Published in final edited form as:

J Mol Cell Cardiol. 2017 November ; 112: 1–7. doi:10.1016/j.yjmcc.2017.08.011.

Novel common variants associated with body mass index and coronary artery disease detected using a pleiotropic cFDR method

Wan-Qiang Lv^a, Xue Zhang^b, Qiang Zhang^a, Jing-Yang He^a, Hui-Min Liu^a, Xin Xia^a, Kun Fan^a, Qi Zhao^{c,d}, Xue-Zhong Shi^a, Wei-Dong Zhang^a, Chang-Qing Sun^a, and Hong-Wen Deng^{a,d,*}

^aCollege of Public Health, Zhengzhou University, No.100 Kexue Road, High-Tech Development Zone of States, Zhengzhou, People's Republic of China

^bDepartment of Geriatrics, Renmin Hospital of Wuhan University, Hubei Zhang Road (Formerly Ziyang Road), Wuchang District No. 99 Jiefang Road 238, Wuhan 430060, People's Republic of China

^cDepartment of Epidemiology, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA 70112, USA

^dCenter for Genomics and Bioinformatics, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA 70112, USA

Abstract

Genome-wide association studies (GWAS) have been successfully applied in identifying single nucleotide polymorphisms (SNPs) associated with body mass index (BMI) and coronary heart disease (CAD). However, the SNPs to date can only explain a small percentage of the genetic variances of traits. Here, we applied a genetic pleiotropic conditional false discovery rate (cFDR) method that combines summary statistic *p* values from different multi-center GWAS datasets, to detect common genetic variants associated with these two traits. The enrichment of SNPs associated with BMI and CAD was assessed by conditional Q-Q plots and the common variants were identified by the cFDR method. By applying the cFDR level of 0.05, 7 variants were identified to be associated with CAD (2 variants being novel), 34 variants associated with BMI (11 variants being novel), and 3 variants associated with both BMI and CAD (2 variants being novel). The SNP rs653178 (*ATXN2*) is noteworthy as this variant was replicated in an independent analysis. SNP rs12411886 (*CNNM2*) and rs794356 (*HIP1*) were of note as the annotated genes may be associated with processes that are functionally important in lipid metabolism. In conclusion, the cFDR method identified novel variants associated with BMI and/or CAD by effectively incorporating different GWAS datasets.

*Corresponding author at: Tulane Center of Bioinformatics and Genomics, Department of Biostatistics and Data Science, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA 70112, USA. hdeng2@tulane.edu (H.-W. Deng).

Competing financial interests

None declared.

Keywords

Body mass index; Conditional FDR; Coronary artery disease; Obesity; Pleiotropy

1. Introduction

Epidemiological studies have estimated that the prevalence of overweight/obesity increased by approximately 41% between 1980 and 2013, making it a major contributor to the rise in coronary heart disease (CAD) [1]. CAD is one of the leading causes of morbidity and mortality worldwide [1]. Risk factors, particularly obesity, have already had well-established associations with CAD [2].

Epidemiological studies supported that obesity is an independent predictor of clinical CAD [3]. A previous study has demonstrated that every 1 kg/m² increase in body mass index (BMI) leads to a 5–7% increase in the incidence of CAD across all BMI categories [4], supporting an inverse relationship between obesity (measured, as conventional, by high BMI) and risk of CAD. Similarly, a large meta-analysis showed that obese participants had a significantly greater risk of CAD (relative risk – RR 1.81, 95% confidence interval – CI 1.56–2.10) after the adjustment for age, sex, physical activities, and smoking [5]. Additional adjustments based on blood pressure and cholesterol levels reduced the RR of obesity to 1.49 (1.32–1.67), but the obesity impact still remained statistically significant. Despite the association between high BMI with an increased occurrence of CAD, studies have also reported the obesity paradox phenomenon that obese patients with established CAD have better clinical outcomes as compared with normal weight patients [6]. Accumulative molecular evidence suggested that obesity might directly be involved in the pathogenesis of CAD [7]. For example, accumulation of body fat could lead to classic metabolic abnormalities [8], including insulin resistance, hyperinsulinemia, hypercholesterolemia, and impaired glucose tolerance, all of which combined would further increase the likelihood of development into CAD [9].

Heritability (h^2) studies demonstrate a substantial genetic contribution to obesity risk ($h^2 \sim 40\text{--}70\%$) [10] and CAD (male twins: $h^2 \sim 45\text{--}69\%$ and female twins: $h^2 \sim 26\text{--}50\%$) [11]. However, the identified genes to date only can explain a small percentage of the variances of BMI and CAD [10,11]. Considering the high degree of heritability in obesity and CAD, more efforts are needed to improve the detection of additional variants that may underlie the “missing heritability”. GWAS have the potential to explain a larger proportion of the heritability, mainly by using enlarged larger sample sizes [12]. However, the additional statistical power gained per subject by increasing recruitments of additional study subjects is limited [13], cost-effective analytical methods are therefore needed to fully utilize the existing GWAS datasets. Such methods have recently been developed and successfully applied, including meta-CCA method [13], Genetic analysis incorporating Pleiotropy and Annotation (GPA) method [14], the method by Zhu X and coworkers [15].

The pleiotropic effect is defined as a single gene or variant being associated with more than two distinct phenotypes [16]. Evidence indicated that pleiotropic effect exists in BMI and CAD, which suggests that they may share common genetic variants [17]. By combining the

independent GWAS from associated trait and disease of BMI and CAD, the samples sizes are effectively enlarged for detection of the pleiotropic genes [18,19]. Based on pleiotropic effect, statistical power and detection of shared variants will be highly improved by leveraging multi-center GWAS datasets of BMI and CAD.

Recently, a pleiotropy-informed cFDR method is developed with the aim to identify some of the missing heritability [20] with GWAS on individual traits/diseases. So far this method has been successfully applied, *e.g.*, in schizophrenia and cardiovascular disease risk factors [20], and blood pressure and other phenotypes [21] by other groups, and by our own group for height and femoral neck bone mineral density [22], type 2 diabetes and birth weight [23], and for CAD and bone mineral density [24]. This method thus theoretically (21) and practically have been shown to have improved the statistical power and improved variants discovery in the studied associated traits or diseases. Here, we applied the genetic pleiotropy-informed cFDR method on summary statistics of two independent meta-GWAS to identify shared variants and pleiotropic effect between BMI and CAD. By using this method, we hypothesized that we could identify more common variants for BMI and/or CAD, and discover some novel etiologic relationship between BMI and CAD.

2. Materials and methods

2.1. GWAS datasets

The dataset for BMI was downloaded from http://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files. This GWAS meta-analysis comprising of 249,796 individuals of white European Ancestry was performed by the Genetic Investigation of Anthropometric Traits (GIANT) Consortium [19]. Two datasets for CAD were downloaded from <http://www.cardiogramplusc4d.org/data-downloads>. The C4D dataset performed by the Coronary Artery Disease (C4D) Genetics Consortium was derived from a meta-analysis of four large GWAS of European and South Asian descent involving 15,420 CAD cases and 15,062 controls [25]. The second dataset conducted by the transatlantic Coronary ARtery Disease Genome-wide Replication and Meta-analysis (CARDIoGRAM) Consortium was derived from a meta-analysis of 22 GWAS of European descent comprising of 22,233 cases and 64,762 controls [18]. All the datasets consist of summary statistics for each SNP based on each meta-analysis publication, providing *p* values after using genomic control (GC) both at the individual study level and after meta-analysis [26]. Further details of the GWAS samples and methods employed within each group were presented in the original references [18,19,25]. Additionally, the CARDIoGRAM dataset for CAD in our analysis was used as the replication dataset. This study used only summary statistics from publically available datasets for previous GWAS. It does not involve human subjects directly. Informed consent was obtained from all participants of contributing studies in the published respective GWAS. Contributing studies received ethical approval from their respective institutional review boards. This study was approved by the Ethical Committee of the Life Sciences of Zhengzhou University.

2.2. Data preparation

Before the analysis, we checked whether there were overlapped samples included in these datasets of the cohorts. We found no individuals were overlapped between C4D and GIANT datasets, and no overlapped individuals between CARDIoGRAM and C4D datasets. However, the datasets used for BMI and CAD had different ancestors, *i.e.*, European Ancestry for BMI and Europeans and South Asians for CAD (45% of the CAD case are Asians).

When dealing with the various datasets, we combined the two phenotypes' summary p values for the common SNPs studied in both datasets. After annotating the common SNPs, we applied a LD-based pruning method to remove the large correlations between pairs of variants. The minor allele frequency (MAF) was used as a criterion in the SNP pruning method, which removed the SNP with smaller MAF for pairs with $R^2 > 0.2$. The datasets were pruned using the HapMap3 genotypes of the corresponding matched ethnicity references. First, this method proceeds by using 50 SNPs as a group where LD is calculated between each pair of SNPs. SNPs with smaller MAF were removed from our analysis if their measured LD had an $R^2 > 0.2$. Then, move forward by 5 SNPs and the process is repeated until there are no pairs of SNPs that are > 0.2 . This pruning method ensures that SNPs are not in LD with each other in the follow-up analysis. It is suggested that complex correlations among the test-statistics may bias the estimate of the conditional FDR, including LD score regression [27] and the effect of correlation in FDR estimation [28]. There are two regions with complex LD structures [29], including the extended major histocompatibility complex (MHC) (chr6:25652429–33368333) and chromosomal region 8p23.1 (chr8:7242715–12483982). Thus, we constructed conditional Q-Q plots and the cFDR analysis after excluding SNPs within these regions to remove potential bias introduced by them.

2.3. Statistical analysis

2.3.1. Genomic control (GC)—Population stratification can be a problem for association studies, where the association could be found due to the underlying structure of the population and not a disease associated locus. GC is one of the most widely used approaches to correct for this problem, which controls the inflation of test statistics due to population structures [30]. GC works by using markers that are not linked with the trait in question to correct for any inflation of the test statistic caused by population stratification [26]. GC has been applied in the original GWAS at the individual study level and for the meta-analysis, there is no need for us to repeat it here in our analyses.

2.3.2. Conditional Q-Q plots for accessing pleiotropic effect enrichment—Q-Q plots are a descriptive tool for visualizing the difference between observed distribution and theoretical distribution. In the analysis of GWAS, quantiles of the observed p -values (nominal), denoted by 'p', are plotted on the y-axis, and quantiles of the theoretical null distribution (the uniform distribution estimated by the empirical cumulative distribution function), here denoted by 'q', are plotted on the x-axis. Commonly, the negative log transformation was used, we denoted the y-axis as nominal $-\log_{10}(p)$, and the x-axis as empirical $-\log_{10}(q)$. The enrichment of pleiotropic effect is graphically accessed by

conditional Q-Q plots [20]. Under the null hypothesis, enrichment of pleiotropic effect can be reflected by leftward deflections of the observed distribution from the null line. If the Q-Q plot falls on the line $x = y$, with no deviation between lines, this would indicate no enrichment of genetic pleiotropic effect. If pleiotropic enrichment does exist, an earlier leftward shift from the null line will be present. Larger spacing in the Q-Q plots is interpreted as a greater extent of pleiotropic effect shared between traits [20].

Here, we presented conditional Q-Q plots of BMI conditioned on summary statistics p values of CAD, and *vice versa*. Specifically, we constructed different categories based on p -values of SNPs for the second condition phenotype ($-\log_{10}(p) \leq 0$, $-\log_{10}(p) \in (0, 1]$, $-\log_{10}(p) \in (1, 2]$, and $-\log_{10}(p) > 2$ corresponding to $p \leq 1$, $p \in (0.1, 1]$, $p \in (0.01, 0.1]$, and $p \leq 0.001$, respectively).

In order to check whether the pleiotropic effect enrichment was consistent, we conducted a replication analysis. C4D dataset for CAD was used as a discovery dataset for cFDR and conjunction FDR analyses with BMI. As there were no overlapping individuals between datasets CARDIoGRAM and C4D, the dataset of CAD performed by CARDIoGRAM was independent of C4D for the replication analysis. If the pleiotropic effect enrichment does exist in these two phenotypes, these two datasets of CAD conditioned on BMI or *vice versa* would generate similar QQ plots.

2.3.3. Conditional and conjunction FDR for identifying shared variants—To identify the variants, we computed conditional FDR for either BMI or CAD, and conjunction FDR for both BMI and CAD. Conditional and conjunction FDR methods are both extensions of the FDR method [31].

FDR method is based on the assumption that SNPs are associated with the phenotype, which has been extensively applied to control the type I errors in multiple hypothesis testing [32]. The FDR for SNPs is defined as a posterior probability that a random SNP from this set is null given that the nominal p values are smaller than the pre-defined cutoffs ($P\text{-values} < p_i$).

$$FDR(p_i) = \Pr(H_0^{(i)} | P_i \leq p_i) \quad (1)$$

where P_i is the random variable that represents the associations between all SNPs and trait i , and p_i represents an example of this random variable corresponding to the certain SNP of interest. $H_0^{(i)}$ indicates the null hypothesis that there is no association between a particular SNP of interest and trait i . Given a set of observed association p -values, the FDR is estimated as the nominal p -value obtained from the GWAS summary statistics divided by the observed quantiles.

The cFDR extends FDR method to two phenotypes situation and is defined as the posterior probability that a given SNP is null for the first phenotype given that the p -values for both phenotypes are as small or smaller as their observed p -values. The cFDR is expressed as cFDR ($p_i | p_j$), where p_i represents the strength of association for a particular variant with the

phenotype i and p_j represents the strength of association for that same SNP with the phenotype j . The cFDR is given by

$$\text{cFDR}(p_i|p_j) = \Pr(H_0^{(1)} | p_i \leq p_i, p_j \leq p_j) \quad (2)$$

Low values of cFDR indicate that the SNPs are associated with both phenotypes or with the primary phenotype only. Therefore, to identify whether SNPs are genetically pleiotropic, we computed the conjunction FDR value, defined as the probability that an SNP is null for either phenotype or for both phenotypes simultaneously given its P -value in both phenotypes are as small or smaller as the observed ones. Low values of conjunction FDR indicate that the SNPs are associated with both phenotypes. The conjunction FDR was calculated as the maximum conditional FDR value in the two cFDR values, and given by

$$\text{Conjunction FDR}_{i \& j} = \max(\text{cFDR}_{i|j}, \text{cFDR}_{j|i}) \quad (3)$$

In the overall analysis of this study, the cFDR value of each SNP was computed in the form that where BMI is the principal phenotype conditioned on the association with CAD(BMI|T2D) and *vice versa* (T2D|BMI). We pre-defined a cutoff threshold of 0.05, which means that 5 of the SNPs are falsely positive per 100 SNPs. If the conditional and conjunction FDR values of SNPs were smaller than 0.05, they were identified to be significantly associated with the phenotype.

2.3.4. Manhattan plots for localizing genetic variants—Additionally, we constructed Manhattan plots based on the conditional and conjunction FDR values of SNPs to illustrate the genetic markers' genomic locations. The Manhattan plots present all SNPs within an LD block in relation to their chromosomal locations. The 22 chromosomal locations are plotted on the x-axis, and the $-\log_{10}$ the SNPs' values are plotted on the y-axis. If the $-\log_{10}$ cFDR value of a certain SNP is > 1.3 , this SNP is determined to be associated with the principal phenotype given the conditional phenotype. And if the $-\log_{10}$ conjunction FDR value is > 1.3 , this SNP is determined to be associated with both the phenotypes.

2.3.5. Functional term enrichment analysis—When given a set of genes, the GO defines concepts used to describe gene function, and relationships between these concepts. It classifies functions along three aspects, including molecular function, cellular component and biological process. Using the GO term enrichment analysis, an enrichment analysis will find which GO terms are overrepresented (or under-represented) using annotations for that gene set. In the GO enrichment analysis, implicated genes (not SNPs) are used and all identified genes are treated the same without referring to their association strength identified in the GWAS or this analysis. Enrichment analysis may provide further insight into the mechanism why individuals with a high BMI may suffer from CAD simultaneously.

3. Results

In order to identify common variants between BMI and CAD, we undertook a two-step analysis strategy. First, by using conditional Q-Q plots, we inspected the enrichment of pleiotropic effect between the two phenotypes. Next, by adopting the cFDR method, we identified those potential common variants.

3.1. Enrichment of pleiotropic effect between BMI and CAD by Q-Q plots

The conditional Q-Q plot for CAD conditioned on BMI showed much enrichment across different levels of significance thresholds, shown in the upper panel (left) of Fig. 1. Enrichment was also observed for BMI conditioned on CAD, as evidenced by a greater degree of departure from the null line across increasing different significance levels, shown in the upper panel (right) of Fig. 1. These earlier deflections from the null line indicated a large proportion of true associations for any given CAD nominal p -values.

3.2. CAD and BMI variants identified by conditional FDR

Based on the enrichment of pleiotropic effect between BMI and CAD in step one, we performed the conditional FDR analysis on them to investigate which variants were associated with CAD conditioned on BMI, and *vice versa* for BMI.

Conditional on the association with BMI, we identified 7 significant SNPs for CAD. Among these SNPs, three reached the standard threshold of genome-wide significance ($p < 5 \times 10^{-8}$), one had p values smaller than 1×10^{-5} but larger than 5×10^{-8} , and the remaining SNPs had the p values between 1×10^{-5} and 1×10^{-2} . There are 6 genes annotated by these 7 SNPs. Details on these identified CAD SNPs were shown in the Supplementary Fig. S1 and Table S1. Conditional on the association with CAD, we identified 34 significant SNPs for BMI. Among these 34 SNPs, 9 reached the standard threshold of genome-wide significance ($p < 5 \times 10^{-8}$), and 19 had p -values smaller than 1×10^{-5} but larger than 5×10^{-8} , the p values of rest SNPs were between 1×10^{-5} and 1×10^{-2} . These SNPs were annotated to 39 genes. Detailed information of these identified BMI SNPs was shown in the Supplementary Fig. S2 and Table S2.

3.3. Pleiotropic variants for both CAD and BMI identified by conjunction FDR

The conjunction FDR values were computed to identify genetic pleiotropic variants which were associated with both BMI and CAD. At the significance level of 0.05, three variants were identified to be associated with both BMI and CAD. Detailed information of these identified SNPs was shown in the Table 1 and Fig. 2. Of the 3 pleiotropic variants, SNPs rs12411886 (*CNNM2*) was previously reported to be associated with both BMI and CAD [18,33,34]. SNP rs653178 (*ATXN2*) influenced blood pressure and cardiovascular disease risk [33]. SNP rs794356 was not identified by previous GWAS but is located at gene *HIP1* which has been implicated in the development of primary metabolic processes [35].

3.4. Functional term enrichment analysis

When the genes annotated by the variants which were associated with BMI or CAD or both were used as the gene sets for the GO term enrichment analysis, several functional terms

were identified to be enriched in the development of BMI or CAD or both. Detailed information was showed in the Supplementary Table S3. Some GO term enrichment analysis may contribute to the etiology of both BMI and CAD. For example, the GO terms of “positive regulation of metabolic process” and “regulation of biological quality” influence the cardiovascular disease risk factors, including fasting glucose [36], low-density lipoprotein, total cholesterol, and triglyceride [37].

3.5. Replication analysis

To address the possibility that the observed pattern of enrichment may result from spurious associations, we conducted a replication analysis. First, we investigated whether the general pleiotropic enrichment pattern observed earlier with the discovery analyses of the C4D dataset with the BMI dataset could be replicated by the conditional Q-Q plot analyses in the second independent CAD GWAS dataset with the BMI dataset. As a result, we observed a similar pleiotropic enrichment pattern which was illustrated by the conditional Q-Q plots, shown in the lower panel of Fig. 1. We also identified the variants which were associated with BMI or CAD or both in the replication analysis. If the R^2 value is > 0.6 , it represents that these two SNPs (one is in the discovery phase, another one is in the replication analysis) are in high LD, this SNP is also considered to be replicated in our analysis. In the discovery phase of analysis, we identified 7 and 34 variants associated with CAD and BMI, respectively. As a result, we replicated 4 and 26 variants associated with CAD and BMI, respectively. For the variants which were with both BMI and CAD, one of the three variants was replicated in the replication analysis. Detailed information was shown in the Table 1 and Supplementary Tables S1–S2. These results indicated that the pleiotropic effect enrichment between BMI and CAD was largely consistent and some of them can be replicated across studies.

4. Discussion

In our study, we investigated the degree of genetic pleiotropy and identify common genetic variants between BMI and CAD by applying the cFDR method. We used a two-step analysis strategy. First, we constructed and assessed conditional Q-Q plots in CAD and BMI for different strata of statistical significance of the conditional trait. The Q-Q plots provided a visualization of pleiotropic effect enrichment. Second, we computed each SNP's conditional FDR value to identify the variants of pleiotropic enrichment. As a result, we found and replicated the enrichment of pleiotropic effect between BMI and CAD, using two independent GWAS datasets of CAD. Applying the cFDR method at the significance level of 0.05, 7 variants were identified to be associated with CAD (5 variants were associated with CAD according to the PubMed and 2 variants were novel), 34 variants were associated with BMI (23 variants were associated with BMI earlier and 11 were novel), and three variants were associated with both CAD and BMI (one variants was associated with both CAD and BMI earlier, and 2 were novel).

Among the SNPs and genes identified in this study, a number of variants and annotated genes were reported to be associated with BMI or CAD in previous studies. For example, 5 SNPs were reported to be associated with CAD [18,38], 14 SNPs were reported to be

associated with BMI [10,39] in the previous studies. Moreover, two SNPs showed pleiotropic effects, SNP rs12411886 was associated with the cardiovascular risk factors [18,33,34], and SNP rs11672660 was associated with BMI and glycemic traits [10,40]. Out of the 39 genes annotated by BMI SNPs, 18 genes were associated with BMI in the previous GWAS [19,41]. Three out of 6 genes annotated by CAD SNPs were identified to play an important role in the development of CAD [18,42]. Detailed information of the SNPs and genes were shown in the Table 1 and Supplementary S1-S2. The novel loci on 12q24.12 (rs653178, *ATXN2*) and 7q11.23 (rs794356, *HIP1*) are particularly interesting.

SNP rs653178 (12q24.12) is located in the intronic region of the *ATXN2* gene, which encodes proteins that are involved in endocytosis and mitochondrial function [43]. GWAS indicated that loss-of-function mutations in this gene may be associated with susceptibility to high BMI [44]. A recently published study documented that *ATXN2* regulates numerous nutrient enzymes in the mitochondrial matrix as well as key factors that may respond to altered Ca^{2+} storage in mitochondria, *ATXN2* mutations lead to a profound mitochondrial dysfunction [45]. Mouse studies showed the knock-out of *ATXN2* to lead to obesity, insulin resistance, and dyslipidemia [46].

SNP rs794356 (7q11.23) is located in the intronic region of the *HIP1* gene, which encodes proteins correlate with increased epidermal growth factor receptor levels in certain tumors [47]. *HIP1* protein is necessary for fundamental cellular and organismal homeostasis *in vivo* phenotypes, deficiency of *HIP1* leads to adult weight loss and early death [35]. *HIP1* protein has also been found to be up-regulated in multiple tumor types, widely implicated in tumorigenesis [48]. Based on the close relationship between obesity and certain cancers, *HIP1* gene might be involved in primary metabolic processes and cellular metabolic processes that are important in the development of high BMI.

Based on our results, we may speculate that some of the underlying shared etiology mechanisms between high BMI and CAD. The potential explanation why high BMI and CAD often co-occur may be partially explained by increased downstream development of metabolic abnormalities, including hypertension, impaired fasting glucose, and dyslipidemia. This explanation was partially supported by the GO term enrichment analysis. The GO terms of “positive regulation of metabolic process”, “regulation of biological quality”, “regulation of signaling”, and “regulation of receptor-mediated endocytosis” have significant impact on the metabolism of blood sugar and lipids, including fasting glucose [36], low-density lipoprotein, total cholesterol, and triglyceride [37], the dysfunction of them would lead to the increased risk of CAD and high BMI. “Single-organism metabolic process” is involved in many disease processes, particularly “regulation of high-density lipoprotein particle assembly” may contribute to the etiology of both BMI and CAD [49].

Applying the cFDR method by leveraging two large GWAS datasets, the statistical power was increased, for the effective sample size for those pleiotropic loci. We applied the second dataset to test whether the enrichment of pleiotropic effect was consistent. In the replication analysis, we replicated a large number of variants which were associated with BMI or CAD. However, we did not replicate all the variants, possibly due to the following potential reasons: The C4D GWAS is a meta-analysis of GWAS studies of European and South Asian

descent, while CARDIoGRAM GWAS included studies of European descent only. Whether this fact would present a potential source of bias for cFDR analysis is not clear and worth further theoretical analyses perhaps by simulations, since the original meta-analyses would yield genetic loci consistent across studied various ethnic groups. However, we did not identify all the variants which were reported to be associated with BMI and CAD in previous studies. It may be because that this study only enhanced the power to detect those loci which had pleiotropic effects. Another limitation of this study is that these findings could not relate to individual clinical outcomes, as we only used the summary statistics and had no access to original individual clinical outcome measures.

In summary, this study showed the usefulness of cFDR method by taking advantage of the pleiotropic effect of two complex related traits. We observed a strong pleiotropic enrichment between BMI and CAD, and identified numbers of common genetic variants. This study may facilitate the discovery of new markers for early identification and new treatment targets which could lead to improved prevention and treatment regimens in CAD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We appreciate the support from Zhengzhou University. HWD was partially supported by grants from the National Institutes of Health [R01AR057049, R01AR059781, D43TW009107, P20 GM109036, R01MH107354, R01MH104680, R01GM109068], and the Edward G. Schlieder Endowment fund to Tulane University. Wan-Qiang Lv as the first author performed data analysis and wrote the manuscript. Jing-Yang He, Qiang Zhang, Xin Xia, and Hui-Min Liu all gave constructive suggestions during the whole process. Qi Zhao provided critical revisions. H.W.D conceived and initiated this project, revised and finalized the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.ymcc.2017.08.011>.

References

1. Writing Group M, Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S, et al. Heart disease and stroke statistics—2010 update: a report from the American Heart Association. *Circulation*. 2010; 121:e46–e215. [PubMed: 20019324]
2. Chen G, Levy D. Contributions of the Framingham Heart Study to the epidemiology of coronary heart disease. *JAMA Cardiol*. 2016; 1:825–830. [PubMed: 27464224]
3. Hubert HB, Feinleib M, McNamara PM, Castelli WP. Obesity as an independent risk factor for cardiovascular disease: a 26-year follow-up of participants in the Framingham Heart Study. *Circulation*. 1983; 67:968–977. [PubMed: 6219830]
4. Cho E, Manson JE, Stampfer MJ, Solomon CG, Colditz GA, Speizer FE, et al. A prospective study of obesity and risk of coronary heart disease among diabetic women. *Diabetes Care*. 2002; 25:1142–1148. [PubMed: 12087011]
5. Bogers RP, Bemelmans WJ, Hoogenveen RT, Boshuizen HC, Woodward M, Knekt P, et al. Association of overweight with increased risk of coronary heart disease partly independent of blood pressure and cholesterol levels: a meta-analysis of 21 cohort studies including more than 300 000 persons. *Arch Intern Med*. 2007; 167:1720–1728. [PubMed: 17846390]

6. Lin GM, Li YH, Yin WH, Wu YW, Chu PH, Wu CC, et al. The obesity-mortality paradox in patients with heart failure in Taiwan and a collaborative meta-analysis for East Asian patients. *Am J Cardiol.* 2016; 118:1011–1018. [PubMed: 27521221]
7. Shen GQ, Li L, Wang QK. Genetic variant R952Q in LRP8 is associated with increased plasma triglyceride levels in patients with early-onset CAD and MI. *Ann Hum Genet.* 2012; 76:193–199. [PubMed: 22404453] Ozaki K, Tanaka T. Molecular genetics of coronary artery disease. *J Hum Genet.* 2016; 61:71–77. [PubMed: 26134515]
8. Krauss RM, Winston M, Fletcher RN, Grundy SM. Obesity: impact of cardiovascular disease. *Circulation.* 1998; 98:1472–1476.
9. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature.* 2006; 444:840–846. [PubMed: 17167471]
10. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature.* 2015; 518:197–206. [PubMed: 25673413]
11. Zdravkovic S, Wienke A, Pedersen NL, Marenberg ME, Yashin AI, De Faire U. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. *J Intern Med.* 2002; 252:247–254. [PubMed: 12270005]
12. Stahl EA, Wegmann D, Trynka G, Gutierrez-Achury J, Do R, Voight BF, et al. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat Genet.* 2012; 44:483–489. [PubMed: 22446960]
13. Cichonska A, Rousu J, Marttinen P, Kangas AJ, Soininen P, Lehtimäki T, et al. metaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis. *Bioinformatics.* 2016; 32:1981–1989. [PubMed: 27153689]
14. Chung D, Yang C, Li C, Gelernter J, Zhao H. GPA: a statistical approach to prioritizing GWAS results by integrating pleiotropy and annotation. *PLoS Genet.* 2014; 10:e1004787. [PubMed: 25393678]
15. 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; 96:21–36. [PubMed: 25500260]
16. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet.* 2013; 14:483–495. [PubMed: 23752797]
17. Cole CB, Nikpay M, Stewart AF, McPherson R. Increased genetic risk for obesity in premature coronary artery disease. *Eur J Hum Genet.* 2016; 24:587–591. [PubMed: 26220701]
18. Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet.* 2011; 43:333–338. [PubMed: 21378990]
19. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet.* 2010; 42:937–948. [PubMed: 20935630]
20. Andreassen OA, Djurovic S, Thompson WK, Schork AJ, Kendler KS, O'Donovan MC, et al. Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *Am J Hum Genet.* 2013; 92:197–209. [PubMed: 23375658]
21. Andreassen OA, McEvoy LK, Thompson WK, Wang Y, Reppe S, Schork AJ, et al. Identifying common genetic variants in blood pressure due to polygenic pleiotropy with associated phenotypes. *Hypertension.* 2014; 63:819–826. [PubMed: 24396023]
22. Greenbaum J, Wu K, Zhang L, Shen H, Zhang J, Deng HW. Increased detection of genetic loci associated with risk predictors of osteoporotic fracture using a pleiotropic cFDR method. *Bone.* 2017; 99:62–68. [PubMed: 28373146]
23. Zeng, CP., Chen, YC., Lin, X., Greenbaum, J., Chen, YP., Peng, C., et al. Increased identification of novel variants in type 2 diabetes, birth weight and their pleiotropic loci. *J Diabetes.* 2016. <http://dx.doi.org/10.1111/1753-0407.12510>
24. Peng C, Shen J, Lin X, Su KJ, Greenbaum J, Zhu W, et al. Genetic sharing with coronary artery disease identifies potential novel loci for bone mineral density. *Bone.* 2017; 103:70–77. [PubMed: 28651948]

25. Coronary Artery Disease Genetics C. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet.* 2011; 43:339–344. [PubMed: 21378988]
26. Bacanu SA, Devlin B, Roeder K. Association studies for quantitative traits in structured populations. *Genet Epidemiol.* 2002; 22:78–93. [PubMed: 11754475]
27. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet.* 2015; 47:291–295. [PubMed: 25642630]
28. Schwartzman A, Lin X. The effect of correlation in false discovery rate estimation. *Biometrika.* 2011; 98:199–214. [PubMed: 23049127]
29. Price AL, Weale ME, Patterson N, Myers SR, Need AC, Shianna KV, et al. Long-range LD can confound genome scans in admixed populations. *Am J Hum Genet.* 2008; 83:132–135. (author reply 135–139). [PubMed: 18606306]
30. Devlin B, Roeder K, Wasserman L. Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol.* 2001; 60:155–166. [PubMed: 11855950]
31. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res.* 2001; 125:279–284. [PubMed: 11682119]
32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple. *J R Stat Soc.* 1995; 57:289–300.
33. International Consortium for Blood Pressure Genome-Wide Association S. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature.* 2011; 478:103–109. [PubMed: 21909115]
34. Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014; 511:421–427. [PubMed: 25056061]
35. Bradley SV, Hyun TS, Oravec-Wilson KI, Li L, Waldorff EI, Ermilov AN, et al. Degenerative phenotypes caused by the combined deficiency of murine HIP1 and HIP1r are rescued by human HIP1. *Hum Mol Genet.* 2007; 16:1279–1292. [PubMed: 17452370]
36. Nielson C, Lange T, Hadjokas N. Blood glucose and coronary artery disease in nondiabetic patients. *Diabetes Care.* 2006; 29:998–1001. [PubMed: 16644627]
37. Karalis DG, Victor B, Ahedor L, Liu L. Use of lipid-lowering medications and the likelihood of achieving optimal LDL-cholesterol goals in coronary artery disease patients. *Cholesterol.* 2012; 2012:861924. [PubMed: 22888414]
38. 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; 47:1121–1130. [PubMed: 26343387]
39. Randall JC, Winkler TW, Kutalik Z, Berndt SI, Jackson AU, Monda KL, et al. Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genet.* 2013; 9:e1003500. [PubMed: 23754948]
40. Scott RA, Lagou V, Welch RP, Wheeler E, Montasser ME, Luan J, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet.* 2012; 44:991–1005. [PubMed: 22885924]
41. Vazquez G, Duval S, Jacobs DR Jr, Silventoinen K. Comparison of body mass index, waist circumference, and waist/hip ratio in predicting incident diabetes: a meta-analysis. *Epidemiol Rev.* 2007; 29:115–128. [PubMed: 17494056]
42. Larson MG, Atwood LD, Benjamin EJ, Cupples LA, D'Agostino RB Sr, Fox CS, et al. Framingham Heart Study 100K project: genome-wide associations for cardiovascular disease outcomes. *BMC Med Genet.* 2007; 8(Suppl 1):S5. [PubMed: 17903304] Myocardial Infarction Genetics C. Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet.* 2009; 41:334–341. [PubMed: 19198609]
43. Sellier C, Campanari ML, Corbier C Julie, Gaucherot A, Kolb-Cheynel I, Oulad-Abdelghani M, et al. Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *EMBO J.* 2016; 35:1276–1297. [PubMed: 27103069]

44. Figueroa KP, Farooqi S, Harrup K, Frank J, O'Rahilly S, Pulst SM. Genetic variance in the spinocerebellar ataxia type 2 (ATXN2) gene in children with severe early onset obesity. *PLoS One*. 2009; 4:e8280. [PubMed: 20016785]
45. Halbach MV, Gispert S, Stehning T, Damrath E, Walter M, Auburger G. Atxn2 knockout and CAG42-knock-in cerebellum shows similarly dysregulated expression in calcium homeostasis pathway. *Cerebellum*. 2017; 16:68–81. [PubMed: 26868665]
46. Meierhofer D, Halbach M, Sen NE, Gispert S, Auburger G. Ataxin-2 (Atxn2)-knock-out mice show branched chain amino acids and fatty acids pathway alterations. *Mol Cell Proteomics*. 2016; 15:1728–1739. [PubMed: 26850065]
47. Bradley SV, Holland EC, Liu GY, Thomas D, Hyun TS, Ross TS. Huntingtin interacting protein 1 is a novel brain tumor marker that associates with epidermal growth factor receptor. *Cancer Res*. 2007; 67:3609–3615. [PubMed: 17440072]
48. Rao DS, Hyun TS, Kumar PD, Mizukami IF, Rubin MA, Lucas PC, et al. Huntingtin-interacting protein 1 is overexpressed in prostate and colon cancer and is critical for cellular survival. *J Clin Invest*. 2002; 110:351–360. [PubMed: 12163454]
49. Sattler K, Graler M, Keul P, Weske S, Reimann CM, Jindrova H, et al. Defects of high-density lipoproteins in coronary artery disease caused by low sphingosine-1-phosphate content: correction by sphingosine-1-phosphate-loading. *J Am Coll Cardiol*. 2015; 66:1470–1485. [PubMed: 26403344]

Fig. 1.

Stratified QQ plots were shown for BMI and CAD. Upper Panel: The plots of nominal *versus* empirical $-\log_{10} p$ -values in (left) BMI as a function of significance of the association with CAD, and in (right) CAD as a function of significance of the association with BMI. Lower Panel: The plots showed the pleiotropic enrichment pattern replicated in the CARDIoGRAM GWAS dataset.

Conjunction Manhattan plot for CAD and BMI. The figure shows the genomic locations of pleiotropic loci and further details are provided in Table 1. The red line in the figure corresponds to the recommended conjunction $-\log$ FDR value of 1.3. The plot has 10 SNPs with conjunction $-\log$ FDR values greater than the threshold value. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1

Conjunction FDR: pleiotropic variants in CAD and BMI (cFDR < 0.05).

Chr	SNP	Nearest gene	Allele	Role	Discovery analysis			Replication analysis		Marker reported	
					P _{CAD}	P _{BMI}	Conjunction FDR	SNP	Conjunction FDR	Trait	Study name
10	rs12411886	CNNM2 ^a	C/A	Intronic	5.E-05	4.E-05	0.002	rs4290163 (R ² = 0.23)	0.030	CAD BMI SBP	21378990 25673413 21909115
12	rs653178	ATXN2	C/T	Intronic	1.E-04	4.E-04	0.009	rs10774625 (R ² = 0.93)	0.002	BP	21909110
7	rs794356	HIP1	G/A	Intronic	5.E-05	4.E-04	0.018			AD	21383967

^aGenes identified in our study have been reported to be associated with BMI and CAD in the previous GWAS. The R² is the measure of linkage disequilibrium between the identified SNP and the SNP which is significant in the replication analysis. If the R² value is > 0.6, it represents that these two SNPs are in high linkage disequilibrium, this SNP is also considered to be replicated. Study name refers to the study which reports the SNP to be associated with the trait in the PubMed. CAD, coronary artery disease; SBP, systolic blood pressure; BMI, body mass index; BP, blood pressure; AD, Autoimmune Diseases; FDR, false discovery rate.