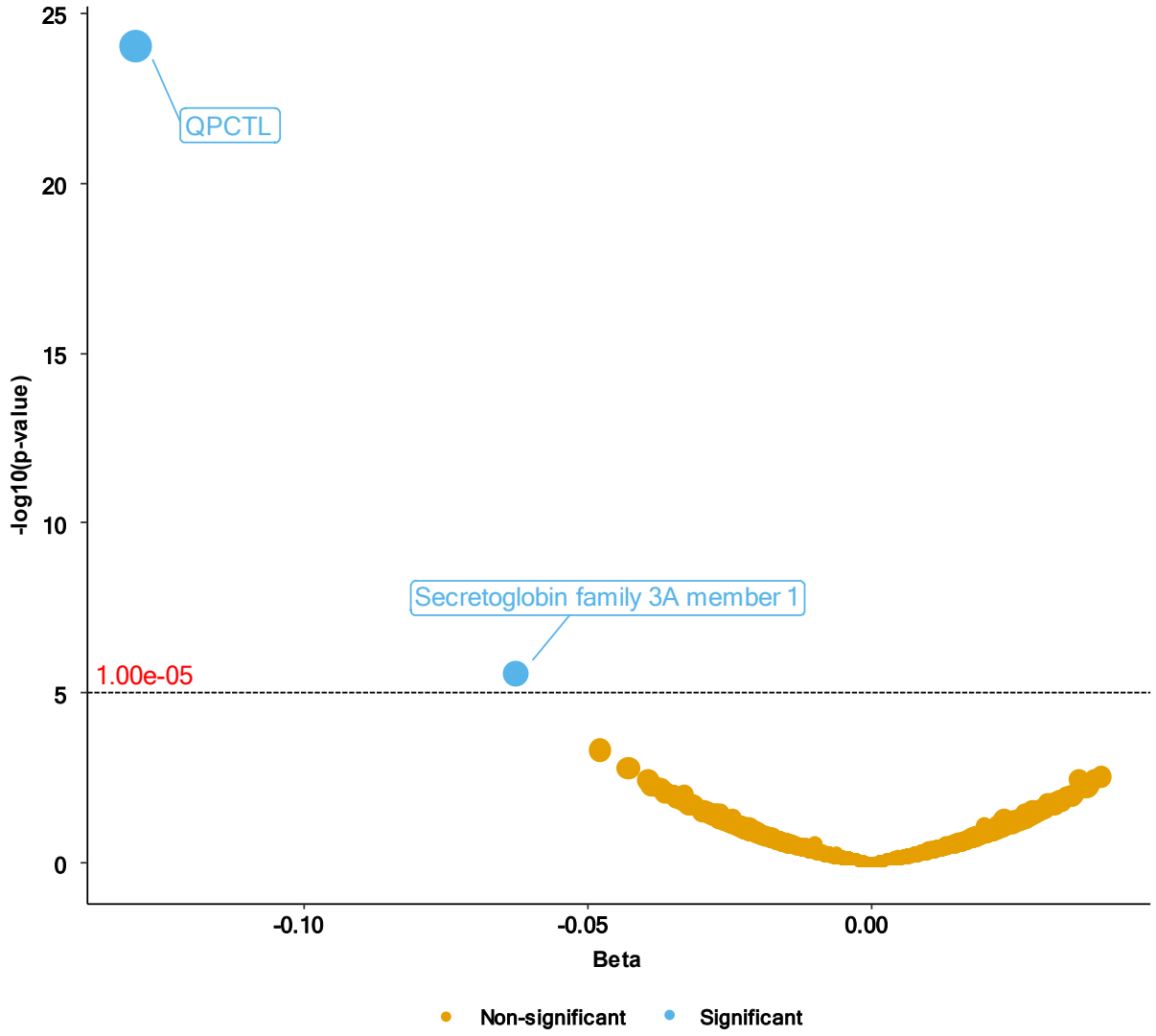


Fig. S10. Volcano plot showing the associations between rs1964272 and 4,979 human protein levels. All estimates are adjusted for age, sex, sample collection site and 10 genetic principal components. The dashed line indicates the Bonferroni significance threshold $P \leq 1x10^{-5}$. The point size for each protein is proportional to its effect size. Significant protein associations with rs1964272 are shown in blue, non-significant proteins are shown in yellow.

Abbreviations: QPCTL, Glutaminyl-peptide cyclotransferase like



References

1 van der Harst P, Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ Res* 2018; **122**: 433–43.

2 Allara E, Morani G, Carter P, *et al.* Genetic Determinants of Lipids and Cardiovascular Disease Outcomes: A Wide-Angled Mendelian Randomization Investigation. *Circ Genomic Precis Med* 2019; **12**: 543–51.

3 Almgren P, Lindqvist A, Krus U, *et al.* Genetic determinants of circulating GIP and GLP-1 concentrations. *JCI Insight* 2017; **2**. DOI:10.1172/jci.insight.93306.

4 Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants: Fig. 1. *Bioinformatics* 2015; **31**: 3555–7.