



# Weighted Co-Expression Network Analysis Identifies RNF181 as a Causal Gene of Coronary Artery Disease

Ruoyu Dang<sup>1</sup>, Bojian Qu<sup>1,2</sup>, Kaimin Guo<sup>3</sup>, Shuiping Zhou<sup>4</sup>, He Sun<sup>3</sup>, Wenjia Wang<sup>3</sup>, Jihong Han<sup>5</sup>, Ke Feng<sup>5</sup>, Jianping Lin<sup>1,2\*</sup> and Yunhui Hu<sup>3\*</sup>

<sup>1</sup>State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy and Tianjin Key Laboratory of Molecular Drug Research, Nankai University, Tianjin, China, <sup>2</sup>Pharmaceutical Intelligence Platform, Tianjin International Joint Academy of Biomedicine, Tianjin, China, <sup>3</sup>GeneNet Pharmaceuticals Co. Ltd., Tianjin, China, <sup>4</sup>The State Key Laboratory of Core Technology in Innovative Chinese Medicine, Tasly Academy, Tasly Holding Group Co., Ltd, Tianjin, China, <sup>5</sup>College of Life Sciences, State Key Laboratory of Medicinal Chemical Biology, Key Laboratory of Bioactive Materials of Ministry of Education, Nankai University, Tianjin, China

## OPEN ACCESS

### Edited by:

Yan Wang,  
Jilin University, China

### Reviewed by:

Pu-Feng Du,  
Tianjin University, China  
Dan Shao,  
Changchun University, China

### \*Correspondence:

Jianping Lin  
jianpinglin@nankai.edu.cn  
Yunhui Hu  
tsi-huyunhui@tasly.com

### Specialty section:

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

Received: 20 November 2021

Accepted: 24 December 2021

Published: 10 February 2022

### Citation:

Dang R, Qu B, Guo K, Zhou S, Sun H, Wang W, Han J, Feng K, Lin J and Hu Y (2022) Weighted Co-Expression Network Analysis Identifies RNF181 as a Causal Gene of Coronary Artery Disease. *Front. Genet.* 12:818813. doi: 10.3389/fgene.2021.818813

**Background:** Coronary artery disease (CAD) exerts a global challenge to public health. Genetic heritability is one of the most vital contributing factors in the pathophysiology of CAD. Co-expression network analysis is an applicable and robust method for the interpretation of biological interaction from microarray data. Previous CAD studies have focused on peripheral blood samples since the processes of CAD may vary from tissue to blood. It is therefore necessary to find biomarkers for CAD in heart tissues; their association also requires further illustration.

**Materials and Methods:** To filter for causal genes, an analysis of microarray expression profiles, GSE12504 and GSE22253, was performed with weighted gene co-expression network analysis (WGCNA). Co-expression modules were constructed after batch effect removal and data normalization. The results showed that 7 co-expression modules with 8,525 genes and 1,210 differentially expressed genes (DEGs) were identified. Furthermore, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were conducted. Four major pathways in CAD tissue and hub genes were addressed in the Hybrid Mouse Diversity Panel (HMDP) and Human Protein Atlas (HPA), and isoproterenol (ISO)/doxycycline (DOX)-induced heart toxicity models were used to validate the hub genes. Lastly, the hub genes and risk variants were verified in the CAD cohort and in genome-wide association studies (GWAS).

**Results:** The results showed that RNF181 and eight other hub genes are perturbed during CAD in heart tissues. Additionally, the expression of RNF181 was validated using RT-PCR and immunohistochemistry (IHC) staining in two cardiotoxicity mouse models. The association was further verified in the CAD patient cohort and in GWAS.

**Conclusion:** Our findings illustrated for the first time that the E3 ubiquitination ligase protein RNF181 may serve as a potential biomarker in CAD, but further *in vivo* validation is warranted.

**Keywords:** WGCNA (weighted gene co-expression network analyses), coronary artery disease, heart tissue, GWAS meta-analysis, diagnostic biomarkers

## INTRODUCTION

Coronary artery disease (CAD) has been noted as a challenge to public health faced by most industrialized and developing countries (Naghavi et al., 2017). Cardiovascular diseases (CVDs) contribute to one-third of the total deaths in the overall population (Joseph et al., 2017). According to the American Heart Association (AHA), in 2021, about 18.2 million (approx. 6.7%) adults (20 years or older) suffered from CAD; meanwhile, the threat has reached middle-aged patients, with about 2 in 10 deaths from CAD occurring in adults younger than 65 years (Virani et al., 2021). The immeasurable cost and huge burden of fatality are detrimental to the economy and society.

Vast endeavors have been put into understanding the pathology of CAD and its therapeutic strategies. For the past two decades, the discovery of novel targets, effective diagnostics, and new treatments for CAD has led to an over 50% decrease in mortality rate in the United States (Yahagi et al., 2016). CAD is a progressive cardiovascular disease that develops following atherosclerotic plaque formation or atherosclerotic occlusion of major arteries in the heart (Fuster et al., 1992). In this pathological process, fatty acid metabolism, glucose oxidation, mitochondrial fission, and oxidative stress largely impact the prognosis of the disease. Strategic treatments in clinical practice include carnitine palmitoyltransferase-1 (CPT-1) inhibitors (etomoxir, oxfenicine, etc.), malonyl-CoA inhibitors (trimetazidine),  $\beta$ -blockers, anti-ischemic agents, and the novel mitochondrial dynamic modulators, such as the Drp1 inhibitor. Nevertheless, these drugs have various disadvantages: inhibition of CPT-1 causes lipotoxicity in the pathological heart, leading to cardiac exacerbation; malonyl CoA inhibitor impedes the synthesis of fatty acid, but causes neuronal and cognition side effects to the brain; Drp1 is able to inhibit mitochondrial fission under oxidative stress conditions—however, its long-term effectiveness and safety remain to be discussed. For most treatments with a  $\beta$ -blocker or an anti-ischemic agent, they were only symptomatic; thus, further insight into identifying the risk factor for CAD is required.

The risk of CAD is a mixed consequence of genetic and lifestyle factors (Khera et al., 2016), such as smoking, physical inactivity, and a high-lipid diet, and hypertension, diabetes and obesity, or a family history (Joseph et al., 2017). In addition, the transcriptional and epigenetic regulation of macrophages and posttranslational modifications have been reported to be correlated with this complex disease (Kuznetsova et al., 2020). For example, the discovery of an abnormality in the hypermethylated region at HIF3A or the expression level of JCAD/KIAA1462 was reported to promote CAD (Xu et al., 2019). So far, the causative relationship of the perturbed genes in CAD has not been fully illustrated.

One significant contributing factor to CAD is genetic heritability. About 50% of genetic heritability was reported to be influential in the progression of CAD (Marenberg et al., 1994; Won et al., 2015). Advanced microarray and high-throughput sequencing technology have changed the research of CAD genetics. Large biobanks or shared data sources now provide

huge amounts of genetic and clinical information to facilitate the discovery of risk genes. Recently, numerous genome-wide association study (GWAS) meta-analyses have identified the risk variant or mutation associated with CAD. In 2015, Nikpay et al. reported 58 susceptible loci, including rs180803 and rs12976411, that were involved in 185,000 CAD and control cases ( $p < 5 \times 10^{-8}$ ) (Nikpay et al., 2015). Braenne et al. reported 159 loci, including rs1137524 and rs1060407, that may be single nucleotide polymorphisms (SNPs) with genome-wide significance in CAD (Bræne et al., 2015). The GWAS conducted by the CARDIoGRAMplusC4D Consortium reported 15 genome-wide significant loci out of 63,746 CAD cases and 130,681 controls (Deloukas et al., 2013). However, this only accounted for 10.6% of the total CAD heritability. Given that most of the locations of the loci or SNP variants identified from GWAS were outside the protein-coding regions, 40% of the risk variants were suspected to correlate with CAD, while robust associations for coding variants were only shown in four (Genetics et al., 2016). To this end, GWAS linked the associated locus or SNP to CAD pathogenesis, but further understanding of CAD heart tissue is needed to interpret the underlying mechanisms and associated genetic heredity.

Apart from GWAS, there are several other methods for the interpretation of significant genes and their associations with clinical traits. Many genetic-based prediction tools with computation scoring functions, such as the genetic risk score (GRS), gene set enrichment analysis (GSEA), differential expression analysis (DEA) and weighted gene co-expression network analysis (WGCNA), have been developed to facilitate the analysis of the genetics of CAD, either to estimate the probability of CAD or prioritize novel risk genes (Ntalla et al., 2019) based on the gene expression profiles.

WGCNA is a widely used computational method based on the gene expression profiles and clinical traits. WGCNA outperforms other analysis methods in detecting correlated gene modules. The hierarchy clustering function entails module finding, which consists of highly correlated genes, and identifies gene module-trait relationship, extracting significant genes from biologically meaningful modules (Goh et al., 2007; Horvath and Dong, 2008). WGCNA has been applied to identify modules and hub genes associated with clinical traits (Zhang and Horvath, 2005) and then to explore causal genes of diseases (Zheng et al., 2015; Wang et al., 2017; Jiao et al., 2020). WGCNA not only avoids the problems of multi-testing inherent in microarray data analysis but also provides means to bridge the gap from individual genes to systems oncology. Giulietti et al. (2018) reported the expression of two long non-coding RNAs (lncRNAs), LINC00675 and LINC01133, associated with the development and progression of pancreatic cancer using WGCNA. Zhou et al. (2018) found that hsa-miR-125b-5p, hsa-miR-145-5p, hsa-let-7c-5p, hsa-miR-218-5p, and hsa-miR-125b-2-3p were hub microRNAs (miRNAs) related to the prognosis of colon cancer. Yet, so far, the causative relationship between risk genes and CAD has not been fully illustrated.

This result further concluded the correlation between the ubiquitin-proteasome system (UPS) molecules and CAD. The

**TABLE 1** | Primer sequences for qRT-PCR analysis

Gene	Sense	Antisense
<i>Mus DKK3</i>	CTCGGGGGTATTTTGTGTGT	TCCTCCTGAGGGTAGTTGAGA
<i>Mus HP</i>	GCTATGTGGAGCACTTGGTTC	CACCCATTGCTTCTCGTCGTT
<i>Mus NME7</i>	AGATTTCGCTTTCATTGCAGAGT	GATCCGCTGTGGGGTAAAC
<i>Mus OXSM</i>	GGGTTATGGACTCTCGGGTGAT	TGGAAGTGGCATGTGCGTTGAC
<i>Mus PIGF</i>	TCCTTCTTCGTGGACAACCTTCT	AGAGGACACATTTCGGTTTCACTA
<i>Mus RNF181</i>	TTTGAGGACCTGGGATTGGTA	TTGGCGTACTGATGACTGTT

significance of UPS molecules, especially E3 ligases, has not been overwhelmingly discussed. Recent studies have emphasized E3 ligases and their contribution to cardiac diseases in CVD physiology. Chen et al. (2012) illustrated that the loss of E3 ligase activity promoted the impaired protein degradation in hypertrophic cardiomyopathy. Xu et al. (2020) reported that the ubiquitin-conjugating enzyme E2 variant 1 (Ube2v1) positively regulated protein aggregation by modulating UPS in cardiomyocytes, partially by enhancing K63 ubiquitination during a proteotoxic stimulus, supporting the hypothesis that UPS-mediated proteotoxic intracellular protein aggregation and degradation may lead to the progression of cardiac disease. From this point of view, targeting the association between UPS and CAD may be important in facilitating the understanding of the underlying mechanism and causal gene identification in CAD.

In this study, WGCNA was conducted in order to identify causal genes associated with CAD. With module clustering, pathway enrichment analysis, and cardiotoxicity mouse models, one E3 ubiquitin ligase gene, *RNF181*, was identified as the causal gene for CAD in the genome, messenger RNA (mRNA) and protein levels. Furthermore, by GWAS meta-analysis, two risk variants located at the *RNF181* locus were identified as associated with coronary heart disease. Our results revealed for the first time that *RNF181* may be a causal gene for CAD, possibly through decreasing the degradation of VEGF2 mediated by the NEDD4 and ERK/MAPK signaling pathways. Thus, targeting *RNF181* might be beneficial for the treatment of CAD.

## MATERIALS AND METHODS

### Microarray Data Resources and Preprocessing

The general workflow chart for this study is shown in **Supplementary Figure S1**. A total of 5 expression profiles from the Gene Expression Omnibus (GEO) database of NCBI, namely, GSE12504 (Ghorbel et al., 2010), GSE22253, GSE77263, GSE20681 (Beineke et al., 2012), and GSE49925 (Kim et al., 2014), were acquired, including expression profiles from CAD and control heart tissue and peripheral whole blood samples from human CAD cases. These datasets included the mRNA expression matrix, probe annotation table, and corresponding clinical features. Annotation and normalization for gene symbol and expression values were performed, respectively. The genetic profiles from the CAD cohort of the PREDICT trial (GSE20681) were collected to examine the association between causal genes in patients.

### Removal of Batch Effect

Normalization and batch effect removal were done for the combined dataset with the R package *sva* (Leek et al., 2012). The gene expression levels of both datasets were log<sub>2</sub> transformed. Normalization was performed to obtain clean data by removing the background variance between samples. The proportions of males in both healthy control and CAD groups were calculated and compared with the pairwise proportion test in R. Furthermore, genes with an expression level of the lowest 25% proportion were pruned.

### Identification of DEGs

The differentially expressed genes (DEGs) between the groups of healthy controls and CAD cases were calculated with the *limma* package (v3.42.2) (Smyth, 2005). The *p*-value was adjusted with the Benjamini–Hochberg (BH) method, and the fold change of all genes was log<sub>2</sub> transformed for normalization of the expression level. Moreover, the DEGs are shown in a volcano plot (**Supplementary Figure S3A**). A total of 1,210 DEGs with absolute log<sub>2</sub>FoldChange over 1 and *p* < 0.05 were identified.

### Construction of the Weighted Gene Co-Expression Network

A co-expression network is a widely used concept in biological interactions. It allows the interpretation of biological functionality in a system level. The conception of network construction is intuitive: some nodes (genes) are connected and co-expressed as a network across samples.

The most popular analysis pipeline for the construction of a co-expression network is WGCNA. WGCNA finds clusters of highly correlated genes (with hierarchical clustering) and summarizes these clusters by module eigengene (ME) or hub gene, in a way to liaise with external sample traits and assign module membership (MM) to genes. The branches of the hierarchical clustering dendrogram represent the modules and are refined with the dynamic tree cut method. The resulting gene clusters are often biologically meaningful. One of the advantages is that WGCNA losses less information on gene correlations and avoids the problem of multiple testing.

In WGCNA, a beta parameter is selected as a soft threshold power (SFT) to construct a co-expression network that achieves a scale-free connectivity. By adjusting the SFT, a Pearson's correlation matrix is established for calculating the pairwise correlation matrix between genes. Then, to reach a scale-free connectivity between genes ( $R^2 > 0.85$ ), the Pearson's correlation matrix is subsequently transformed to a weighted adjacency

matrix by setting a series of beta values (Ponsuksili et al., 2013). Next, a conversion of the weighted adjacency matrix to the topological overlapping matrix (TOM) is made using the block-wise module function of WGCNA. Based on the TOM, gene modules with similar expression patterns are identified with the hierarchy average lineage clustering in response to the dissimilarity of genes. For any module, module significance (MS) is defined as the Pearson's correlation coefficient between the ME and CAD traits. The module with the highest MS is selected as the causal gene module with high association with the disease condition. Then, the relationship between the ME of the module and the CAD trait is established to reveal its significance in CAD (Langfelder and Horvath, 2014).

## Pathway Enrichment Analysis With Gene Ontology and Kyoto Encyclopedia of Gene and Genomes

For the purpose of exploring the biological function of the module and hub genes correlated with the CAD phenotype, we performed Gene Ontology (GO) (Consortium, 2015) and Kyoto Encyclopedia of Gene and Genomes (KEGG) (Kanehisa et al., 2016) analyses with the R package clusterProfiler (Yu et al., 2012). To describe gene functions, the enriched terms were assigned into three GO categories: biological process (BP), cellular component (CC), and molecular function (MF). KEGG pathway enrichment analysis was performed to determine the significant KEGG pathway terms enriched by genes (Kanehisa and Goto, 2000). Each term was calculated with a *p*-value using Fisher's exact test, with a significance level of  $p < 0.05$ .

## PPI and Hub Gene Identification

In the study, the turquoise module was analyzed. Furthermore, the protein-protein interactions (PPIs) between the hub genes were queried from the STRING database and a cutoff threshold of the combined score  $>0.4$  (as default) was set. Genes with intramodular gene significance (GS) over 0.3 and MM over 0.8 were defined as having robust correlation. For hub gene identification, overlapping genes with GO/KEGG enriched pathways or DEGs for the significant module were often considered as hub genes.

## Animals

The Animal Ethics Committee of Nankai University approved the protocol of *in vivo* studies. C57BL/6J wild-type mice were purchased from Vital River Laboratory Animal Technology Co., Ltd. (Beijing, China). These animals were placed in the animal center of Nankai University in Tianjin, China, under specific pathogen-free conditions and had free access to water and food. All animal experiments were conducted in accordance with the ARRIVE guidelines and in accordance with the National Institutes of Health Guidelines for the Care and Use of Laboratory Animals (NIH publication no. 8023, revised in 1978).

## DOX/ISO-Induced Cardiotoxicity Mouse Experiment

Eight-week-old male C57BL/6 mice were assigned into a control group and a doxycycline (DOX) group (3 per group). The following treatments were given: mice in the control group were fed with normal diet and injected with normal saline once a week, a total of 4 times; mice in the DOX group were fed with normal diet, and DOX was injected intraperitoneally at 5 mg/kg per body weight, once a week for four times in total. Four weeks after treatment initiation, mice in both groups were sacrificed and heart tissue samples were collected.

For the isoproterenol (ISO) group, 8-week-old male C57BL/6 mice in both control and ISO group (3 mice/group) received similar procedures, except for the subcutaneous (s.c.) injection with saline or ISO at 3 mg/kg per body weight daily for 18 days.

## Echocardiography and Electrocardiogram Tests

The operation procedures for the echocardiography and electrocardiogram (ECG) have been described in a previous study (Feng et al., 2021).

## Real-Time Quantitative Polymerase Chain Reaction

Heart tissue total RNA of the sacrificed mice was collected from the control, DOX, and ISO groups using the RNA extraction and purification kit as per the manufacturer's protocol. Complementary DNA (cDNA) was synthesized by reverse transcription with the same amount of total RNA in each group (Ma et al., 2018), followed by quantitative real-time PCR (qRT-PCR) with the SYBR green PCR master mix purchased from Vazyme (Nanjing, China). The sequences of the primer templates are listed in **Table 1**. The expression levels of mRNAs, such as *DKK3*, *HP*, *NME7*, *OSXM*, *PIGF*, and *RNF181*, were normalized with the level of *GAPDH*.

## Hematoxylin-Eosin and Immunohistochemistry Staining

Heart tissues were collected from sacrificed mice of the control, DOX, and ISO groups, fixed in 4% paraformaldehyde for 24 h, and embedded in paraffin. After preparation, hematoxylin-eosin (HE) and immunohistochemistry (IHC) stainings were performed on paraffin-embedded 5- $\mu$ m heart sections. The epitopes of the slices were extracted in 10 mmol/L citric acid buffer and heated at pH 7.2 in a microwave. The slides were then incubated with mouse rnf181 primary antibody overnight at 4°C and incubation performed with a horseradish peroxidase (HRP) binding secondary antibody for 1 h at room temperature. The substrate diaminophenyl guanidine (DAB) was used to detect the antibody, and the slides were counterstained with hematoxylin. Immunostained areas of the IHC stains were evaluated and positive ratios were used for statistical analysis.

## GWAS Meta-Analysis

A total of 9 GWAS results were included. Results from the large CAD population-wise studies CARDIoGRAMplusC4D (Nikpay et al., 2015; Schunkert et al., 2011; Peden et al., 2011; Deloukas et al., 2013; Nelson et al., 2017), the Framingham Heart Study 100K Project (Larson et al., 2007), Bivariate Genome-Wide Association Scan (Siewert and Voight, 2018), and the CAD data in the UK Biobank (Fall et al., 2018) were included in our study. SNP variants located near the chromosome region of *RNF181* were analyzed. The prognostic value of each SNP in CAD groups was quantified using inverse variance weighted effect size (meta-analysis method implemented in METASOFT software, <http://genetics.cs.ucla.edu/meta/>) (Han and Eskin, 2011).

## Statistical Analysis

Values are represented as the mean  $\pm$  SEM. All assays were done in triplicate independently. Initially, all data were analyzed with the software GraphPad Prism. The workflow of the bioinformatics analyses, including WGCNA and DEGs, was performed with R version 3.6.0. The default statistics tests and cutoff values were specified in the corresponding sections in *Materials and Methods*. The statistical significance between the two groups of experimental data was assessed with the Student's *t*-test. The linear relationship between the gene expression levels was evaluated with Pearson's correlation coefficient. A significant difference was considered at the following levels of *p*-values: \**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001 (*n*  $\geq$  3).

## RESULTS

### Defining of a List of Marker Genes Associated With CAD Implicated by GWAS and the Literature

In this study, we focused on the hub genes and the risk factors for CAD in heart tissue with in-depth insight. We began by generating a list of marker genes located within the CAD risk loci from GWAS and the literature. The keywords “coronary artery disease” and “risk gene” were searched in the literature and 158 risk gene loci were collected (**Supplementary Table S1-1**). As another source of risk genes, 182 independent associations with a cutoff threshold of  $p < 5.0 \times 10^{-8}$  for CAD identified by GWAS meta-analysis were used from the 1000 Genomes Project (discovery and replication cohort with an enrolled population of  $\sim 185,000$ ; **Supplementary Table S1-2**) (Nikpay et al., 2015). For each of the 15 non-redundant lead GWAS SNPs, all RefSeq genes located within or overlapped with the region of the risk loci defined by linkage disequilibrium (LD) ( $r^2 \geq 0.7$ ) were included. The resulting CAD marker gene list contained 231 genes (**Supplementary Table S1-3**). Each association contained reported gene loci ranging from 1 to 16, with an average of  $4.5 \pm 4.7$ . The CAD marker gene list was enriched for GO terms such as “triglyceride homeostasis” ( $p = 1.7 \times 10^{-5}$ ), “lipoprotein metabolic process” ( $p = 4.3 \times 10^{-7}$ ), “vascular endothelial growth factor receptor signaling pathway” ( $p = 3.5 \times 10^{-5}$ ), and “angiogenesis” ( $p = 5.7 \times 10^{-4}$ ) (list shown in **Supplementary**

**Table S1-4**), suggesting that it may actually contain causal genes associated with CAD.

### WGCNA Identifying Gene Modules From the Expression Profiles of Heart Tissues From CAD Patients

Firstly, a total of 50 heart tissue samples, including 20 CAD cases and 30 healthy controls from GSE12504 and GSE22253, respectively, were used to perform WGCNA. Implementation of quality controls is essential to prevent batch effects in different sequencing datasets. Thus, batch effect was removed with the package *sva*, as shown in **Supplementary Figure S2**. The batch effect variation analysis results showed that the sample-wise mean and *p*-value of variance were 0.472 and 0.7246, respectively. Sample-wise skewness *p*-values were also calculated, with a value of 0.3656 suggesting batch effect has been removed.

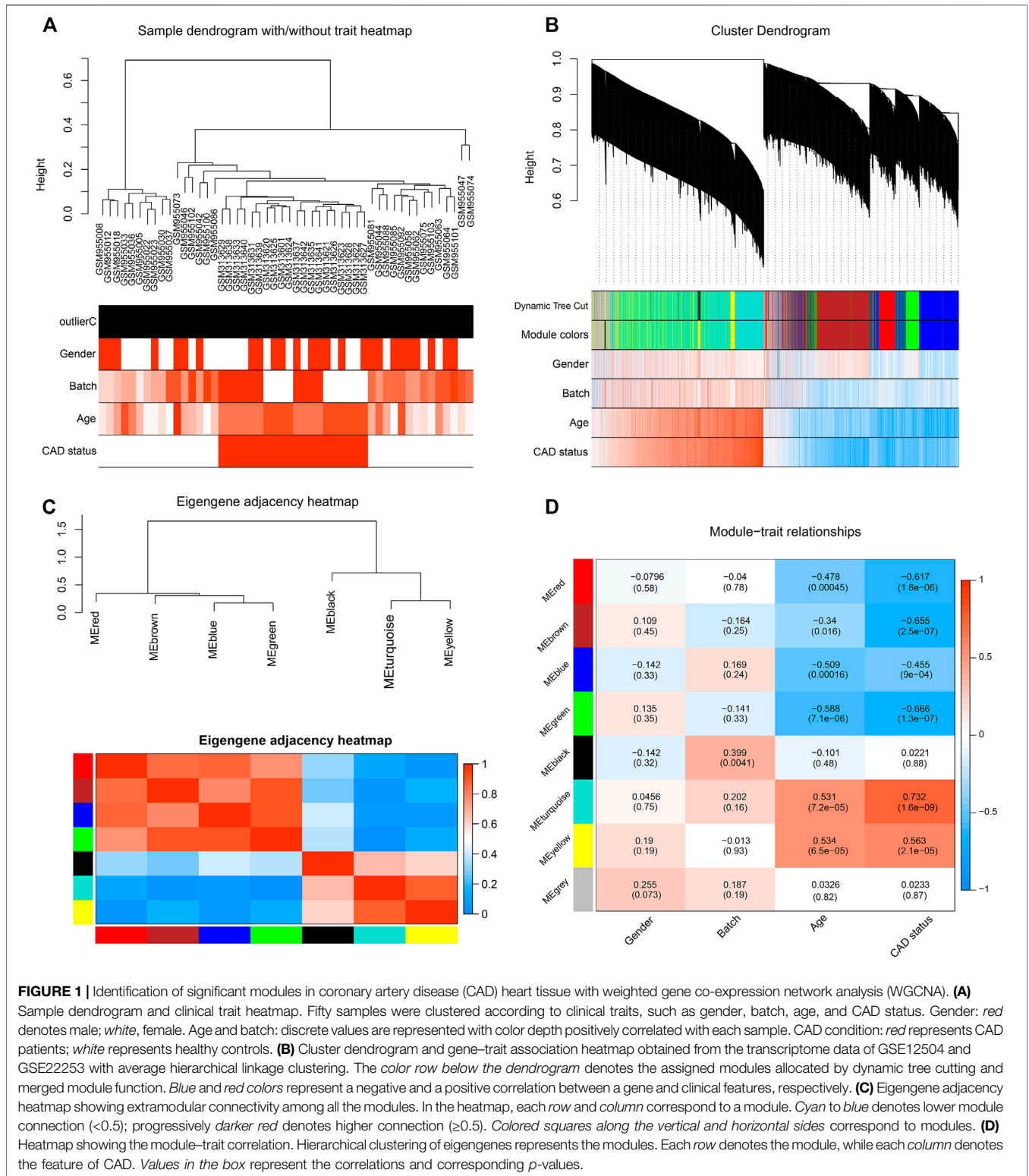
After data preprocessing, the expression profiles of 13,081 genes were gathered from the 50 samples. The DEGs between the CAD cases and healthy controls were analyzed with the *limma* package; the result showed that 1,210 DEGs were identified (**Supplementary Figure S3A**). Subsequently, WGCNA was performed for cluster analysis, and the quality of the dataset was evaluated with the *flashClust* function by sample clustering. A hierarchical clustering tree for all 50 samples in both groups was included (**Figure 1A**), based on which, all samples were included from the current study. With sample clustering, the 50 samples were assigned into two clusters: one containing 20 CAD samples and another containing 30 healthy controls.

After pruning low expression genes, a total of 8,525 genes in the 50 samples were included to construct the co-expression network using WGCNA. An appropriate soft thresholding power of 12 was set for the balance between the scale independence and mean connectivity between genes. Thus, a scale independence equal to 0.85 was achieved (**Supplementary Figure S3B**). As a result, the hierarchical clustering tree showed 7 co-expressed gene modules identified for further analysis (**Figure 1B**, genes not assigned to any module are shown in gray).

A TOM was built, based on which the independence among the 7 co-expression modules was analyzed. The results showed no significant overlap between module genes (**Figure 1C**), which suggested that a higher content of MM was achieved. The modules with the highest coefficient of correlation emerged as having the most significance in CAD. Each module was ranked by the correlation coefficient values to the CAD status.

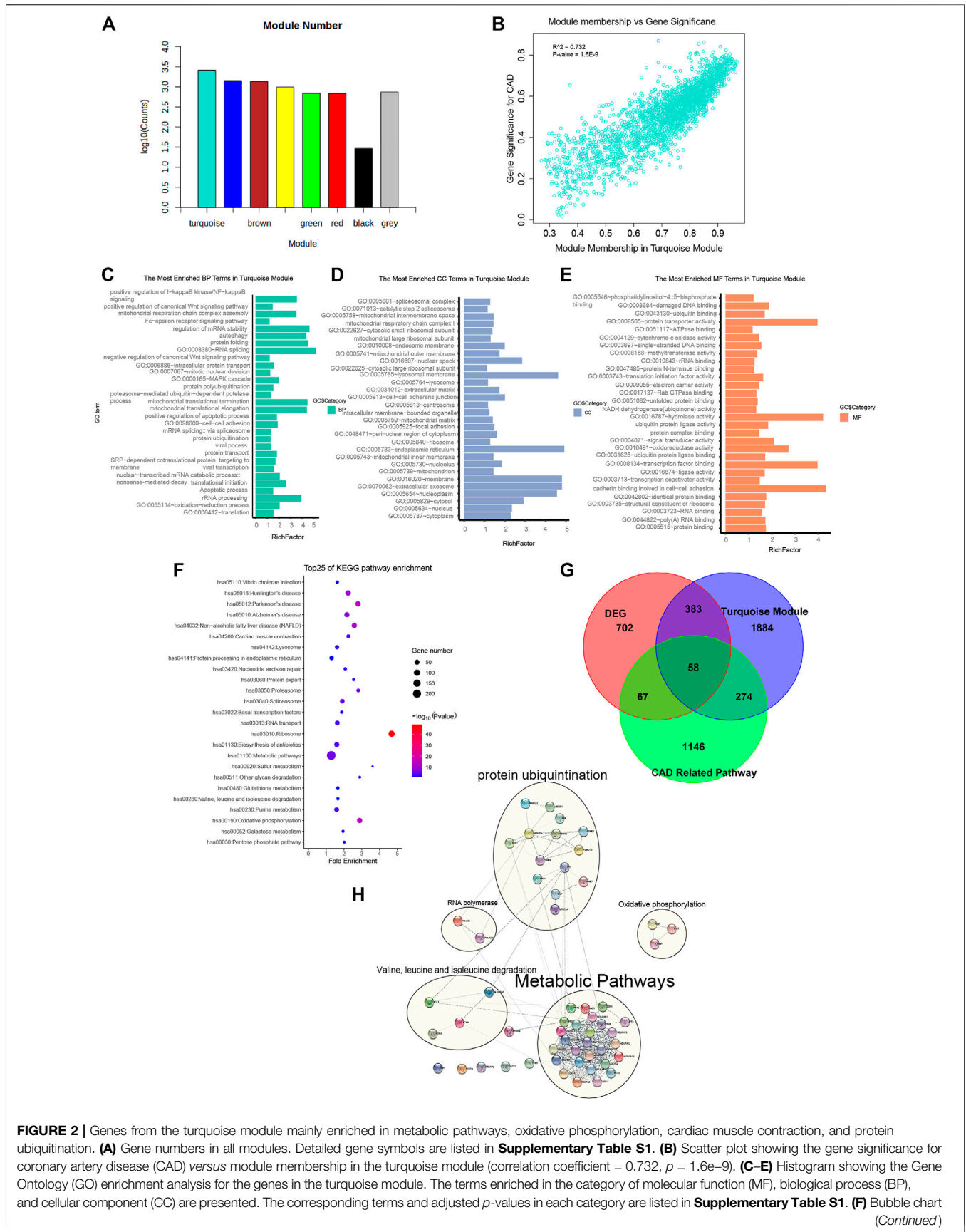
### Measurement of Module–CAD Associations and Functional Enrichment Analysis for the Causal Gene Module

To quantify the module–clinical feature associations, the eigengene expression dendrogram and eigengene *versus* clinical feature adjacency heatmap were constructed (**Figure 1D**). Seven co-expression modules with gene number ranging from 29 to 2,599 were identified (**Figure 2A**). Firstly, for the 7 co-expression modules, connectivity and cluster analysis was performed. The



degree of association was assessed between the eigengene and CAD traits, and the corresponding GS was deemed as the correlation between each gene and traits. The arithmetical mean GS of all genes in a module was regarded as MS. For

each module, the MS represents the association between its genes and CAD. The module–trait correlation heatmap showed the turquoise module as the most associated with the features of CAD, with a correlation coefficient of 0.732 (Figure 2B). Thus,



**FIGURE 2 |** Genes from the turquoise module mainly enriched in metabolic pathways, oxidative phosphorylation, cardiac muscle contraction, and protein ubiquitination. **(A)** Gene numbers in all modules. Detailed gene symbols are listed in **Supplementary Table S1**. **(B)** Scatter plot showing the gene significance for coronary artery disease (CAD) versus module membership in the turquoise module (correlation coefficient = 0.732,  $p = 1.6 \times 10^{-9}$ ). **(C–E)** Histogram showing the Gene Ontology (GO) enrichment analysis for the genes in the turquoise module. The terms enriched in the category of molecular function (MF), biological process (BP), and cellular component (CC) are presented. The corresponding terms and adjusted  $p$ -values in each category are listed in **Supplementary Table S1**. **(F)** Bubble chart (Continued)

**FIGURE 2** | showing the enriched pathway terms referring to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database for the genes of the turquoise module. **(G)** Venn diagram representing the intersection between the differentially expressed genes (DEGs), the turquoise module genes, and the genes of 4 CAD-related pathways. The intersecting 58 hub gene symbols in these three groups are listed in **Supplementary Table S1**. **(H)** Subnetworks constructed by 58 CAD causal genes with protein–protein interaction. Genes were divided into 4 main subnets of biological functions based on the degree of connection.

the turquoise module, including a total of 2,599 genes, was identified as a causal module that positively correlated with CAD ( $R^2 = 0.732$ ,  $p = 1.6E-09$ ) (**Supplementary Table S2-6**).

To interpret the biological content of the genes in the turquoise module, GO and KEGG enrichment analyses were performed. In the turquoise module, 258 GO terms were enriched. The GO terms enriched in the BP category included translation ( $p = 2.80E-40$ ), oxidation–reduction process ( $p = 9.50E-06$ ), apoptotic process ( $p = 3.90E-02$ ), and protein ubiquitination ( $p = 4.60E-02$ ). Other important biological processes reported to be associated with CAD, such as MAPK cascade ( $p = 3.30E-02$ ), Wnt signaling pathway ( $p = 2.76E-07$ ), and NF- $\kappa$ B signaling pathway ( $p = 1.10E-02$ ), were also enriched in the BP terms (**Figure 2C** and **Supplementary Table S2-2**). The GO terms of the turquoise module were enriched in the CC category of cytoplasm ( $n = 899$ ,  $p = 9.2E-21$ ), extracellular exosome ( $n = 524$ ,  $p = 1.40E-17$ ), and membrane ( $n = 384$ ,  $p = 1.20E-08$ ) (**Figure 2D** and **Supplementary Table S2-3**). The GO terms in the MF category were enriched in protein binding ( $p = 3E-31$ ), cell–cell adhesion ( $p = 3.50E-03$ ), ligase activity ( $p = 1.80E-02$ ), and ubiquitin protein ligase binding ( $p = 9.10E-02$ ) (**Figure 2E** and **Supplementary Table S2-4**). Referring to the KEGG database, the noted pathways enriched in genes of the turquoise module included metabolic pathways ( $p = 4.5E-6$ ), oxidative phosphorylation ( $p = 3.7E-15$ ), and cardiac muscle contraction ( $p = 8.0E-5$ ) (**Figure 2F** and **Supplementary Table S2-1**).

The significance of the biological functions of the causal gene modules from CAD were implicated from the GO and KEGG analyses. The first subset was enriched in metabolic pathways such as metabolic disturbances, involving the biosynthesis and degradation of cholesterol, triglycerides, and lipoproteins, which influences the presence of CAD. The second was enriched in oxidative phosphorylation, mainly including *PIGF*, *PIGP*, and *PIGT*, which are associated with CAD due to mitigating the increased production of reactive oxygen species in the mitochondria, accumulation of mitochondrial DNA damage, and progressive respiratory chain dysfunction. The third subset was cardiac muscle contraction, impairing mitochondrial integrity predisposed by vascular cell growth in CAD. The last subset was enriched in protein ubiquitination, whereas the dysfunction of the UPS deteriorates foam cell maintenance and mitigates low-density lipoprotein (LDL) aggregation *via* mediating the ubiquitination and degradation of p53.

In summary, we identified the turquoise module as a causal module for CAD. GO and KEGG enrichment analyses revealed that the modular genes shared a high association with biological functions such as metabolic pathways, oxidative phosphorylation, cardiac muscle contraction, and protein ubiquitination.

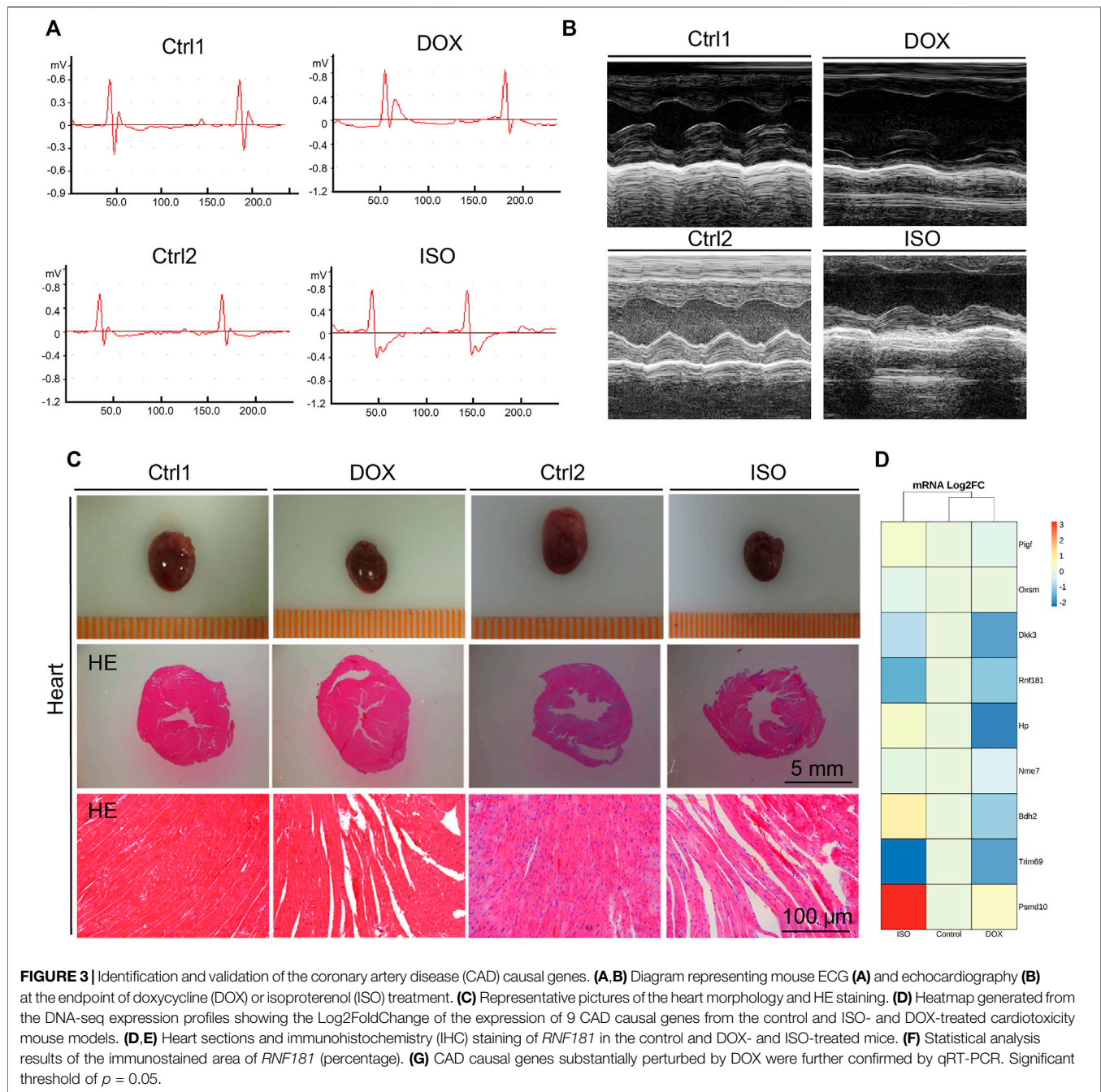
## Identification of Hub Genes From the CAD Casual Module With DEGs and PPI

Inferred from the above-mentioned results, the turquoise gene module was identified as a casual module associated with CAD. To filter for hub genes, genes in the turquoise module were compared with 1,210 DEGs (**Supplementary Table S2-5**), and 441 overlapping genes were deemed as significant genes (**Figure 2G** and **Supplementary Table S2-7**). Further analysis of pathways, including metabolic pathways, oxidative phosphorylation, cardiac muscle contraction, and protein ubiquitination (a total of 1,545 genes), yielded 58 genes (**Supplementary Table S2-8**). Moreover, the list of 57 causal CAD genes was queried in the STRING for PPI network (**Figure 2H** and **Supplementary Table S2-9**) and to further distinguish subnetworks. Genes were divided into 4 main subnets of biological functions based on the degree of connection. Subnetwork analysis indicated that most of the genes participated in metabolic pathways, oxidative phosphorylation, cardiac muscle contraction, and protein ubiquitination.

## Validation of the Hub Genes With Gene and Protein Expressions in HMDP and HPA

To improve the performance of the gene co-expression analysis, we incorporated prior knowledge for the purpose of extracting modules with biological meanings. Thus, further validation of the expression of hub genes was performed on the mouse heart tissue samples from the Hybrid Mouse Diversity Panel (HMDP). Another analysis framework with WGCNA to discover the intrinsic differences between similar tissues (Abbassi-Daloui et al., 2020) was performed on the murine heart profiling of HMDP (GSE77263). To select an optimal set of WGCNA parameters, a total of 270 combinations of parameters, including power, minimum cluster size, deep split value, and tree cutting height, were tested for co-expression network construction. The selection of the best parameters considered prior knowledge of gene–gene interactions, including the enrichment terms of pathway databases and the minimum size of the gray module (genes not assigned with any module) (**Supplementary Figure S3C**). The soft thresholding power was set as 6 and the minimum cluster size as 15; the deepSplit parameter was set as 3 and the tree cutoff height as 0.15 (**Supplementary Figure S3C**). With the optimal parameters, WGCNA generated 57 modules, which included 7,211 genes (5,186 non-redundant gene symbols), including the gray module with a total of 675 genes. Of which, 12 significant modules were identified with a threshold of false discovery rate (FDR)  $< 0.05.99$  of the 231 marker genes that were assigned to at least one module, suggesting that most of the CAD-related genes were enriched





(Supplementary Figure S3D, 4A). Of the 58 hub genes, 35 were assigned to at least one module.

To verify the heart tissue specificity, we further investigated the 58 hub genes based on the protein and mRNA expression levels in the Human Protein Atlas (HPA). The following criteria were applied for screening CAD causal genes from the 58 genes: 1) mRNA expression was detectable in heart tissue; 2) the protein expression of the genes was reported in human heart tissues; and 3) medium to high levels of IHC staining in the left ventricle heart slices in HPA. The genes meeting these criteria included *DKK3*, *HP*, *BDH2*, *NME7*, *PIGF*, *OXSM*, *PSMD10*, *RNF181*, and *TRIM69*

(Supplementary Figure S3E). Therefore, these 9 hub genes were speculated to possibly play critical roles during the pathogenesis of CAD.

### Determination of Hub Gene Perturbation in the DOX- and ISO-Induced CAD Mouse Model

To verify the expression of the hub genes derived from the WGCNA microarray in CAD heart tissue, DOX- and ISO-induced cardiotoxicity mouse models were introduced.

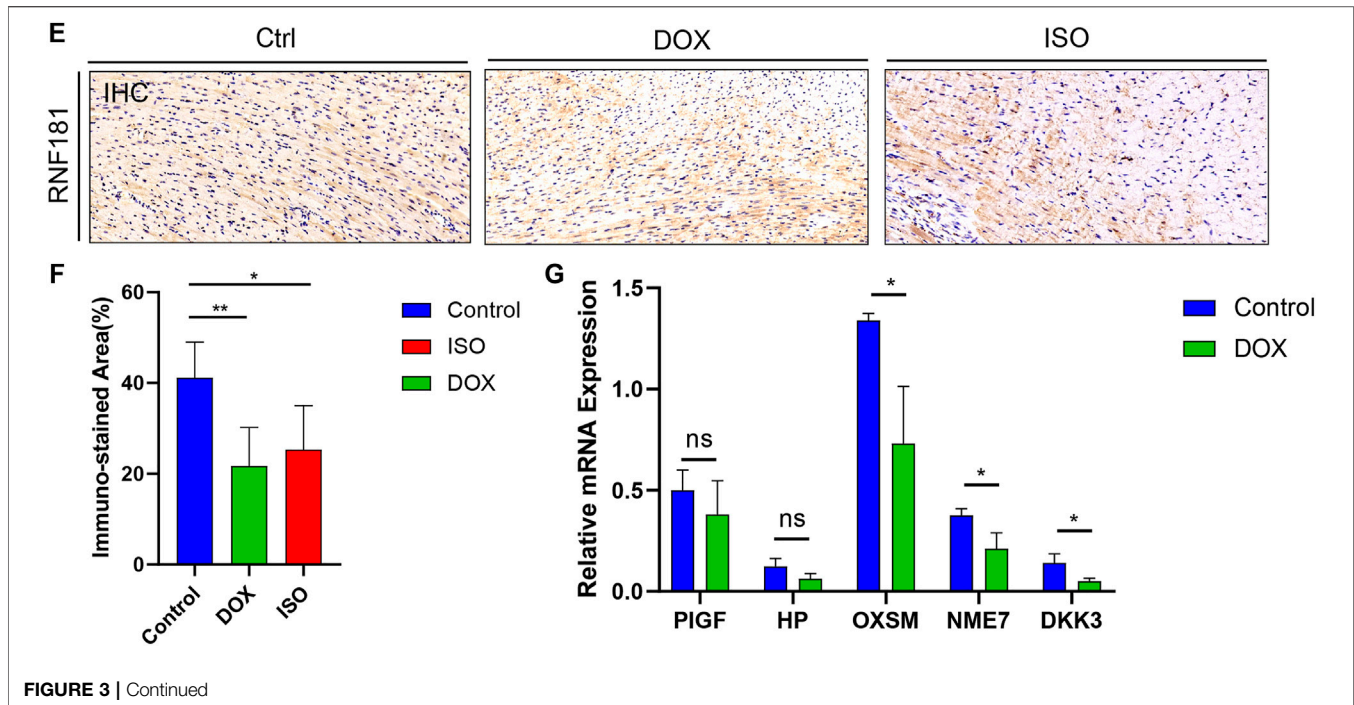


FIGURE 3 | Continued

DOX and ISO are two typical agents that can cause myocardial damage by mitigating oxidative stress, necrosis, and cardiac-related metabolic disorders. To investigate the disturbance of the CAD causal genes during the myocardial damage induced by ISO or DOX, mice were treated with s.c. injection of both agents, as previously described (Feng et al., 2021). Both ECG and echocardiography tests were conducted at the end of the treatment. Compared with the control group, an increased ST segment in the ECG was observed in both DOX and ISO treatment groups (Figure 3A), which was related to the decrease of the ejection fraction (EF) and fractional shortening (FS) (Figure 3B), indicating that DOX and ISO caused heart dysfunction.

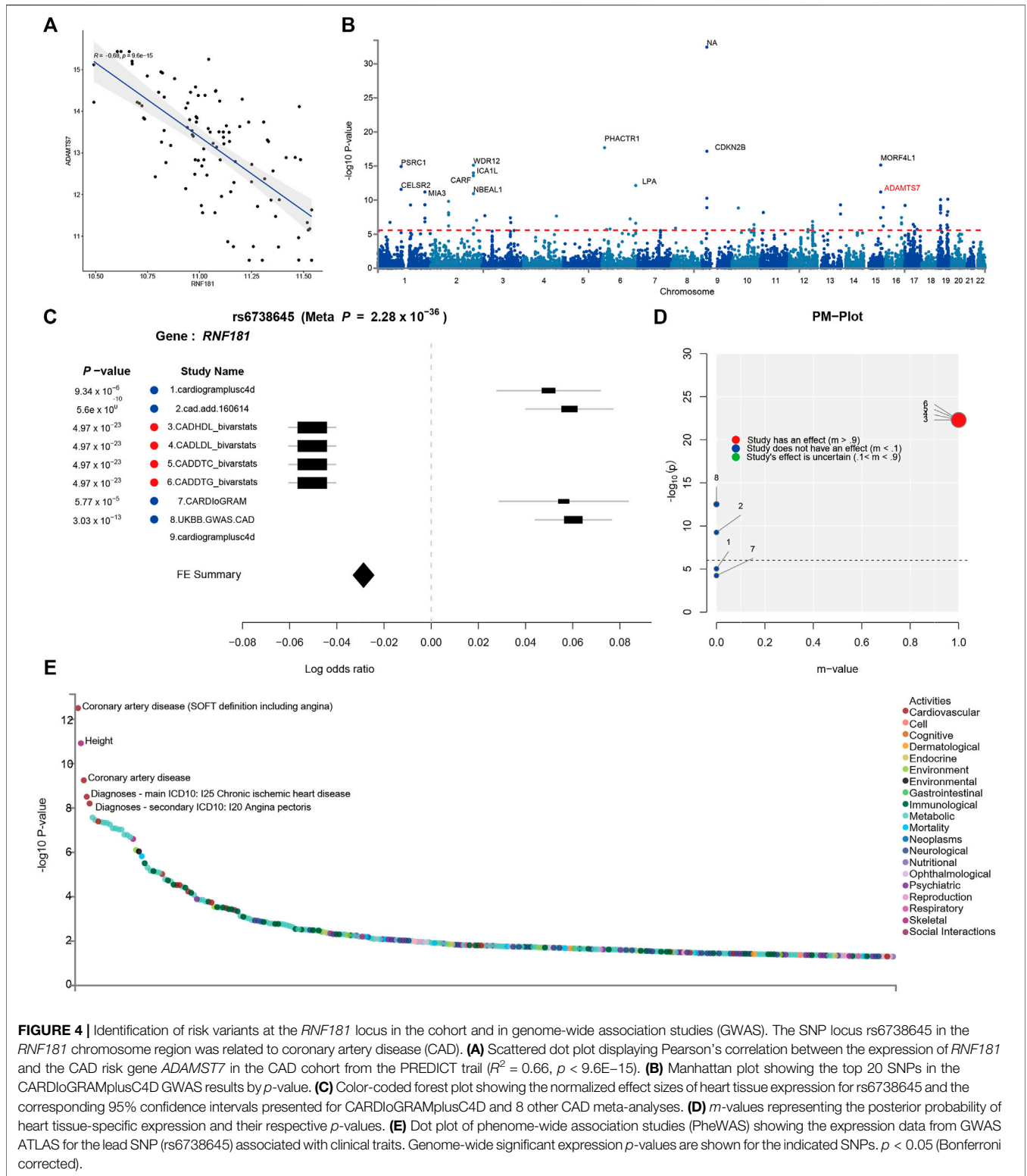
Subsequently, to determine the heart morphology affected by DOX or ISO treatment, HE staining was conducted on the heart sections. The results showed that DOX or ISO caused disordered arrangement of myocardial cells, and the myocardial structure of lysed muscle fibers was unclear (Figure 3C). Taken together, the results demonstrate that DOX or ISO can significantly induce heart atrophy and its related cardiac morphology in mice, suggesting the cardiac dysfunction in DOX- or ISO-induced heart injury.

To validate the change of the CAD causal genes in DOX- or ISO-induced heart injury, an RNA sequencing (RNA-seq) assay was performed on RNA extracted from mouse heart. The results indicated that the expression of several CAD causal genes associated with heart function were changed by DOX or ISO (Figure 3D). For instance, the gene expression levels of *RNF181* and *DKK3* were significantly decreased in both DOX- and ISO-induced cardiotoxicity models. Dickkopf-3 is a key vascular progenitor of the atherosclerotic plaque phenotype that mitigates the differentiation of fibroblasts into functional

endothelial cells and is encoded by *DKK3* (Karamariti et al., 2018). Circulating Dickkopf-3 (He et al., 2016) and the fibroblast–endothelial cell transition are associated with CAD development. Besides, the expression levels of several other causal genes identified were also changed in heart function or cardiac diseases, such as *HP*, *BDH2*, *NME7*, *PIGF*, *OXSM*, *PSMD10*, and *TRIM69*. DOX decreased the expression of most hub genes; particularly, the decrease of *RNF181* expression in both DOX and ISO cardiotoxicity models followed a similar manner.

Subsequently, an IHC staining experiment for *RNF181* was conducted (Figure 3E). The *RNF181* immunostained area in both heart toxicity models showed a decrease of *RNF181* protein expression appearing in areas of CAD heart compared with the control heart tissue, suggesting that, at the protein level, the downregulation of *RNF181* may be correlated with CAD progression (Figure 3F). The effect of DOX on the expression of some CAD causal genes was further verified at the mRNA level by qRT-PCR. Consistent with the results of the RNA-seq assay, the mRNA levels of *OXSM*, *NME7*, and *DKK3* were also inhibited by DOX (Figure 3G). These results also demonstrate that *RNF181*, as well as *OXSM*, *NME7*, and *DKK3*, deteriorated during the heart failure process by mitigating multiple pathways.

Furthermore, to understand the causal relationship between *RNF181* and CAD, genes with a PPI with *RNF181* were identified in the STRING and INTACT databases. In both databases, the E3 ubiquitination ligase protein family members *UBE2N*, *UBC*, *UBE2D2*, *UBB*, *UBE2D1*, *UBE2D3*, *UBA52*, *BCL10*, *UBE2E1*, and *RPS27A* were shown to directly interact with *RNF181*. Moreover, 9 of the 10 aforementioned E3 ubiquitination proteins interacting with *RNF181* were reported to be related with heart-associated diseases (Supplementary Figure S4B).



**FIGURE 4 |** Identification of risk variants at the *RNF181* locus in the cohort and in genome-wide association studies (GWAS). The SNP locus rs6738645 in the *RNF181* chromosome region was related to coronary artery disease (CAD). **(A)** Scattered dot plot displaying Pearson’s correlation between the expression of *RNF181* and the CAD risk gene *ADAMTS7* in the CAD cohort from the PREDICT trail ( $R^2 = 0.66, p < 9.6E-15$ ). **(B)** Manhattan plot showing the top 20 SNPs in the CARDIoGRAMplusC4D GWAS results by  $p$ -value. **(C)** Color-coded forest plot showing the normalized effect sizes of heart tissue expression for rs6738645 and the corresponding 95% confidence intervals presented for CARDIoGRAMplusC4D and 8 other CAD meta-analyses. **(D)**  $m$ -values representing the posterior probability of heart tissue-specific expression and their respective  $p$ -values. **(E)** Dot plot of phenome-wide association studies (PheWAS) showing the expression data from GWAS ATLAS for the lead SNP (rs6738645) associated with clinical traits. Genome-wide significant expression  $p$ -values are shown for the indicated SNPs.  $p < 0.05$  (Bonferroni corrected).

**TABLE 2** | Lead genome-wide association studies from GWASATLAS for the SNP rs6738645 associated with CAD clinical traits

	PMID	Year	Domain	Trait	p-value	N	EA	NEA
3925	28714975	2017	Cardiovascular	Coronary artery disease (SOFT definition including angina)	3.14E-13	148,815	G	T
4043	30124842	2018	Skeletal	Height	1.2E-11	693,529	G	T
108	26343387	2015	Cardiovascular	Coronary artery disease	5.6E-10	184,305	G	T
3668	31427789	2019	Cardiovascular	Chronic ischemic heart disease	3.102E-09	300,791	G	T
3692	31427789	2019	Cardiovascular	Angina pectoris	6.244E-09	244,890	G	T
3470	31427789	2019	Metabolic	Trunk fat-free mass	2.666E-08	379,507	G	T
3471	31427789	2019	Metabolic	Trunk fat-free mass	3.528E-08	379,507	G	T

EA, effect allele; NEA, non-effect allele

To further verify the function of *RNF181*, the expression data from GSE143947 with selective knockdown of *RNF181* in MCF-7 cells were analyzed. The results showed that si*RNF181* induced the downregulation of CDK4, as well as MAPK and NEDD4 (**Supplementary Figure S4C**).

Recent pieces of evidence have shown that the E3 ubiquitin ligase Skp2/fbx1 might regulate the proliferation of vascular smooth muscle cells (Wu et al., 2006) and cardiomyocytes (Tamamori-Adachi et al., 2004), which suggests that impaired UPS-mediated degradation may impede cardiomyocyte proliferation, while the silencing of *RNF181* reduced the cell proliferation induced by the expression of CDK4 and worsened proliferation and deteriorated cardiac efficiency post-ischemia *in vivo*.

Wang et al. reported that *RNF181* exhibits an inhibitory role in the ERK/MAPK signaling pathway, the mechanism of which was through the control of the activity of the cyclin D1-CDK4 signaling cascade. Its consequent regulatory role on the cell fate from the G1 to the S phase in gastric cancer led to a deteriorated prognosis (Wang et al., 2019). In our results, the silencing of *RNF181* was associated with the downregulation of the expressions of MAPK8 and MAPK14.

In addition, vital signaling pathways in the cardiovascular system, such as the vascular endothelial growth factor (VEGF) pathway, are also regulated by UPS. The VEGF signaling pathway mitigates angiogenesis, as well as multiple cellular activities such as cell permeability, proliferation, and survival (Pagan et al., 2013). Selective silencing of *RNF181* resulted to a significant downregulation of the expression of NEDD4, suggesting a hypothesis that the silencing of *RNF181* may increase the degradation of VEGFR2 *via* NEDD4 in CAD.

## Correlation Between *RNF181* Expression and CAD Progression in the Cohort and GWAS Studies

To further verify the causal relationship between *RNF181* and CAD heredity, we investigated the cohort and GWAS in the CAD population. The PREDICT trial (GSE20681) is a prospective multicenter coronary artery catheterization laboratory study conducted in the United States to identify biomarkers related to CAD (Baim et al., 2001). Pearson's correlation was performed for the CAD cohort to investigate the expression of *RNF181* and CAD marker genes. The results showed that *RNF181* correlated with the expression levels of *ADAMTS7* and *S100A6* (**Figure 4A**

and **Supplementary Figure S5A**), two well-established variant genes associated with CAD (Bauer et al., 2015; Chan et al., 2017; Cai et al., 2011; Mofid et al., 2017).

Large-scale GWAS such as the 1000 Genomes Project have identified 46 independent genome-wide significant SNPs for CAD. These SNPs included rs3918226 in *NOS3* on 7q36.1, rs10455872 and rs3798220 in *LPA* ( $p = 5.7 \times 10^{-39}$ ,  $4.7 \times 10^{-9}$ ), and rs7412 in *APOE* ( $p = 8.2 \times 10^{-11}$ ) (Nikpay et al., 2015) (**Figure 4B**). To identify risk variants in the *RNF181* locus related to CAD, a random-effects meta-analysis was conducted as a sensitivity analysis using the Han and Eskin method in METASOFT. To avoid significant heterogeneity, a function that mitigates power loss was utilized in the Han-Eskin method (Han and Eskin, 2011). The meta-analysis results showed that the fixed effects of gene polymorphisms on *RNF181* outcomes across CAD cohorts yielded a single SNP, rs6738645, which was therefore considered significant (with a  $p$ -value of  $2.28 \times 10^{-36}$ ). rs6738645 was considered as significantly related to CAD across the four treatment cohorts (*RNF181*: meta-analysis hazard ratio = 1.26,  $p = 3.03E-105$ , FDR = 1.13E-73). This effect was likely driven by the results of the Biovarstats study, where variant alleles were associated with high-density lipoprotein (HDL), LDL, and changes in total cholesterol (DTC) and triglyceride (DTG) (**Figure 4C**). Additionally, we also identified another association for gene variance, rs1562322, with CAD (*RNF181*: OR = 0.76,  $p = 0.290648$ , FDR = 0.168875) (**Supplementary Figure S5B-E**), suggesting that risk variants rs6738645 and rs1562322 in the *RNF181* locus may be associated with coronary heart disease.

PM analysis from the variant-trait association map of rs6738645 showed that the 4 studies had an effect on the association between rs6738645 and CAD by the  $m$ -value, suggesting the probability of a causal relationship (**Figure 4D**). Analysis of phenome-wide association studies (PheWAS) suggested that genetic variation rs6738645 is significantly associated with coronary heart disease (Pearson's  $r = 2.437e-32$ ,  $p = 1.19E-06$ ) in several large GWAS. These results can be explained by the associations involving the variant rs6738645 of *RNF181*, which showed genome-wide significance in the meta-analysis for all these studies ( $p < 0.05$ ) (**Figure 4E**). *RNF181* is located on chromosome 2 at chr2: 85722848-85924831. In the GWAS ATLAS, the mutation of T to G at chr2:85783128 was reported in two independent studies as associated with the increase in the incidence of CAD, with

enrolled populations of 148,815 and 184,305 ( $p = 3.14E-13$  and  $5.6E-10$ , respectively) (Table 2).

Taken together, *RNF181* may serve as a causal gene of prognostic or therapeutic value targeting coronary heart disease.

## DISCUSSION

The current study indicated 7 co-expression modules, one of which was significantly associated with CAD. From the causal gene module, the significance of 4 pathways, including protein ubiquitination during CAD pathogenesis, was revealed. *RNF181* and 8 other CAD causal genes were further identified from the causal gene module. The protein expression of *RNF181* was verified in both human and mouse heart toxicity samples. The gene expression of *RNF181* correlated with the expressions of *ADMAST7* and *S100A6* in the CAD cohort study. Further evidence in GWAS indicated that the SNP variants rs6738645 and rs1562322 in the *RNF181* locus were related to the risk of CAD. These results confirm the association between the causal gene *RNF181* and CAD progression.

Complex diseases such as CAD may be the results not only of the accumulative effects of individual genetic factors but also of gene interaction *via* biological pathway/networks. Using WGCNA on the tissue-level expression data, we found that *RNF181* was co-expressed with 7 other genes and that many of them were validated in previous studies to be associated with CAD.

Integration with the computational-based method for the discovery of novel targets in complicated diseases has been applied successfully. WGCNA and GWAS are two representative methods for identifying risk genes from genetic data. They are different in that one is based on the gene module–disease trait correlations while the other provides mutation–phenotype information from a population cohort.

A standard WGCNA workflow includes finding modules of highly correlated genes, summarizing the modules with ME or intramodular hub genes, and utilizing the eigengene network methodology to liaise modules to external traits, finally calculating MM.

One of the major shortcomings of WGCNA is that Pearson's correlation conveys only the linear dependencies in a theoretical network. However, the true relationships observed in a biological system are sophisticated, involving both linear and nonlinear dependencies. Another is that the function of hierarchical clustering is irreversible; thus, it is impossible to readjust or relocate the uncategorized genes during module identification. Several attempts have been made in improving WGCNA. Greenfest-Allen et al. (2017) sought to improve robustness by a process of pruning uncategorized genes and performed a re-clustering step to obtain a precise module. Dai et al. (2018) introduced a modified method called cusWGCNA, which combined both the signed and unsigned network functions. Botia et al. (2017) proposed an additional *k*-means clustering step to improve the performance of WGCNA. Abbassi-Daloui et al. (2020) suggested another analysis framework with WGCNA to discover the intrinsic differences between similar tissues, which incorporated pathway knowledge and a combination of parameter selection to acquire refined gene modules with bio-meaningful content.

GWAS has been the predominant approach to the genetic analysis of complex diseases in the last decade and had demonstrated its usefulness in prioritizing over 150 novel risk loci associated with CAD. The most prominent example was *PCSK9*. Early GWAS identified that the locus rs11591147 on *PCSK9* was significantly associated with CVD ( $p = 7.5 \times 10^{-6}$ ), which suggested that the Arg46Leu substitution is associated with the cholesterol levels of LDL and has favorable prognosis in cardiology. Thus far, two *PCSK9* inhibitors approved by the US Food and Drug Administration (FDA) were immediately proven as effective, which decreased the risk of cardiovascular events by decreasing ~50% of circulating LDL cholesterol in clinical trials (Robinson et al., 2015; Sabatine et al., 2015). Thus, an integrated method of WGCNA and GWAS summary analysis was used to prioritize novel causal genes associated with CAD. The results may be suggestive in that a certain analogy could be a paradigm in prioritizing causal genes in cardiovascular diseases.

Moreover, we further discussed the association between the UPS molecule *RNF181* and CAD. As a member of the RING finger protein family, *RNF181* is a novel type of E3 ubiquitin ligase that regulates biological activities such as protein dimerization, PPI, and ubiquitin ligase activities. Thus, *RNF181* and its role in the regulation of phenotypic change, proliferation, migration, and apoptosis have been discussed in several types of malignances. *RNF181* exhibits its E3 ubiquitin ligase activity *via* binding to the integrin alpha-IIb (ITGA2B)/beta-3 (ITGB3) complex (Allen, 2015). The interaction between *RNF181* and *CARD11* may enhance the NF- $\kappa$ B signaling pathway in lymphoma (Bedsaul et al., 2018). Besides, *RNF181* also facilitates cell viability and angiogenesis in colon carcinoma (Xiong, 2017). In breast cancer, *RNF181* prolongs the stability of *Era* associated with AF1 *via* its RING domain binding to the domain of *Era*, and enhancement of the gene expression of *Era* promotes breast cancer progression (Zhu et al., 2020). In the incurable triple-negative breast cancer (TNBC), *RNF181* inhibits the K48-linked poly-ubiquitination of YAP, thus promoting YAP stability. In this process, it mediates the activation of Hippo/YAP signaling in a positive way that prohibits the treatment of TNBC (Zhou et al., 2020). As discussed in previous studies, the UPS molecule *RNF181* played an important role in tumor biology. Our results further support the idea of *RNF181* and its association with CAD progression.

We are aware that certain limitations are present in our study. Initially, in the preprocessed CAD data, only 8,525 genes were incorporated in this study. It is likely that the coverage of related genes is incomplete and sparse, which led to only 8 positive findings with WGCNA. Certain low expressed genes that may have a causal relationship with CAD may not have been included. Also, for *RNF181* and other causal genes, bioactivity and function tests, including silencing or overexpression experiments on cardiomyocytes or assessments on druggability and pharmacology, were not performed. Even it is evolved in the DOX/ISO-induced heart toxicity model, GWAS, and the cohort study, there is no sufficient supporting evidence for clinical practice. Despite this, our results support *RNF181* being another putative novel causal gene for CAD. Subsequently, the cell origin of the RNA-seq profiles from GEO contained only tissues from the left ventricle and did not collect sorted subtypes of cells with various biological functions, including muscular

cells, endothelial cells, or macrophages. Thus, we cannot immediately conclude in which specific type of cell the *RNF181*-mediated phenotype change may have functions. Lastly, WGCNA may lose power compared to the standard gene-based analysis in circumstances where the true biological mechanism is independent of gene expression.

In summary, for the first time, our findings illustrate that the E3 ubiquitination ligase *RNF181* may serve as a causal gene affecting CAD through its downregulation. *RNF181* may play an important role in CAD progression. Further efforts are required to verify the potent interactions and the regulatory mechanism of *RNF181* in CAD and other cardiovascular diseases.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by CardioDx. The patients/participants provided written informed consent to participate in this study. The animal study was reviewed and approved by The Animal Ethics Committee of Nankai University.

## AUTHOR CONTRIBUTIONS

RD contributed to the methodology, software, manuscript drafting, project administration and manuscript revision. BQ performed the investigation, data analysis, software, and manuscript revision. KG performed the formal analysis, validation and manuscript revision. SZ contributed to the data interpretation, manuscript revision, final approval and integrity Check. HS and WW conducted the investigation. JH and KF helped with validation. JL contributed to conceptualization, Validation and Final approval. YH helped with the methodology, validation, manuscript revision, and final approval.

## FUNDING

This study was supported by the National Key Research and Development Program of China (2017YFC1104400 to JL).

## REFERENCES

- Abbassi-Daloi, T., Kan, H. E., Raz, V., and 't Hoen, P. A. C. (2020). Recommendations for the Analysis of Gene Expression Data to Identify Intrinsic Differences between Similar Tissues. *Genomics* 112, 3157–3165. doi:10.1016/j.ygeno.2020.05.026
- Allen, S. (2015). *Characterization of a Protein: Protein Interaction between RNF181 and the Platelet Integrin,  $\alpha$ IIb $\beta$ 3*. Dublin: Royal College of Surgeons in Ireland.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.818813/full#supplementary-material>

**Supplementary Figure S1** | The General Flowchart showing the analysis process of the study: Data collection, Batch effect correction, data analysis, and validation. Expression data of CAD samples from GSE12504 and normal samples from GSE22253 were analyzed after batch effect removal and preprocessing. Then, 1,210 differentially expression genes (DEGs) were screened, and the turquoise module was identified through WGCNA. GO functional and KEGG pathway enrichment analyses were performed the significant modules. CAD causal genes were identified from the module with PPI and verified with HMDB and HPA. The RNF181 and other 8 hub genes were recognized and validated using IHC, RT-PCR and gene expression in two heart toxicity mouse models. Two risk variants rs6738645 and rs1562322 at the RNF181 locus identified in GWAS meta-analysis

**Supplementary Figure S2** | Batch effect removal for the expression profile of CAD and healthy control heart tissue. Hierarchy clustering tree showing samples before (A) and after (B) batch effect removal with sva, each branch represents a sample, Red, CAD, and blue, healthy Control. Principle component analysis for the expression data of samples before (C) and after (D) batch effect removal. Red, CAD, blue, healthy Control. (E, F) Batch effect variation analysis results for the expression data of GSE12504(CAD) and GSE22253(Control).

**Supplementary Figure S3** | Validation of the hub genes with gene and protein expressions from HMDB and HPA. (A) Volcano plot showing the DEGs between CAD and Control. (B) Network topology analysis for adjacency matrix with a set of soft-thresholding powers. The red numbers in the drawing indicate soft-thresholding power in keeping with the correlation coefficient value. (C) Scattered dot plot showing the selection of parameter combinations of power, minClustersize, deepsplit and cutoff height; (D) Heatmap showing the gene module-traits correlation, FDR are represented as depth of color; (E) Cluster heatmap generated from expression profiles and clinical information of these 57 genes from the HMDB and HPA database. Gene expression (Log2FoldChange), >3, red <-2, blue. LV tissue expression, 0, white; >3.5, blue; HPA protein expression, 0, white; >3.5, purple. mRNA expression, 0, white; >3, green.

**Supplementary Figure S4** | siRNF181 is closely related to multiple biological processes in CAD. RNF181 PPI profiles from (A) STRING and (B) INTACT network. (C) Chord diagram constructed from siRNF181 data on the correlation between RNF181 downregulation and some cardiovascular process mediators or UPS molecules. Red denotes positive correlations, while the green, negative.

**Supplementary Figure S5** | Identification of CAD-associated risk variants rs1562322 at RNF181 locus in cohort and GWAS studies. (A) Scattered dot plot displaying the Pearson's correlation between the genes expression of RNF181 and the CAD-causal gene S100A6 in the CAD cohort from the CAD cohort of predict trail. N=99,  $R^2 = 0.78$ , p-value < 2.2E-16; (B) Identification of lead SNP rs1562322 in the CARDIoGRAMplusC4D and other 9 meta-analysis. (C) Color-coded forest plot showing the heart-tissue expression normalized effect sizes for rs1562322 and corresponding 95% confidence intervals are presented for CARDIoGRAMplusC4D and other 8 CAD meta-analysis. (D) m-values representing the posterior probability of heart tissue-specific expression and their respective expression P-values are also shown. (E) PheWAS showing the expression data from GWASATLAS for the lead GWA studies SNP rs1562322 associated with clinical traits. Genome-wide significant expression P-values are shown for indicated SNPs. P-value, Bonferroni corrected < 0.05.

- Baim, D. S., Flatley, M., Caputo, R., O'Shaughnessy, C., Low, R., Fanelli, C., et al. (2001). Comparison of PRE-dilatation vs Direct Stenting in Coronary Treatment Using the Medtronic AVE S670 Coronary Stent System (The PREDICT Trial). *Am. J. Cardiol.* 88, 1364–1369. doi:10.1016/s0002-9149(01)02114-2
- Bauer, R. C., Tohyama, J., Cui, J., Cheng, L., Yang, J., Zhang, X., et al. (2015). Knockout of *Adams7*, a Novel Coronary Artery Disease Locus in Humans, Reduces Atherosclerosis in Mice. *Circulation* 131, 1202–1213. doi:10.1161/circulationaha.114.012669

- Bedsaul, J. R., Carter, N. M., Deibel, K. E., Hutcherson, S. M., Jones, T. A., Wang, Z., et al. (2018). Mechanisms of Regulated and Dysregulated CARD11 Signaling in Adaptive Immunity and Disease. *Front. Immunol.* 9, 2105. doi:10.3389/fimmu.2018.02105
- Beineke, P., Fitch, K., Tao, H., Elashoff, M. R., Rosenberg, S., Kraus, W. E., et al. (2012). A Whole Blood Gene Expression-Based Signature for Smoking Status. *BMC Med. Genomics* 5, 58–59. doi:10.1186/1755-8794-5-58
- Botía, J. A., Vandrovčova, J., Forabosco, P., Guelfi, S., D'Sa, K., Hardy, J., et al. (2017). An Additional K-Means Clustering Step Improves the Biological Features of WGCNA Gene Co-expression Networks. *BMC Syst. Biol.* 11, 47–16. doi:10.1186/s12918-017-0420-6
- Brænne, I., Civelek, M., Vilne, B., Di Narzo, A., Johnson, A. D., Zhao, Y., et al. (2015). Prediction of Causal Candidate Genes in Coronary Artery Disease Loci. *Arterioscler Thromb. Vasc. Biol.* 35, 2207–2217. doi:10.1161/ATVBAHA.115.306108
- Cai, X. Y., Lu, L., Wang, Y. N., Jin, C., Zhang, R. Y., Zhang, Q., et al. (2011). Association of Increased S100B, S100A6 and S100P in Serum Levels with Acute Coronary Syndrome and Also with the Severity of Myocardial Infarction in Cardiac Tissue of Rat Models with Ischemia-Reperfusion Injury. *Atherosclerosis* 217, 536–542. doi:10.1016/j.atherosclerosis.2011.05.023
- Chan, K., Pu, X., Sandesara, P., Poston, R. N., Simpson, I. A., Quyyumi, A. A., et al. (2017). Genetic Variation at the ADAMTS7 Locus Is Associated with Reduced Severity of Coronary Artery Disease. *J. Am. Heart Assoc.* 6, e006928. doi:10.1161/JAHA.117.006928
- Chen, S. N., Czernuszewicz, G., Tan, Y., Lombardi, R., Jin, J., Willerson, J. T., et al. (2012). Human Molecular Genetic and Functional Studies Identify TRIM63, Encoding Muscle RING Finger Protein 1, as a Novel Gene for Human Hypertrophic Cardiomyopathy. *Circ. Res.* 111, 907–919. doi:10.1161/circresaha.112.270207
- Consortium, G. O. (2015). Gene Ontology Consortium: Going Forward. *Nucleic Acids Res.* 43, D1049–D1056. doi:10.1093/nar/gku1179
- Dai, R., Xia, Y., Liu, C., and Chen, C. (2018). csuWGCNA: a Combination of Signed and Unsigned WGCNA to Capture Negative Correlations. *bioRxiv*, 288225. doi:10.1101/288225
- Deloukas, P., Deloukas, P., Kanoni, S., Willenborg, C., Farrall, M., Assimes, T. L., et al. (2013). Large-scale Association Analysis Identifies New Risk Loci for Coronary Artery Disease. *Nat. Genet.* 45, 25–33. doi:10.1038/ng.2480
- Fall, T., Gustafsson, S., Orho-Melander, M., and Ingelsson, E. (2018). Genome-wide Association Study of Coronary Artery Disease Among Individuals with Diabetes: the UK Biobank. *Diabetologia* 61, 2174–2179. doi:10.1007/s00125-018-4686-z
- Feng, K., Liu, Y., Sun, J., Zhao, C., Duan, Y., Wang, W., et al. (2021). Compound Danshen Dripping Pill Inhibits Doxorubicin or Isoproterenol-Induced Cardiotoxicity. *Biomed. Pharmacother.* 138, 111531. doi:10.1016/j.biopha.2021.111531
- Fuster, V., Badimon, L., Badimon, J. J., and Chesebro, J. H. (1992). The Pathogenesis of Coronary Artery Disease and the Acute Coronary Syndromes (2). *N. Engl. J. Med.* 326, 310–318. doi:10.1056/NEJM199201303260506
- Genetics, M. I., Stitzel, N. O., Stirrups, K. E., Masca, N. G., Erdmann, J., Ferrario, P. G., et al. (2016). Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *N. Engl. J. Med.* 374, 1134–1144. doi:10.1056/NEJMoa1507652
- Ghorbel, M. T., Cherif, M., Mokhtari, A., Bruno, V. D., Caputo, M., and Angelini, G. D. (2010). Off-pump Coronary Artery Bypass Surgery Is Associated with Fewer Gene Expression Changes in the Human Myocardium in Comparison with On-Pump Surgery. *Physiol. genomics* 42, 67–75. doi:10.1152/physiolgenomics.00174.2009
- Giulietti, M., Righetti, A., Principato, G., and Piva, F. (2018). LncRNA Co-expression Network Analysis Reveals Novel Biomarkers for Pancreatic Cancer. *Carcinogenesis* 39, 1016–1025. doi:10.1093/carcin/bgy069
- Goh, K.-I., Cusick, M. E., Valle, D., Childs, B., Vidal, M., and Barabasi, A.-L. (2007). The Human Disease Network. *Proc. Natl. Acad. Sci.* 104, 8685–8690. doi:10.1073/pnas.0701361104
- Greenfest-Allen, E., Cartailier, J.-P., Magnuson, M. A., and Stoekert, C. J. (2017). iterativeWGCNA: Iterative Refinement to Improve Module Detection from WGCNA Co-expression Networks. *bioRxiv*, 234062.
- Han, B., and Eskin, E. (2011). Random-effects Model Aimed at Discovering Associations in Meta-Analysis of Genome-wide Association Studies. *Am. J. Hum. Genet.* 88, 586–598. doi:10.1016/j.ajhg.2011.04.014
- He, X. W., Wang, E., Bao, Y. Y., Wang, F., Zhu, M., Hu, X. F., et al. (2016). High Serum Levels of Sclerostin and Dickkopf-1 Are Associated with Acute Ischaemic Stroke. *Atherosclerosis* 253, 22–28. doi:10.1016/j.atherosclerosis.2016.08.003
- Horvath, S., and Dong, J. (2008). Geometric Interpretation of Gene Coexpression Network Analysis. *Plos Comput. Biol.* 4, e1000117. doi:10.1371/journal.pcbi.1000117
- Jiao, M., Li, J., Zhang, Q., Xu, X., Li, R., Dong, P., et al. (2020). Identification of Four Potential Biomarkers Associated with Coronary Artery Disease in Non-diabetic Patients by Gene Co-expression Network Analysis. *Front. Genet.* 11, 542. doi:10.3389/fgene.2020.00542
- Joseph, P., Leong, D., McKee, M., Anand, S. S., Schwalm, J.-D., Teo, K., et al. (2017). Reducing the Global burden of Cardiovascular Disease, Part 1: the Epidemiology and Risk Factors. *Circ. Res.* 121, 677–694. doi:10.1161/circresaha.117.308903
- Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30. doi:10.1093/nar/28.1.27
- Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a Reference Resource for Gene and Protein Annotation. *Nucleic Acids Res.* 44, D457–D462. doi:10.1093/nar/gkv1070
- Karamariti, E., Zhai, C., Yu, B., Qiao, L., Wang, Z., Potter, C. M. F., et al. (2018). DKK3 (Dickkopf 3) Alters Atherosclerotic Plaque Phenotype Involving Vascular Progenitor and Fibroblast Differentiation into Smooth Muscle Cells. *Arterioscler Thromb. Vasc. Biol.* 38, 425–437. doi:10.1161/ATVBAHA.117.310079
- Khera, A. V., Emdin, C. A., Drake, I., Natarajan, P., Bick, A. G., Cook, N. R., et al. (2016). Genetic Risk, Adherence to a Healthy Lifestyle, and Coronary Disease. *N. Engl. J. Med.* 375, 2349–2358. doi:10.1056/nejmoa1605086
- Kim, J., Ghasemzadeh, N., Eapen, D. J., Chung, N. C., Storey, J. D., Quyyumi, A. A., et al. (2014). Gene Expression Profiles Associated with Acute Myocardial Infarction and Risk of Cardiovascular Death. *Genome Med.* 6, 40–13. doi:10.1186/gm560
- Kuznetsova, T., Prange, K. H. M., Glass, C. K., and De Winther, M. P. J. (2020). Transcriptional and Epigenetic Regulation of Macrophages in Atherosclerosis. *Nat. Rev. Cardiol.* 17, 216–228. doi:10.1038/s41569-019-0265-3
- Langfelder, P., and Horvath, S. (2014). *Tutorials for the WGCNA Package*. Los Angeles: UCLA.
- Larson, M. G., Atwood, L. D., Benjamin, E. J., Cupples, L. A., D'Agostino, R. B., Fox, C. S., et al. (2007). Framingham Heart Study 100K Project: Genome-wide Associations for Cardiovascular Disease Outcomes. *BMC Med. Genet.* 8 Suppl 1, S5–S9. doi:10.1186/1471-2350-8-S1-S5
- Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., and Storey, J. D. (2012). The Sva Package for Removing Batch Effects and Other Unwanted Variation in High-Throughput Experiments. *Bioinformatics* 28, 882–883. doi:10.1093/bioinformatics/bts034
- Ma, C., Zhang, W., Yang, X., Liu, Y., Liu, L., Feng, K., et al. (2018). Functional Interplay between Liver X Receptor and AMP-Activated Protein Kinase  $\alpha$  Inhibits Atherosclerosis in Apolipoprotein E-Deficient Mice – a New Anti-atherogenic Strategy. *Br. J. Pharmacol.* 175, 1486–1503. doi:10.1111/bph.14156
- Marenberg, M. E., Risch, N., Berkman, L. F., Floderus, B., and de Faire, U. (1994). Genetic Susceptibility to Death from Coronary Heart Disease in a Study of Twins. *N. Engl. J. Med.* 330, 1041–1046. doi:10.1056/nejm199404143301503
- Mofid, A., Newman, N. S., Lee, P. J., Abbasi, C., Matkar, P. N., Rudenko, D., et al. (2017). Cardiac Overexpression of S100A6 Attenuates Cardiomyocyte Apoptosis and Reduces Infarct Size after Myocardial Ischemia-Reperfusion. *J. Am. Heart Assoc.* 6, e004738. doi:10.1161/JAHA.116.004738
- Naghavi, M., Abajobir, A. A., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abera, S. F., et al. (2017). Global, Regional, and National Age-Sex Specific Mortality for 264 Causes of Death, 1980–2016: a Systematic Analysis for the Global Burden of Disease Study 2016. *Lancet* 390, 1151–1210. doi:10.1016/S0140-6736(17)32152-9
- Nelson, C. P., Goel, A., Goel, A., Butterworth, A. S., Kanoni, S., Webb, T. R., et al. (2017). Association Analyses Based on False Discovery Rate Implicate New Loci for Coronary Artery Disease. *Nat. Genet.* 49, 1385–1391. doi:10.1038/ng.3913

- 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
- Ntalla, I., Kanoni, S., Zeng, L., Giannakopoulou, O., Danesh, J., Watkins, H., et al. (2019). Genetic Risk Score for Coronary Disease Identifies Predispositions to Cardiovascular and Noncardiovascular Diseases. *J. Am. Coll. Cardiol.* 73, 2932–2942. doi:10.1016/j.jacc.2019.03.512
- Pagan, J., Seto, T., Pagano, M., and Cittadini, A. (2013). Role of the Ubiquitin Proteasome System in the Heart. *Circ. Res.* 112, 1046–1058. doi:10.1161/circresaha.112.300521
- Peden, J. F., Hopewell, J. C., Saleheen, D., Chambers, J. C., Hager, J., Soranzo, N., et al. (2011). A Genome-wide Association Study in Europeans and South Asians Identifies Five New Loci for Coronary Artery Disease. *Nat. Genet.* 43, 339–344. doi:10.1038/ng.782
- Ponsuksili, S., Du, Y., Hadlich, F., Siengdee, P., Murani, E., Schwerin, M., et al. (2013). Correlated mRNAs and miRNAs from Co-expression and Regulatory Networks Affect Porcine Muscle and Finally Meat Properties. *BMC genomics* 14, 1–17. doi:10.1186/1471-2164-14-533
- Robinson, J. G., Farnier, M., Krempf, M., Bergeron, J., Luc, G., Averna, M., et al. (2015). Efficacy and Safety of Alirocumab in Reducing Lipids and Cardiovascular Events. *N. Engl. J. Med.* 372, 1489–1499. doi:10.1056/nejmoa1501031
- Sabatine, M. S., Giugliano, R. P., Wiviott, S. D., Raal, F. J., Blom, D. J., Robinson, J., et al. (2015). Efficacy and Safety of Evolocumab in Reducing Lipids and Cardiovascular Events. *N. Engl. J. Med.* 372, 1500–1509. doi:10.1056/nejmoa1500858
- Schunkert, H., König, I. R., Kathiresan, S., Reilly, M. P., Assimes, T. L., Holm, H., et al. (2011). Large-scale Association Analysis Identifies 13 New Susceptibility Loci for Coronary Artery Disease. *Nat. Genet.* 43, 333–338. doi:10.1038/ng.784
- Siewert, K. M., and Voight, B. F. (2018). Bivariate Genome-wide Association Scan Identifies 6 Novel Loci Associated with Lipid Levels and Coronary Artery Disease. *Circ. Genom. Precis. Med.* 11, e002239. doi:10.1161/CIRCGEN.118.002239
- Smyth, G. K. (2005). *Limma: Linear Models for Microarray Data. Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Berlin: Springer, 397–420.
- Tamamori-Adachi, M., Hayashida, K., Nobori, K., Omizu, C., Yamada, K., Sakamoto, N., et al. (2004). Down-regulation of p27Kip1 Promotes Cell Proliferation of Rat Neonatal Cardiomyocytes Induced by Nuclear Expression of Cyclin D1 and CDK4: Evidence for Impaired Skp2-dependent Degradation of P27 in Terminal Differentiation. *J. Biol. Chem.* 279, 50429–50436. doi:10.1074/jbc.m403084200
- Virani, S. S., Alonso, A., Aparicio, H. J., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., et al. (2021). Heart Disease and Stroke Statistics-2021 Update: A Report from the American Heart Association. *Circulation* 143, e254–e743. doi:10.1161/CIR.0000000000000950
- Wang, S., Wang, X., Gao, Y., Peng, Y., Dong, N., Xie, Q., et al. (2019). RN181 Is a Tumour Suppressor in Gastric Cancer by Regulation of the ERK/MAPK-cyclin D1/CDK4 Pathway. *J. Pathol.* 248, 204–216. doi:10.1002/path.5246
- Wang, T., He, X., Liu, X., Liu, Y., Zhang, W., Huang, Q., et al. (2017). Weighted Gene Co-expression Network Analysis Identifies FKBP11 as a Key Regulator in Acute Aortic Dissection through a NF- $\kappa$ B Dependent Pathway. *Front. Physiol.* 8, 1010. doi:10.3389/fphys.2017.01010
- Won, H.-H., Natarajan, P., Dobbyn, A., Jordan, D. M., Roussos, P., Lage, K., et al. (2015). Disproportionate Contributions of Select Genomic Compartments and Cell Types to Genetic Risk for Coronary Artery Disease. *Plos Genet.* 11, e1005622. doi:10.1371/journal.pgen.1005622
- Wu, Y.-J., Bond, M., Sala-Newby, G. B., and Newby, A. C. (2006). Altered S-phase Kinase-Associated Protein-2 Levels Are a Major Mediator of Cyclic Nucleotide-Induced Inhibition of Vascular Smooth Muscle Cell Proliferation. *Circ. Res.* 98, 1141–1150. doi:10.1161/01.res.0000219905.16312.28
- Xiong, C. (2017). RNF181 and TPX2 Expression in colon Cancer Tissue and Their Correlation with Cancer Cell Viability and Angiogenesis. *J. Hainan Med. Univ.* 23, 95–98.
- Xu, N., Gulick, J., Osinska, H., Yu, Y., McLendon, P. M., Shay-Winkler, K., et al. (2020). Ube2v1 Positively Regulates Protein Aggregation by Modulating Ubiquitin Proteasome System Performance Partially through K63 Ubiquitination. *Circ. Res.* 126, 907–922. doi:10.1161/circresaha.119.316444
- Xu, S., Xu, Y., Liu, P., Zhang, S., Liu, H., Slavina, S., et al. (2019). The Novel Coronary Artery Disease Risk Gene JCAD/KIAA1462 Promotes Endothelial Dysfunction and Atherosclerosis. *Eur. Heart J.* 40, 2398–2408. doi:10.1093/eurheartj/ehz303
- Yahagi, K., Kolodgie, F. D., Otsuka, F., Finn, A. V., Davis, H. R., Joner, M., et al. (2016). Pathophysiology of Native Coronary, Vein Graft, and In-Stent Atherosclerosis. *Nat. Rev. Cardiol.* 13, 79–98. doi:10.1038/nrcardio.2015.164
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omic: a J. Integr. Biol.* 16, 284–287. doi:10.1089/omi.2011.0118
- Zhang, B., and Horvath, S. (2005). A General Framework for Weighted Gene Co-expression Network Analysis. *Stat. Appl. Genet. Mol. Biol.* 4. doi:10.2202/1544-6115.1128
- Zheng, X., Xue, C., Luo, G., Hu, Y., Luo, W., and Sun, X. (2015). Identification of Crucial Genes in Intracranial Aneurysm Based on Weighted Gene Coexpression Network Analysis. *Cancer Gene Ther.* 22, 238–245. doi:10.1038/cgt.2015.10
- Zhou, R., Ding, Y., Xue, M., Xiong, B., and Zhuang, T. (2020). RNF181 Modulates Hippo Signaling and Triple Negative Breast Cancer Progression. *Cancer Cell Int* 20, 1–11. doi:10.1186/s12935-020-01397-3
- Zhou, X.-g., Huang, X.-l., Liang, S.-y., Tang, S.-m., Wu, S.-k., Huang, T.-t., et al. (2018). Identifying miRNA and Gene Modules of colon Cancer Associated with Pathological Stage by Weighted Gene Co-expression Network Analysis. *Ott* 11, 2815–2830. doi:10.2147/ott.s163891
- Zhu, J., Li, X., Su, P., Xue, M., Zang, Y., and Ding, Y. (2020). The Ubiquitin Ligase RNF181 Stabilizes ER $\alpha$  and Modulates Breast Cancer Progression. *Oncogene* 39, 6776–6788. doi:10.1038/s41388-020-01464-z

**Conflict of Interest:** KG, HS, WW, and YH were employed by GeneNet Pharmaceuticals Co., Ltd. SZ was employed by Tasly Holding Group Co., Ltd.

The remaining 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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Dang, Qu, Guo, Zhou, Sun, Wang, Han, Feng, Lin and Hu. 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.