EXPERIMENTAL STUDY

Potential Regulatory Role of lncRNA-miRNA-mRNA in Coronary Artery Disease (CAD)

Liyuan Zhu,1 BM, Shuiping Zhao,1 MD and Wang Zhao,1 MD

Summary
Coronary artery disease (CAD) is a high-incidence of heart disease. We aimed to identify potential biomarkers linked to the progression of CAD using multiple sets of data mining analysis methods. The long noncoding RNA (lncRNA) + messenger RNA (mRNA) data set GSE113079 and microRNA (miRNA) data set GSE28858 were downloaded from Gene Expression Omnibus. After data preprocessing, differentially expressed mRNA, lncRNA, and miRNA were identified using limma software. In addition, weighted gene co-expression network analysis (WGCNA) was used for the construction and screening of modules related to disease states. Besides, key mRNAs and lncRNAs were extracted for protein-protein interaction (PPI) network construction and lncRNA-mRNA co-expression analysis. Additionally, the final integration resulted in the lncRNA-miRNA-mRNA relationship pairs (competing endogenous RNA (ceRNA) network). Finally, CTD 2020 update database was used for the verification of the expression level of the candidate genes. A total of 1319 differentially expressed mRNAs and 1983 lncRNAs were screened. After WGCNA, a total of 234 mRNAs and 546 lncRNAs were identified. A PPI network including 127 mRNA corresponding proteins was constructed. The ceRNA network included 24 up-regulated lncRNAs, 16 down-regulated miRNAs, and 42 up-regulated mRNAs. Through the validation of CTD 2020 update database, 21 CAD related mRNAs, and four important ceRNAs those may be participated in the pathogenesis of CAD were obtained. In this study, through multiple sets of data mining methods, the regulatory relationship of lncRNA, miRNA, and mRNA was comprehensively analyzed, and the important role of lncRNA-miRNA-mRNA in the pathogenesis of CAD was emphasized.

Key words: Competing endogenous RNA network

Coronary artery disease (CAD) is a chronic inflammatory disease and the leading cause of death in the world. The pathogenesis of CAD is extremely complex. According to previous reports, the phenotypic transition of vascular smooth muscle cells (VSMC) plays an extremely important role in atherosclerosis and vascular stenosis and ultimately leads to myocardial infarction due to vascular stenosis and obstruction. In addition, the destruction of calcium ion homeostasis in VSMC leads to increased vasoconstriction, which is also a key factor leading to the pathogenesis of CAD.

Although the diagnosis and treatment of CAD have made great progress in the past few decades, the pathophysiological mechanisms of CAD need to be further studied. Molecular biology studies have shown that several genes like long noncoding RNA (IncRNA) and messenger RNA (mRNA) are related to the occurrence and development of CAD. mRNA is generally considered to be regulated by microRNA (miRNA), which is commonly used in studying various diseases (including CAD). In addition, IncRNA was more stable than mRNA. IncRNA is a class of noncoding transcripts with > 200 nucleotides, which has been increasingly recognized as a contributor to cardiac pathology, and its role in coronary heart disease has been extensively studied more frequently. For example, ANRIL is a lncRNA gene and mediates the molecular mechanism of CAD development.

Candidate genes that involve the pathogenesis of CAD and have clinical significance will be used for targeted treatment of CAD as potential new biomarkers. In the current research, the lncRNA platform and miRNA data sets were downloaded from public databases, and multiple data mining methods were used to further screen and analyze them to identify genes that were potentially related to CAD. This study may contribute to revealing the relevance between candidate genes with altered expression and the occurrence and development of CAD.

From the Department of Cardiovascular Medicine, The Second Xiangya Hospital, Central South University, Changsha, China.
This study was supported by the National Natural Science Foundation of China (No. 81500234 and 81270956) and Fundamental Research Funds for the Central Universities of Central South University (No. 2020zzts882).
Address for correspondence: Wang Zhao, MD, Department of Cardiovascular Medicine, The Second Xiangya Hospital, Central South University, No.139 Middle Renmin Road, Changsha, Hunan 410011, China. E-mail: zhaowang920@csu.edu.cn
Received for publication March 15, 2021. Revised and accepted August 2, 2021.
Released in advance online on J-STAGE November 17, 2021.
doi: 10.1536/ihj.21-156
All rights reserved by the International Heart Journal Association.
Methods

Data source and preprocessing: For the data screening of our study, the following criteria were established: 1) *Homo sapiens*, coronary heart disease, and expression profiling by array were set as species, keywords, and study types, respectively, and related data sets were retrieved through NCBI Gene Expression Omnibus;\(^\text{9,10}\) the retrieved data set must be an lncRNA platform data set or miRNA data set; and 3) the normal control and disease groups must be included in the grouping.

According to the aforementioned standards, the lncRNA + miRNA data sets GSE113079, GSE69587, and GSE42148 met the requirements. However, GSE42148 was subsequently eliminated because the number of annotated lncRNAs was too small compared with that of the other two sets. For miRNA data sets, GSE28858, GSE59421, and GSE105449 met the requirements. After preliminary difference analysis, only 17 and 0 miRNAs were respectively obtained in GSE59421 and GSE105449, respectively, according to the threshold. Because too few miRNA results would miss important information, GSE28858 was finally selected for analysis.

Through data screening, a total of three sets of data were used for analysis as follows; GSE113079 contains a total of 141 samples, including 93 CAD samples and 48 healthy samples. The chip platform is a GPL20115 Agilent-067406 Human CBC lncRNA + mRNA microarray V4.0 (Probe name version). The tissue source is the human peripheral blood mononuclear cell. This data set was mainly used as a test set for key gene mining.

GSE69587 contains 6 samples, including 3 CAD samples and 3 healthy samples. The chip platform is a GPL15314 Arraystar Human LncRNA microarray V2.0 (Agilent_033010 Probe Name version). The tissue source is the human peripheral blood mononuclear cell. This data set was mainly used as a verification set to verify key gene sets.

GSE28858 contains 24 samples, including 12 CAD samples and 12 healthy samples. The chip platform is a GPL8179 Illumina Human v2 MicroRNA expression beadchip. The tissue source is human platelet.

Data preprocessing: For GSE113079 and GSE69587, the standardized probe expression matrix was first downloaded, and then, the probe sequences of their respective platforms were downloaded. According to the genocode.v25 provided by the GENCODE database,\(^\text{9,10}\) lncRNA transcripts and mRNA transcript sequences were obtained. Besides, by using SeqMap software,\(^\text{11}\) the respective probe sequences were aligned to lncRNA and mRNA sequences. And we retained the unique aligned sequence. If different probes were aligned to the same lncRNA or mRNA, the average value of distinct probes was taken as the ultimate expression value of this gene. In this way, lncRNA and mRNA expression matrices were obtained for the next step of differential expression analysis.

For GSE28858, the standardized probe expression matrix was also downloaded and used directly for the next step of differential expression analysis.

RNA differential expression analysis: Using the linear regression and empirical Bayes method provided in the limma software package (version 3.10.3),\(^\text{12}\) P value and logFC were obtained using differential expression analysis of mRNA, lncRNA, and miRNA between the patients with CAD and healthy groups. In addition, the Benjamini and Hochberg method was used to perform multiple test corrections, and the corrected P value, namely, adj.P value, was obtained.

For miRNA, miRSystem can convert miRNA probes into official miRNA names (from version 21 included in miRBase) to facilitate subsequent miRNA target gene prediction.\(^\text{13,14}\)

Weighted gene co-expression network analysis (WGCNA): The mRNA and lncRNA obtained from the differential expression analysis were combined, and the differentially expressed (DE) genes were obtained. Moreover, R package WGCNA (version 1.61) was used to analyze the differential genes to identify highly coordinated gene set modules.\(^\text{9,10}\) The screening steps of disease-related modules are shown as follows: 1) The squared values of the correlation coefficients for the degree of connection K and P (k) and the average degree of connection at each value of powers were calculated by setting a series of powers; the power value was selected when the square value of the correlation coefficient reached above 0.8 for the first time; 2) on the basis of clustering and dynamic pruning, the parameters (minModuleSize = 30, MEDissThres = 0.1) were set, and highly correlated genes were clustered into modules; and 3) by calculating the correlation between the module and the disease state, the gene set modules closely related to the disease state were identified. And then, the modules with the largest correlation coefficients with “CAD” and “Healthy” were selected as the key modules for subsequent analysis.

Protein-protein interaction (PPI) network analysis: On the basis of the key modules obtained, the mRNA was first extracted. Then, to predict and analyze whether there was an interaction relationship between the proteins encoded by mRNA, the STRING\(^\text{16}\) (version 10.0) database was used. The PPI score was set to 0.4 (medium confidence). After obtaining the PPI relationship pair, CYTOSCAPE software (version 3.4.0) was used in constructing a network diagram.\(^\text{17}\) Finally, DAVID\(^\text{18}\) (version 6.8) was used for the analysis of the GO BP and KEGG pathways enriched by genes in the aforementioned PPI network.\(^\text{20-22}\)

The standard of significant enrichment was the significance threshold P value < 0.05 and the enrichment number ≥ 2.

Analysis of lncRNA-mRNA co-expression: After the lncRNA in the key module was extracted, the Pearson correlation coefficients of lncRNA and mRNA were calculated separately and the correlation test was performed. In addition, through the “BH” correction, the corrected adj.P value was obtained. With regard to the subsequent construction of the competing endogenous RNA (ceRNA) network, the lncRNA-mRNA relationship pairs were selected from those with correlation coefficient r > 0.9 and adj.P value < 0.01. Then, these relationships were integrated with the PPI network, resulting in the formation of a co-expression interaction network of lncRNA-miRNA-mRNA. Besides, the lncRNA-miRNA-mRNA co-expression interaction network was constructed by using...
the CytoNCA plug-in (version 2.1.6), and the topological characteristics of the node network were analyzed. The parameter was set to without weight. Moreover, closeness centrality (CC), betweenness centralities (BC), and degree centrality (DC) were included in the results. Then, according to the ranking of the network topology properties of each node, the intersection of TOP50 lncRNAs in each topological property was selected as the key lncRNA.

Finally, the R package clusterProfiler \(^{25}\) (version: 3.8.1) was adopted. As a potential target gene of lncRNA, miRNA was analyzed for KEGG pathway enrichment. The items with corrected \(P\) values < 0.05 were regarded as significant enrichment results.

**Construction of the ceRNA network:** First, after the lncRNA-miRNA relationship pairs were selected, the key lncRNA-related lncRNA-miRNA relationship pairs were obtained.

Second, for the lncRNA in the lncRNA-miRNA relationship pairs, combined with the DE miRNA, the local tool Miranda v3.3a \(^{24}\) was used to perform binding site prediction analysis. By setting the parameters “-sc 140,” “-strict,” and “-en -20,” the lncRNA-miRNA relationship pairs were screened out and obtained.

Again, according to the mRNA in lncRNA-mRNA, miRWalk2.0 \(^{20}\) database was used, and the results of seven commonly used databases, namely, miRDB, miRWalk, RNA22, miRanda, Pictar2, PITA, and TargetScan, were integrated. If the predicted miRNA appeared in the aforementioned four databases, it was considered that the miRNA regulated the corresponding mRNA. After obtaining the predicted miRNA-mRNA relationship pairs, the miRNA and the differential miRNA were intersected to obtain the miRNA-mRNA relationship pairs. Simultaneously, because miRNA was generally believed to inhibit mRNA expression, the mRNA-mRNA relationship pairs with the completely opposite regulatory relationship between the miRNA and mRNA were further selected for the next analysis.

Finally, on the basis of the obtained lncRNA-miRNA and miRNA-mRNA relationship pairs, lncRNA and mRNA regulated by the same miRNA were screened out. Then, combined with the forward co-expressed lncRNA-mRNA, the relationship pairs of lncRNA-miRNA-mRNA were integrated to obtain the ceRNA relationship. Finally, the ceRNA network was constructed via Cytoscape software.

**Gene expression verification:** The CTD 2020 update (Comparative Toxicogenomics Database) \(^{20}\) database was used for verification, which provided information on chemical-gene/protein interactions, chemical-disease, and genetic-information about disease relationships. “Coronary heart disease” was used as a keyword to search for disease-related genes. According to inference score, \(^{27,28}\) genes with a score of ≥10 were regarded as disease-related genes. Then, these genes and the mRNAs in the ceRNA network were intersected to verify whether the mRNAs we obtained were indeed closely related to the disease, which further showed that the ceRNA network was closely related to the disease.

**Results**

**Analysis of differential expression of mRNA, lncRNA, and miRNA:** After comparison and annotation, 7368 mRNAs and 11,895 lncRNAs in GSE113079, 10,217 mRNAs and 7693 lncRNAs in GSE69587, and 858 miRNA probes in GSE28858 were obtained.

According to the threshold set by the method, a total of 1,319 DE mRNAs (662 up-regulated and 657 down-regulated), 1983 DE lncRNAs (1448 up-regulated and 535 down-regulated), and 103 DE miRNAs (58 up-regulated and 45 down-regulated) were analyzed. The volcano map is shown in Figure 1A.

On the basis of DE mRNA, lncRNA, and miRNA, the bidirectional hierarchical clustering heat map is shown in Figure 1B. Among them, because there were too many DE mRNAs and lncRNAs, we only selected RNAs with fold change > 2 for drawing.

**Screened disease status module:** The 3302 DE genes (1319 DE mRNAs and 1983 DE lncRNAs) were analyzed using WGCNA, and the soft threshold value was 9. The result is shown in Figure 2A. Then, the highly correlated genes were clustered into modules according to clustering and dynamic pruning. Besides, these modules were clustered and merged with correlation coefficients > 0.9 and < 0.1 and finally integrated into seven modules, as shown in Figure 2B.

Further two methods were used to mine the modules related to the disease state: 1) the correlation was calculated between the feature vector gene of each module and the disease state, as shown in Figure 3A. 2) During the correlation analysis, the disease state was shown as follows: 1 for CAD and 0 for healthy. Then, the absolute mean value of the correlation between the expression of each gene in each module and the disease state was taken as the correlation between the module and the disease state (shown in Figure 3B). It can be seen that the blue module significantly had positive correlation with “CAD,” and the yellow module significantly had positive correlation with “Healthy.”

Then, by calculating the significant Module Membership (MM) of the genes and modules in the blue and yellow modules and the correlation between genes and disease states (GS), the corresponding scatter diagram was drawn as shown in Figure 3B. Moreover, those points that were highly correlated to the module (Figure 3C, D) were also highly correlated with traits. Therefore, the thresholds were set to MM > 0.8 and GS > 0.6, and from the two models, important hub genes were screened out and finally integrated. A total of 234 mRNAs and 546 lncRNAs were included.

**PPI network construction and GO BP/KEGG pathway enrichment analysis results:** For the aforementioned 234 mRNAs, the STRING database was used to construct a protein interaction network, and a total of 143 PPI relationship pairs composed of 127 mRNA corresponding proteins were obtained. As shown in Figure 4A, the size of the node in the figure represented the size of the connection. In addition, the enrichment analysis of the GO BP and KEGG pathway was performed on these 127 mRNAs. The results were shown in Figure 4B. A total of
Figure 1. Analysis of differential expression of mRNA, lncRNA, and miRNA. A: Differentially expressed genes mRNA (left), lncRNA (middle), and miRNA (right). Red triangles indicate up-regulation, blue squares indicate down-regulation, and gray dots indicate that the difference is not significant. The labels in Figure 1A represent up and down genes (TOP10) according to the log2FC ranking. The up-regulation/down-regulation of gene expression levels involved in this study indicates changes in gene expression in patients with CAD compared with the healthy group. B: Heat map of differentially expressed genes of mRNA (left), lncRNA (middle), and miRNA (right). The red horizontal bar at the top represents CAD samples, and the green horizontal bar represents healthy samples.

10 GO BP and 5 KEGG pathways were enriched. The 10 enriched GO BP pathways are as follows: 1) G protein-coupled receptor signaling pathway (GO: 0007186); 2) detection of chemical stimulus involved in sensory perception of smell (GO: 0050911); 3) extracellular matrix organization (GO: 0030198); 4) sensory perception of smell (GO: 007608); 5) potassium ion transmembrane transport (GO: 0071805); 6) regulation of ion transmembrane transport (GO: 0034765); 7) negative regulation of T cell proliferation (GO: 0042130); 8) negative regulation of neural precursor cell proliferation (GO: 2000178); 9) lung epithelium development (GO: 0060428); and 10) collagen-activated tyrosine kinase receptor signaling pathway (GO: 0038063). Besides, the 5 KEGG pathways include 1) olfactory transduction (hsa04740); 2) neuroactive ligand-receptor interaction (hsa04080); 3) amoebiasis (hsa05146); 4) ECM-receptor interaction (hsa04512); and 5) protein digestion and absorption (hsa04974).

Analysis of lncRNA-miRNA-mRNA co-expression interaction network: On the basis of the 234 mRNAs and 546 lncRNAs selected by WGCNA analysis, the correlation coefficient between the mRNAs and lncRNAs was calculated. According to the threshold, a total of 5117 lncRNA-mRNAs were screened. In addition, through integration with the aforementioned PPI relationship pairs, a total of 5260 relationship pairs were finally obtained, and the integrated construction of the lncRNA-miRNA-mRNA network is shown in Figure 5A.

Moreover, the CytoNCA plug-in was used to analyze the DC, BC, and CC of the network nodes. And the results are shown in Figure 5B. It can be known that the DC and BC of most nodes were small, and only a few nodes were larger, in line with the characteristics of the network. To obtain key lncRNAs, we selected lncRNAs ranked TOP50 in each topological property. Finally, 39 lncRNAs, such as AC005624.2, KB-208E9.1, RP3-402G11.27, were obtained, which were considered to be essential lncRNAs (Figure 5C).

In order to understand the possible regulatory pathways of the 39 lncRNAs obtained above, the mRNA co-expressed with lncRNA was used for the purpose of a potential target gene of the relevant lncRNA, and the enrichment analysis of KEGG pathway was performed. A total of 12 pathway enrichment analysis results were obtained, and the results are shown in Figure 5D. Among the 12 pathways, the olfactory transduction pathway was enriched by the largest number of genes.

ceRNA network: By screening IncRNA and mRNA that
Figure 2. The power value and clustering and merging of WGCNA modules. A: Graph of WGCNA power value. The abscissa soft threshold (power) represents the weight, and the ordinate represents the square value of the correlation coefficient between the degree of connection k and p (k) (left); soft threshold (power) represents the weight, and the ordinate represents the average connectivity. Generally, the power when the square value of the correlation coefficient of k and p (k) reaches 0.8 for the first time is required as the β value, and it can be seen that the β value is 9 (right). B: Diagram of clustering and merging of WGCNA modules. Module clustering result graph, the vertical axis represents the difference coefficient, and the blue horizontal line represents the difference coefficient of 0.1 (left); the gene modules generated using the hierarchical clustering tree and dynamic shearing method. The different colors in the figure subpart represent different gene modules, and the gray parts represent genes that cannot be merged into any other modules (right).

... were under the control of the same miRNA and combining the lncRNA-mRNA relationship pairs that were forward co-expressed, 234 lncRNA-miRNA-mRNA relationship pairs were finally obtained, including 24 up-regulated lncRNAs, 16 down-regulated miRNAs, and 42 up-regulated mRNAs. The network is shown in Figure 6A. In this network, there were a total of 46 lncRNA-miRNA relationship pairs, 105 miRNA-mRNA relationship pairs, and 188 lncRNA-mRNA co-expression relationships.

The CTD database was used to verify the mRNA in the aforementioned network. We retrieved a total of 5769 annotated as the coronary heart disease-related genes from the CTD database. In addition, through the intersection with the mRNA in the ceRNA network, we learned that 21 mRNAs including Fas-apoptotic inhibitory molecule 2 (FAIM2), solute carrier family 46 member 1 (SLC46A1), microtubule affinity-regulating kinase 4 (MARK4), and NADH dehydrogenase (ubiquinone) 1 alpha subcomplex (NDUFA10) were reported disease-related genes (Figure 6 B). The close relationship between the network and disease has been proved.

It can be seen from the constructed ceRNA network that FAIM2 showed the highest degree of connection, followed by miR-665 and SLC46A1. They can form ceRNA with AC005624.2 (one of 39 important lncRNAs): AC005624.2-miR-665-FAIM2 and AC005624.2-miR-665-SLC46A1. In addition to FAIM2, miR-665, and SLC46A1, the degree value of MARK4 was closely followed, and KB-208E9.1-miR-339-5p-MARK4 was formed. Among them, the degree value of miR-339-5p was the second highest among miRNAs and KB-208E9.1 was also one of 39 important lncRNAs. Besides, miR-143-5p with high degree of connection also participated in networks such as RP3-402G11.27-miR-143-5p-NDUFA10.
Figure 3. Correlation analysis between WGCNA module and CAD status. A, B: Results of correlation analysis between WGCNA modules and traits. Exploring the trait-related modules based on the correlation between the trait and the module feature vector gene and the P value (the upper number indicates the correlation coefficient, and the lower bracket number indicates the significant P value) (left); The mean value of the correlation in each module is regarded as the significance of the trait in the module (right). C, D: Two-dimensional scatter plot of correlation. The blue (left) and yellow (right) modules represent gene–module correlation (MM) and gene–disease state correlation (GS), respectively.

Discussion

The purpose of this study was to identify potential biomarkers closely related to the progression of CAD by using multiple sets of data mining analysis methods. These biomarkers may exist in lncRNA, miRNA, or mRNA. On the basis of the analysis of the lncRNA-miRNA-mRNA ceRNA network, 24 DE lncRNAs, 16 DE miRNAs, and 21 DE mRNAs were related to the pathogenesis of CAD. Moreover, combined with current research reports on related genes in CAD, we mainly focused on the 4 ceRNAs composed of genes with high degree of connection in the ceRNA network, including 665-AC005624.2-miR-FAIM2, AC005624.2-miR-665-SLC46A1, KB-208E9.1-miR-339-5p-MARK4, and RP3-402G11.27-miR-143-5p-NDUFA10.

The longitudinal EWASs in the previous Japanese cohort study have been shown to be significantly related to
the incidence of CAD. According to the database, SNP rs7132908 of FAIM2 may be a new susceptibility locus for CAD. By comparing the incidence of CAD with this SNP, it is shown that the minor alleles were related to the reduced sensitivity to CAD.30) Actually, hypertension, diabetes, or dyslipidemia has been considered to be a risk factor for CAD.30) A previous study has revealed 225 DE lncRNAs and 473 DE mRNAs in the submandibular gland of rats with spontaneous hypertension compared with that of Wistar-Kyoto rats.31) As an obesity-related gene, FAIM2 is regulated by nutritional state and the methylation levels of the FAIM2 promoter are significantly associated with obesity.32) And it has been reported that obesity-related loci in FAIM2 is associated with obesity and type 2 diabetes in Han Chinese patients.33) Moreover, in a high glucose condition, FAIM2 is directly targeted by miR-3202 to promote the apoptosis of H5V cells.34) In the present study, because in the ceRNA network mRNA FAIM2 showed the highest degree of connectivity, which was co-expressed with AC005624.2, and also regulated by miR-665, we speculate that AC005624.2-miR-665-FAIM2 may be involved in the pathogenesis of CAD.

Similar to FAIM2, due to the high degree of connection of mRNA SLC46A1 in the ceRNA network, AC005624.2-miR-665-SLC46A1 was found as a good predictor for the pathogenesis of CAD. As for SLC46A1, the single nucleotide polymorphisms in SLC46A1 were important predictors of plasma high-density lipoprotein in healthy adults, which was considered a protective factor to prevent CAD.35,36) In the ceRNA network, SLC46A1 and AC005624.2 were co-expressed, and the former was also regulated by miR-665. Therefore, ceRNA AC005624.2-miR-665-SLC46A1 can also be involved in the pathogenesis of CAD.

Recently, MARK4 was identified as a potential drug target for several complex diseases, including cancer, diabetes, diet-induced obesity, and neurodegenerative diseases.37,38) Although there is no evidence that MARK4 and CAD are directly related, by examining the role of MARK4 in hematopoietic cells during atherosclerosis, it is found that MARK4-dependent NLRP3 (NLR family pyrin domain containing 3) inflammasome activation in the hematopoietic cells regulates the development of atherosclerosis.39) It has been reported that atherosclerosis can cause
Figure 5. Construction of lncRNA–mRNA–mRNA network. A: lncRNA–mRNA–mRNA network. The red hexagon represents the up-regulated lncRNA; the blue hexagon represents the down-regulated lncRNA; the yellow circle represents the up-regulated mRNA; the green circle represents the down-regulated mRNA; the gray dashed line represents the co-expression relationship between lncRNA and mRNA, and the dark blue line represents the PPI relationship. B: The distribution of nodes in the lncRNA–mRNA–mRNA network. Left, middle, and right represent degree centrality, betweenness centrality, and closeness centrality, respectively. C: Venn diagram of the intersection of TOP50 key lncRNA topological properties. D: KEGG pathway enrichment analysis results of 39 lncRNAs. The size of the bubble indicates the number of enriched genes, and the redder the bubble color, the more significant it is.

Figure 6. ceRNA network construction and gene analysis. A: ceRNA network. Different colors indicate different modules. Red hexagons represent up-regulated lncRNA; yellow circles represent up-regulated mRNA; blue inverted triangles represent down-regulated miRNA; green dotted lines indicate the co-expression of lncRNA and mRNA; gray arrows indicate miRNA regulation of mRNA; pink T-shaped connection indicates that lncRNA competes to bind miRNA; the size of the node indicates the degree of connection. B: Venn diagram of the intersection of coronary heart disease-related genes in the ceRNA–mRNA network and CTD database.
CAD. In our ceRNA network, the expression level of MARK4 was up-regulated in patients with CAD. MARK4 and KB-208E9.1 were co-expressed, and MARK4 was also regulated by miR-339-5p. Therefore, KB-208E9.1-miR-339-5p-MARK4 may also be involved in the pathogenesis of CAD.

It is generally believed that mRNA is regulated by miRNA, which is the key to gene regulation and a conserved part of evolution. It is shown that the NDUFA10 was regulated by miR-143-5p via the ceRNA network. By constructing the ceRNA network, it was also found that the expression level of miR-143-5p in patients with CAD was down-regulated compared with that in the normal group. Because miR-143-5p has previously been proven to be a useful cancer biomarker, current studies on its relationship with coronary heart disease are relatively rare. The latest research shows that miR-143-5p can directly target HIF-1α to promote the phenotypic transformation, proliferation, and migration of pulmonary artery smooth muscle cells under hypoxic conditions. Overexpression of miR-143-5p significantly reduced the level of vascular smooth muscle specific contraction marker protein and apoptosis of pulmonary artery smooth muscle cells under hypoxia stimulation, whereas the low expression of miR-143-5p significantly reduced the level of vascular muscle cells under hypoxic conditions. However, the mechanism of action needs further research.

It is shown that the ceRNA network that NDUFA10 was regulated by miR-143-5p. NDUFA10 is a key gene in the mitochondrial electron transport chain. Studies have shown that mitochondrial gene mutations eventually cause vascular endothelial cell damage, one of the root causes of coronary atherosclerosis. Moreover, the gene expression of NDUFA10 correlates positively with glucose-stimulated insulin secretion in patients with type 2 diabetes. And NDUFA10 is found to have special expression modification patterns in spontaneously hypertensive rat strains. In this study, the novel ceRNA RP3-402G11.27-miR-143-5p-NDUFA10 can provide new ideas and in-depth research possibilities for the study of CAD pathogenesis. Because of the lack of large-scale clinical research, the dynamic regulation mechanism of IncRNA-miRNA-mRNA is still unclear. In general, the crossstalk among IncRNA, miRNA, and mRNA seems to be complicated. In this study, through multiple sets of data mining analysis methods, we discussed the regulation mechanism of IncRNA and miRNA in CAD and emphasized the role of IncRNA-miRNA-mRNA in the CAD process. IncRNA can act as a regulator or effector of miRNA through a variety of post-transcriptional mechanisms, thereby regulating gene expression and ultimately the process of CAD. However, the journey of exploring the function and mechanism of IncRNA-miRNA-mRNA in the physiological and pathological regulation of CAD is still long.

Conclusions

In this research, through multiple sets of data mining and analysis methods, we constructed a ceRNA network to investigate the possible pathogenesis of CAD. The ceRNA axes AC005624.2-miR-665-FAIM2, AC005624.2-miR-665-SLC4A1, KB-208E9.1-miR-339-5p- MARK4, and RP3-402G11.27-miR-143-5p-NDUFA10 may be associated with the progression of CAD. The genes involved in these ceRNAs may serve as diagnostic biomarkers or therapeutic targets of CAD.

Disclosure

Conflicts of interest: The authors declare that they have no conflict of interest.

Authors’ contributions: Conception and design of the research and analysis and interpretation of data: LZ; acquisition of data and statistical analysis: SZ; drafting the manuscript: LZ and WZ; revision of manuscript for important intellectual content: WZ. All authors read and approved the final manuscript.

References

1. Bautista-Sánchez D, Arriaga-Canon C, Pedrera-Torres A, et al. The promising role of miR-21 as a cancer biomarker and its importance in RNA-Based Therapeutics. Mol Ther Nucleic Acids 2020; 20: 409-20.
2. Aarabi G, Zeller T, Seedorf H, et al. Genetic susceptibility contributing to periodontal and cardiovascular disease. J Dent Res 2017; 96: 610-7.
3. Climent B, Santiago E, Sánchez A, et al. Metabolic syndrome inhibits store-operated Ca(2+) entry and calcium-induced calcium-release mechanism in coronary artery smooth muscle. Biochem Pharmacol 2020; 182: 114222.
4. Dai X, Wiernek S, Evans JP, Runge MS. Genetics of coronary artery disease and myocardial infarction. World J Cardiol 2016; 8: 1-23.
5. Wang LA-O, Hu J, Zhou J, Guo F, Yao T, Zhang L. Weighed gene coexpression network analysis screens the potential long noncoding RNAs and genes associated with progression of coronary artery disease. Comput Math Methods Med 2020: 2020: 813420.
6. Zhang YH, Pan X, Zeng T, Chen L, Huang T, Cai YD. Identifying the RNA signatures of coronary artery disease from combined IncRNA and mRNA expression profiles. Genomics 2020; S0888-7543: 30508-5.
7. Cho H, Shen GQ, Wang X, et al. Long noncoding RNA ANRIL regulates endothelial cell activities associated with coronary artery disease by up-regulating CLIP1, EZR, and LYVE1 genes. J Biol Chem 2019; 294: 3881-98.
8. Cho HA-O, Li YA-O, Archacki S, et al. Splice variants of IncRNA RNA ANRIL exert opposing effects on endothelial cell activities associated with coronary artery disease. RNA Biol 2020; 17: 1391-401.
9. Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: the reference human genome annotation for the ENCODE Project. Genome Res 2012; 22: 1760-74.
10. Yan E, Kai W, Yi-Zhou J, Xin-Wen C, Cai-Feng D, Jing Z. Etiology of familial breast cancer with undetected BRCA1 and BRCA2 mutations: clinical implications. Cell Oncol 2014; 37: 1-8.
11. Smyth GK. Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W, eds. Bioinformatics and Computational Biology Solutions Using R and Bioconductor. New York: Springer; 2015: 397-420.
13. Griffiths-Jones S. miRBase: the microRNA sequence database. Methods Mol Biol 2006; 342: 129-38.
14. Lu TP, Lee CY, Tsai MH, et al. miRSystem: an integrated system for characterizing enriched functions and pathways of microRNA targets. PLOS ONE 2012; 7: e42390.
15. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008; 9: 559.
16. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015; 43: D447-52.
17. Catalano C, Cogoni C, Zardo G. MicroRNA in control of gene expression: an overview of nuclear functions. Int J Mol Sci 2016; 17.
18. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4: 44-57.
19. Huang DW, Sherman BT, Tan Q, et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms for better extract biology from large gene lists. Nucleic Acids Res 2007; 35: W169-75.
20. Ashburner M, Ball CA, Blake JA, et al. Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000; 25: 25-29.
21. Kanehisa M, Furumichi M, Sato Y, Ishiguo-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. Nucleic Acids Res 2021; 49: D545-51.
22. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000; 28: 27-30.
23. Lu TP, Lee CY, Tsai MH, et al. miRSystem: an integrated system for characterizing enriched functions and pathways of microRNA targets. PLOS ONE 2012; 7: e42390.
24. Zhang F, Zhang R, Zhang X, et al. microRNA targets. PLOS ONE 2012; 7: e42390.
25. Catalanotto C, Cogoni C, Zardo G. MicroRNA in control of gene expression: an overview of nuclear functions. Int J Mol Sci 2016; 17.
26. Wang J, Wang LJ, Zhong Y, et al. CETP gene polymorphisms and risk of coronary atherosclerosis in a Chinese population. Lipids Health Dis 2013; 12: 176.
27. Naz F, Shahtabaz M, Khan S, et al. PKR-inhibitor binds efficiently with human microtubule affinity-regulating kinase 4. J Mol Graph Model 2015; 62: 245-52.
28. Mohammad T, Khan FI, Lobb KA, Islam A, Ahmad F, Hassan ML. Identification and evaluation of bioactive natural products as potential inhibitors of human microtubule affinity-regulating kinase 4 (Mark4). J Biomol Struct Dyn 2019; 37: 1813-29.
29. Clement M, Chen X, Chenoweth HL, et al. MARK4 (Microtubule affinity-regulating kinase 4)-dependent inflammasome activation promotes atherosclerosis-brief report. Arterioscler Thromb Vasc Biol 2019; 39: 1645-51.
30. Weber C, Noels H. Atherosclerosis: current pathogenesis and therapeutic options. Nat Med 2011; 17: 1410-22.
31. Peterson KJ, Dietrich MR, McPeek MA. MicroRNAs and metazoan macroevolution: insights into canalization, complexity, and the Cambrian explosion. BioEssays 2009; 31: 736-47.
32. Caritag O, Navarro A, Moreno I, et al. Identifying high-risk Stage II colon cancer patients: A three-microRNA-based score as a prognostic biomarker. Clin Colorectal Cancer 2016; 15: e175-82.
33. Jia Z, Zhang Y, Li Q, et al. A coronary artery disease-associated tRNAThr mutation altered mitochondrial function, apoptosis and angiogenesis. Nucleic Acids Res 2019; 47: 2056-74.
34. Panet R, Velasquez S, Durrani S, et al. MicroRNA-1825 induces proliferation of adult cardiomyocytes and promotes cardiac regeneration post ischemic injury. Am J Transl Res 2017; 9: 3120-37.
35. Olsson AH, Yang BT, Hall E, et al. Decreased expression of genes involved in oxidative phosphorylation in human pancreatic islets from patients with type 2 diabetes. Eur J Endocrinol 2011; 165: 589-95.
36. Meng C, Jin X, Xia L, et al. Alterations of mitochondrial enzymes contribute to cardiac hypertrophy before hypertension development in spontaneously hypertensive rats. J Proteome Res 2009; 8: 2463-75.