



# Improved Detection of Potentially Pleiotropic Genes in Coronary Artery Disease and Chronic Kidney Disease Using GWAS Summary Statistics

Haimiao Chen<sup>1†</sup>, Ting Wang<sup>1†</sup>, Jinna Yang<sup>2</sup>, Shuiping Huang<sup>1,3\*</sup> and Ping Zeng<sup>1,3\*</sup>

<sup>1</sup> Department of Epidemiology and Biostatistics, School of Public Health, Xuzhou Medical University, Xuzhou, China,

<sup>2</sup> Department of Infectious Diseases, People's Hospital of Zhuji, Shaoxing, China, <sup>3</sup> Center for Medical Statistics and Data Analysis, School of Public Health, Xuzhou Medical University, Xuzhou, China

## OPEN ACCESS

### Edited by:

Sheng Yang,  
Nanjing Medical University, China

### Reviewed by:

Can Yang,  
Hong Kong University of Science  
and Technology, Hong Kong  
Le Shu,  
University of California, Los Angeles,  
United States

### \*Correspondence:

Shuiping Huang  
sph@xzhmu.edu.cn  
Ping Zeng  
zpstat@xzhmu.edu.cn

<sup>†</sup>These authors share first authorship

### Specialty section:

This article was submitted to  
Computational Genomics,  
a section of the journal  
Frontiers in Genetics

**Received:** 07 August 2020

**Accepted:** 17 November 2020

**Published:** 03 December 2020

### Citation:

Chen H, Wang T, Yang J,  
Huang S and Zeng P (2020) Improved  
Detection of Potentially Pleiotropic  
Genes in Coronary Artery Disease  
and Chronic Kidney Disease Using  
GWAS Summary Statistics.  
*Front. Genet.* 11:592461.  
doi: 10.3389/fgene.2020.592461

The coexistence of coronary artery disease (CAD) and chronic kidney disease (CKD) implies overlapped genetic foundation. However, the common genetic determination between the two diseases remains largely unknown. Relying on summary statistics publicly available from large scale genome-wide association studies ( $n = 184,305$  for CAD and  $n = 567,460$  for CKD), we observed significant positive genetic correlation between CAD and CKD ( $r_g = 0.173$ ,  $p = 0.024$ ) via the linkage disequilibrium score regression. Next, we implemented gene-based association analysis for each disease through MAGMA (Multi-marker Analysis of GenoMic Annotation) and detected 763 and 827 genes associated with CAD or CKD ( $FDR < 0.05$ ). Among those 72 genes were shared between the two diseases. Furthermore, by integrating the overlapped genetic information between CAD and CKD, we implemented two pleiotropy-informed informatics approaches including cFDR (conditional false discovery rate) and GPA (Genetic analysis incorporating Pleiotropy and Annotation), and identified 169 and 504 shared genes ( $FDR < 0.05$ ), of which 121 genes were simultaneously discovered by cFDR and GPA. Importantly, we found 11 potentially new pleiotropic genes related to both CAD and CKD (i.e., *ARHGEF19*, *RSG1*, *NDST2*, *CAMK2G*, *VCL*, *LRP10*, *RBM23*, *USP10*, *WNT9B*, *GOSR2*, and *RPRML*). Five of the newly identified pleiotropic genes were further repeated via an additional dataset CAD available from UK Biobank. Our functional enrichment analysis showed that those pleiotropic genes were enriched in diverse relevant pathway processes including quaternary ammonium group transmembrane transporter, dopamine transport. Overall, this study identifies common genetic architectures overlapped between CAD and CKD and will help to advance understanding of the molecular mechanisms underlying the comorbidity of the two diseases.

**Keywords:** coronary artery disease, chronic kidney disease, pleiotropy-informed integrative analysis, gene-based association analysis, pleiotropic gene, genome-wide association study

## INTRODUCTION

Both coronary artery disease (CAD) and chronic kidney disease (CKD) are the leading causes of death and disability worldwide, representing serious global public health threats (Kessler et al., 2013; Inrig et al., 2014; Ene-Iordache et al., 2016; Levin et al., 2017; Musunuru and Kathiresan, 2019). In practice, it is often observed that CKD patients encounter an increased risk of CAD and CAD is in turn a major cause of death for CKD patients (Tonelli et al., 2012). Pathologically, the endothelial dysfunction is closely related to cardiovascular diseases and plays an important role in all stages of atherosclerosis (Ross, 1999). On the other hand, the role of CAD in CKD is also widely studied; for example, the endothelial dysfunction in the development of CKD was also well documented (Moody et al., 2012). As originally proposed by Lindner et al. (1974), CKD patients with an estimated glomerular filtration rate (eGFR)  $< 60$  ml/min per  $1.73$  m<sup>2</sup> have 2~16 times higher risk of major adverse cardiovascular events (MACE) compared to those with an eGFR  $> 60$  ml/min per  $1.73$  m<sup>2</sup> (Go et al., 2004). Moreover, for CKD patients not yet requiring renal replacement therapy, the probability of developing MACE is much higher than reaching end-stage renal disease (ESRD) and requiring renal replacement therapy (Foley et al., 2005).

All those empirical observations suggest that there exist a common susceptible mechanism underlying these two complex diseases. As part of efforts to understand their genetic foundation, in the past few years many large scale genome-wide association studies (GWASs) have been implemented for CAD (Nikpay et al., 2015) and CKD (Wuttke et al., 2019). It is found that a lot of genes and single nucleotide polymorphisms (SNPs) exhibit pleiotropic effects and are associated with both the two diseases (Solovieff et al., 2013; **Supplementary Table 1**). This genetic overlap partly contributes to the co-existence of CAD and CKD. The understanding of common genetic determinants has significant implication for identifying important biomarkers and developing novel therapeutic strategies for joint prediction, prevention, and intervention of CAD and CKD.

However, like many other diseases/traits (Manolio et al., 2009; Eichler et al., 2010; Gusev et al., 2013; Girirajan, 2017; Kim et al., 2017; Young, 2019), CAD- or CKD-associated SNPs identified by GWAS only explain a very small fraction of phenotypic variance of CKD (Wuttke et al., 2019) and CAD (Nikpay et al., 2015), implying that a large number of genetic variants with small to modest effect sizes (but still important) have yet been discovered and that more pleiotropic genes would be found if increasing sample sizes (Wang et al., 2005; Altshuler et al., 2008; Tam et al., 2019). However, the increase of sample sizes is generally not feasible since the recruiting and genotyping of additional participants are time consuming and expensive. Therefore, it is a promising way to leverage genetic computational methods that can efficiently analyze information contained in the existing pool of available GWAS summary statistics for identifying loci with pleiotropic effects.

To achieve this aim, many pleiotropy-informed approaches have been proposed (Andreassen et al., 2013; Chung et al., 2014; Zeng et al., 2018). Those previous studies were focused

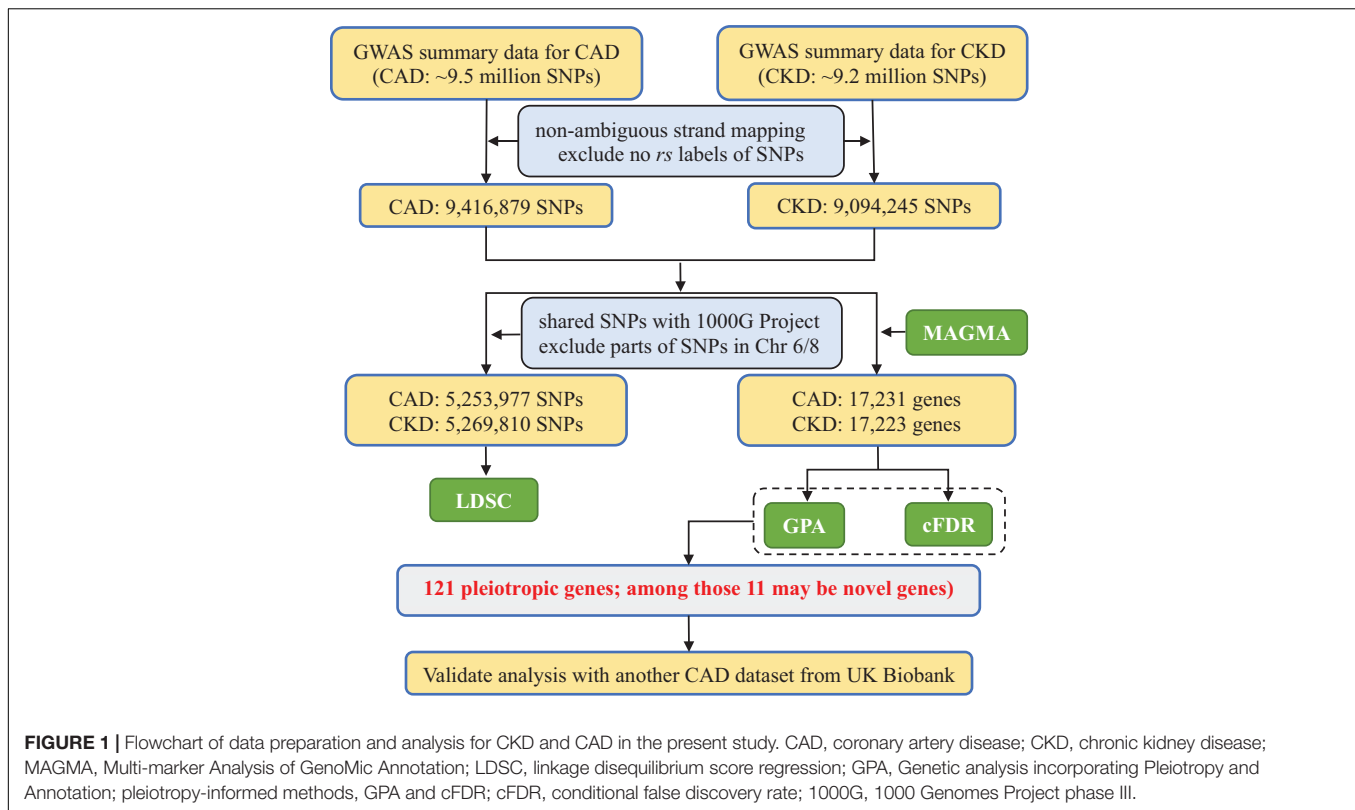
on individual SNP associations and fine-mapping was further needed to find causal genes once newly novel genetic variants were detected (Hormozdiari et al., 2014, 2015; Wen et al., 2015; Kichaev et al., 2016). In addition, those methods cannot effectively handle the correlation among genetic variants due to linkage disequilibrium (LD) (Zeng et al., 2018). As a result, pruning [e.g., using PLINK (Purcell et al., 2007)] has to be employed to keep less dependent SNPs in their analysis, which inevitably leads to the loss of useful information included in correlated SNPs. Compared with the traditional single SNP analysis which only considers only one SNP each time and often suffers from power reduction (Zeng et al., 2015), the gene-based association study is another popular supplementary analysis, which examines the joint significance of a group of SNPs and has the potential to aggregate weak association signals across multiple genetic variants and is thus more powerful (Zeng et al., 2014). Moreover, gene-based associations are easily to interpret because gene is a more meaningfully biological unit compared with individual genetic variant.

Given the potential pleiotropy between CAD and CKD that was widely implied in previous work (Go et al., 2004; Liu et al., 2012; Ene-Iordache et al., 2016), we hypothesize that shared genes identified by different pleiotropy-informed methods should have a higher probability to be candidate pleiotropic genes. To do so, in the present study we first evaluated the overall genetic correlation between CAD and CKD with summary statistics available from large scale GWASs through cross-trait LDSC (linkage disequilibrium score regression) (Bulik-Sullivan B. et al., 2015). We next conducted a gene-based association analysis using MAGMA (Multi-marker Analysis of GenoMic Annotation) (de Leeuw et al., 2015) to integrate association signals from SNP level into gene level. We thus obtained *P*-value for each protein coding gene. Depending on those gene-level *P*-values, we detected pleiotropic genes with two pleiotropy-informed association methods including cFDR (conditional false discovery rate) (Andreassen et al., 2013; Smeland et al., 2020) and GPA (Genetic analysis incorporating Pleiotropy and Annotation) (Chung et al., 2014). We also attempted to validate our results in another CAD dataset available from the UK Biobank (UKB) cohort. The framework of our data analysis is demonstrated in **Figure 1**.

## MATERIALS AND METHODS

### GWAS Summary Statistics

We obtained summary statistics (e.g., effect allele, effect size, and *P*-values) for CKD from the latest GWAS of the CKDGen consortium (Wuttke et al., 2019). In this study the creatinine value obtained with a Jaffé assay before 2009 was calibrated by multiplying by 0.95, and glomerular filtration rate (GFR) was estimated with the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation for adults (larger than 18 years age) while using the Schwartz formula for individuals less than 18 years old and was winsorized at 15–200 ml min<sup>-1</sup> per 1.73 m<sup>2</sup>. CKD was defined as an eGFR below 60 ml min<sup>-1</sup> per 1.73 m<sup>2</sup>. After stringent quality control, a total



of 567,460 (64,164 cases and 502,296 controls;  $N_{eff} = 227,584$ ) individuals of European ancestry and  $\sim 9.6$  million SNPs for CKD were left. We yielded summary statistics of CAD from the CARDIoGRAMplusC4D Consortium (Nikpay et al., 2015), which included 184,305 (60,801 cases and 123,504 controls;  $N_{eff} = 162,972$ ) individuals of European ancestry and  $\sim 9.4$  million SNPs after quality control.

We further validated our results using another summary statistic of CAD obtained from the UKB cohort<sup>1</sup>. The UKB-CAD dataset included 405,940 individuals of European ancestry (23,888 cases and 382,052 controls;  $N_{eff} = 89,929$ ) and 23,861,747 SNPs after quality control (i.e., INFO scores  $> 0.8$ , allele count at least 20 and minor allele count less than 20). The association in the UKB-CAD dataset was analyzed through the SAIGE method (Zhou et al., 2018), which implemented the logistic mixed model with a kinship matrix as random effects and age, sex, age  $\times$  sex, age<sup>2</sup>, age<sup>2</sup>  $\times$  sex as well as the first ten principal components as fixed-effects covariates.

## Estimated Overall Genetic Correlation With LDSC

We applied the cross-trait LDSC (Bulik-Sullivan B. et al., 2015) to assess the overall genetic correlation  $r_g$  between CKD and CAD using all available SNPs. The software of LDSC (version v1.0.1) was downloaded at <https://github.com/bulik/ldsc> and our analysis was conducted with default settings. Following

prior studies (Bulik-Sullivan B. et al., 2015), we performed stringent quality control procedures during the LDSC analysis: (1) excluded non-biallelic SNPs and those with strand-ambiguous alleles; (2) excluded duplicated SNPs and those having no rs labels; (3) excluded SNPs that were located within two genetic regions including major histocompatibility complex (chr6: 28,500,000–33,500,000) (Bulik-Sullivan B. et al., 2015) and chr8: 7,250,000–12,500,000 (Price et al., 2008); (4) kept SNPs that were included in the 1000 Genomes Project phase III; (5) removed SNPs whose allele did not match that in the 1000 Genomes Project phase III (The 1000 Genomes Project Consortium, 2015).

The LD scores  $\ell_j$  were computed using genotypes of 7,120,251 common SNPs (minor allele frequency  $> 0.01$  and the  $P$ -value of Hardy Weinberg equilibrium test  $> 1E-5$ ) with a 10 Mb window on 503 European individuals in the 1000 Genomes Project phase III (The 1000 Genomes Project Consortium, 2015); and then regressed on the product of  $Z$ -score statistics of the two diseases

$$E(z_{1j}z_{2j}) = \frac{\sqrt{N_1 N_2} \ell_j}{M} \times r_g + \frac{\rho N_s}{\sqrt{N_1 N_2}} \quad (1)$$

where  $N_1$  and  $N_2$  are the sample sizes for CAD and CKD, respectively;  $N_s$  is the number of individuals shared by the two GWASs, and  $\rho$  is the disease correlation among the  $N_s$  overlapping individuals. Theoretically, SNPs with high LD will have higher  $\chi^2$  statistics on average than those with low LD provided that the disease has a polygenic genetic foundation (Bulik-Sullivan B. K. et al., 2015). In terms of LDSC shown in

<sup>1</sup><https://pan.ukbb.broadinstitute.org/>

(1), the regression slope provides an unbiased estimate for genetic correlation  $r_g$  and is in general not influenced by sample overlap (Bulik-Sullivan B. et al., 2015).

## Summary Statistics-Based Gene-Level Association With MAGMA

Many gene-based association approaches with only summary statistics have been developed recently; among those MAGMA is a fast and flexible method and widely employed (de Leeuw et al., 2015). During the implementation of MAGMA, we defined the set of SNPs that were located within a given gene in terms of the annotation file provided in VAGIS (Liu et al., 2010). For numerical stability, we only focused on protein coding genes with at least ten SNPs (note that, this threshold was to some extent chosen arbitrarily). The genotypes of 503 European individuals in the 1000 Genomes Project phase III (The 1000 Genomes Project Consortium, 2015) were exploited as reference panel for calculating the LD matrix to incorporate the correlation structure among SNPs. After the implementation of MAGMA, the  $P$ -value for each gene can be available in the CAD or CKD GWAS. Depending on those  $P$ -values we attempted to discover significant genes that were related to CAD or CKD as well as potentially pleiotropic genes that were associated with both the two types of disease. To detect newly novel association signals, we ruled out identified genes located within 1 Mb on each side of previously reported CAD- or CKD associated genes or SNPs from the GWAS Catalog<sup>2</sup> as done similarly in other studies (Bis et al., 2020). Of note, doing this was a conservative strategy and might miss potentially important association signals although false discoveries were well controlled.

## Pleiotropy-Informed Association Methods With Summary Statistics

To further leverage the pleiotropic information shared between CAD and CKD to identify gene association signals more efficiently, we employed two novel statistical genetic methods in the following. First, we utilized the cFDR method (Andreassen et al., 2013) which extended the unconditional FDR (Benjamini et al., 2001) from an empirical Bayes perspective. The cFDR measures the probability of the association of the principal disease conditioned on the strength of association with the conditional disease (Andreassen et al., 2013)

$$\text{cFDR}(p_i || P_i \leq p_i, P_j \leq p_j) \quad (2)$$

where  $p_i$  and  $p_j$  are the observed  $P$ -values of a particular gene of the principal and conditional diseases, respectively;  $H_0^{(i)}$  denotes the null hypothesis that there does not exist association between the gene and the principal disease.

Besides cFDR, we also carried out the GPA analysis (Chung et al., 2014), which was constructed as

$$\pi_{00} = \text{Prob}(Z_{j00} = 1) : (P_{j1} | Z_{j00} = 1) \sim$$

$$U[0, 1], (P_{j2} | Z_{j00} = 1) \sim U[0, 1]$$

$$\pi_{10} = \text{Prob}(Z_{j10} = 1) : (P_{j1} | Z_{j10} = 1) \sim$$

$$\text{Beta}(\alpha_1, 1), (P_{j2} | Z_{j10} = 1) \sim U[0, 1]$$

$$\pi_{01} = \text{Prob}(Z_{j01} = 1) : (P_{j1} | Z_{j01} = 1) \sim$$

$$U[0, 1], (P_{j2} | Z_{j01} = 1) \sim \text{Beta}(\alpha_2, 1)$$

$$\pi_{11} = \text{Prob}(Z_{j11} = 1) : (P_{j1} | Z_{j11} = 1) \sim$$

$$\text{Beta}(\alpha_1, 1), (P_{j2} | Z_{j11} = 1) \sim \text{Beta}(\alpha_2, 1) \quad (3)$$

where the latent variables  $Z_j = (Z_{j00}, Z_{j10}, Z_{j01}, Z_{j11})$  indicates the association between the  $j$ -th gene and the two diseases:  $Z_{j00} = 1$  denotes the  $j$ -th gene is associated with neither of them (with probability  $\pi_{00}$ ),  $Z_{j10} = 1$  denotes the  $j$ -th gene is only associated with the first one (with probability  $\pi_{10}$ ),  $Z_{j01} = 1$  denotes the  $j$ -th gene is only associated with the second one (with probability  $\pi_{01}$ ), and  $Z_{j11} = 1$  denotes the  $j$ -th gene is associated with both the diseases (with probability  $\pi_{11}$ ), indicating the extent of common biological pathways to which the two diseases may share (Chung et al., 2014). In addition,  $\alpha_1$  and  $\alpha_2$  ( $0 < \alpha_k < 1, k = 1, 2$ ) are unknown shape parameters of the Beta distribution.

## Functional Analysis

To explore functional features of newly discovered pleiotropic genes, we performed functional enrichment analysis [e.g., Gene Ontology (GO) and KEGG pathway analysis] with DAVID 6.8<sup>3</sup> (Huang da et al., 2009). Enrichment analysis allows us to validate our findings by determining functional annotations for those genes with pleiotropic effects. We also conducted the protein-protein interaction analysis to detect interaction and association in terms of the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING 11.0 at <https://string-db.org/>) database (Szklarczyk et al., 2019). We implemented the signaling pathways of these significant genes by Cytoscape software and visualized them by CluePedia (Bindea et al., 2009).

## RESULTS

### Estimated Overall Genetic Correlation Between CAD and CKD

After quality control, a total of 5,253,977 and 5,269,810 genetic variants are reserved for CAD or CKD, respectively. The genome-wide SNP-based heritability is estimated to be 4.69% (SE = 0.35%) for CAD and 0.53% (SE = 0.12%) for CKD with LDSC. The genomic inflation factor (i.e., the ratio of the observed median  $\chi^2$  statistic to the expected median) is 1.015 for CAD and 1.143 for CKD, which, together the LDSC intercept [i.e., 0.903 (SE = 0.005) for CAD and 1.134 (SE = 0.007) for CKD], suggests that the weak inflation of the  $\chi^2$  statistic of CKD is primarily due to polygenicity rather than population stratification or cryptic

<sup>2</sup>[www.ebi.ac.uk](http://www.ebi.ac.uk)

<sup>3</sup><https://david.ncifcrf.gov/>

relatedness. In terms of those results, the adjustment of genomic control is also not necessary.

Next, based on all overlapped genetic variants (i.e., 5,117,020 SNPs), using LDSC we observe there exists a positive genetic correlation between the two types of diseases [ $\hat{r}_g = 0.173$ , 95% confidence interval (CI) 0.023 ~ 0.332,  $P = 0.024$ ], providing empirical evidence that the two diseases share common genetic components. We further quantify genetic correlation between CAD and CKD separately in six functional categories (Gusev et al., 2014), including coding, UTR (untranslated region), promoter, DHS (DNaseI hypersensitivity sites), intronic and intergenic. It is found that all the estimates of  $r_g$  in those categories are positive, again supporting the statement that CAD and CKD have overlapped genetic foundation. In particular, there exists a significantly positive genetic correlation in the regions of DHS ( $\hat{r}_g = 0.197$ , 95% CI 0.074 ~ 0.319,  $P = 1.60E-3$ ) and intergenic ( $\hat{r}_g = 0.264$ , 95% CI 0.153 ~ 0.375,  $P = 3.03E-6$ ) (Supplementary Table S2 and Figure 2).

Overall, through genetic correlation analysis we reveal that CAD and CKD are genetically similar and share moderate overlap in genetic etiology, especially at some certain regions. Therefore, it is worthy of additional investigation into shared genetic mechanisms through pleiotropy-informed statistical tools.

## Associated Genes Identified With MAGMA, cFDR, and GPA

In our gene-based association analysis, we assign a set of genetic variants to predefined genes and obtain a total of 17,231 and 17,223 protein coding genes for CAD or CKD, respectively. Using MAGMA, we identify 763 CAD-associated genes and 827 CKD-associated genes (FDR < 0.05) (Supplementary Tables 3, 4 and Supplementary Figure 1). Importantly, 25.8% (=197/763) CAD-associated genes (e.g., *ACER2*, *ACSS2*, *ARHGEF19*, and *BBS10*) and 60.7% (=503/827) CKD-associated genes (e.g., *BAG6*, *BAK1*, *BTNL2*, and *C4BPB*) are likely novel genes because those genes are not nearby (within 1 Mb upstream and downstream) any previous GWAS index SNPs or associated genes in terms of the GWAS catalog (McMahon et al., 2019).

In our cFDR analysis the Q-Q plot of CAD conditional on the nominal  $P$ -value of CKD illustrates the existence of enrichment at different significance thresholds of CKD (Supplementary Figure 2A). The presence of leftward shift suggests that the proportion of true associations for a given CKD  $P$ -value would increase when the analysis is limited to include more significant SNPs. On the other hand, in terms of the Q-Q plot of CKD conditional on the nominal  $P$ -value of CAD (Supplementary Figure 2B), we observe a more pronounced separation in different curves, implying that there exists a stronger enrichment for CKD given CAD than that for CAD given CKD. We further formally analyze the two diseases jointly using cFDR and show the results in Supplementary Tables 5, 6 and Supplementary Figure 3. Briefly, with cFDR we identify 875 CAD-associated genes and 1,062 CKD-associated genes (cFDR < 0.05). Among those genes, 243 CAD-associated and 639 CKD-associated genes are possibly novel (Supplementary Tables 5, 6). More interesting,

all CAD-associated genes identified by MAGMA are replicated and 111 additional genes are discovered (Supplementary Figure 4); and all CKD-associated genes identified by MAGMA are also verified and 234 more genes are newly discovered (Supplementary Figure 5).

We next employ GPA to implement another integrative analysis for the two diseases. In terms of the GPA result we discover 504 and 1395 significant genes that are related to CAD or CKD (Supplementary Tables 7, S8 and Supplementary Figure 6). Among those, 17.3% (=87/504) novel CAD-associated genes (e.g., *ACVR2A*, *AP3M1*, *ARHGEF19*, and *BACH1*) and 61.2% (=854/1395) CKD-associated genes (e.g., *ABCA4*, *ABCC2*, *ABCF3*, and *ACOX1*) may be newly novel genes because they are not nearby (within 1 Mb upstream and downstream) any previous GWAS index SNPs or associated genes in terms of the GWAS catalog (McMahon et al., 2019). Furthermore, we find 504 CAD-associated and 770 CKD-associated genes that are identified simultaneously by GPA and MAGMA (Supplementary Figures 7, 8).

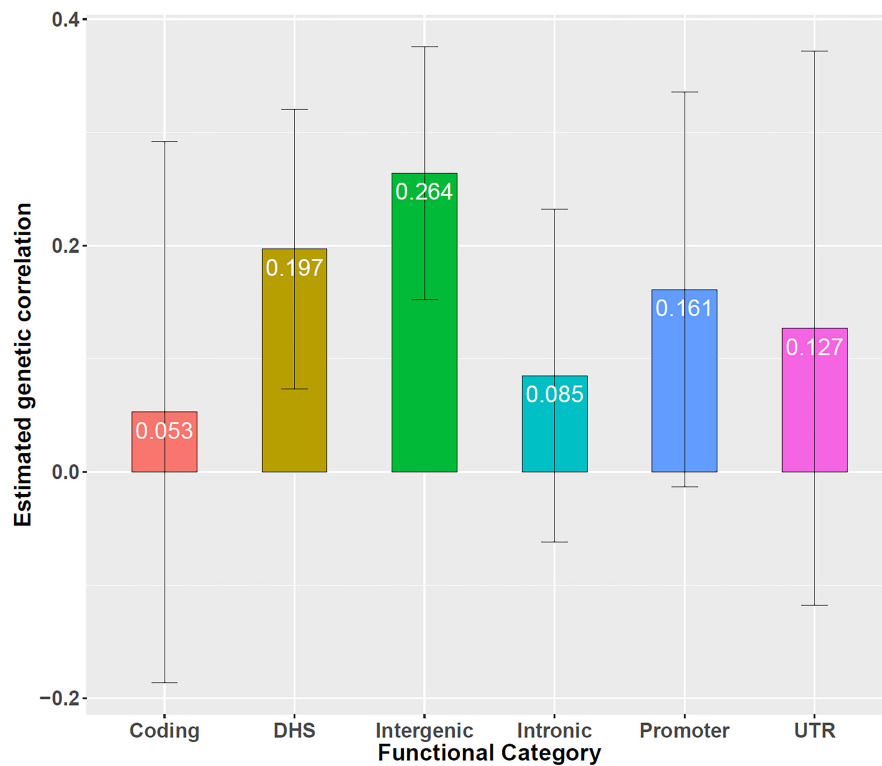
## Identified Pleiotropic Gene With Both cFDR and GPA

According to the result of MAGMA, 72 genes are related to both CAD and CKD (Supplementary Table 9 and Figure 3A). Based on the two integrative analyses, 169 genes are shared between CAD and CKD when using cFDR (Supplementary Table 10 and Figure 3B) and 504 genes are shared between CAD and CKD when using GPA (Supplementary Table 11 and Figure 3C). In addition, through GPA we observe that a substantial fraction of genes that are simultaneously related to CAD and CKD, with  $\pi_{11}$  estimated to be 8.2% (SE = 0.1%), offering additional statistical evidence supporting the existence of pleiotropy between CAD and CKD [the statistic of the likelihood ratio test is 225.6 and  $P = 5.35E-51$  (Chung et al., 2014)].

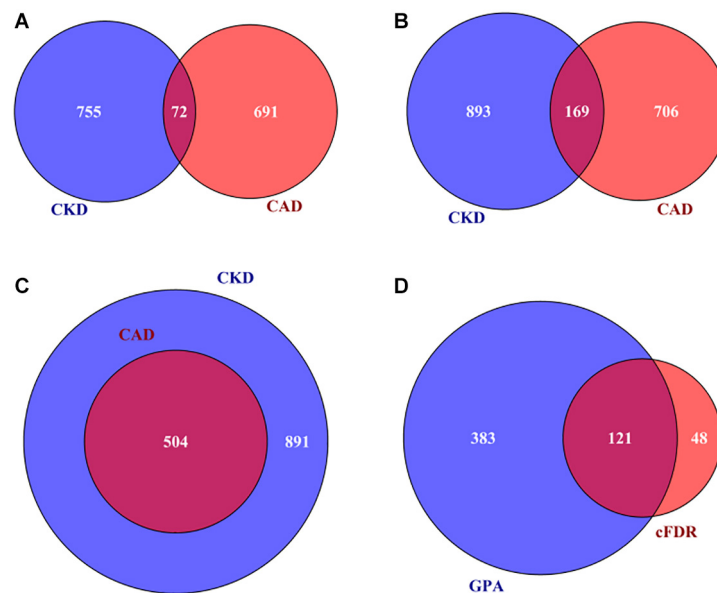
Due to the difference of power in identifying pleiotropic genes via cFDR or GPA, we expect that a gene would be more likely to have pleiotropic effect if it is discovered by cFDR and GPA simultaneously. Relying on this principle we define a set of 121 genes that are associated with CAD and CKD and are jointly detected by cFDR and GPA to be pleiotropic genes (Supplementary Table 12 and Figure 3D), among which five (i.e., *IGF2R*, *LPA*, *BCAS3*, *SLC22A2*, and *ATXN2*) were identified in previous studies (Supplementary Table 1). Furthermore, after ruling out genes located within 1 Mb on each side of previously reported genes or SNPs, we ultimately discover 11 newly novel pleiotropic genes associated with both CAD and CKD (i.e., *RHGEF19*, *RSG1*, *NDST2*, *CAMK2G*, *VCL*, *LRP10*, *RBM23*, *USP10*, *WNT9B*, *GOSR2*, and *RPRML*) (Table 1 and Supplementary Figures 9–13).

## Validation the Results in a Latest GWAS From the UK Biobank

We further validate the main results using the UKB-CAD summary statistics and show the results in Supplementary Tables 13–16. The genome-wide SNP-based heritability is estimated to be 2.42% (SE = 0.20%) for UKB-CAD with LDSC.



**FIGURE 2 |** Genetic correlation between CAD and CKD in six functional categories, including coding, UTR, promoter, DHS, intronic, and intergenic. Error bars show  $1.96 \times SE$ . Besides DHS and intergenic, the genetic correlation is  $\hat{r}_g = 0.053$  ( $SE = 0.122$ ,  $P = 6.64E-1$ ) for coding,  $\hat{r}_g = 0.127$  ( $SE = 0.125$ ,  $P = 3.10E-1$ ) for UTR,  $\hat{r}_g = 0.161$  ( $SE = 0.089$ ,  $P = 7.00E-2$ ) for promoter,  $\hat{r}_g = 0.089$  ( $SE = 0.075$ ,  $P = 2.55E-1$ ) for intronic.



**FIGURE 3 | (A)** A total of 72 associated genes shared by CAD and CKD using MAGMA; **(B)** 169 associated genes shared by CAD and CKD using cFDR; **(C)** a total of 504 genes shared by CAD and CKD using GPA; **(D)** a total of 121 pleiotropic genes of CAD and CKD simultaneously discovered by cFDR and GPA. CAD, coronary artery disease; CKD, chronic kidney disease; MAGMA, Multi-marker Analysis of GenoMic Annotation; GPA, Genetic analysis incorporating Pleiotropy and Annotation; cFDR, conditional false discovery rate.

**TABLE 1** | Pleiotropic genes associated with CAD and CKD identified by cFDR and GPA jointly.

Gene	CHR	Position	cFDR		GPA	
			CAD	CKD	CAD	CKD
<i>ARHGEF19</i>	1	16,424,598–16,639,104	2.55E-03	4.16E-02	1.19E-02	1.76E-03
<i>RSG1</i>	1	16,458,181–16,663,659	2.83E-03	3.38E-02	1.28E-02	1.47E-03
<i>NDST2</i>	10	75,461,668–75,671,589	1.09E-02	4.13E-03	4.69E-02	8.63E-04
<i>CAMK2G</i>	10	75,472,258–75,734,349	5.06E-03	3.02E-03	2.51E-02	4.13E-04
<i>VCL</i>	10	75,657,871–75,979,914	6.07E-03	1.52E-02	2.75E-02	1.58E-03
<i>LRP10</i>	14	23,240,959–23,447,291	8.22E-03	2.21E-02	3.35E-02	2.42E-03
<i>RBM23</i>	14	23,269,853–23,488,396	7.62E-03	2.80E-02	3.13E-02	2.78E-03
<i>USP10</i>	16	84,633,554–84,913,527	6.48E-03	2.63E-04	4.08E-02	1.01E-04
<i>WNT9B</i>	17	44,828,967–45,054,437	4.20E-04	3.37E-02	1.84E-03	2.47E-04
<i>GOSR2</i>	17	44,900,485–45,118,733	8.13E-04	3.22E-02	4.08E-03	5.39E-04
<i>RPRML</i>	17	44,955,521–45,156,614	1.90E-03	2.16E-02	8.30E-03	6.50E-04

We also do not observe a substantial inflation in the UKB-CAD summary statistics [the estimated  $\lambda = 1.178$  with the intercept = 1.057 (SE = 0.005)].

According to the result of MAGMA, 184 genes are related to both UKB-CAD and CKD (Supplementary Table 13 and Supplementary Figure 14). Based on the two pleiotropy-informed integrative analyses, 373 genes are shared between UKB-CAD and CKD using cFDR (Supplementary Table 14 and Supplementary Figure 15) and 371 genes are shared between UKB-CAD and CKD using GPA (Supplementary Table 15 and Supplementary Figure 16). All the 11 pleiotropic genes described above are also analyzed here and five (i.e., *RSG1*, *LRP10*, *RBM23*, *WNT9B*, and *GOSR2*) are replicated (Supplementary Table 16).

## Functional Analyses for Pleiotropic Genes

We now undertake functional analyses for the 121 pleiotropic genes. Among these, most are located within chr 17 (20.7% = 25/121), followed by chr 1 (15.7% = 19/121) and chr 11 (12.4% = 15/121) (Supplementary Figure 17). In terms of the DAVID analysis, these genes are enriched in 34 GO terms (Supplementary Table 17). The top five candidate pathways include “dopamine transmembrane transporter activity” ( $P = 2.28E-04$ ), “quaternary ammonium group transport” ( $P = 3.54E-04$ ), “quaternary ammonium group transmembrane transporter activity” ( $P = 3.79E-04$ ), “dopamine transport” ( $P = 7.38E-04$ ), and “organic cation transmembrane transporter activity” ( $P = 7.89E-04$ ). These pathways offer part of evidence supporting common genetic foundations between CAD and CKD. For instance, it has been shown that CKD patients had higher levels for some quaternary ammonium salts (e.g., choline) (Rennick et al., 1976), which were also risk factors for CAD (Guo et al., 2020). In our PPI analysis (Supplementary Figure 18), strong interactions are found among pleiotropic genes, such as *NDST2*, *CAMK2G*, *RASGRF1*, *IGF2R*, *SORT1*, and *TRIB1*. These genes were reported to be associated with organic cation transmembrane transporter, such as organic anion transporters *oat1* and *oat3*, and organic cation transporters *oct1* and *oct2*,

which was also altered with chronic kidney failure in rats (Komazawa et al., 2013).

## DISCUSSION

It has been widely observed that CAD and CKD share common pathological and clinical feature (Go et al., 2004; Liu et al., 2012; Tonelli et al., 2012; Ene-Iordache et al., 2016). However, the underlying genetic overlap between the two diseases remains unclear and a large proportion of genes related to CAD and CKD are yet discovered (Manolio et al., 2009). Large-scale GWAs undertaken for CAD and CKD offer an unprecedented opportunity to answer this question. In the present study a positive genetic correlation was found between CAD and CKD, implying genetic variants that were associated with the risk of CKD would be also related to the risk of CAD. This finding also partly explained the observed comorbidity of the two diseases (Go et al., 2004; Ene-Iordache et al., 2016).

Using existing well-established statistical approaches, we ultimately identified 11 novel pleiotropic genes shared by CAD and CKD, including *ARHGEF19*, *RSG1*, *NDST2*, *CAMK2G*, *VCL*, *LRP10*, *RBM23*, *USP10*, *WNT9B*, *GOSR2*, and *RPRML*, some of which were previously reported to play important roles in the pathogenesis of CAD or CKD (Agosti, 2002; Sivapalaratnam et al., 2012; Zanders, 2015). Furthermore, we also validated our main finding in an independent UKB-CAD dataset and replicated five genes.

Specifically, prior studies showed that *ARHGEF19* (Klarin et al., 2018) and *LRP10* (Sugiyama et al., 2000) were associated with total cholesterol and low-density lipoprotein (LDL) cholesterol, which were in turn related to CAD (Nissen et al., 2005) and CKD (Baigent et al., 2011). *RSG1* is involved in targeted membrane trafficking, and further involved in cilium biogenesis by regulating the transportation of cargo proteins to the basal body and apical tips of cilia with its protein (Agbu et al., 2018). Mice and humans with abnormal primary cilia can exhibit defects in cardiac morphogenesis, and also can cause kidney disease (Agbu et al., 2018).

*NDST2* encodes a member of the N-deacetylase/N-sulfotransferase subfamily, which has dual functions (N-deacetylation and N-sulfation) in processing heparin polymers (Humphries et al., 1998). Inactivation of *NDST2* may impact the atherosclerosis by altering the structure of monocytes/macrophages heparan sulfate (HS) (Gordts et al., 2014), while also alter the glomerular HS to impact the primary kidney diseases (Goode et al., 1995). *CAMK2G* belongs to the  $Ca^{2+}$ /calmodulin-dependent protein kinase subfamily (Moyers et al., 1997). Vascular calcification correlates with the vessel stiffening and hypertension, and further increases the risk of atherosclerosis and myocardial infarction. It also exhibits a hugely elevated risk of cardiovascular mortality in CKD patients (Shanahan Catherine et al., 2011).

Vinculin (*VCL*) is a membrane-cytoskeletal protein, which associated with the linkage of integrin adhesion molecules to the actin cytoskeleton (Burrige and Feramisco, 1980), and the cell-cell and cell-matrix junctions, where it is thought to function in anchoring F-actin to the membrane (Geiger, 1979). Endothelial dysfunction caused by F-actin cytoskeleton disorder is a well-recognized instigator of cardiovascular diseases and CKD (Ding et al., 2016). *USP10* encodes a member of the ubiquitin-specific protease family of cysteine proteases (Wang et al., 2015; Lim et al., 2019). Inactivation of *USP10* can diminish Notch-induced target gene expression in endothelial cells. Importantly, tight quantitative and temporal control of Notch activity is essential for vascular development (Wang et al., 2015; Lim et al., 2019).

*WNT9B*, encodes the secreted signaling proteins (Garriock et al., 2007), is significantly associated with systolic blood pressure (Hoffmann et al., 2017), which is further related to the risk of CAD (Turner et al., 1998) and CKD (Jafar et al., 2003). *GOSR2* encodes a trafficking membrane protein which transports proteins among the medial- and trans-Golgi compartments (Bui et al., 1999). Due to its chromosomal location and trafficking function, *GOSR2* may be involved in familial essential hypertension (Boissé Lomax et al., 2013), and also was reported to be relevant to systolic blood pressure (Ehret et al., 2011) and CAD (van der Harst and Verweij, 2018).

The major strength of our work is that multiple pleiotropy-informed methods were implemented to detect pleiotropic genes by combining existing GWAS summary results without requiring individual-level datasets. Unlike previous studies (Andreassen et al., 2013; Chung et al., 2014; Zeng et al., 2018), we perform MAGMA methods to enrich a group of SNPs which may be likely associated with CAD or CKD but cannot reach genome-wide significance because of modest effects if using single marker analysis. Moreover, to minimize possible false discovery, we only reported pleiotropic genes that were simultaneously discovered by GPA and cFDR and thus were more likely to be related to both CAD and CKD. Therefore, our findings are robust.

Nevertheless, there are some limitations needed to state. First, we cannot replicate all these genes via *in vivo* and *in vitro* experiments. Second, the individuals involved in our study are of European ancestry, it is not clear whether the finding can be generalized to other populations because of ethnic diversity in genetics. Third, although empirical evidence shown above indicates that the newly identified pleiotropic genes may underlie

certain aspects of the pathogenesis of CAD and CKD in a direct or indirect way, the causally biological mechanisms of those genes are still largely unclear; therefore, further studies are needed to completely delineate their functions on CAD and CKD.

## CONCLUSION

This study identifies common genetic architectures overlapped between CAD and CKD and will help to advance understanding of the molecular mechanisms underlying the comorbidity of the two diseases.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

## AUTHOR CONTRIBUTIONS

PZ and SH conceived the idea for the study. PZ, TW, and HC obtained the data. PZ and HC performed the data analyses and wrote the manuscript with the participation of all authors. PZ, JY, TW, and HC interpreted the results of the data analyses. All authors contributed to the article and approved the submitted version.

## FUNDING

The research of PZ was supported in part by the Youth Foundation of Humanity and Social Science funded by Ministry of Education of China (18YJC910002), the Natural Science Foundation of Jiangsu Province of China (BK20181472), the China Postdoctoral Science Foundation (2018M630607 and 2019T120465), the QingLan Research Project of Jiangsu Province for Outstanding Young Teachers, the Six-Talent Peaks Project in Jiangsu Province of China (WSN-087), the Training Project for Youth Teams of Science and Technology Innovation at Xuzhou Medical University (TD202008), the Postdoctoral Science Foundation of Xuzhou Medical University, the National Natural Science Foundation of China (81402765), Postgraduate Research & Practice Innovation Program of Jiangsu (KYCX20\_2501), and the Statistical Science Research Project from National Bureau of Statistics of China (2014LY112). The research of SH was supported in part by the Social Development Project of Xuzhou City (KC19017). The research of TW was supported in part by the Social Development Project of Xuzhou City (KC20062).

## ACKNOWLEDGMENTS

We thank the CKDGen Consortium and the CARDIoGRAMplusC4D Consortium who made the summary data of CKD and CAD publicly available. The genetic and phenotypic UK Biobank data were



available at <https://pan.ukbb.broadinstitute.org/>. We are also grateful to all the investigators and participants contributed to those studies. The data analyses in the present study were supported by the high-performance computing cluster at Xuzhou Medical University.

## REFERENCES

- Agbu, S. O., Liang, Y., Liu, A., and Anderson, K. V. (2018). The small GTPase RSG1 controls a final step in primary cilia initiation. *J. Cell Biol.* 217, 413–427. doi: 10.1083/jcb.201604048
- Agosti, J. (2002). *Biotherapeutic Approaches to Asthma*. Boca Raton, FL: CRC Press.
- Altshuler, D., Daly, M., and Lander, E. (2008). Genetic mapping in human disease. *Science* 322, 881–888. doi: 10.1126/science.1156409
- Andreassen, O. A., Djurovic, S., Thompson, W. K., Schork, A. J., Kendler, K. S., O'Donovan, M. C., et al. (2013). Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *Am. J. Hum. Genet.* 92, 197–209. doi: 10.1016/j.ajhg.2013.01.001
- Baigent, C., Landray, M. J., Reith, C., Emberson, J., Wheeler, D. C., Tomson, C., et al. (2011). The effects of lowering LDL cholesterol with simvastatin plus ezetimibe in patients with chronic kidney disease (Study of Heart and Renal Protection): a randomised placebo-controlled trial. *Lancet* 377, 2181–2192. doi: 10.1016/S0140-6736(11)60739-3
- Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* 125, 279–284. doi: 10.1016/s0166-4328(01)00297-2
- Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., et al. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25, 1091–1093. doi: 10.1093/bioinformatics/btp101
- Bis, J. C., Jian, X., Kunkle, B. W., Chen, Y., Hamilton-Nelson, K. L., Bush, W. S., et al. (2020). Whole exome sequencing study identifies novel rare and common Alzheimer's-Associated variants involved in immune response and transcriptional regulation. *Mol. Psychiatry* 25, 1859–1875. doi: 10.1038/s41380-018-0112-7
- Boissé Lomax, L., Bayly, M. A., Hjalgrim, H., Møller, R. S., Vlaar, A. M., Aaberg, K. M., et al. (2013). 'North Sea' progressive myoclonus epilepsy: phenotype of subjects with GOSR2 mutation. *Brain* 136(Pt 4), 1146–1154. doi: 10.1093/brain/awt021
- Bui, T. D., Levy, E. R., Subramaniam, V. N., Lowe, S. L., and Hong, W. (1999). cDNA characterization and chromosomal mapping of human golgi SNARE GS27 and GS28 to chromosome 17. *Genomics* 57, 285–288. doi: 10.1006/geno.1998.5649
- Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P. R., et al. (2015). An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* 47, 1236–1241. doi: 10.1038/ng.3406
- Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., et al. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* 47:291. doi: 10.1038/ng.3211
- Burridge, K., and Feramisco, J. R. (1980). Microinjection and localization of a 130K protein in living fibroblasts: a relationship to actin and fibronectin. *Cell* 19, 587–595. doi: 10.1016/s0092-8674(80)80035-3
- Chung, D., Yang, C., Li, C., Gelernter, J., and Zhao, H. (2014). GPA: a statistical approach to prioritizing GWAS results by integrating pleiotropy and annotation. *PLoS Genet.* 10:e1004787. doi: 10.1371/journal.pgen.1004787
- de Leeuw, C. A., Mooij, J. M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* 11:e1004219. doi: 10.1371/journal.pcbi.1004219
- Ding, N., Liu, B., Song, J., Bao, S., Zhen, J., Lv, Z., et al. (2016). Leptin promotes endothelial dysfunction in chronic kidney disease through AKT/GSK3 $\beta$  and  $\beta$ -catenin signals. *Biochem. Biophys. Res. Commun.* 480, 544–551. doi: 10.1016/j.bbrc.2016.10.079
- Ehret, G. B., Munroe, P. B., Rice, K. M., Bochud, M., Johnson, A. D., Chasman, D. I., et al. (2011). Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 478, 103–109. doi: 10.1038/nature10405

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.592461/full#supplementary-material>

- Eichler, E. E., Flint, J., Gibson, G., Kong, A., Leal, S. M., Moore, J. H., et al. (2010). Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genet.* 11, 446–450. doi: 10.1038/nrg2809
- Ene-Iordache, B., Perico, N., Bikbov, B., Carminati, S., Remuzzi, A., Perna, A., et al. (2016). Chronic kidney disease and cardiovascular risk in six regions of the world (ISN-KDDC): a cross-sectional study. *Lancet Glob. Health* 4, e307–e319. doi: 10.1016/s2214-109x(16)00071-71
- Foley, R. N., Murray, A. M., Li, S., Herzog, C. A., McBean, A. M., Eggers, P. W., et al. (2005). Chronic kidney disease and the risk for cardiovascular disease, renal replacement, and death in the United States Medicare population, 1998 to 1999. *J. Am. Soc. Nephrol.* 16, 489–495. doi: 10.1681/asn.2004030203
- Garriock, R. J., Warkman, A. S., Meadows, S. M., D'Agostino, S., and Krieg, P. A. (2007). Census of vertebrate Wnt genes: isolation and developmental expression of *Xenopus* Wnt2, Wnt3, Wnt9a, Wnt9b, Wnt10a, and Wnt16. *Dev. Dyn.* 236, 1249–1258. doi: 10.1002/dvdy.21156
- Geiger, B. (1979). A 130K protein from chicken gizzard: its localization at the termini of microfilament bundles in cultured chicken cells. *Cell* 18, 193–205. doi: 10.1016/0092-8674(79)90368-4
- Girirajan, S. (2017). Missing heritability and where to find it. *Genome Biol.* 18:89. doi: 10.1186/s13059-017-1227-x
- Go, A. S., Chertow, G. M., Fan, D., McCulloch, C. E., and Hsu, C. Y. (2004). Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N. Engl. J. Med.* 351, 1296–1305. doi: 10.1056/NEJMoa041031
- Goode, N. P., Shires, M., Crellin, D. M., Aparicio, S. R., and Davison, A. M. (1995). Alterations of glomerular basement membrane charge and structure in diabetic nephropathy. *Diabetologia* 38, 1455–1465. doi: 10.1007/bf00400607
- Gordts, P. L., Foley, S. M., Erin, M., Lawrence, R., Sinha, R., Lameda-Diaz, C., et al. (2014). Reducing macrophage proteoglycan sulfation increases atherosclerosis and obesity through enhanced Type I interferon signaling. *Cell Metab.* 20, 813–826. doi: 10.1016/j.cmet.2014.09.016
- Guo, F., Zhou, J., Li, Z., Yu, Z., and Ouyang, D. (2020). The association between trimethylamine N-Oxide and its predecessors choline, L-carnitine, and betaine with coronary artery disease and artery stenosis. *Cardiol. Res. Pract.* 2020:5854919. doi: 10.1155/2020/5854919
- Gusev, A., Bhatia, G., Zaitlen, N., Vilhjalmsson, B. J., Diogo, D., Stahl, E. A., et al. (2013). Quantifying missing heritability at known GWAS loci. *PLoS Genet.* 9:e1003993. doi: 10.1371/journal.pone.1003993
- Gusev, A., Lee, S. H., Trynka, G., Finucane, H., Vilhjalmsson, B. J., Xu, H., et al. (2014). Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. *Am. J. Hum. Genet.* 95, 535–552. doi: 10.1016/j.ajhg.2014.10.004
- Hoffmann, T. J., Ehret, G. B., Nandakumar, P., Ranatunga, D., Schaefer, C., Kwok, P.-Y., et al. (2017). Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat. Genet.* 49, 54–64. doi: 10.1038/ng.3715
- Hormozdiari, F., Kichaev, G., Yang, W.-Y., Pasaniuc, B., and Eskin, E. (2015). Identification of causal genes for complex traits. *Bioinformatics* 31, 206–213. doi: 10.1093/bioinformatics/btv240
- Hormozdiari, F., Kostem, E., Kang, E. Y., Pasaniuc, B., and Eskin, E. (2014). Identifying causal variants at loci with multiple signals of association. *Genetics* 198, 497–U484. doi: 10.1534/genetics.114.167908
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211
- Humphries, D. E., Lanciotti, J., and Karlinsky, J. B. (1998). cDNA cloning, genomic organization and chromosomal localization of human heparan glucosaminyl N-deacetylase/N-sulphotransferase-2. *Biochem. J.* 332(Pt 2), 303–307. doi: 10.1042/bj3320303
- Inrig, J. K., Califf, R. M., Tasneem, A., Vegunta, R. K., Molina, C., Stanifer, J. W., et al. (2014). The landscape of clinical trials in nephrology: a systematic review

- of Clinicaltrials.gov. *Am. J. Kidney Dis.* 63, 771–780. doi: 10.1053/j.ajkd.2013.10.043
- Jafar, T. H., Stark, P. C., Schmid, C. H., Landa, M., Maschio, G., de Jong, P. E., et al. (2003). Progression of chronic kidney disease: the role of blood pressure control, Proteinuria, and Angiotensin-converting enzyme inhibition: a patient-level meta-analysis. *Ann. Intern. Med.* 139, 244–252. doi: 10.7326/0003-4819-139-4-200308190-00006
- Kessler, T., Erdmann, J., and Schunkert, H. (2013). Genetics of coronary artery disease and myocardial infarction–2013. *Curr. Cardiol. Rep.* 15:368. doi: 10.1007/s11886-013-0368-0
- Kichaev, G., Roytman, M., Johnson, R., Eskin, E., Lindstrom, S., Kraft, P., et al. (2016). Improved methods for multi-trait fine mapping of pleiotropic risk loci. *Bioinformatics* 2020:btw615.
- Kim, H., Grueneberg, A., Vazquez, A. I., Hsu, S., and de los Campos, G. (2017). Will big data close the missing heritability gap? *Genetics* 207, 1135–1145. doi: 10.1534/genetics.117.300271
- Klarin, D., Damrauer, S. M., Cho, K., Sun, Y. V., Teslovich, T. M., Honerlaw, J., et al. (2018). Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nat. Genet.* 50, 1514–1523. doi: 10.1038/s41588-018-0222-9
- Komazawa, H., Yamaguchi, H., Hidaka, K., Ogura, J., Kobayashi, M., and Iseki, K. (2013). Renal uptake of substrates for organic anion transporters Oat1 and Oat3 and organic cation transporters Oct1 and Oct2 is altered in rats with adenine-induced chronic renal failure. *J. Pharm. Sci.* 102, 1086–1094. doi: 10.1002/jps.23433
- Levin, A., Tonelli, M., Bonventre, J., Coresh, J., Donner, J. A., Fogo, A. B., et al. (2017). Global kidney health 2017 and beyond: a roadmap for closing gaps in care, research, and policy. *Lancet* 390, 1888–1917. doi: 10.1016/s0140-6736(17)30788-2
- Lim, R., Sugino, T., Nolte, H., Andrade, J., Zimmermann, B., Shi, C., et al. (2019). Deubiquitinase USP10 regulates Notch signaling in the endothelium. *Science* 364, 188–193. doi: 10.1126/science.aat0778
- Lindner, A., Charra, B., Sherrard, D. J., and Scribner, B. H. (1974). Accelerated atherosclerosis in prolonged maintenance hemodialysis. *N. Engl. J. Med.* 290, 697–701.
- Liu, H., Yan, L., Ma, G. S., Zhang, L. P., Gao, M., Wang, Y. L., et al. (2012). Association of chronic kidney disease and coronary artery disease in 1,010 consecutive patients undergoing coronary angiography. *J. Nephrol.* 25, 219–224. doi: 10.5301/jn.2011.8478
- Liu, J. Z., McRae, A. F., Nyholt, D. R., Medland, S. E., Wray, N. R., Brown, K. M., et al. (2010). A versatile gene-based test for genome-wide association studies. *Am. J. Hum. Genet.* 87, 139–145. doi: 10.1016/j.ajhg.2010.06.009
- Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorf, L. A., Hunter, D. J., et al. (2009). Finding the missing heritability of complex diseases. *Nature* 461, 747–753. doi: 10.1038/nature08494
- McMahon, A., Malangone, C., Suveges, D., Sollis, E., Cunningham, F., Riat, H. S., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* 47, D1005–D1012. doi: 10.1093/nar/gky1120
- Moody, W. E., Edwards, N. C., Madhani, M., Chue, C. D., Steeds, R. P., Ferro, C. J., et al. (2012). Endothelial dysfunction and cardiovascular disease in early-stage chronic kidney disease: cause or association? *Atherosclerosis* 223, 86–94. doi: 10.1016/j.atherosclerosis.2012.01.043
- Moyers, J. S., Bilan, P. J., Zhu, J., and Kahn, C. R. (1997). Rad and Rad-related GTPases interact with calmodulin and calmodulin-dependent protein kinase II. *J. Biol. Chem.* 272, 11832–11839. doi: 10.1074/jbc.272.18.11832
- Musunuru, K., and Kathiresan, S. (2019). Genetics of common, complex coronary artery disease. *Cell* 177, 132–145. doi: 10.1016/j.cell.2019.02.015
- Nikpay, M., Goel, A., Won, H. H., Hall, L. M., Willenborg, C., Kanoni, S., et al. (2015). A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.* 47, 1121–1130. doi: 10.1038/ng.3396
- Nissen, S. E., Tuzcu, E. M., Schoenhagen, P., Crowe, T., Sasiela, W. J., Tsai, J., et al. (2005). Statin therapy, LDL cholesterol, C-reactive protein, and coronary artery disease. *New Engl. J. Med.* 352, 29–38. doi: 10.1056/NEJMoa042000
- Price, A. L., Weale, M. E., Patterson, N., Myers, S. R., Need, A. C., Shianna, K. V., et al. (2008). Long-range LD can confound genome scans in admixed populations. *Am. J. Hum. Genet.* 83, 132–135. doi: 10.1016/j.ajhg.2008.06.005
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Rennick, B., Acara, M., Hysert, P., and Mookerjee, B. (1976). Choline loss during hemodialysis: homeostatic control of plasma choline concentrations. *Kidney Int.* 10, 329–335. doi: 10.1038/ki.1976.116
- Ross, R. (1999). Atherosclerosis—an inflammatory disease. *N. Engl. J. Med.* 340, 115–126. doi: 10.1056/nejm199901143400207
- Shanahan Catherine, M., Crouthamel Matthew, H., Kapustin, A., Giachelli Cecilia, M., and Towler Dwight, A. (2011). Arterial calcification in chronic kidney disease: key roles for calcium and phosphate. *Circ. Res.* 109, 697–711. doi: 10.1161/CIRCRESAHA.110.234914
- Sivapalaratnam, S., Basart, H., Watkins, N. A., Maiwald, S., Rendon, A., Krishnan, U., et al. (2012). Monocyte gene expression signature of patients with early onset coronary artery disease. *PLoS One* 7:e32166. doi: 10.1371/journal.pone.0032166
- Smeland, O. B., Frei, O., Shadrin, A., O’Connell, K., Fan, C.-C., Bahrami, S., et al. (2020). Discovery of shared genomic loci using the conditional false discovery rate approach. *Hum. Genet.* 139, 85–94. doi: 10.1007/s00439-019-02060-2
- Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M., and Smoller, J. W. (2013). Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* 14, 483–495. doi: 10.1038/nrg3461
- Sugiyama, T., Kumagai, H., Morikawa, Y., Wada, Y., Sugiyama, A., Yasuda, K., et al. (2000). A novel low-density lipoprotein receptor-related protein mediating cellular uptake of apolipoprotein E-enriched beta-VLDL in vitro. *Biochemistry* 39, 15817–15825. doi: 10.1021/bi001583s
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613. doi: 10.1093/nar/gky1131
- Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* 20, 467–484. doi: 10.1038/s41576-019-0127-1
- The 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. *Nature* 526, 68–74. doi: 10.1038/nature15393
- Tonelli, M., Muntner, P., Lloyd, A., Manns, B. J., Klarenbach, S., Pannu, N., et al. (2012). Risk of coronary events in people with chronic kidney disease compared with those with diabetes: a population-level cohort study. *Lancet* 380, 807–814. doi: 10.1016/s0140-6736(12)60572-8
- Turner, R. C., Millns, H., Neil, H. A. W., Stratton, I. M., Manley, S. E., Matthews, D. R., et al. (1998). Risk factors for coronary artery disease in non-insulin dependent diabetes mellitus: United Kingdom prospective diabetes study (UKPDS: 23). *BMJ* 316:823. doi: 10.1136/bmj.316.7134.823
- van der Harst, P., and Verweij, N. (2018). Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. *Circ. Res.* 122, 433–443. doi: 10.1161/circresaha.117.312086
- Wang, W., Huang, X., Xin, H. B., Fu, M., Xue, A., and Wu, Z. H. (2015). TRAF family member-associated NF- $\kappa$ B activator (TANK) inhibits genotoxic nuclear factor  $\kappa$ B activation by facilitating deubiquitinase USP10-dependent Deubiquitination of TRAF6 ligase. *J. Biol. Chem.* 290, 13372–13385. doi: 10.1074/jbc.M115.643767
- Wang, W. Y., Barratt, B. J., Clayton, D. G., and Todd, J. A. (2005). Genome-wide association studies: theoretical and practical concerns. *Nat. Rev. Genet.* 6, 109–118. doi: 10.1038/nrg1522
- Wen, X., Luca, F., and Pique-Regi, R. (2015). Cross-population joint analysis of eQTLs: fine mapping and functional annotation. *PLoS Genet.* 11:e1005176. doi: 10.1371/journal.pgen.1005176
- Wuttke, M., Li, Y., Li, M., Sieber, K. B., Feitosa, M. F., Gorski, M., et al. (2019). A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat. Genet.* 51, 957–972. doi: 10.1038/s41588-019-0407-x
- Young, A. I. (2019). Solving the missing heritability problem. *PLoS Genet.* 15:e1008222. doi: 10.1371/journal.pgen.1008222
- Zanders, E. D. (2015). *Human Drug Targets: A Compendium for Pharmaceutical Discovery*. Hoboken, NJ: Wiley.
- Zeng, P., Hao, X., and Zhou, X. (2018). Pleiotropic mapping and annotation selection in genome-wide association studies with penalized Gaussian mixture models. *Bioinformatics* 34, 2797–2807. doi: 10.1093/bioinformatics/bty204

- Zeng, P., Zhao, Y., Liu, J., Liu, L., Zhang, L., Wang, T., et al. (2014). Likelihood ratio tests in rare variant detection for continuous phenotypes. *Ann. Hum. Genet.* 78, 320–332. doi: 10.1111/ahg.12071
- Zeng, P., Zhao, Y., Qian, C., Zhang, L., Zhang, R., Gou, J., et al. (2015). Statistical analysis for genome-wide association study. *J. Biomed. Res.* 29, 285–297. doi: 10.7555/jbr.29.20140007
- Zhou, W., Nielsen, J. B., Fritsche, L. G., Dey, R., Gabrielsen, M. E., Wolford, B. N., et al. (2018). Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat. Genet.* 50, 1335–1341. doi: 10.1038/s41588-018-0184-y

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Chen, Wang, Yang, Huang and Zeng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.