

Genetic and Causal Relationship Between Coffee Intake and Cardiometabolic Risks: Cross-Phenotype Association And Mendelian Randomization Analysis

Xinpei Wang

Peking University

Jinzhu Jia (✉ jzjia@math.pku.edu.cn)

Peking University

Tao Huang

Peking University

Original investigation

Keywords: cardiometabolic risks, coffee intake, genetic and causal relationship, cross-phenotype association, Mendelian randomization analysis, Transcriptome-wide association analysis

Posted Date: October 12th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-86424/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Many epidemiological studies have shown that there is a significant association between coffee intake and cardiometabolic diseases, which may be due to the common genetic structure or causal relationship.

Methods: We used linkage disequilibrium score regression analysis to calculate the genetic correlation between coffee intake and 23 cardiometabolic traits (diseases), and then used cross-phenotype association analysis to identify the shared genetic loci for the trait pairs with significant genetic correlation. Besides, a bi-directional Mendelian Randomization analysis was used to explore the causal relationship between coffee intake and 23 cardiometabolic traits (diseases).

Results: Coffee intake has a significant genetic correlation (after Bonferroni correction) with body mass index (BMI) ($R_g = 0.3713$, $P\text{-value} = 4.13 \times 10^{-64}$), body fat percentage (BF%) ($R_g = 0.2810$, $P\text{-value} = 1.81 \times 10^{-13}$), type 2 diabetes (T2D) (unadjusted for BMI) ($R_g = 0.1189$, $P\text{-value} = 8.80 \times 10^{-6}$), heart failure (HF) ($R_g = 0.2626$, $P\text{-value} = 6.00 \times 10^{-9}$), atrial fibrillation (AF) ($R_g = 0.1007$, $P\text{-value} = 4.30 \times 10^{-5}$). There are 203, 18, 86, 13, 38 independent shared loci between coffee intake and BMI, BF%, T2D, HF, AF, respectively, among which 22, 2, 23, 4, 13 loci do not achieve genome-wide significance in single trait GWAS. Coffee intake has significant causal effect on BMI ($b = 0.0717$, $P\text{-value} = 2.33 \times 10^{-5}$), T2D (unadjusted for BMI, $OR = 1.27$, $P\text{-value} = 1.46 \times 10^{-7}$), and intracerebral haemorrhage (ICH) (all types ICH: $OR = 1.86$, $P\text{-value} = 3.37 \times 10^{-4}$; deep ICH: $OR = 2.12$, $P\text{-value} = 2.93 \times 10^{-4}$). And BMI ($b = 0.3694$, $P\text{-value} = 3.64 \times 10^{-154}$), BF% ($b = 0.5500$, $P\text{-value} = 1.68 \times 10^{-4}$), T2D (adjusted for BMI, $b = -0.0252$, $P\text{-value} = 4.83 \times 10^{-6}$) and triglycerides (TG) ($b = -0.1209$, $P\text{-value} = 4.56 \times 10^{-15}$) have significant causal effect on coffee intake.

Conclusions: Our study identified the shared genetic structure and causal relationship between coffee intake and several cardiometabolic traits (diseases), providing a new insight into the mechanism of coffee intake and cardiometabolic traits (diseases).

Introduction

Coffee is one of the most widely consumed beverages in the world. In North America, each person consumes about two cups of coffee a day. In many European countries, the average is four cups (1). Coffee contains more than one thousand compounds. In addition to caffeine, there are diterpene, chlorogenic acid and other antioxidant and anti-inflammatory substances (2). Thus, it is thought to have a long-term effect on human health (2).

Many observational studies have examined the association between coffee intake and cardiometabolic traits (diseases). For example, both P Mirmirmiran (3) and Yejee Lim (4) found a negative association between coffee intake and the prevalence of type 2 diabetes (T2D). Vijaykumar Bodar and P Bazal found that moderate coffee intake was associated with a reduced risk of atrial fibrillation (AF) (5, 6), while Long Mo stated that drinking more than three cups of coffee a day was significantly associated with an

increased risk of myocardial infarction (MI) (7). Moreover, previous Genome-wide association studies (GWAS) have identified a number of genetic variants that have significant effects on coffee intake as well as some cardiometabolic traits (diseases) (8–11). This suggests that the association between coffee intake and cardiometabolic traits (diseases) may be due to a common genetic structure or causal relationship.

Based on this, we performed this large-scale genome-wide cross-phenotype association analysis and bi-directional Mendelian randomization (MR) analysis. We analysed the genetic and causal relationship between coffee intake and 23 cardiometabolic traits (diseases). Our results provide new insights into the association between coffee intake and cardiometabolic traits (diseases).

Research Design And Methods

Study Design

Our study is composed of two main parts: the common genetic structure and causal relationship between coffee intake and 23 cardiometabolic traits (diseases). These 23 cardiometabolic traits (diseases) mainly include three aspects: obesity (body mass index (BMI), body fat percentage (BF%)), blood glucose and lipid homeostasis (T2D (adjusted and unadjusted for BMI), fasting glucose (FG), fasting insulin (FI), insulin resistance (HOMA-IR), beta-cell function (HOMA- β), low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol, triglycerides (TG), total cholesterol (TC)), and cardiovascular disease (heart failure (HF), AF, coronary artery disease (CAD), MI, intracerebral haemorrhage (ICH) and its subtypes (lobar ICH and deep ICH), as well as ischaemic stroke and its subtypes (cardioembolic stroke, large-vessel stroke and small-vessel stroke)).

First, we calculated the genetic correlation between coffee intake and 23 cardiometabolic traits (diseases). For the trait pairs with significant genetic correlations, we carried out cross-phenotype association analysis to identify the shared genetic loci. Then, we conducted a bi-directional MR analysis to explore whether there is a causal relationship between coffee intake and these 23 traits (diseases).

Data Sources

For each trait (disease), we used the summary statistics from a recent GWAS. The data of coffee intake came from GWAS of UK Biobank 501,625 participants, who were asked how many cups of coffee they drank per day to obtain the association results of coffee intake and each single nucleotide polymorphism (SNP) (12). BMI data was obtained from a meta-analysis of ~700,000 individuals (13). BF% data came from a GWAS meta-analysis of more than 100,000 participants (14). The T2D (adjusted and unadjusted for BMI) data came from the largest T2D GWAS meta-analysis, with a total of 898,130 participants (15). The data of FG, FI, HOMA-IR, HOMA - β were obtained from the meta-analysis of 21 GWAS by Jos é e Dupuis et al (16). The data of LDL, HDL, TG, TC were from the meta-analysis by cristen J willer et al (17). HF data was derived from a large-scale GWAS meta-analysis involving 47,309 cases and 930,014 controls (18). AF data came from a meta-analysis of more than 500,000 participants (19). Both CAD and

MI data were from a large-scale GWAS meta-analysis conducted by the CARDIoGRAMplusC4D Consortium (20). Data of ICH and its subtypes were derived from a meta-analysis including people of European descent over 55 years of age (21). Data of ischemic stroke and its subtypes were derived from a meta-analysis of 15 ischemic stroke cohorts (22). For details of these studies, please refer to the corresponding literature.

Statistical Analysis

Linkage disequilibrium score regression analysis

We used linkage disequilibrium score regression (LDSC) (23) to calculate the genetic correlation between coffee intake and 23 cardiometabolic traits (diseases) and Bonferroni correction was performed on the obtained P-value. Trait pair with $P\text{-value} < 0.05/23$ was considered to have a significant genetic correlation. And trait pair with $0.05/23 < P\text{-value} < 0.05$ was considered to have a suggestive genetic correlation.

cross-phenotype association analysis

After calculating the genetic correlation between coffee intake and 23 cardiometabolic traits (diseases), we selected trait pairs with significant genetic correlation for cross-phenotype association (24) analysis to identify the shared genetic loci of these trait pairs. The SNPs with cross-phenotype association analysis $P\text{-value} < 5 \times 10^{-8}$ and single trait GWAS analysis $P\text{-value} < 0.01$ were thought to have significant influence on both traits (diseases). Then we used PLINK 1.9 (25) software to divide these SNPs into different independent clumping regions. SNPs with a distance less than 10000 kb and linkage disequilibrium score $R^2 > 0.001$ were divided into the same clumping region. And the SNP with the lowest P-value in each region was taken as the index SNP.

Co-localization analysis

In order to further calculate the probability that the shared loci we identified containing the common genetic causal variant(s) of two traits (diseases), we conducted co-localization analysis by R package coloc (26–28). Locus ($H4 > 0.5$) was considered to be related to two traits (diseases) simultaneously in our co-localization analysis.

Fine-mapping credible set analysis

For each shared locus, we gave a SNP list which has a 99% probability of containing the common genetic causal variant(s) by Fine-Mapping (FM) summary method. This method maps the primary signal and uses flat prior with steepest descent approximation (29, 30).

Transcriptome-wide association analysis

In order to gain more insights into the genetic basis of each trait (disease), we combined the genotype, gene expression and phenotype data for large-scale transcription wide association studies (TWAS) using FUSION software (31). We used False Discovery Rate (FDR) correction and gene-tissue pairs with $FDR <$

0.05 were thought to be significant. And then, on the basis of TWAS results, we used R package RHOGE (32) to estimate the correlation of effect sizes across traits (diseases) in each tissue. Coffee intake and cardiometabolic traits (diseases) were thought to have a significant gene expression correlation in the tissues with P-value < 0.05/48.

Mendelian Randomization analysis

To explore the causal relationship between coffee intake and these 23 cardiometabolic traits (diseases), we conducted a bi-directional MR analysis (33, 34) using TwoSampleMR package (35, 36). Coffee intake and 23 cardiometabolic traits (diseases) were analysed as exposure for two-sample MR analysis in turn. For example, when taking coffee intake as exposure for MR analysis, we first selected SNPs with P-value < 5×10^{-8} in GWAS summary statistics of coffee intake, and then selected independent SNPs as instrumental variables. We used inverse variance weighting method (37) to integrate the causal effect estimation of different instrumental variables and tested whether the MR-Egger regression (36) intercept item is zero to check whether the horizontal pleiotropy is balanced. Besides, we also carried out a series of sensitivity analysis, including single SNP analysis and leave one out analysis.

Results

Genetic correlation

We calculated the genetic correlation between coffee intake and 23 cardiometabolic traits (diseases) (Fig. 1, Supplementary Table 1). Among them, coffee intake has significant genetic correlation (after Bonferroni correction) with BMI ($R_g = 0.3713$, P-value = 4.13×10^{-64}), BF% ($R_g = 0.2810$, P-value = 1.81×10^{-13}), T2D (unadjusted for BMI) ($R_g = 0.1189$, P-value = 8.80×10^{-6}), HF ($R_g = 0.2626$, P-value = 6.00×10^{-9}), AF ($R_g = 0.1007$, P-value = 4.30×10^{-5}), and suggestive genetic correlation with HOMA-IR ($R_g = 0.1387$, P-value = 4.05×10^{-2}), CAD ($R_g = 0.0757$, P-value = 3.44×10^{-2}), MI ($R_g = 0.0949$, P-value = 1.51×10^{-2}).

cross-phenotype association analysis

Based on the significant genetic correlation between coffee intake and BMI, BF%, T2D, HF, AF, we further carried out cross-phenotype association analysis to identify the shared genetic loci. There are 203, 18, 86, 13, 38 independent shared loci between coffee intake and BMI, BF%, T2D, HF, AF, respectively, among which 22, 2, 23, 4, 13 loci do not achieve genome-wide significance in single trait GWAS (Supplementary Table 2).

The most significant locus in the cross-phenotype association analysis of coffee intake and BMI/BF% is rs3751813 (P-value = 3.42×10^{-306}) on chromosome 16 and rs936227 (P-value = 3.97×10^{-49}) on chromosome 15, respectively. The most significant locus in the cross-phenotype association analysis of coffee intake and T2D is rs2472297 (P-value = 6.38×10^{-221}) on chromosome 15. This SNP has been repeatedly demonstrated to have a significant association with coffee intake, but there has been no

previous study found its association with T2D (38). In the cross-phenotype association analysis of coffee intake and HF, rs476828 (P-value = 4.15×10^{-22}) on chromosome 18 showed the strongest significance, so as rs12898551 (P-value = 2.19×10^{-70}) to coffee intake and AF.

We mapped 5895, 47, 1044, 75, 382 shared genes in the shared regions of coffee intake and BMI, BF%, T2D, HF, and AF, respectively (Supplementary Table 2, Fig. 2). Interestingly, some of these mapped genes are shared by more than two trait pairs. Particularly, *TMEM18*, *FTO* and *MC4R* genes are significant in all five cross-phenotype association analysis, indicating that these genes have significant influence on all six traits (diseases). *TMEM18*, whose full name is the transmembrane protein 18, is a protein encoding gene that is more expressed in ovaries, thyroid, endometrium, and other tissues. Previous GWAS studies have found significant associations between *TMEM18* and obesity (39), age of menarche (40), etc. *FTO* is called *FTO* alpha-ketoglutarate dependent dioxygenase, which is a protein coding gene and is mainly expressed in brain, adrenal gland, and other tissues. This gene is a nuclear protein of the AlkB related non-haem iron and 2-oxoglutarate-dependent oxygenase superfamily, but the exact physiological function is unknown (38). *MC4R* is fully known as melanocortin 4 receptor. This gene encodes a membrane-bound receptor and is a member of the melanocortin receptor family. A defect in this gene is responsible for autosomal dominant obesity (38).

Co-localization and fine-mapping analysis

We also used co-localization analysis to calculate the probability that the shared locus identified in cross-phenotype association analysis contains the common pathogenic variant(s) of two traits (diseases). The main results are presented in Supplementary Table 3. To identify the pathogenic mutations more accurately in each region, we carried out fine mapping. The Supplementary Table 4 shows a SNP list for each region that has a 99% probability of containing the common genetic causal variant(s).

TWAS

To identify genes whose cis-regulated expression is significantly associated with complex traits and identify important expression - trait associations, we conducted TWAS analysis on coffee intake, BMI, BF%, T2D, HF and AF. A total of

7368, 60749, 355, 22365, 749, 6457 gene-tissue pairs (FDR < 0.05) are significantly associated with coffee intake, BMI, BF%, T2D, HF, AF, respectively (Supplementary Table 5–10).

Then, based on the results of TWAS, we calculated the genetic correlation between coffee intake and these five traits (diseases) in each tissue (Fig. 3, Supplementary Table 11). After Bonferroni correction, coffee intake and BMI are associated positively in all 48 tissues. There is a significant and positive genetic correlation between coffee intake and BF% in thyroid, brain cerebellar hemisphere, cells transformed fibroblasts, skin sun exposed lower leg; with T2D in oesophagus mucosa; with AF in brain cortex. Unfortunately, we did not find any significant gene expression association between coffee intake and HF in all 48 tissues.

MR

To explore the causal relationship between coffee intake and cardiometabolic traits (diseases), we conducted a bi-directional MR analysis with coffee intake and 23 cardiometabolic traits (diseases) as exposure in turn (Fig. 4, Supplementary Table 12).

In MR analysis with coffee intake as exposure, we found that coffee intake had significant causal effect on BMI ($b = 0.0717$, $P\text{-value} = 2.33 \times 10^{-5}$), T2D (unadjusted for BMI, $OR = 1.27$, $P\text{-value} = 1.46 \times 10^{-7}$), and ICH (all types ICH: $OR = 1.86$, $P\text{-value} = 3.37 \times 10^{-4}$; deep ICH: $OR = 2.12$, $P\text{-value} = 2.93 \times 10^{-4}$). For BF%, T2D, FI, HOMA-IR, HOMA – β and CAD, there is a suggestive causal effect (T2D (adjusted for BMI): $OR = 1.27$, $P\text{-value} = 3.12 \times 10^{-3}$; FI: $b = 0.0405$, $P\text{-value} = 2.72 \times 10^{-2}$; HOMA-IR: $b = 0.0464$, $P\text{-value} = 1.26 \times 10^{-2}$; HOMA – β : $b = 0.0379$, $P\text{-value} = 3.42 \times 10^{-3}$; BF%: $b = 0.0790$, $P\text{-value} = 2.71 \times 10^{-2}$; CAD: $OR = 1.08$, $P\text{-value} = 1.54 \times 10^{-2}$).

Similarly, we also found some significant causal association in MR analysis with cardiometabolic traits (diseases) as exposure. BMI ($b = 0.3694$, $P\text{-value} = 3.64 \times 10^{-154}$), BF% ($b = 0.5500$, $P\text{-value} = 1.68 \times 10^{-4}$), T2D (adjusted for BMI, $b = -0.0252$, $P\text{-value} = 4.83 \times 10^{-6}$) and TG ($b = -0.1209$, $P\text{-value} = 4.56 \times 10^{-15}$) have significant causal effect on coffee intake. And T2D (unadjusted for BMI, $b = -0.0112$, $P\text{-value} = 6.14 \times 10^{-3}$), HOMA – β ($b = -0.1344$, $P\text{-value} = 2.76 \times 10^{-2}$), TC ($b = -0.0286$, $P\text{-value} = 7.66 \times 10^{-3}$), AF ($b = -0.0100$, $P\text{-value} = 2.97 \times 10^{-2}$) have suggestive causal effect on coffee intake. All the above results were obtained at the equilibrium of horizontal pleiotropy (Supplementary Table 13). In addition to the above results, we also conducted several sensitivity analyses to examine the effect of a single SNP on the results (Supplementary Table 14, 15).

Discussion

This is the first study to date to simultaneously analyse the shared genetic structure and causal direction between coffee intake and cardiometabolic traits (diseases). Using persuasive GWAS summary statistics, we found significant positive genetic correlation and identified the shared genetic loci between coffee intake and several cardiometabolic traits/diseases (BMI, BF%, T2D, HF, AF). Furthermore, we found that coffee intake could significantly increase BMI and the risk of T2D and ICH. Higher BMI and BF% could increase the level of coffee intake while T2D (adjusted for BMI) and higher TG reduce coffee intake.

Our results show that there is a significant and positive genetic correlation between coffee intake and T2D, but this significance no longer exists after adjusting for BMI. We also found that there was a positive genetic correlation between coffee intake and BMI/BF%, which is consistent with the results of Gene ATLAS (12). These results suggest that BMI may mediate the observed genetic correlation between coffee intake and T2D. In addition, we also found a significant genetic correlation between coffee intake and HF/AF, indicating that the association between coffee intake and HF/AF may be due to the shared genetic structure.

After that, we identified the shared genes between coffee intake and BMI, BF%, T2D, HF, AF by cross-phenotype association analysis. We found 6316 shared genes totally. *TMEM18*, *FTO* and *MC4R* were found in all five cross-phenotype association analysis, indicating that these three genes can simultaneously affect coffee intake and BMI, BF%, T2D, HF, AF. These three genes are all closely related to obesity, which once again highlights the importance of obesity in the association between coffee intake and cardiometabolic traits (diseases).

Our cross-phenotype association analysis revealed many novel genetic regions. Previous studies have not found a significant association between these regions and coffee intake or cardiometabolic traits (diseases). Many of these regions are associated with cognitive and intellectual performance (41–44), and some are associated with mental disorders such as anxiety(45), schizophrenia(46), depression(47), and immune system symptoms such as systemic lupus erythematosus(48), and digestive diseases such as Barrett's esophagus and esophageal adenocarcinoma (49). These findings suggest that the genetic structure shared by coffee intake and cardiometabolic traits (diseases) may also play an important role in these diseases, which needs further research and exploration.

We also integrated the genotype, gene expression and phenotype for TWAS analysis to show the expression of each phenotype in 48 tissues and calculated the gene expression association between coffee intake and five other phenotypes in each tissue. Our results indicate that there is a significant gene expression correlation between coffee intake and these five cardiometabolic traits (diseases) in nervous, digestive system and endocrine system. This is consistent with our findings about novel genetic regions. These results suggest that some biological metabolic processes in these tissues may affect coffee intake and these five cardiometabolic traits (diseases) at the same time.

In addition to identifying the shared genetic structure, we also conducted a bi-directional MR analysis to explore the causal relationship between coffee intake and cardiometabolic traits (diseases). We got a lot of significant results. A-SD increase in coffee intake (cups/day) can increase 0.0717-SD BMI (kg/m²), increase the risk of T2D by 0.27 times, increase the risk of deep ICH by 1.12 times and increase the risk of all types ICH by 0.86 times. These results highlight the harmful effects of coffee intake on health. We did not find any causal effect of coffee intake on AF or stroke, which is consistent with previous MR studies (50, 51).

It should be noted that our MR results show that coffee intake can increase the risk of T2D, which is contrary to the results of most previous observational studies (52). The reasons may include the following aspects: First, our study found that coffee intake can increase BMI, which is the same as previous MR studies on coffee intake and obesity (53). BMI is a recognized risk factor for T2D (54), so BMI is likely to be a mediator between coffee intake and T2D. However, we have observed that previous observational studies on coffee intake and T2D have adjusted BMI in the analysis (52, 55), which will inevitably lead to underestimate or misestimate the impact of coffee intake on T2D. Moreover, our MR results also showed that when BMI was adjusted for T2D data, T2D (adjusted for BMI) had a negative causal effect on coffee intake. Therefore, we speculate that the negative correlation between coffee

intake and T2D found in many observational studies may be due to improper adjustment of BMI and reverse causality. Second, our coffee intake data is based on a survey of coffee intake among UK Biobank participants. Among all coffee drinkers, the participants who drank instant coffee accounted for the majority (54%). Instant coffee usually contains a lot of dairy creamer and sugar, and previous studies on instant coffee have found an important association with the increased risk of metabolic diseases (56). Although studies on coffee ingredients such as chlorogenic acids and caffeine have found their beneficial effects on the health (57), such beneficial effects may be neutralized by other substances added to instant coffee. Third, our MR results reflect the effect of lifetime exposure to coffee intake on cardiometabolic traits (diseases), while experimental studies focus on the effect of short-term exposure, which may be different.

Our study has many advantages. First, this is the first systematic study on the shared genetic structure and causal relationship between coffee intake and multiple cardiometabolic traits (diseases). We identified many shared genes and found significant causal relationship, which provide new insights into the mechanism of their association. Second, our research scale is very large. We have analysed a total of 23 cardiometabolic traits (diseases), involving obesity, blood glucose and lipid homeostasis and cardiovascular disease, which is helpful to comprehensively understand the relationship between coffee intake and cardiometabolic diseases. Third, the GWAS summary statistics selected for our study are all from large-scale and high-quality GWAS. We identified many shared loci and found many significant causal relationships that were not found in previous less-powerful MR studies. Fourth, our research is very comprehensive, which uses phenotypic information, gene information and expression information to comprehensively explain the relationship between coffee intake and cardiometabolic traits (diseases).

At the same time, our research also has some areas that can be further improved. First, in our analysis, we found that obesity played an important role in the association between coffee intake and cardiometabolic traits (diseases), but unfortunately, for some traits (diseases), we do not have the data adjusted for obesity. If appropriate data is available later, more interesting results may be found. Second, UK Biobank does not limit the types of coffee in the survey of coffee intake, and different types of coffee may have different effects on health. Studies on specific coffee type may bring more specific result.

Conclusion

In this study, we found significant positive genetic correlation between coffee intake and several cardiometabolic traits/diseases (BMI, BF%, T2D, HF, AF) and identified many shared genetic loci among them, including some novel genetic regions. By MR analysis, we found that coffee intake can significantly increase BMI and the risk of T2D and ICH. Our results highlighted the important role of BMI in the relationship between coffee intake and cardiometabolic diseases, which has been overlooked in most previous observational studies. This study provides new insights and evidence for the health effect of coffee intake, which is of great significance.

Abbreviations

AF

atrial fibrillation

BF%

body fat percentage

BMI

body mass index

CAD

coronary artery disease

FDR

False Discovery Rate

FG

fasting glucose

FI

fasting insulin

FM

Fine-Mapping

GWAS

Genome-wide association study

HDL

high-density lipoprotein cholesterol

HF

heart failure

HOMA-IR

insulin resistance

HOMA- β

beta-cell function

ICH

intracerebral haemorrhage

LDL

low-density lipoprotein cholesterol

LDSC

linkage disequilibrium score regression

MI

myocardial infarction

MR

Mendelian randomization

SNP

single nucleotide polymorphism

T2D

type 2 diabetes

TC

total cholesterol

TG

triglycerides

TWAS

transcription wide association study

Declarations

- Ethics approval and consent to participate

Not applicable.

- Consent for publication

Not applicable.

- Availability of data and materials

All data generated or analysed during this study are included in these published reference articles:

Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. *Nat Genet.* 2018;50(11):1593-9. <http://geneatlas.roslin.ed.ac.uk/downloads/?traits=753>

Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum Mol Genet.* 2018;27(20):3641-9.

https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files;
<http://cnsgenomics.com/data.html>

Lu Y, Day FR, Gustafsson S, Buchkovich ML, Na J, Bataille V, et al. New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk. *Nat Commun.* 2016;7:10495.
<https://doi.org/10.1038/ncomms10495>

Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet.* 2018;50(11):1505-13. <http://diagram-consortium.org/>; <http://www.type2diabetesgenetics.org/>

Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet.* 2010;42(2):105-16. <https://doi.org/10.1038/ng.520>

Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013;45(11):1274-83. <https://doi.org/10.1038/ng.2797>

Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, et al. Genome-wide association study provides new insights into the genetic architecture and pathogenesis of heart failure. *bioRxiv*. 2019. <https://doi.org/10.1101/682013>

Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet*. 2018;50(9):1225-33. <https://doi.org/10.1038/s41588-018-0133-9>

Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47(10):1121-30. <https://doi.org/10.1038/ng.3396>

Woo D, Falcone GJ, Devan WJ, Brown WM, Biffi A, Howard TD, et al. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. *Am J Hum Genet*. 2014;94(4):511-21. <https://doi.org/10.1016/j.ajhg.2014.02.012>

Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol*. 2012;11(11):951-62. [https://doi.org/10.1016/S1474-4422\(12\)70234-X](https://doi.org/10.1016/S1474-4422(12)70234-X)

- Competing interests

The authors declare that they have no competing interests.

- Funding

The study was supported by grants from the Peking University Start-up Grant (BMU2018YJ002), High-performance Computing Platform of Peking University and Beijing Technology and Business University Grant(88442Y0033). The funding organization had no role in the preparation of the manuscript.

- Authors' contributions

XW, JJ, and TH designed the study. XW performed the statistical analysis. XW wrote the manuscript. All authors helped interpret the data, reviewed and edited the final paper, and approved the submission.

- Acknowledgements

Not applicable.

References

1. Cornelis MC, Byrne EM, Esko T, Nalls MA, Ganna A, Paynter N, et al. Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. *Mol Psychiatry*.

- 2015;20(5):647–56.
2. Gonzalez de Mejia E, Ramirez-Mares MV. Impact of caffeine and coffee on our health. *Trends Endocrinol Metab.* 2014;25(10):489–92.
 3. Mirmiran P, Carlström M, Bahadoran Z, Azizi F. Long-term effects of coffee and caffeine intake on the risk of pre-diabetes and type 2 diabetes: Findings from a population with low coffee consumption. *Nutr Metab Cardiovasc Dis.* 2018;28(12):1261–6.
 4. Lim Y, Park Y, Choi SK, Ahn S, Ohn JH. The Effect of Coffee Consumption on the Prevalence of Diabetes Mellitus: The 2012–2016 Korea National Health and Nutrition Examination Survey. *Nutrients.* 2019;11(10).
 5. Bodar V, Chen J, Gaziano JM, Albert C, Djoussé L. Coffee Consumption and Risk of Atrial Fibrillation in the Physicians' Health Study. *J Am Heart Assoc.* 2019;8(15):e011346.
 6. Bazal P, Gea A, Navarro AM, Salas-Salvadó J, Corella D, Alonso-Gómez A, et al. Caffeinated coffee consumption and risk of atrial fibrillation in two Spanish cohorts. *Eur J Prev Cardiol.* 2020:2047487320909065.
 7. Mo L, Xie W, Pu X, Ouyang D. Coffee consumption and risk of myocardial infarction: a dose-response meta-analysis of observational studies. *Oncotarget.* 2018;9(30):21530–40.
 8. Zhong VW, Kuang A, Danning RD, Kraft P, van Dam RM, Chasman DI, et al. A genome-wide association study of bitter and sweet beverage consumption. *Hum Mol Genet.* 2019;28(14):2449–57.
 9. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, et al. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am J Hum Genet.* 2019;104(1):65–75.
 10. Kanai M, Akiyama M, Takahashi A, Matoba N, Momozawa Y, Ikeda M, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet.* 2018;50(3):390–400.
 11. Jia H, Nogawa S, Kawafune K, Hachiya T, Takahashi S, Igarashi M, et al. GWAS of habitual coffee consumption reveals a sex difference in the genetic effect of the 12q24 locus in the Japanese population. *BMC Genet.* 2019;20(1):61.
 12. Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. *Nat Genet.* 2018;50(11):1593–9.
 13. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum Mol Genet.* 2018;27(20):3641–9.
 14. Lu Y, Day FR, Gustafsson S, Buchkovich ML, Na J, Bataille V, et al. New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk. *Nat Commun.* 2016;7:10495.
 15. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet.* 2018;50(11):1505–13.

16. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet.* 2010;42(2):105–16.
17. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013;45(11):1274–83.
18. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, et al. Genome-wide association study provides new insights into the genetic architecture and pathogenesis of heart failure. *bioRxiv.* 2019.
19. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet.* 2018;50(9):1225–33.
20. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet.* 2015;47(10):1121–30.
21. Woo D, Falcone GJ, Devan WJ, Brown WM, Biffi A, Howard TD, et al. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. *Am J Hum Genet.* 2014;94(4):511–21.
22. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol.* 2012;11(11):951–62.
23. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet.* 2015;47(11):1236–41.
24. Zhu X, Feng T, Tayo BO, Liang J, Young JH, Franceschini N, et al. Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. *Am J Hum Genet.* 2015;96(1):21–36.
25. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4:7.
26. Plagnol V, Smyth DJ, Todd JA, Clayton DG. Statistical independence of the colocalized association signals for type 1 diabetes and RPS26 gene expression on chromosome 12q13. *Biostatistics.* 2009;10(2):327–34.
27. Wallace C. Statistical testing of shared genetic control for potentially related traits. *Genet Epidemiol.* 2013;37(8):802–13.
28. 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. *PLoS Genet.* 2014;10(5):e1004383.
29. Gormley P, Anttila V, Winsvold BS, Palta P, Esko T, Pers TH, et al. Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nat Genet.* 2016;48(8):856–66.
30. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014;511(7510):421-7.

31. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet.* 2016;48(3):245–52.
32. Integrating Gene Expression with Summary Association. *Statistics to Identify Genes Associated with 30 Complex Traits.* *The American Journal of Human Genetics.* 2017;100(3):473–87.
33. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* 2003;32(1):1–22.
34. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet.* 2014;23(R1):R89–98.
35. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol.* 2013;178(7):1177–84.
36. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol.* 2015;44(2):512–25.
37. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. *Stat Med.* 2016;35(11):1880–906.
38. O'Leary NA, Wright MW, Brister JR, Ciufu S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 2016;44(D1):D733-45.
39. Pei YF, Zhang L, Liu Y, Li J, Shen H, Liu YZ, et al. Meta-analysis of genome-wide association data identifies novel susceptibility loci for obesity. *Hum Mol Genet.* 2014;23(3):820–30.
40. Elks CE, Perry JR, Sulem P, Chasman DI, Franceschini N, He C, et al. Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat Genet.* 2010;42(12):1077–85.
41. Sniekers S, Stringer S, Watanabe K, Jansen PR, Coleman JRI, Krapohl E, et al. Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat Genet.* 2017;49(7):1107–12.
42. Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun.* 2018;9(1):2098.
43. Lam M, Trampush JW, Yu J, Knowles E, Davies G, Liewald DC, et al. Large-Scale Cognitive GWAS Meta-Analysis Reveals Tissue-Specific Neural Expression and Potential Nootropic Drug Targets. *Cell Rep.* 2017;21(9):2597–613.
44. Hill WD, Marioni RE, Maghzian O, Ritchie SJ, Hagenaars SP, McIntosh AM, et al. A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. *Mol Psychiatry.* 2019;24(2):169–81.
45. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J, et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet.* 2018;50(7):920–7.

46. Lam M, Hill WD, Trampush JW, Yu J, Knowles E, Davies G, et al. Pleiotropic Meta-Analysis of Cognition, Education, and Schizophrenia Differentiates Roles of Early Neurodevelopmental and Adult Synaptic Pathways. *Am J Hum Genet.* 2019;105(2):334–50.
47. Baselmans BML, Jansen R, Ip HF, van Dongen J, Abdellaoui A, van de Weijer MP, et al. Multivariate genome-wide analyses of the well-being spectrum. *Nat Genet.* 2019;51(3):445–51.
48. Langefeld CD, Ainsworth HC, Cunninghame Graham DS, Kelly JA, Comeau ME, Marion MC, et al. Transancestral mapping and genetic load in systemic lupus erythematosus. *Nat Commun.* 2017;8:16021.
49. Levine DM, Ek WE, Zhang R, Liu X, Onstad L, Sather C, et al. A genome-wide association study identifies new susceptibility loci for esophageal adenocarcinoma and Barrett's esophagus. *Nat Genet.* 2013;45(12):1487–93.
50. Qian Y, Ye D, Huang H, Wu DJH, Zhuang Y, Jiang X, et al. Coffee Consumption and Risk of Stroke: A Mendelian Randomization Study. *Ann Neurol.* 2020;87(4):525–32.
51. Yuan S, Larsson SC. No association between coffee consumption and risk of atrial fibrillation: A Mendelian randomization study. *Nutr Metab Cardiovasc Dis.* 2019;29(11):1185–8.
52. van Dam RM, Hu FB, Willett WC. Coffee, Caffeine, and Health. *N Engl J Med.* 2020;383(4):369–78.
53. Nicolopoulos K, Mulugeta A, Zhou A, Hyppönen E. Association between habitual coffee consumption and multiple disease outcomes: A Mendelian randomisation phenome-wide association study in the UK Biobank. *Clin Nutr.* 2020.
54. Haslam DW, James WP. Obesity *Lancet.* 2005;366(9492):1197–209.
55. Ding M, Bhupathiraju SN, Chen M, van Dam RM, Hu FB. Caffeinated and decaffeinated coffee consumption and risk of type 2 diabetes: a systematic review and a dose-response meta-analysis. *Diabetes Care.* 2014;37(2):569–86.
56. Kim HJ, Cho S, Jacobs DR Jr, Park K. Instant coffee consumption may be associated with higher risk of metabolic syndrome in Korean adults. *Diabetes Res Clin Pract.* 2014;106(1):145–53.
57. Hu GL, Wang X, Zhang L, Qiu MH. The sources and mechanisms of bioactive ingredients in coffee. *Food Funct.* 2019;10(6):3113–26.

Figures

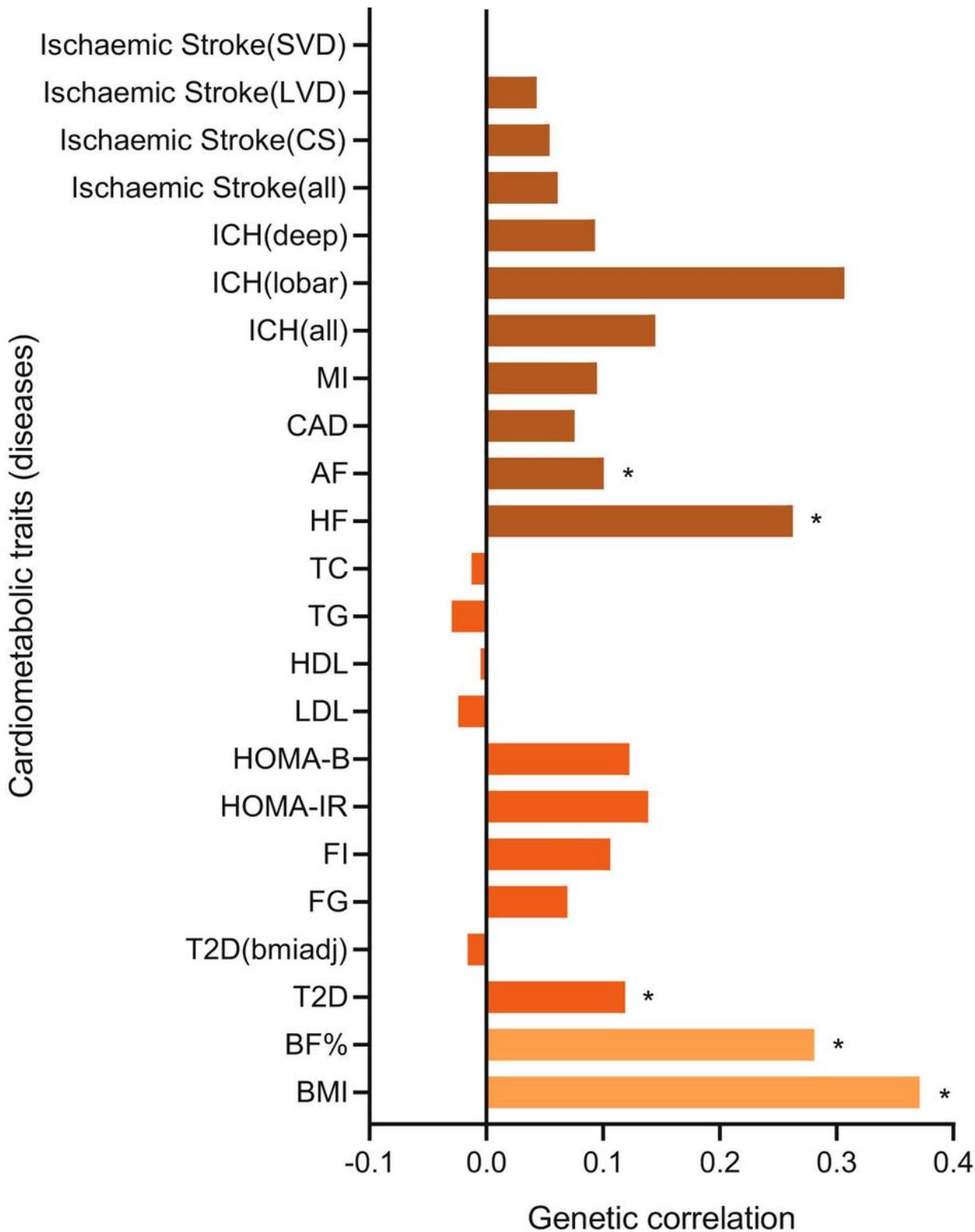


Figure 1

Genetic correlation of coffee intake and cardiometabolic traits (diseases) The vertical axis represents 23 cardiometabolic traits (diseases); the horizontal axis represents the genetic correlation; asterisk indicates that the genetic correlation between the trait/disease and coffee intake is significant (P-value < 0.05/23); the three different colours of the bars represent the three aspects of the 23 traits (diseases).

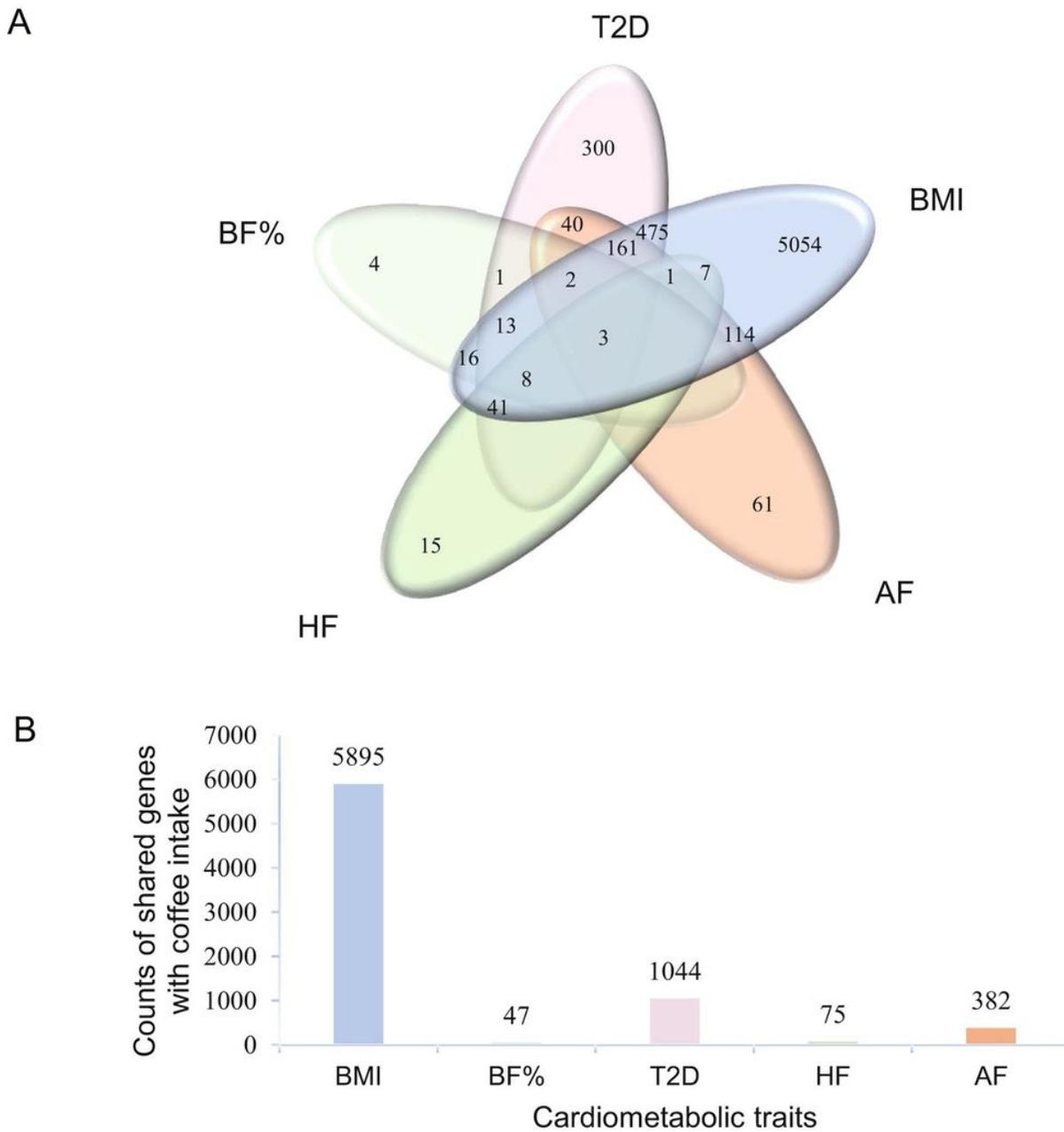


Figure 2

Number of shared genes between coffee intake and cardiometabolic traits (diseases) A. The circles with different colours represent the cross-phenotype association analysis of different cardiometabolic traits (diseases) and coffee intake; the overlapped parts between different circles represent the number of genes jointly identified in two/several cross-phenotype association analysis; the unique part of each circle represents the number of genes only identified in this cross-phenotype association analysis. B. The horizontal axis represents five cardiometabolic traits (diseases) that have significant genetic correlation with coffee intake; the vertical axis represents the number of genes shared by each cardiometabolic trait (disease) and coffee intake.

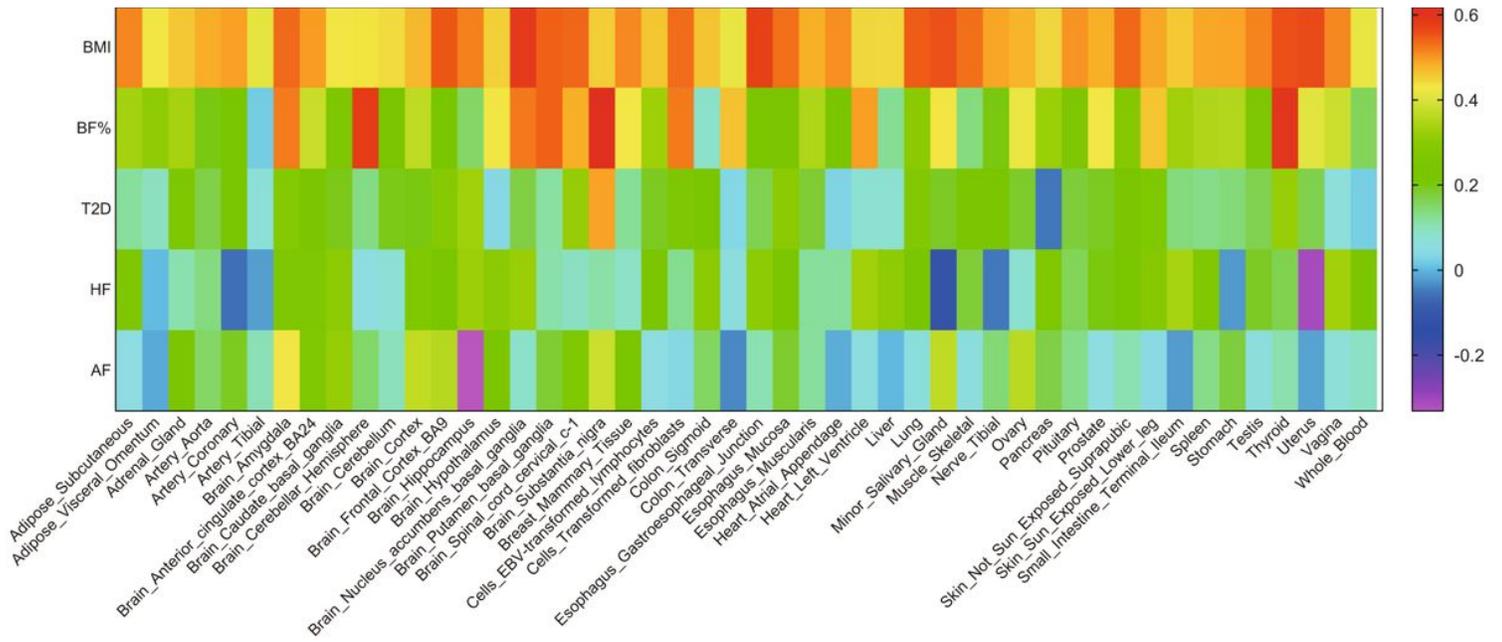


Figure 3

Gene-expression correlation of coffee intake and BMI/BF%/T2D/HF/AF The horizontal axis represents 48 GTEx tissues; the vertical axis represents traits/diseases that have significant genetic correlation with coffee intake; different colours indicate different intensity of gene-expression correlation.

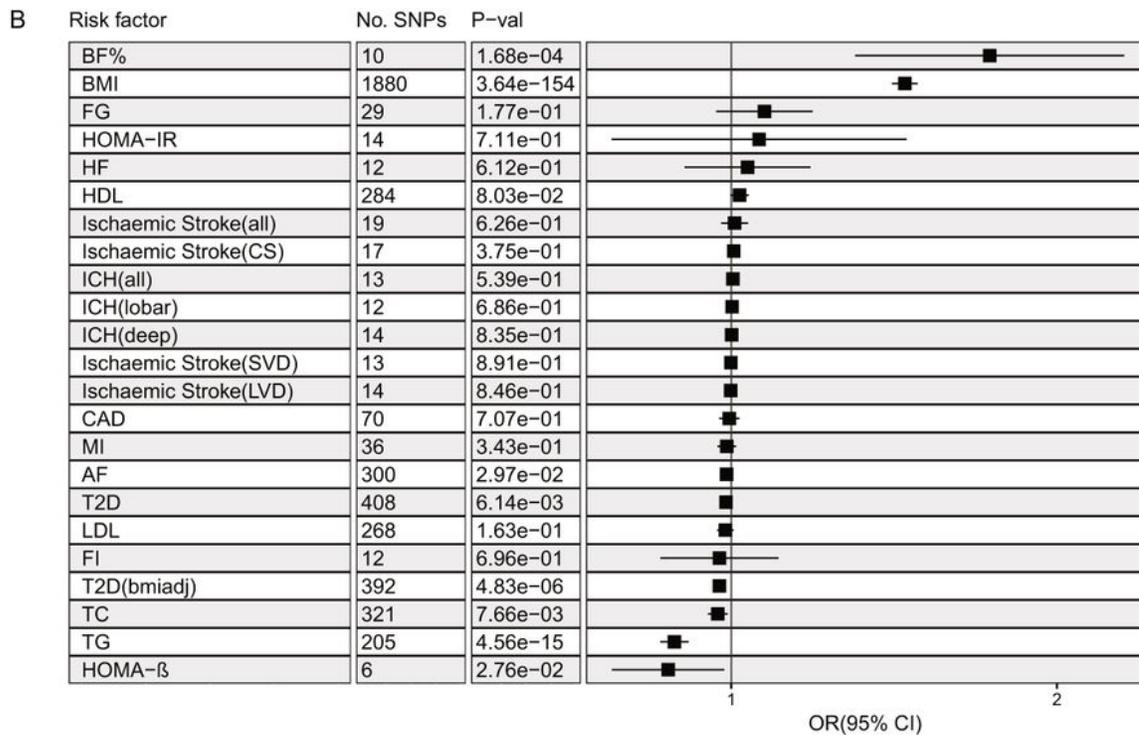
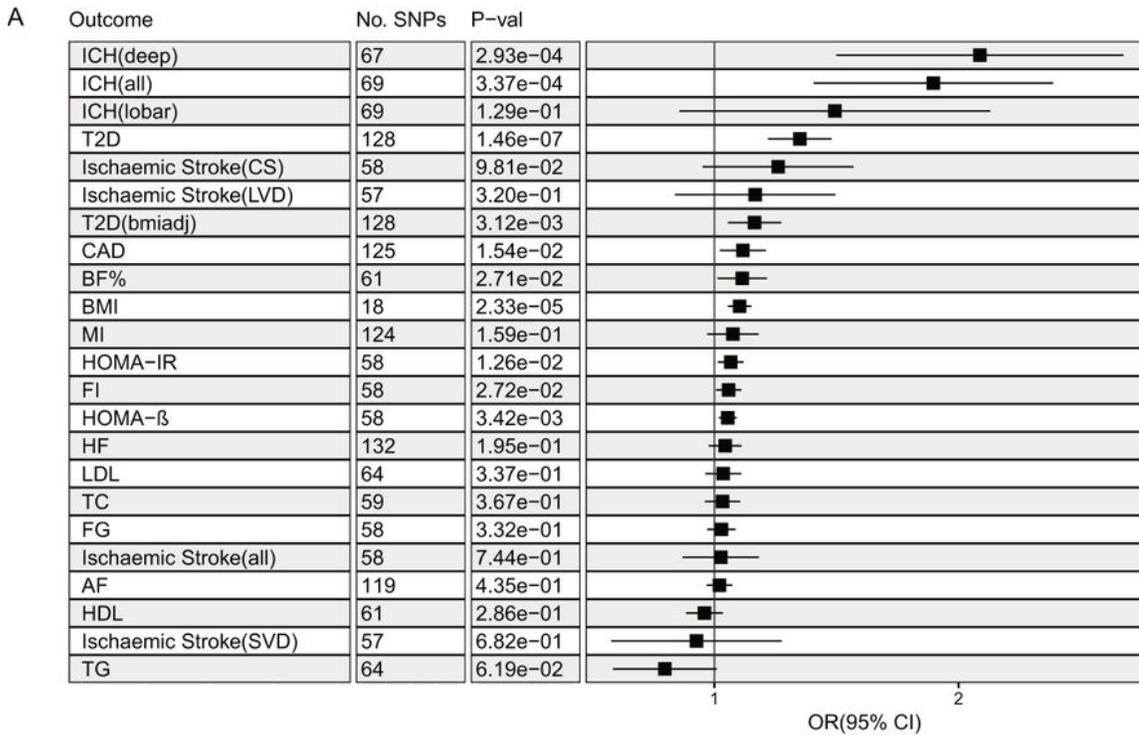


Figure 4

Mendelian Randomization analysis A. The causal effect of coffee intake on cardiometabolic traits (diseases). B. The causal effect of cardiometabolic traits (diseases) on coffee intake

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial.xlsx](#)