

# Combined Analysis of ceRNA Regulatory Network and DNA Methylation in CAD

**Yue Zhao**

wu han da xue zhong nan yi yuan: Wuhan University Zhongnan Hospital

**Chen Wang**

wu han da xue zhong nan yi yuan: Wuhan University Zhongnan Hospital

**Wangxia Li**

Wuhan Blood Center

**Bingyu Jin**

wu han da xue zhong nan yi yuan: Wuhan University Zhongnan Hospital

**Yang Xiang**

wu han da xue zhong nan yi yuan: Wuhan University Zhongnan Hospital

**Bin Liang**

wu han da xue zhong nan yi yuan: Wuhan University Zhongnan Hospital

**Xiaokang Zhang**

wu han da xue zhong nan yi yuan: Wuhan University Zhongnan Hospital

**Zhibing Lu**

wu han da xue zhong nan yi yuan: Wuhan University Zhongnan Hospital

**Fang Zheng** (✉ [zhengfang@whu.edu.cn](mailto:zhengfang@whu.edu.cn))

wu han da xue zhong nan yi yuan: Wuhan University Zhongnan Hospital <https://orcid.org/0000-0002-2024-2424>

---

## Research article

**Keywords:** coronary artery disease, competitive endogenous RNAs, DNA methylation, differentially expressed gene, AGPAT4, FAM169A

**Posted Date:** September 30th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-917459/v1>

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

[Read Full License](#)

---

# Abstract

## Background

The morbidity and mortality of coronary artery disease (CAD) is increasing year by year. Hence it is urgent to probe into the molecular mechanism of CAD and seek new therapeutic strategies. The purpose of our study was to screen genes associated with the development of CAD by using bioinformatics tools and clinical samples.

## Methods

Microarray datasets from the Gene Expression Omnibus (GEO) database of peripheral blood cells (PBLs) were chosen for this study, and candidate differentially expressed microRNAs (DEMs) were screened using the limma and weighted co-expression network analysis (WGCNA) packages in R (v4.0). Subsequently, we construct a competitive endogenous RNAs (ceRNA) network and perform enrichment analysis of genes in the network. Meanwhile, differentially methylated genes (DMGs) in PBLs were identified using the "ChAMP" package in a DNA methylation chip. We then constructed the methylation-associated ceRNA network in CAD. Eventually, the methylation levels of genes and the relationship with the expression of genes in ceRNA were validated in PBLs samples using the Illumina Methylation 850K chip and transcriptome sequencing, while gene expressions were verified by qRT-PCR. And the regulation of DNA methylation on gene expression was verified in the THP-1 cells treated with 5-Aza-2'-deoxycytidine (5-AZA).

## Results

A total of 71 differentially expressed miRNAs were screened by both WGCNA and limma. Then the ceRNA network in CAD was constructed with 269 nodes and 705 edges, which were significantly enriched in the chemokine-mediated signaling pathway and so on. Furthermore, from 4354 identified DMGs in a methylation data, 34 methylation-associated differentially expressed genes (DEGs) and 1 differentially expressed lncRNA (DEL) were obtained. After verification of methylation experiments in study population A, three genes were found to have altered methylation consistent with the bioinformatics results. And these genes were correlated in terms of methylation and expression levels. Corresponding with the bioinformatics results, qRT-PCR results in validation set B also showed that the expression of AGPAT4 and FAM169A were significantly lower in CAD. In addition, 5-AZA treatment could increase the expression of AGPAT4 and FAM169A in THP-1 cells.

## Conclusions

Our study deepens the understanding of the molecular mechanisms underlying the pathogenesis of CAD and provides new ideas for its treatment.

# Introduction

Coronary artery atherosclerotic heart disease (CAD) is a type of heart disease caused by atherosclerosis of the coronary arteries narrowing or blocking the lumen, thereby causing myocardial ischemia, hypoxia, and is also referred to as coronary heart disease. Cardiovascular diseases (CVD), are the leading cause of all global deaths, with CAD accounting for 32.7% of the global CVD burden and their severe deprivation of patients' quality of life and heavy burden on the social and medical economy (Roth et al., 2017; Bauersachs et al., 2019). Therefore, it is necessary to study the pathogenesis of CAD to find effective treatment options to improve clinical management.

Recently, epigenetic elements, as the emerging risk factors for CAD, including long non-coding RNAs (lncRNAs), microRNAs (miRNAs) and DNA methylation have gotten more and more attention. (Archer et al., 2015). And the ceRNA hypothesis elaborates the non-coding RNAs and mRNAs regulatory mechanism, that lncRNAs act as miRNAs molecular sponges and compete with miRNAs at the post-transcriptional level to regulate the transcriptional expression of mRNAs (Salmena et al., 2011; Karreth et al., 2015; Li et al., 2017). With the deepening of the study, mounting evidence shows that ceRNA regulatory networks play an essential role in the development of CAD (Song et al., 2016), so does the DNA methylation. Our previous study has demonstrated that the global hypomethylation status of blood cells, defined predominantly by monocyte DNA hypomethylation, is independently associated with the risk of CAD (Deng et al., 2018). Thus, we speculate that a combined analysis of ceRNAs and DNA methylation associated with gene expressions could provide new insights to the pathogenesis of CAD. Our study combined ceRNA networks with methylation profiles and performed regression analysis to find potential markers in CAD. This may provide new clues to further understand the pathogenesis of CAD and may offer a new theoretical basis for targeted therapy in CAD.

## Materials And Methods

### Microarray data

The miRNA microchip dataset GSE31568, mRNA and lncRNA microchip dataset GSE113079 and methylation microchip dataset GSE107143 were downloaded from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/query/>) for bioinformatic analysis. The GSE31568 expression profile contained peripheral blood profiles of healthy control individuals and patients of 14 different diseases (different cancer types, autoimmune diseases, cardiovascular, and chronic inflammatory diseases). In our study, we chose healthy subjects and patients with acute myocardial infarction. Data were analyzed by the GPL9040 febit Homo Sapiens miRBase 13.0 platform. GSE113079 dataset contained PBLs from patients with CAD and healthy controls. Data was performed using the platform GPL20115 Agilent-067406 Human CBC lncRNA+mRNA microarray V4.0. And GSE107143, whole-genome DNA methylation analysis array contains DNA methylation information from blood samples from clinical atherosclerosis patients and healthy donors. This dataset was analyzed by the platform GPL13534 Illumina HumanMethylation450 BeadChip platform analyzed data. The downloaded data were pre-processed with R software for normalization and summarization, respectively.

## Differential Expression Analysis

The limma package was used to analyze the differences in the processed expression data (Ritchie et al., 2015), in which  $|\logFC| > 0.9986$  and  $p$  values  $< 0.05$  was set as the threshold for screening differentially expressed genes (DEGs);  $|\logFC| > 0.85$  and  $p$  values  $< 0.05$  was set as the threshold for screening differentially expressed miRNAs (DEMs);  $|\logFC| > 1.112$  and  $p$  values  $< 0.05$  were set as the cut-off points for screening differentially expressed

lncRNAs (DELs) (**Additional file 1**). The logFC-cutoff was calculated using the formula:

$$\mathbf{logFC\_cutoff} = \mathbf{with(DEG, mean(abs(logFC)) + 2 * sd(abs(logFC)) )}$$

## WGCNA Analysis

The WGCNA package in R exploits soft thresholds to construct scale-free gene co-expression networks, and it is often used to perform bioinformatics analysis (Langfelder and Horvath, 2008). This package is used to identify the modules that are highly associated with the disease according to the gene expression level in the data sets GSE31568 and GSE113079. The pickSoftThreshold function was used to determine the appropriate soft threshold ( $\beta$ ) to obtain an approximate scale-free network distribution with a fit index  $R^2 > 0.8$ . The blockwiseModules function was used for one-step network construction and module detection to generate the co-expressed gene module and topological overlap matrix (TOM) with a minimum module size of 30 and a combined cut height of 0.25. Principal component analysis was performed for each module, and the eigenvalues (module eigengenes, MEs) of the gene modules were calculated using the first central component. The correlation coefficients between MEs and each clinical phenotype were calculated by introducing the above clinical phenotypes, and the gene modules that were significantly associated with CAD were identified. The gene significance (GS) and module membership (MM) were analyzed within modules to screen key genes (**Additional file 1**). The screened essential genes were selected as the significant differential genes by taking the intersection with the respective differential genes using Venn plots.

## CeRNA regulatory network construction

The previous DEMs filtered by the WGCNA were identified as the candidate DEMs by drawing Venn diagrams. The RNA22 online ([miRNA Target Visualization \(jefferson.edu\)](http://miRNA.TargetVisualization.jefferson.edu)) prediction tool can perform miRNA target prediction with high reliability by identifying miRNA binding sites and corresponding heterodimers, and was used to predict the target lncRNAs of the candidate DEMs. MultiMiR (Ru et al., 2014), an R package that can be used to predict interactions between miRNA-target genes, can simultaneously query 11 available public databases, namely validated (miRecords, miRTarBase, TarBase databases) and predicted (DIANA-microT, EIMMo, MicroCosm, miRanda, miRDB, PicTar, PITA, TargetScan databases). The multiMiR package was used to predict the target mRNAs of the candidate DEMs. The predicted results were intersected with the DELs and DEGs, and filtered by WGCNA, respectively. Then the

lncRNA-miRNA and miRNA-mRNA interactions were integrated, and the ceRNA network was constructed (**Additional file 2**) and used with Cytoscape (v3.5.1) for visualization.

## GO and KEGG Pathway Enrichment Analyses

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses for the coregulated genes in the ceRNA network were performed using ClusterProfiler and enrichplot. Three aspects of GO terms were used, including cellular compartment (CC), biological process (BP) and molecular function (MF). Statistical significance of GO terms and KEGG pathways were set at  $p$  values  $<0.05$ .

## Methylation analysis of ceRNA

The DNA methylation dataset GSE107143 was downloaded from the GEO database and imported into R. Differentially methylated genes (DMGs) were obtained using the Chip Analysis Methylation Pipeline (ChAMP) package with cut-off values taken as  $|\log_{2}FC| > 0.0514$  and  $p$  value  $< 0.05$  (**Additional file 3**). The genes in DMGs and ceRNA were then overlapped and the intersection set was taken.

## Study populations

A total of 71 individuals, including 34 patients with CAD and 37 control subjects, were recruited at Zhongnan Hospital of Wuhan University (Hubei, China) from July 2020 to April 2021, which were divided to validation population A and validation population B. DNA methylation chip analysis and transcriptome sequencing were performed in validation population A of 12 CAD cases and 12 sex-age matched normal controls with 10 mL peripheral blood samples. The RT-qPCR detection was carried on in validation population B of 22 CAD cases and 25 sex-age matched normal controls with 1 ml peripheral blood samples. This study was authorized by the Medical Ethics Committee of Zhongnan Hospital of Wuhan University.

## Methylation analysis and transcriptome sequencing in validation population A

Total DNA and RNA was extracted from PBLs of validation population A, and the concentration and purity of DNA and RNA were determined by NanoDrop 2000C (Thermo Scientific, USA). DNA methylation analysis was performed via the Infinium Methylation EPIC BeadChips (Illumina, San Diego, CA, USA). The general procedure is as follows: DNA samples were first converted to DNA sulfite using the EZ DNA Methylation Kit (Illumina, USA) optimization method, followed by amplification - break - precipitation - re-suspension, and hybridization - wash - extension - staining scan on an Illumina Methylation 850K chip. The DNA methylation signal was obtained by washing- extension- staining scan on Illumina Methylation 850K chip. Signal scans were performed using iScan software (Illumina, USA) to get raw signal values for each gene locus. The raw signal values were normalized using Genome Studio software (Illumina, USA) to obtain genome-wide methylation profiles, which were then screened for DMGs by the ChAMP in R (3.6.0) (**Additional file 3**). Transcriptome sequencing analysis was performed by first creating libraries

from RNA samples, and subsequently qualified libraries were sequenced using a PacBio Sequel platform (Pacific Biosciences, USA) to obtain transcriptome sequencing data. (**Additional file 3**)

### **Verification of differential expression of hub genes in validation population B and validation on DNA methylation regulation in THP-1 cells**

To further verify the differently expression of hub genes in clinical samples, RNA expressions of hub genes were detected using RT-qPCR in validation population B. Subsequently, we further validated the expression of hub genes regulated by DNA methylation at the cellular level. THP-1 cells in six-well plates were treated with inhibitor 5-AZA (0 $\mu$ mol/L, 1 $\mu$ mol/L, 3 $\mu$ mol/L, 4 $\mu$ mol/L, 5 $\mu$ mol/L). The medium containing 5-AZA (Sigma, USA) was changed every day, and the cells were collected after 72 hours. Total RNA was extracted from PBLs and THP-1 cells using TRIzol reagent (Invitrogen, USA), and the concentration and purity of RNA were determined using NanoDrop 2000C. The RNA was reverse transcribed into cDNA by PrimeScript<sup>TM</sup> RT reagent Kit with gDNA Remover (Takara, Japan). Real-time fluorescent quantitative PCR (qRT-PCR) was performed by using SYBR Green I UltraSYBR mix (Cwbio, China) kit on Bio-Rad CFX96 (BioRad, USA). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as the endogenous reference gene, and each primer pair was synthesized at Tsingke Biotechnology after was designed on primer 3.0 software. The specific sequences of each primer pair can be found in **Additional file 4**. All experiments were performed three times. The relative expression of genes was calculated using the  $2^{\Delta\Delta Cq}$  method, where  $\Delta Cq$  indicates the difference between the average quantization period (CQ) of the target gene and the endogenous reference gene (GAPDH).

## **Results**

### **Identifying the aberrantly expressed genes in CAD**

The data analysis workflow was displayed in **Fig. 1**. A total of 98 DEMs were screened using the LIMMA package, of which 85 were up-regulated, and 13 were down-regulated in the CAD group (**Fig. 2A**). A total of 703 DEGs were screened, of which 452 were expressed up-regulated and 251 were expressed down-regulated in the CAD group (**Fig. 2B**). There were 474 DELs, of which 335 were up-regulated, and 139 were down-regulated in the CAD group (**Fig. 2C**). The top 10 up-regulated and down-regulated lncRNAs, miRNAs and mRNAs were shown in **Fig. 2D**.

### **WGCNA construction and key modules identification**

Regarding the WGCNA, we first selected the appropriate soft thresholds according to the scale-free topology criterion. The appropriate soft thresholds of miRNA, mRNA and lncRNA were 4, 5 and 6 (**Fig. 3A-C**). After calculating the eigenvalues of the module, several respective individual modules in miRNA, mRNA and lncRNAs were generated (**Fig. 3D-F**). By correlation analysis of CAD features, modules and gene expression profiles, we finally confirmed the meaningful module, which has the lowest *p* value and highest correlation index in these three datasets. The most significant modules of miRNA, mRNA and lncRNA were highlighted in blue, turquoise and turquoise respectively, and were selected for the next

analysis (**Fig. 3G-I**). Then, we took a ceRNA network construction approach with core miRNAs to predict their target mRNAs and lncRNAs.

First, we identified the candidate DEMs based on the intersection between DEMs obtained from the limma package and the miRNAs in the most significant module in WGCNA, and then the result was presented in Venn diagrams (**Fig. 3J**).

The DEGs, DELs in the most significant module in WGCNA were stored for the next analysis.

### **CeRNA network construction**

After the name transformation of candidate DEMs in the miRBase online database, 7 DEMs were excluded. Targeted lncRNAs predicted by the RNA22 online database were subsequently intersected with previous DELs and lncRNAs in most significant module of WGCNA, and the identified candidate DELs were presented in a Wayne diagram (**Fig. 4A**). According to ceRNA theory, miRNA-lncRNA interaction pairs, including 50 candidate DEMs and 78 candidate DELs, were obtained (**Additional file 2**), and the expression trends of these DELs and DEMs were reversed. Meanwhile, the target mRNAs of candidate DEMs were predicted using the "multiMiR" package. Then, the candidate DEGs were obtained based on the intersection between DEGs in Fig. 2B, genes in the most significant module of WGCNA and the predicted DEGs, and were presented by drawing Venn diagrams. (**Fig. 4B**). Eventually, the final miRNA-mRNA interaction pairs, including 52 candidate miRNAs and 144 candidate mRNAs, were obtained (**Additional file 2**). These miRNAs and mRNAs expression trends were also opposite. These interaction pairs between lncRNAs and miRNAs and between miRNAs and mRNAs were used to construct a ceRNA network with 269 nodes and 705 edges (78 key lncRNAs, 50 key miRNAs and 141 key mRNAs) (**Fig. 4C,4D**). Green dots represent lncRNA, blue dots represent mRNA, and red dots represent miRNA. Positive triangles represent up-regulated gene expression, and inverted triangles represent down-regulation.

### **GO and KEGG functional enrichment analysis**

ClusterProfiler and enrichplot were used to predict the intrinsic linkage and potential functions of key DEGs in ceRNA. GO analysis results showed key genes largely participated in GO:0070098 (chemokine-mediated signaling pathway) and GO:0032211 (negative regulation of telomere maintenance via telomerase) ( $p$  values  $<0.05$ ) (**Fig. 5A-C**). KEGG analysis results showed key genes largely participated in hsa04137 (Mitophagy) ( $p$  values  $<0.05$ ) (**Fig. 5D**). All the results of the analysis can be observed in **Additional file 5**.

### **Methylated related ceRNA network**

Differential methylation analysis of GSE107143 obtained 4354 DMGs. The key mRNAs, lncRNAs and miRNAs from the ceRNA network were combined with DMGs to take intersections (**Fig. 6A**). There were 34 mRNAs, 1 lncRNA and no miRNAs from ceRNA network overlapped with DMGs. Therefore, we constructed another ceRNA network related to DNA methylation based on this result and performed functional enrichment analysis of the mRNAs in the network. These genes were mainly concentrated in cell

proliferation and transcription from RNA polymerase II promoter (**Fig. 6B**). The methylation-related ceRNA network is exhibited in Sankey Diagram, and it included key 30 miRNAs, 36 mRNAs and 69 lncRNAs related to differential DNA methylation (**Fig. 6C**).

### Validation of hub genes

To verify the results of bioinformatics analysis, we carried out differential methylation analysis and transcriptome sequencing in validation population A to identify hub genes in methylation-related ceRNA. There were 3149 DMGs were screened by differential methylation analysis with  $|\log_{2}FC| \geq 0.03$  and  $p$  values  $< 0.05$  as the standard. There were 1656 DEMRNAs were screened by transcriptome sequencing with  $|\log_{2}FC| \geq 1.19$  and  $p$  values  $< 0.05$  as the standard. After intersection with the methylation-related key genes by the "Venn" package (**Fig. 7A**), a total of 3 hub overlapping genes were found. Correlation analysis showed the DNA methylation of AGPAT4 was negatively related with the expression level (**Fig. 7B**). Positive correlation between CXCR1 and FAM169A methylation and expression were demonstrated in **Fig. 7C** and **Fig. 7D**. RT-qPCR results of hub gene expression further confirmed the differently expression of 3 hub genes in CAD. The expressions of AGPAT4 and FAM169A were found to be significantly lower in CAD cases than controls, however, there was no significant difference in the expression of CXCR1 (**Fig. 7E**). Furthermore, AGPAT4 and FAM169 expressions were significantly upregulated in THP-1 cells after 5-AZA treatment (**Fig. 7F**).

## Discussion

Presently, people pay more and more attention to CAD, since the morbidity and mortality of cardiovascular diseases are still very high, which has become a big medical burden all over the world (Benjamin et al., 2018). Therefore, it is imperative to study and understand the pathogenesis in-depth to find early prediction, effective diagnostic and therapeutic targets.

Coronary angiography (CAG), conventional electrocardiogram (ECG) and dynamic electrocardiogram (DCG) are common clinical methods for CAD diagnosis, among which CAG is the gold diagnostic standard. Though CAG is the standard, it is invasive and its detection stage is almost the end of disease development (Gaede et al., 2019), and it is mostly used for the examination of patients with suspected CAD symptoms. In addition, both ECG and DCG have their own drawbacks for the accurate diagnosis of CAD. As we know, the early diagnosis of coronary atherosclerosis lesions and the establishment of an effective treatment plan has positive implications for improving the prognosis, but because of the low significance of non-invasive cardiac examinations in the early stages of CAD and the completely unavoidable dangers of invasive diagnostic means, the new non-invasive method is needed for early detection with high specificity and sensitivity in the initial stages of CAD.

The rapid growth of biological data and the rapid development of high-throughput sequencing technologies have provided a good platform for studying CAD to further elucidate the pathogenesis. Studies have shown that CAD is a complex disease caused by multiple risk factors such as lifestyle, environment and genetics, among which genetic factors play an important role in its pathogenesis (Li et

al., 2020). In recent years, there has been an increasing number of studies on CAD genetics. In 2016, Dai et al. identified three potential causal genes and their responsible mutations associated with Mendelian autosomal dominant CAD and MI in their study, and functional analysis of ST6GALNAC5, CYP27A1 and MEF2A also supported their possible involvement in CAD and MI pathogenesis of the disease (Dai et al., 2016). Epigenetic alterations, mainly at the level of DNA methylation modification, ncRNA regulation and histone modification, as a "bridge" between environmental and genetic factors, it reflects the link between environment and genes and is involved in the development of CAD (Duan et al., 2013; van Diepen et al., 2016). Now, an increasing number of studies have demonstrated that ceRNAs are involved in CAD and of great value in the diagnosis and prediction of the disease (Boon et al., 2016; Liang et al., 2020). Meanwhile, it has been shown (Hai and Zuo, 2016; Sumi et al., 2021) that alterations in genome-wide methylation status and specific genes with DNA methylation sites are strongly associated with CAD development. However, the relationship between ceRNA and DNA methylation in CAD was seldom explored and the role of a single gene or a single pathway has been very limited because the mechanisms of CAD occurrence and development are too intricate. Therefore, the construction of ceRNA networks and analysis of corresponding methylation variations could probably provide more targeted areas for new perspectives. In this study, we started from the GEO database and firstly performed bioinformatics analysis of differentially expressed ceRNAs in CAD, before analyzed their alterations in methylation, and to construct the DNA methylation related ceRNA network, and finally identified the potential functional genes for CAD in the clinical samples and at the cellular level.

Firstly, we downloaded and analyzed datasets from the GEO database to screen DEGs, DELs and DEMs. Then, a lncRNA-miRNA-mRNA ternary network was constructed using these data, WGCNA results, the RNA22 database and the multiMiR package based on ceRNA theory. Based on GO and KEGG enrichment analysis results, we found that the key DMGs were related to chemokine-mediated signaling pathway, negative regulation of telomere maintenance via telomerase, mitophagy and so on. Subsequently, key DMGs were screened in the methylation data and co-constructed with ceRNAs to construct a methylation-related ceRNA network. The enrichment analysis showed that the key DMGs associated with methylation were mainly concentrated in cell proliferation and transcription from RNA polymerase II promoter. Then, the methylation regulation and differential expression of genes involved in the ceRNA network were verified, two hub genes were identified: AGPAT4 and FAM169A.

In our study, AGPAT4 was reduced and hypermethylated, while FAM169A was downregulated and hypomethylated in CAD. However, in THP1 cells, the expression of both AGPAT4 and FAM169A was elevated after demethylation treatment. The reason for the different relationship between FAM169A methylation and expression in CAD and THP1 cells may be that in CAD we studied PBLs, whereas THP1 cells are a human monocyte cell line. And, no related studies have so far focused on the gene methylation of both. The acylglycerol phosphate acyltransferases/lysophosphatidic acid acyltransferases (AGPATs/LPAATs) family is a lipid-modifying enzyme that produces phosphatidic acid (PA) with lysophosphatidic acid (LPA) and acyl-coenzyme A as substrates, and is also the rate-limiting enzyme in TG synthesis. This might be the reason for that AGPAT4 affects endogenous TG storage and TG metabolism and SNPs in the *AGPAT3* and *AGPAT4* are associated with plasma TG levels(Ouellette et al.,

2013; Yamashita et al., 2014). Eleven AGPAT/LPAAT enzymes have been identified in mice and humans (Zhukovsky et al., 2019). It has been shown that each AGPAT enzyme mediates specific cellular and organelle membrane lipid composition, which in turn affects their physiological functions, and therefore, functional and biochemical alterations in specific AGPATs cannot be replaced by other AGPATs (Bradley et al., 2018). In addition, AGPAT4 stimulates M1-dependent T cell activation and tumor suppression by regulating LPA production in colorectal cancer (CRC) cells (Zhang et al., 2020). All these results from reports on AGPAT4 could probably explained the relevance between AGPAT4 and CAD.

FAM169A, also known as SLAP75 (Soluble Lamina-Associated Protein of 75 kDa), was originally identified as a new nuclear envelope (NE)-associated protein in laminin A (LaA) by the biotin identification (BioID) technique (Roux et al., 2012). Subsequent analysis by et al. showed that FAM169A is highly conserved in vertebrates and is directly or indirectly linked to higher order chromatin organization (Maric et al., 2017). Moreover, DNA methylation can cause chromatin structural changes in the corresponding regions of the genome, facilitating chromatin to become highly helical, condensed into clusters, and lose transcriptional activity, thereby regulating gene expression (Vymetalkova et al., 2019). However, the relationship between DNA methylation, chromatin structure and FAM169A expression in CAD is not known to us, which provides a direction for future studies.

The innovation of our study is the association study between DNA methylation regulated ceRNAs and CAD. However, there are still some shortcomings in our study, for example, the small number of selected GEO data sets may not be representative and lead to errors in the final results; the small amount of CAD and control samples in the validation population may lead to sampling errors and the the underlying regulation mechanism of DNA methylation on differential expressed genes in CAD need to be further investigated.

## Conclusions

In conclusion, in this study, we identified 2 potential hub genes, which provides a new basis for further exploration of the pathogenesis of CAD.

## Abbreviations

CAD: Coronary artery atherosclerotic heart disease; GEO: Gene Expression Omnibus; PBLs: Peripheral blood cells; DEMs: Differentially expressed microRNAs; WGCNA: Weighted co-expression network analysis; DMGs: Differentially methylated genes; ceRNA: Competitive endogenous RNAs; DEGs: Differentially expressed genes; 5-AZA: 5-Aza-2'-deoxycytidine; DELs: Differentially expressed lncRNAs; CVD: Cardiovascular diseases; lncRNAs: long non-coding RNAs; miRNA: MicroRNA; GO: Gene Ontology and; KEGG: Kyoto Encyclopedia of Genes and Genomes; CC: Cellular compartment; BP: Biological process; MF: Molecular function; ChAMP: Chip Analysis Methylation Pipeline; qRT-PCR: Real-time fluorescent quantitative PCR; GAPDH: Glyceraldehyde-3-phosphate dehydrogenase; FAM169A: Family with sequence similarity 169 member A; CXCR1: C-X-C motif chemokine receptor 1; AGPAT4: 1-

acylglycerol-3-phosphate O-acyltransferase 4; CAG: Coronary angiography; ECG: Electrocardiogram; DCG: Dynamic electrocardiogram; SLAP75: Soluble Lamina-Associated Protein of 75 kDa.

## Declarations

### Acknowledgements

None.

### Authors' contributions

BL and ZL collected the samples. YZ, BJ and YX performed the experiments and analyzed the data. YZ, CW and XZ analyzed the data and wrote the manuscript. WL created the figures. FZ conceived and designed the workflow and revised the manuscript. All authors read and approved the final manuscript.

### Funding

This work was supported by Zhongnan Hospital of Wuhan University Science, Technology and Innovation Seed Fund (znp2019049 & znp2019054), Translational Medicine and Interdisciplinary Research Joint Fund of Zhongnan Hospital of Wuhan University (No. ZNJC201932), Hubei Natural Science Foundation, Award (No. 2015CFA082) and Translational Medicine and Interdisciplinary Research Joint Fund of Zhongnan Hospital of Wuhan University (No. ZNLH201907).

### Availability of data and materials

The gene expression profiles of GSE31568, GSE113079 and GSE107143 were downloaded from Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/query>).

### Ethics approval and consent to participate

The studies involving human participants were reviewed and approved by Medical Ethics Committee of Zhongnan Hospital of Wuhan University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interest.

## References

- Archer, K., Broskova, Z., Bayoumi, A.S., Teoh, J.P., Davila, A., Tang, Y., et al. (2015). Long Non-Coding RNAs as Master Regulators in Cardiovascular Diseases. *Int J Mol Sci* 16(10), 23651-23667. doi: 10.3390/ijms161023651.
- Bauersachs, R., Zeymer, U., Brière, J.B., Marre, C., Bowrin, K., and Huelsebeck, M. (2019). Burden of Coronary Artery Disease and Peripheral Artery Disease: A Literature Review. *Cardiovasc Ther* 2019, 8295054. doi: 10.1155/2019/8295054.
- Benjamin, E.J., Virani, S.S., Callaway, C.W., Chamberlain, A.M., Chang, A.R., Cheng, S., et al. (2018). Heart Disease and Stroke Statistics-2018 Update: A Report From the American Heart Association. *Circulation* 137(12), e67-e492. doi: 10.1161/cir.0000000000000558.
- Boon, R.A., Jaé, N., Holdt, L., and Dimmeler, S. (2016). Long Noncoding RNAs: From Clinical Genetics to Therapeutic Targets? *J Am Coll Cardiol* 67(10), 1214-1226. doi: 10.1016/j.jacc.2015.12.051.
- Bradley, R.M., Bloemberg, D., Aristizabal Henao, J.J., Hashemi, A., Mitchell, A.S., Fajardo, V.A., et al. (2018). Lpaatδ/Agpat4 deficiency impairs maximal force contractility in soleus and alters fibre type in extensor digitorum longus muscle. *Biochim Biophys Acta Mol Cell Biol Lipids* 1863(7), 700-711. doi: 10.1016/j.bbalip.2018.04.001.
- Dai, X., Wiernek, S., Evans, J.P., and Runge, M.S. (2016). Genetics of coronary artery disease and myocardial infarction. *World J Cardiol* 8(1), 1-23. doi: 10.4330/wjc.v8.i1.1.
- Deng, Q., Huang, W., Peng, C., Gao, J., Li, Z., Qiu, X., et al. (2018). Genomic 5-mC contents in peripheral blood leukocytes were independent protective factors for coronary artery disease with a specific profile in different leukocyte subtypes. *Clin Epigenetics* 10, 9. doi: 10.1186/s13148-018-0443-x.
- Duan, S., Luo, X., and Dong, C. (2013). Identification of susceptibility modules for coronary artery disease using a genome wide integrated network analysis. *Gene* 531(2), 347-354. doi: 10.1016/j.gene.2013.08.059.
- Gaede, L., Möllmann, H., Rudolph, T., Rieber, J., Boenner, F., and Tröbs, M. (2019). Coronary Angiography With Pressure Wire and Fractional Flow Reserve. *Dtsch Arztebl Int* 116(12), 205-211. doi: 10.3238/arztebl.2019.0205.
- Hai, Z., and Zuo, W. (2016). Aberrant DNA methylation in the pathogenesis of atherosclerosis. *Clin Chim Acta* 456, 69-74. doi: 10.1016/j.cca.2016.02.026.
- Karreth, F.A., Reschke, M., Ruocco, A., Ng, C., Chapuy, B., Léopold, V., et al. (2015). The BRAF pseudogene functions as a competitive endogenous RNA and induces lymphoma in vivo. *Cell* 161(2), 319-332. doi: 10.1016/j.cell.2015.02.043.
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559. doi: 10.1186/1471-2105-9-559.

- Li, C., Ma, R., Zhang, X., Ma, J., Wang, X., He, J., et al. (2020). Risk of coronary heart disease in the rural population in Xinjiang: A nested case-control study in China. *PLoS One* 15(3), e0229598. doi: 10.1371/journal.pone.0229598.
- Li, M.J., Zhang, J., Liang, Q., Xuan, C., Wu, J., Jiang, P., et al. (2017). Exploring genetic associations with ceRNA regulation in the human genome. *Nucleic Acids Res* 45(10), 5653-5665. doi: 10.1093/nar/gkx331.
- Liang, B., Li, M., Deng, Q., Wang, C., Rong, J., He, S., et al. (2020). CircRNA ZNF609 in peripheral blood leukocytes acts as a protective factor and a potential biomarker for coronary artery disease. *Ann Transl Med* 8(12), 741. doi: 10.21037/atm-19-4728.
- Maric, M., Wong, E., and Burke, B. (2017). SLAP75 (Soluble Lamina-Associated Protein 75) – a novel determinant of higher order chromatin organization. *Mechanisms of Development* 145, S138. doi: <https://doi.org/10.1016/j.mod.2017.04.387>.
- Ouellette, C., Cormier, H., Rudkowska, I., Guénard, F., Lemieux, S., Couture, P., et al. (2013). Polymorphisms in genes involved in the triglyceride synthesis pathway and marine omega-3 polyunsaturated fatty acid supplementation modulate plasma triglyceride levels. *J Nutrigenet Nutrigenomics* 6(4-5), 268-280. doi: 10.1159/000357432.
- Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43(7), e47. doi: 10.1093/nar/gkv007.
- Roth, G.A., Johnson, C., Abajobir, A., Abd-Allah, F., Abera, S.F., Abyu, G., et al. (2017). Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015. *J Am Coll Cardiol* 70(1), 1-25. doi: 10.1016/j.jacc.2017.04.052.
- Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* 196(6), 801-810. doi: 10.1083/jcb.201112098.
- Ru, Y., Kechris, K.J., Tabakoff, B., Hoffman, P., Radcliffe, R.A., Bowler, R., et al. (2014). The multiMiR R package and database: integration of microRNA-target interactions along with their disease and drug associations. *Nucleic Acids Res* 42(17), e133. doi: 10.1093/nar/gku631.
- Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P.P. (2011). A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 146(3), 353-358. doi: 10.1016/j.cell.2011.07.014.
- Song, C., Zhang, J., Liu, Y., Pan, H., Qi, H.P., Cao, Y.G., et al. (2016). Construction and analysis of cardiac hypertrophy-associated lncRNA-mRNA network based on competitive endogenous RNA reveal functional lncRNAs in cardiac hypertrophy. *Oncotarget* 7(10), 10827-10840. doi: 10.18632/oncotarget.7312.

Sumi, M.P, Mahajan, B., Sattar, R.S.A., Nimisha, Apurva, Kumar, A., et al. (2021). Elucidation of Epigenetic Landscape in Coronary Artery Disease: A Review on Basic Concept to Personalized Medicine. *Epigenet Insights* 14, 2516865720988567. doi: 10.1177/2516865720988567.

van Diepen, J.A., Thiem, K., Stienstra, R., Riksen, N.P., Tack, C.J., and Netea, M.G. (2016). Diabetes propels the risk for cardiovascular disease: sweet monocytes becoming aggressive? *Cell Mol Life Sci* 73(24), 4675-4684. doi: 10.1007/s00018-016-2316-9.

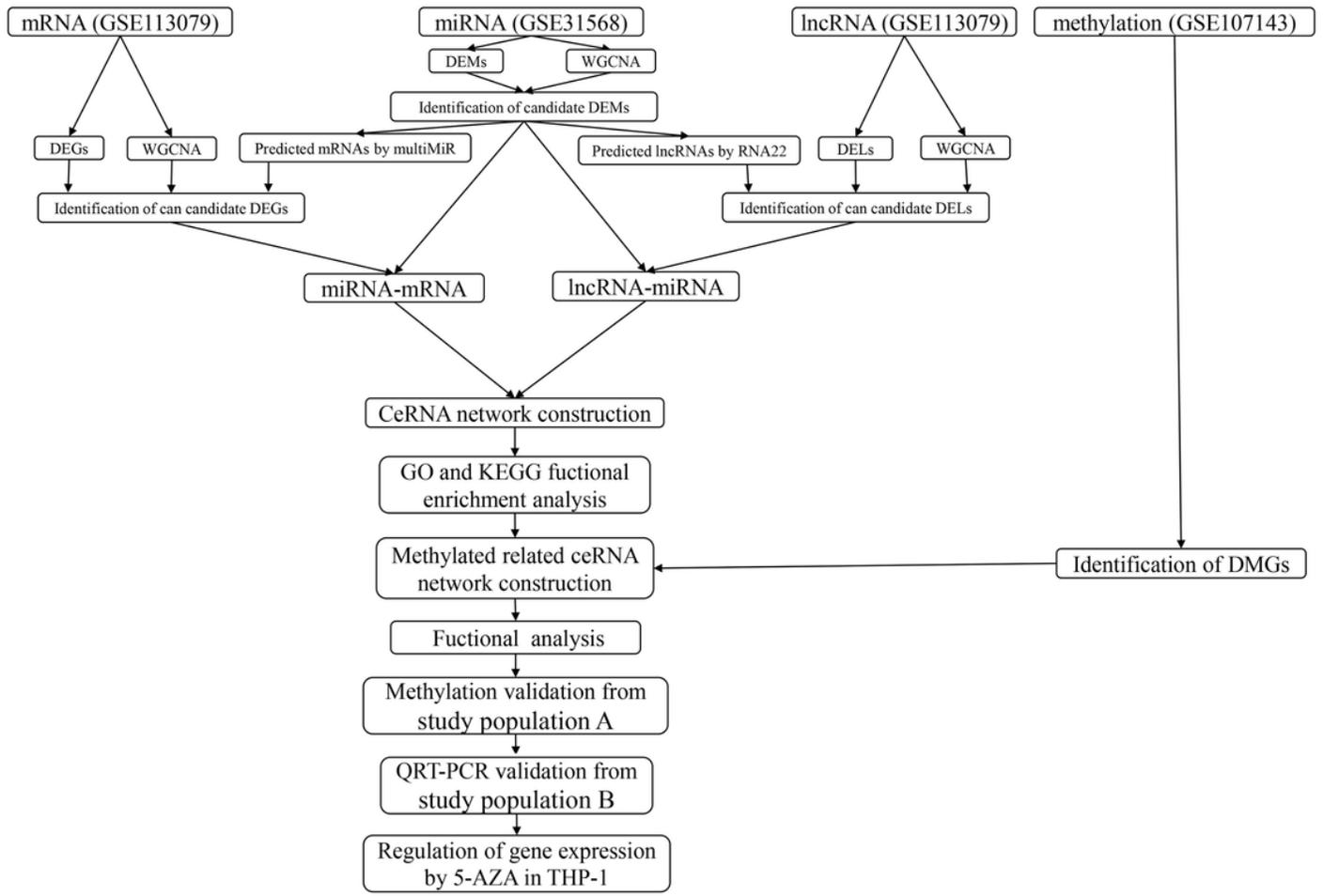
Vymetalkova, V., Vodicka, P., Vodenkova, S., Alonso, S., and Schneider-Stock, R. (2019). DNA methylation and chromatin modifiers in colorectal cancer. *Mol Aspects Med* 69, 73-92. doi: 10.1016/j.mam.2019.04.002.

Yamashita, A., Hayashi, Y., Matsumoto, N., Nemoto-Sasaki, Y., Oka, S., Tanikawa, T., et al. (2014). Glycerophosphate/Acylglycerophosphate acyltransferases. *Biology (Basel)* 3(4), 801-830. doi: 10.3390/biology3040801.

Zhang, D., Shi, R., Xiang, W., Kang, X., Tang, B., Li, C., et al. (2020). The Agpat4/LPA axis in colorectal cancer cells regulates antitumor responses via p38/p65 signaling in macrophages. *Signal Transduct Target Ther* 5(1), 24. doi: 10.1038/s41392-020-0117-y.

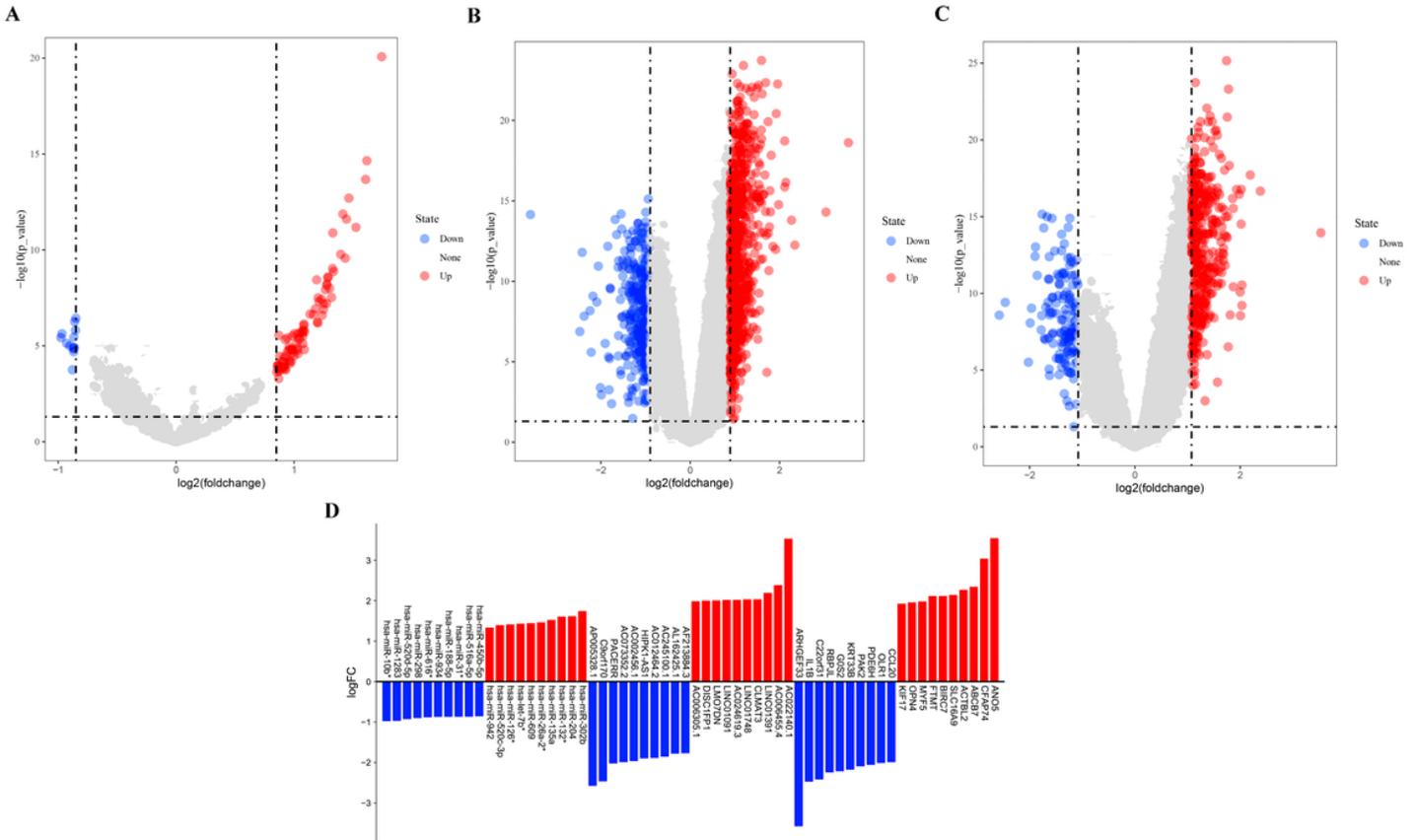
Zhukovsky, M.A., Filograna, A., Luini, A., Corda, D., and Valente, C. (2019). The Structure and Function of Acylglycerophosphate Acyltransferase 4/ Lysophosphatidic Acid Acyltransferase Delta (AGPAT4/LPAAT $\delta$ ). *Front Cell Dev Biol* 7, 147. doi: 10.3389/fcell.2019.00147.

## Figures



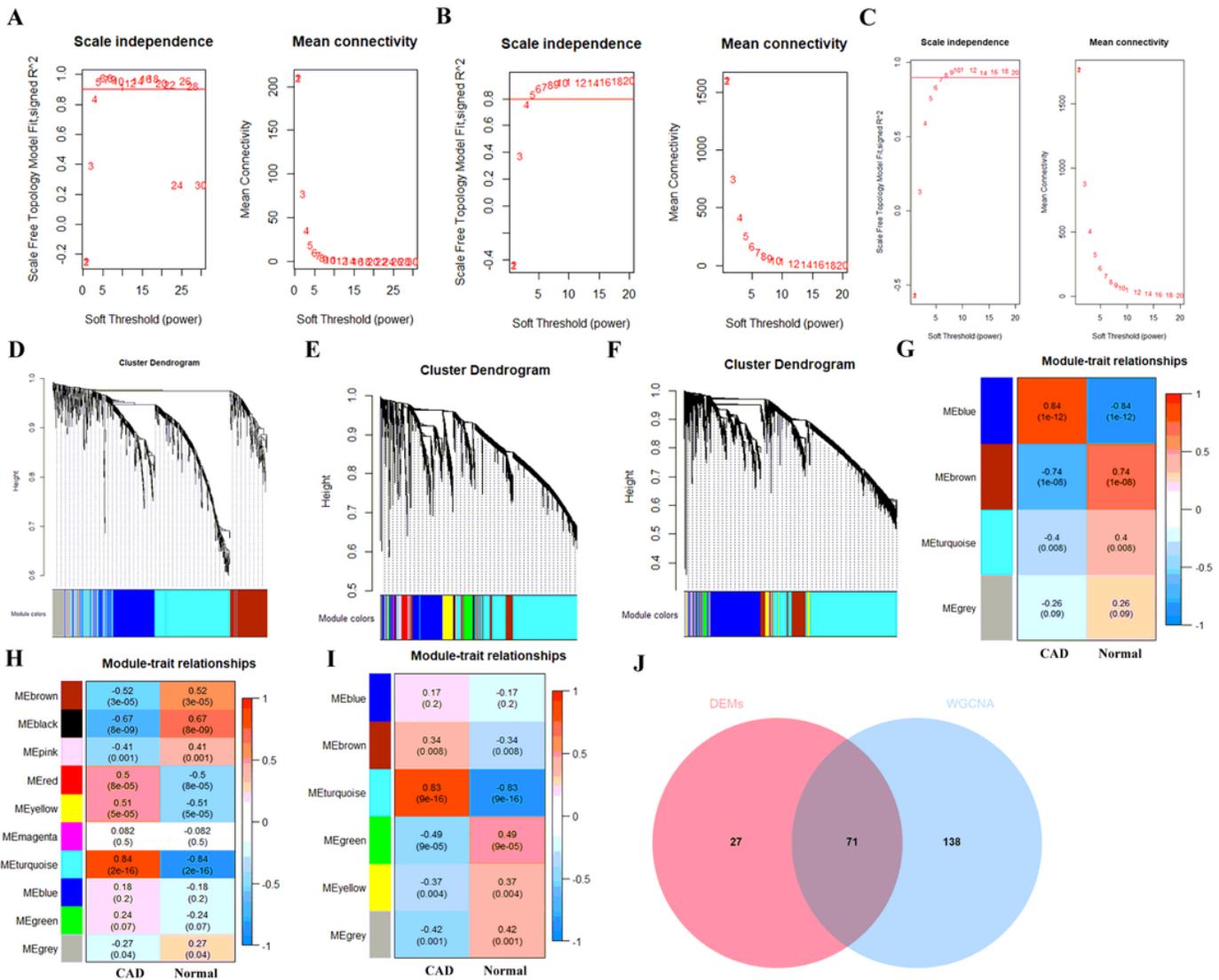
**Figure 1**

The workflow of the research.



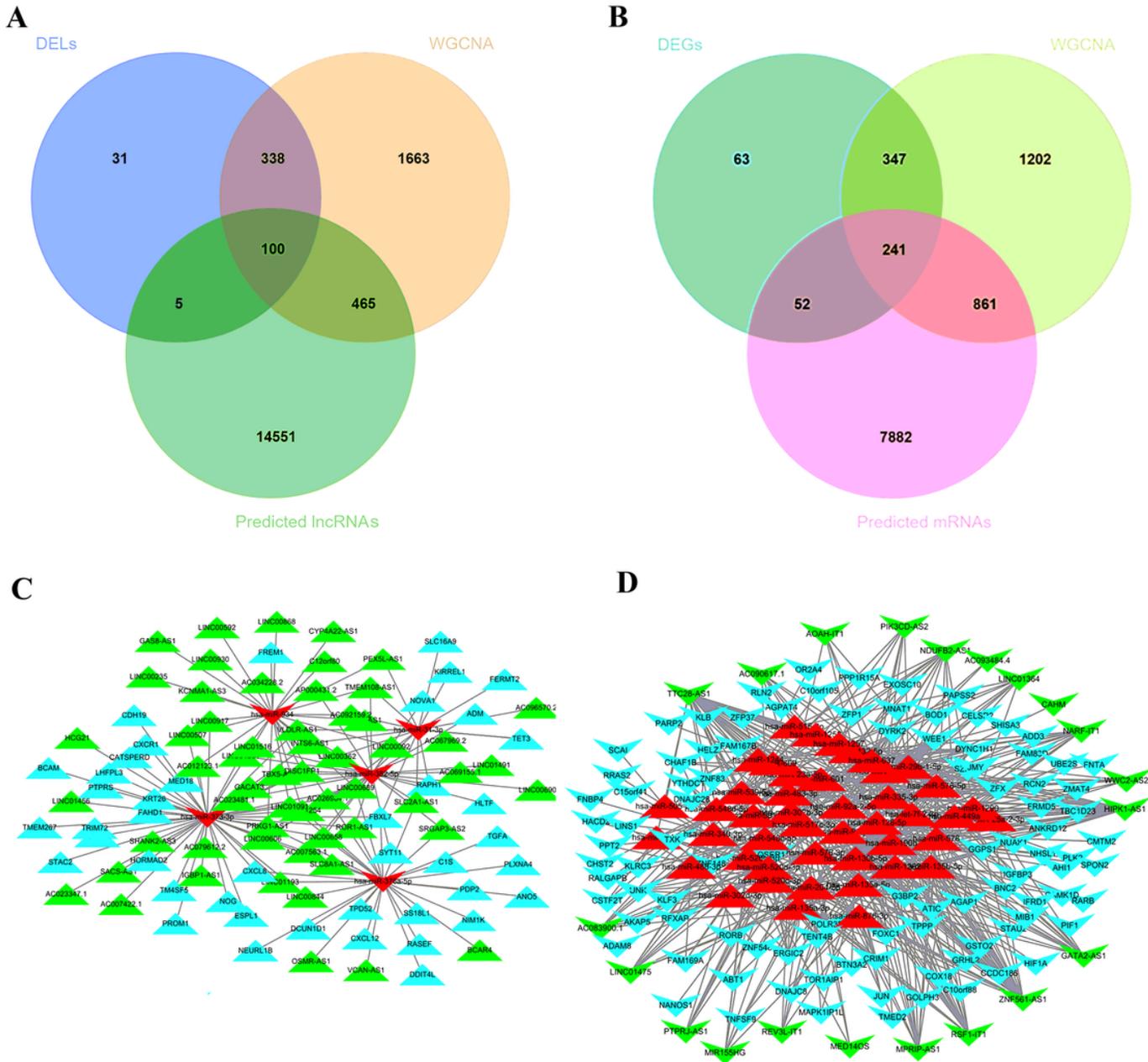
**Figure 2**

Differential expressed genes analysis. (A) Volcano plot presenting the differences of miRNAs in GSE31568. (B) Volcano plot presenting the differences of mRNAs in GSE113079. (C) Volcano plot presenting the differences of lncRNAs in GSE113079. (D) Bar plot showing top 10 upregulated and downregulated miRNAs, lncRNAs and mRNAs.



**Figure 3**

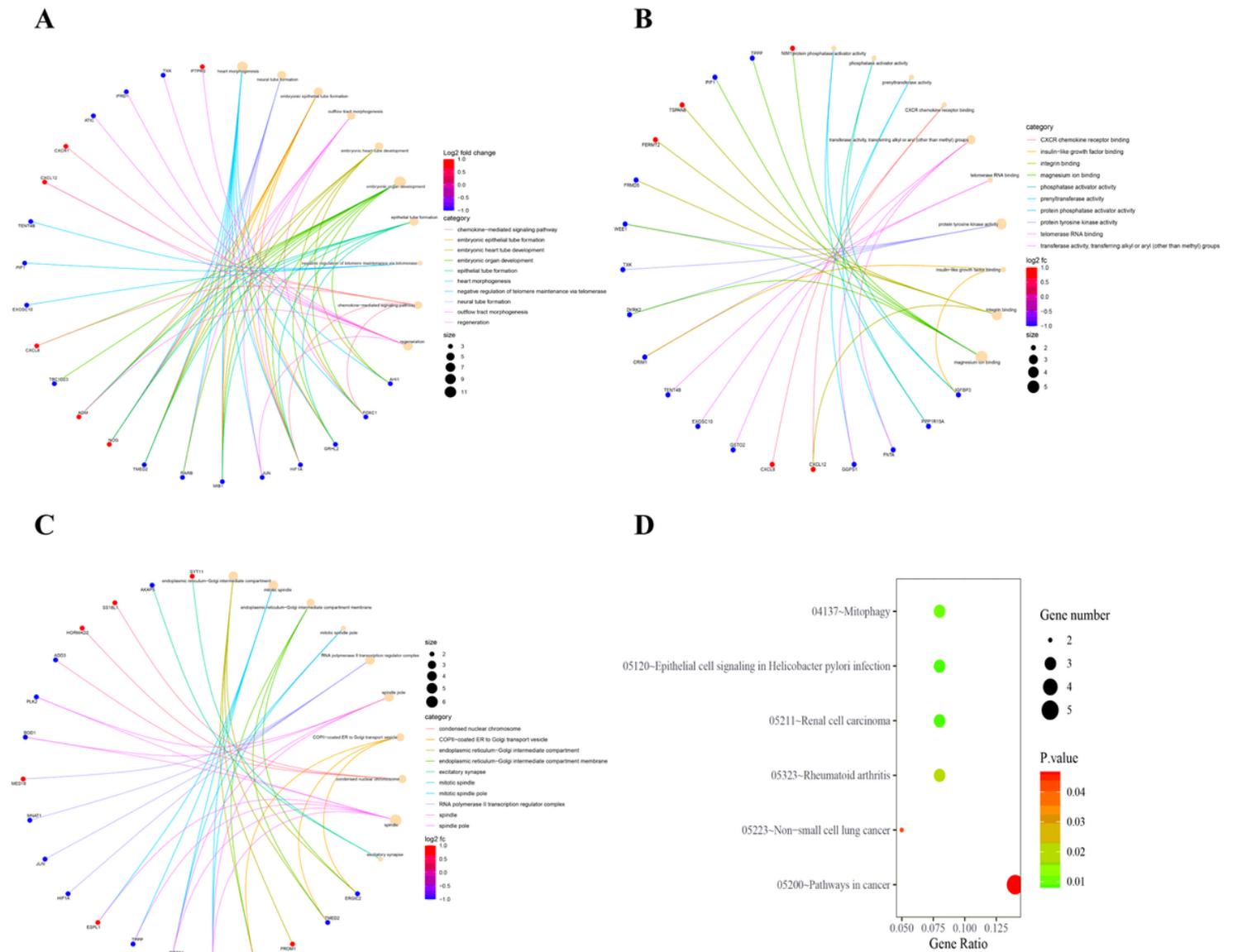
Identification of key genes correlated with CAD through WGCNA. (A), (B) and (C) Soft-thresholding powers analysis of miRNA, mRNA and lncRNA. (D), (E) and (F) Clustering dendrogram of miRNA, mRNA and lncRNA. Each color represents a certain gene module. (G), (H) and (I) The module-trait relationships in miRNA, mRNA and lncRNA. It represents the relationship between each module in CAD and the normal. (J) Candidate DEMs identified using venn diagram of previous DEMs and miRNAs in blue module of WGCNA.



**Figure 4**

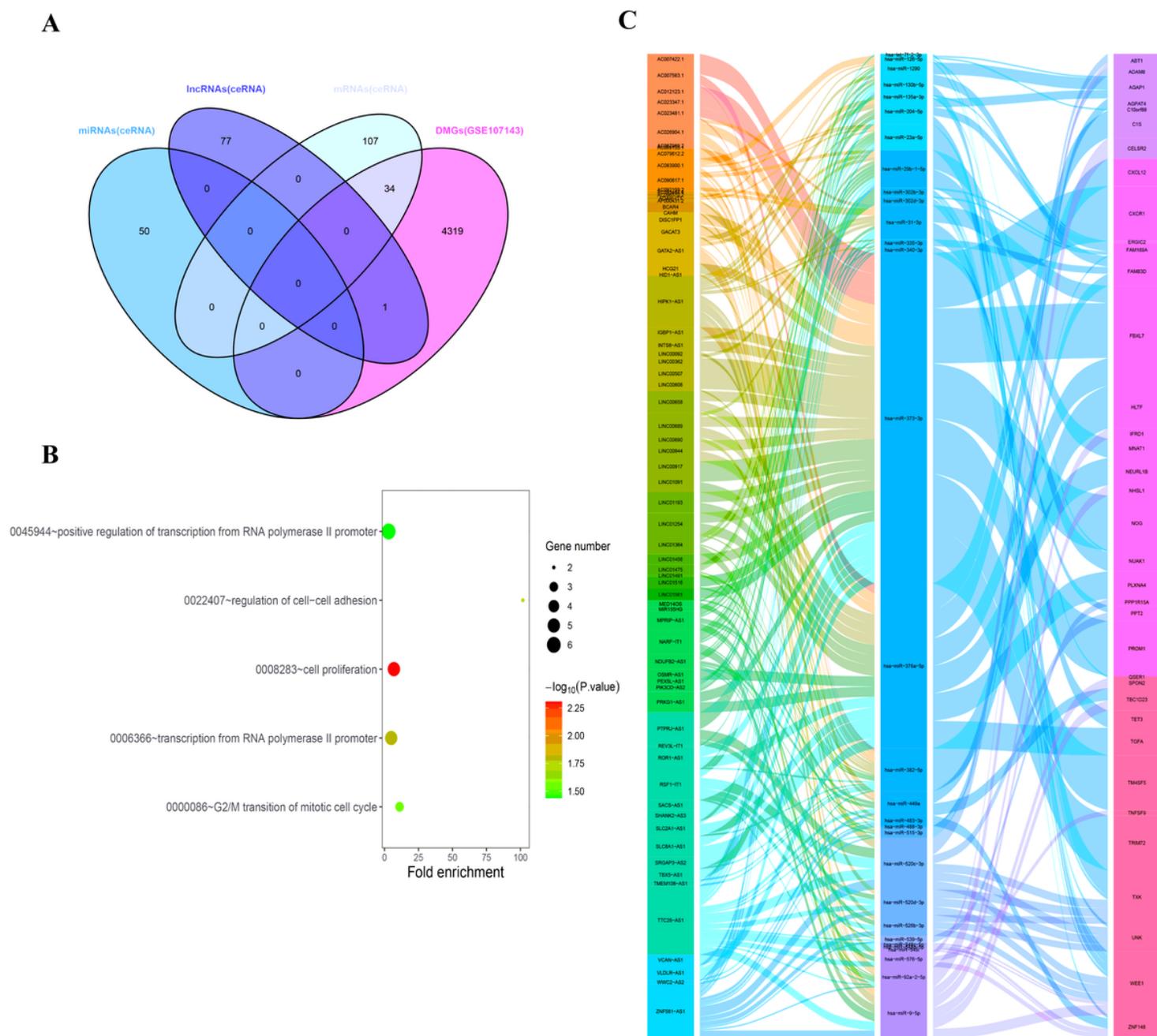
CeRNA network construction. (A) The candidate DELs identified by venn diagram of previous DELs, lncRNAs predicted by candidate DEMs and lncRNAs in the most significant module of WGCNA. (B) The candidate mRNAs identified by venn diagram of DEGs in Fig. 2B, mRNAs predicted by candidate DEMs and mRNAs in the most significant module of WGCNA. (C) and (D) lncRNA-miRNA-mRNA ceRNA interaction network in CAD. Green dots represent key lncRNAs, blue dots represent key mRNAs, and red

dots represent key miRNAs. Positive triangles represent up-regulated gene expression, and inverted triangles represent down-regulation.



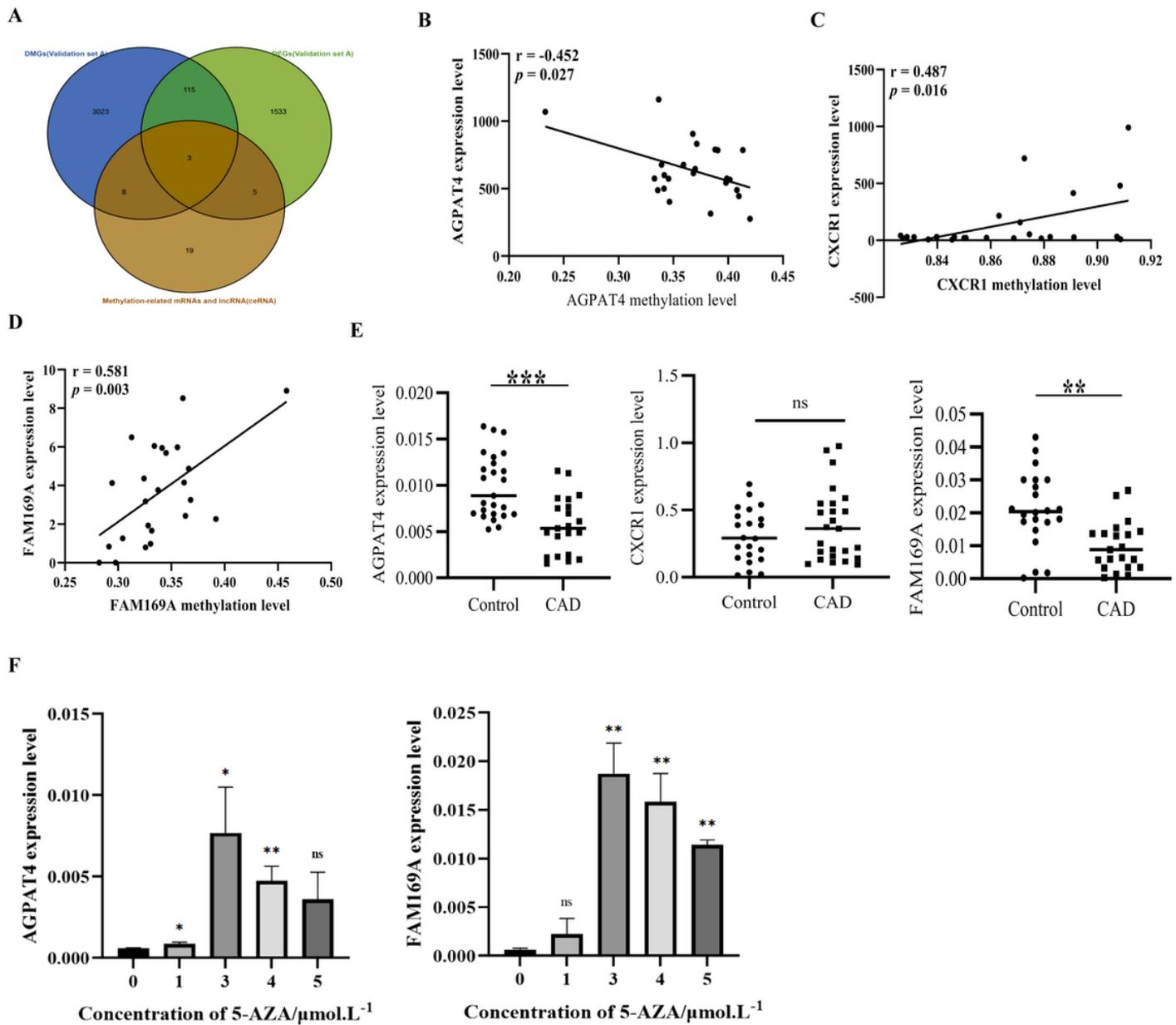
**Figure 5**

Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analysis for the key genes in ceRNA network. (A) Cnetplot plot revealed the relationship between key genes and GO terms of biological process (BP). (B) Cnetplot plot revealed the relationship between genes and GO terms of molecular function (MF). (C) Cnetplot plot revealed the relationship between genes and GO terms of cellular component (CC). (D) Dot plot indicated the relationship between genes and KEGG pathways.



**Figure 6**

DNA methylated related ceRNA network construction and functional enrichment analysis. (A) Venn plot of DMGs in GSE107143, key mRNAs, lncRNAs and miRNAs in ceRNA. (B) Functional enrichment analysis for mRNAs in methylated related ceRNA network. (C) Sankey diagram revealed the relationship between key lncRNAs, miRNAs and mRNAs related to DNA methylation.



**Figure 7**

Validation of the regulation by DNA methylation on differential expression of hub genes. (A) Identified hub DEGs related to DNA methylation using Venn plot. (B) Negative correlations between expressions and methylations of AGPAT4. (C) and (D) Positive correlations between expressions and methylations of CXCR1 and FAM169A. (E) Expressions of the AGPAT4, CXCR1 and FAM169A detected by RT-qPCR in CAD cases and controls. (F) The expression of AGPAT4 and FAM169A in THP-1 before and after 5-AZA (5-Aza-2'-deoxycytidine) treatment for 72hrs. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

## Supplementary Files

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

- [Additionalfile1.xlsx](#)
- [Additionalfile2.xlsx](#)
- [Additionalfile3.xlsx](#)
- [Additionalfile4.docx](#)
- [Additionalfile5.xlsx](#)