Uncovering potential diagnostic biomarkers of acute myocardial infarction based on machine learning and analyzing its relationship with immune cells

Ling Kang[1†], Qiang Zhao[1†], Ke Jiang[1*], Xiaoyan Yu[2], Hui Chao[2], Lijuan Yin[1] and Yueqing Wang[1*]

Abstract

Background Acute myocardial infarction (AMI) is a common cardiovascular disease. This study aimed to mine biomarkers associated with AMI to aid in clinical diagnosis and management.

Methods All mRNA and miRNA data were downloaded from public database. Differentially expressed mRNAs (DEmRNAs) and differentially expressed miRNAs (DEmiRNAs) were identified using the metaMA and limma packages, respectively. Functional analysis of the DEmRNAs was performed. In order to explore the relationship between miRNA and mRNA, we construct miRNA-mRNA negative regulatory network. Potential biomarkers were identified based on machine learning. Subsequently, ROC and immune correlation analysis were performed on the identified key DEmRNA biomarkers.

Results According to the false discovery rate < 0.05, 92 DEmRNAs and 272 DEmiRNAs were identified. GSEA analysis found that kegg_peroxisome was up-regulated in AMI and kegg_steroid_hormone_biosynthesis was down-regulated in AMI compared to normal controls. 5 key DEmRNA biomarkers were identified based on machine learning, and classification diagnostic models were constructed. The random forests (RF) model has the highest accuracy. This indicates that RF model has high diagnostic value and may contribute to the early diagnosis of AMI. ROC analysis found that the area under curve of 5 key DEmRNA biomarkers were all greater than 0.7. Pearson correlation analysis showed that 5 key DEmRNA biomarkers were correlated with most of the differential infiltrating immune cells.

Conclusion The identification of new molecular biomarkers provides potential research directions for exploring the molecular mechanism of AMI. Furthermore, it is important to explore new diagnostic genetic biomarkers for the diagnosis and treatment of AMI.

Keywords Acute myocardial infarction, Biomarker, Machine learning, Functional enrichment, Differentially expressed mRNAs, miRNA-mRNA network

†Ling Kang and Qiang Zhao contributed equally to this paper

*Correspondence:
Ke Jiang
jiangganke@126.com
Yueqing Wang
tafang.2007@163.com
[1] Department of Cardiology, The Second Affiliated Hospital of Shandong First Medical University, No. 706, Taishan Street, Taian 271000, Shandong, China

[2] Coronary Care Unit, The Second Affiliated Hospital of Shandong First Medical University, No. 706, Taishan Street, Taian 271000, Shandong, China

Kang *et al. BMC Cardiovascular Disorders*   (2023) 23:2

Page 2 of 12

Introduction

Acute myocardial infarction (AMI) is a necrotic event caused by an unstable ischemic syndrome with high morbidity and mortality. AMI can result in vascular disintegration and thinning of capillaries in the infarcted area [1]. Plaque rupture, plaque erosion and coronary microvascular dysfunction are common risk factors for myocardial infarction (MI) [2]. In addition to high mortality, AMI poses a severe economic burden. The early diagnosis and treatment of AMI can not only improve the therapeutic effect, but also reduce the economic pressure of patients [3]. Therefore, development of new AMI biomarkers is beneficial to improve diagnosis, treatment and management.

Diagnosis and prognostic detection of diseases at the molecular level are the current trend of medical development. The pathogenesis of AMI is complex, and its specific pathological mechanism remains unclear. However, several genes have been reported to be involved in the pathogenesis of AMI. DNA sequence variants in GATA binding protein 5 promoter can increase susceptibility to AMI [4]. Proprotein convertase subtilisin/kexin type 6 (PCSK6) play a key role in cardiac remodeling after acute myocardial infarction [5]. MiR-382-5p can affect apoptosis of myocardial cells after AMI by regulating stanniocalcin-1 (STC1) expression [6]. Overexpression of miR-147 can inhibit myocardial inflammation and apoptosis after MI and improve cardiac function by targeting homeodomain interacting protein kinase 2 (HIPK2) [7]. In addition, immunity also plays an important regulatory role in the progression of AMI. C-X-C motif chemokine receptor 4 affects tissue repair after MI by regulating regulatory T cells [8]. Activated B cells are involved in the sustained state of myocardial inflammation after AMI, and may affect the metabolism of myocardial collagen after AMI by secreting cytokines [9]. These studies suggest that studying the molecular mechanisms of AMI may be beneficial for the realization of new management approaches.

Machine learning is a growing field of medicine that integrates computer science and statistics into medical problems. It plays an important role in clinical diagnosis, precision treatment and health monitoring of diseases. Machine learning algorithms have been used to analyze the results of biomedical datasets, including random forests (RF), decision tree (DT), and support vector machine (SVM) [10]. Based on machine learning, clinical influencing factors of AMI were studied [11], the 1-year mortality of AMI was predicted [12], and a prediction model of arrhythmia after AMI was established [13]. Furthermore, a RF diagnostic model for AMI based on ferroptosis-related genes in circulating endothelial cells has been developed [14]. To date, machine learning has rarely been used to identify potential biomarkers of AMI. Therefore, in order to identify potential biomarkers of AMI, we performed machine learning algorithm analysis. In this study, all mRNA and miRNA data were downloaded from Gene Expression Omnibus (GEO) database. Subsequently, differentially expressed mRNAs (DEmRNAs) and differentially expressed miRNAs (DEmiRNAs) were identified and mRNA-mRNA network was constructed. Furthermore, 5 key DEmRNA biomarkers (ANPEP, REPS2, TUBB2A, ZNF281 and ARHGEF3) were identified based on machine learning and a classification diagnosis model was constructed. Our study has important value for understanding the pathological mechanism of AMI and exploring new diagnostic genetic biomarkers for the diagnosis and treatment of the disease.

Materials and methods

Source and processing of data

All mRNA and miRNA data were downloaded from the GEO database [15]. The keyword "acute myocardial infarction" and "homo sapiens" were searched in the GEO database. Datasets that meet the following criteria will be included in the study: (1) Dataset must be genome-wide mRNA or miRNA transcriptome data; (2) Data were obtained from blood samples of patients with AMI and

Table 1 Details of the included dataset

GEO accession	Author	Platform	Samples (normal controls: AMI)	Year	Tissue	Data type
GSE66360	Eri Kramer	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	50:49	2015	Blood	mRNA
GSE48060	XING LI	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	21:31	2014	Blood	mRNA
GSE34198	Michal Kolář	GPL6102 Illumina human-6 v2.0 expression beadchip	48:49	2014	Blood	mRNA
GSE97320	Bo Fan Meng	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	3:3	2017	Blood	mRNA
GSE31568	Andreas Keller	GPL9040 febit Homo Sapiens miRBase 13.0	70:20	2011	Blood	miRNA

Kang *et al. BMC Cardiovascular Disorders*   (2023) 23:2

Page 3 of 12

normal controls; (3) Standardized or original datasets were considered. A total of 4 mRNA datasets (GSE66360, GSE48060, GSE34198 and GSE97320) and 1 miRNA dataset (GSE31568) were selected (Table 1). Download raw data from GSE66360, GSE48060, GSE34198 and GSE97320 and remove probes corresponding to multiple genes. For genes corresponding to multiple probes, only one probe with the highest average expression was retained. Scale normalization was performed on 17,044 mRNAs common to the 4 datasets. Subsequently, the datasets were merged and batch effects were removed with the ComBat function from the sva package.

Differential expression analysis of mRNAs and miRNAs
DEmRNAs between AMI and normal controls were analyzed using metaMA package. The screening criterion was false discovery rate (FDR) < 0.05. DEmiRNAs between AMI and normal controls were analyzed using limma package. The screening criterion was FDR < 0.05.

Functional enrichment analysis
To understand the biological functions of DEmRNAs, the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analysis of DEmRNAs was performed using genecodis database (https://genecodis.genyo.es/). GO enrichment analysis includes biological process (BP) terms, cell composition (CC) terms and molecular function (MF) terms. KEGG contains numerous signaling pathways [16–18]. The screening criterion was FDR < 0.05. In addition, "c2. cp.kegg.v7.4.symbols.gmt" was selected as the reference gene set, and GSEA 4.1.0 software was used for gene set enrichment analysis. The screening criterion was $P < 0.05$.

Construction of miRNA-mRNA network
To explore the targeting relationship between DEmiR-NAs and DEmRNAs, miRWalk (http://mirwalk.umm. uni-heidelberg.de/interactions/) was used to perform targeting relationship prediction for DEmiRNA. Relational pairs that were validated in at least one database (TargetScan, miRDB, MiRTarBase) were selected. Negative regulatory pairs involving DEmRNAs were used to construct the miRNA-mRNA network. Cytoscape (www. cytoscape.org/) was used to visualize the miRNA-mRNA network. Subsequently, GO and KEGG functional enrichment analysis was performed on target DEmRNAs in the miRNA-mRNA network. The screening criterion was FDR < 0.05.

Identification of key DEmRNA biomarkers based on machine learning
RF (https://cran.r-project.org/web/packages/rando mForest/) algorithm was used to construct classification model based on GSE66360 dataset. The importance of target DEmRNAs was sorted from large to small according to "Mean Decrease Accuracy" value. According to the sorting order, add one DEmRNAs from top to bottom. Then, RF algorithm was used for classification, and ten-fold cross-validation was used to obtain accuracy and area under curve (AUC). Then, key DEm-RNA biomarkers were selected to construct SVM, RF, and DT classification models. The accuracy, sensitivity, specificity and AUC values in the receiver operating characteristic (ROC) curve of the three classification models obtained by the tenfold cross-validation were used to evaluate the potential diagnostic ability of classification model. To further verify the potential diagnostic value of the classification models constructed by the key DEmRNA biomarkers, we performed hold-out validation. The GSE66360 data was randomly divided into two datasets, test1 (Control:AMI = 25:25) and test2 (Control:AMI = 25:24). The accuracy, sensitivity, specificity and AUC values in the ROC curves of the three classification models were obtained by hold-out validation. In addition, classification models were also constructed using key DEmRNA biomarkers based on the total data (GSE66360, GSE48060, GSE34198 and GSE97320). The ten-fold cross-validation was used to obtain AUC.

Diagnostic and expression of key DEmRNA biomarkers
To further investigate the potential diagnostic value of key DEmRNA biomarkers, we performed diagnostic analyses using GSE66360 dataset. The ROC analysis was performed using pROC package in R language to predict the diagnostic accuracy of key DEmRNA biomarkers. The sensitivity and specificity at the cut-offs were determined according to a previous study [19]. In ROC analysis, the greater the AUC, the higher the diagnostic accuracy [20]. AUC > 0.6 indicates sufficient diagnostic accuracy. Boxplots of key DEmRNA biomarkers in AMI and normal controls were also drawn. In addition, the expression of key DEmRNA biomarkers was also validated on the GSE60993 dataset.

Immune correlation analysis
The CIBERSORT package was used to perform analysis of immune cell infiltration level according to gene expression matrix. Statistical analysis of immune cell infiltration levels was performed using the rank-sum test. To further understand the potential role of key DEmRNA biomarkers, the correlation between immune cells and key DEm-RNA biomarkers was analyzed using Pearson correlation coefficient.

Kang *et al. BMC Cardiovascular Disorders* (2023) 23:2

Page 4 of 12

Statistical analysis

All statistics were performed using R software. The metaMA and limma packages are used to analyze DEmRNAs and DEmiRNAs, respectively. The screening criterion was FDR < 0.05. Rank-sum test was used to analyze the statistical difference of immune cell infiltration level between AMI and normal control groups. $P < 0.05$ was considered to be statistically significant.

Results

Identification of DEmRNAs and DEmiRNAs

According to the screening criteria FDR < 0.05, 92 DEmRNAs that were consistently up-regulated and down-regulated in the 4 datasets were screened (Additional file 1: Fig. 1A). Among them, 80 were up-regulated and 12 were down-regulated. Top 10 up-regulated DEmRNAs were CDA, MMP9, LRG1, MXD1, C19orf59, LY96, ADM, ANPEP, PLBD1 and SMAP2 (Additional file 4: Table S1). Inversely, top 10 down-regulated DEmRNAs were KLRG1, PSMA5, VPS29, CD38, ANAPC5, LIMS1, ARHGEF3, HINT1, RPS24 and GBP5 (Additional file 4: Table S1). At the same time, 272 DEmiRNAs were identified (Additional file 1: Fig. 1B). Among them, 124 were up-regulated and 148 were down-regulated. Top 10 up-regulated DEmiRNAs were hsa-miR-1290, hsa-miR-302b, hsa-miR-126*, hsa-miR-302d, hsa-miR-1468, hsa-miR-1258, hsa-miR-508-3p, hsa-miR-609, hsa-miR-27a and hsa-miR-892b (Additional file 5: Table S2). Inversely, top 10 down-regulated DEmiRNAs were hsa-miR-1283, hsa-miR-31*, hsa-miR-518a-3p, hsa-miR-519e*, hsa-miR-488*, hsa-miR-566, hsa-miR-1278, hsa-miR-1291, hsa-miR-515-5p, hsa-miR-591 (Additional file 5: Table S2).

Functional enrichment analysis of DEmRNAs

In order to explore the biological functions of DEmRNAs, functional enrichment analysis was performed. In GO: BP, DEmRNAs were mainly involved in neutrophil degranulation, signal transduction and inflammatory response (Fig. 1A). In GO: CC, DEmRNAs were mainly distributed in the cytoplasm, membrane and cytosol (Fig. 1B). In GO: MF, DEmRNAs were mainly associated with protein binding, metal ion binding and identical protein binding (Fig. 1C). KEGG analysis showed that DEmRNAs was significantly enriched in metabolic pathways, pathways in cancer and amyotrophic lateral sclerosis (Fig. 1D). Furthermore, GSEA analysis found that gene set kegg_peroxisome ($P = 0.037$) was up-regulated in AMI (Fig. 1E) and gene set kegg_steroid_hormone_biosynthesis ($P = 0.047$) was down-regulated in AMI (Fig. 1F) compared to normal controls.

Construction of miRNA-mRNA network and functional enrichment of target DEmRNAs

MiRWalk was used to predict the target mRNAs of 272 DEmiRNAs. Negative regulatory pairs involving DEmRNAs were used to construct the miRNA-mRNA network. Cytoscape was used to visualize the miRNA-mRNA network. A total of 132 negatively regulated targeting relationship pairs were obtained in the miRNA-mRNA network. In addition, the 132 negative regulatory targeting relationship pairs include 26 DEmRNAs and 35 DEmiRNAs (Fig. 2A). GO and KEGG functional enrichment analysis of 26 target DEmRNAs was performed using the genecodis database. In GO: BP, target DEmRNAs were mainly involved in positive regulation of cell migration and response to manganese ion (Fig. 2B). In GO: CC, target DEmRNAs were mainly distributed in the nucleus and membrane (Fig. 2C). In GO: MF, target DEmRNAs were mainly associated with protein binding and metal ion binding (Fig. 2D). KEGG analysis showed that target DEmRNAs was significantly enriched in metabolic pathways, phagosome and oxytocin signaling pathway (Fig. 2E).

Identification of key DEmRNA biomarkers and construction of classification models

Based on "Mean Decrease Accuracy" value, the importance of 26 target DEmRNAs was sorted from large to small (Fig. 3A). According to the sorting order, add one DEmRNAs from top to bottom. Then, RF algorithm was used for classification, and ten-fold cross-validation was used to obtain accuracy and AUC. The results showed that accuracy and AUC value reached the maximum value when the number of DEmRNA reached 5 (Fig. 3B and C). Therefore, these 5 DEmRNAs (ANPEP, REPS2, TUBB2A, ZNF281 and ARHGEF3) were selected as key biomarkers. SVM, RF and DT classification models were constructed based on 5 key DEmRNA biomarkers. The results show that the RF model has the highest accuracy (0.859) (Table 2). The AUC in the ROC curve of SVM, RF and DT were 0.922, 0.962, and 0.880, respectively (Fig. 3D–F). To further verify the potential diagnostic value of the classification models constructed by the 5 key DEmRNA biomarkers, we performed hold-out validation. The results show that the accuracy of the RF classification model for test1 and test2 datasets was greater than 0.8 (Additional file 6: Table S3). In test1 data, the AUC of SVM, RF and DT were 0.862, 0.934 and 0.863, respectively (Additional file 2: Fig. 2A–C). In test2 data, the AUC of SVM, RF and DT were 0.930, 0.958 and 0.763, respectively (Additional file 2: Fig. 2D–F). In addition, three classification models were also constructed using 5 key DEmRNA biomarkers based on the total data (GSE66360, GSE48060, GSE34198 and GSE97320). The
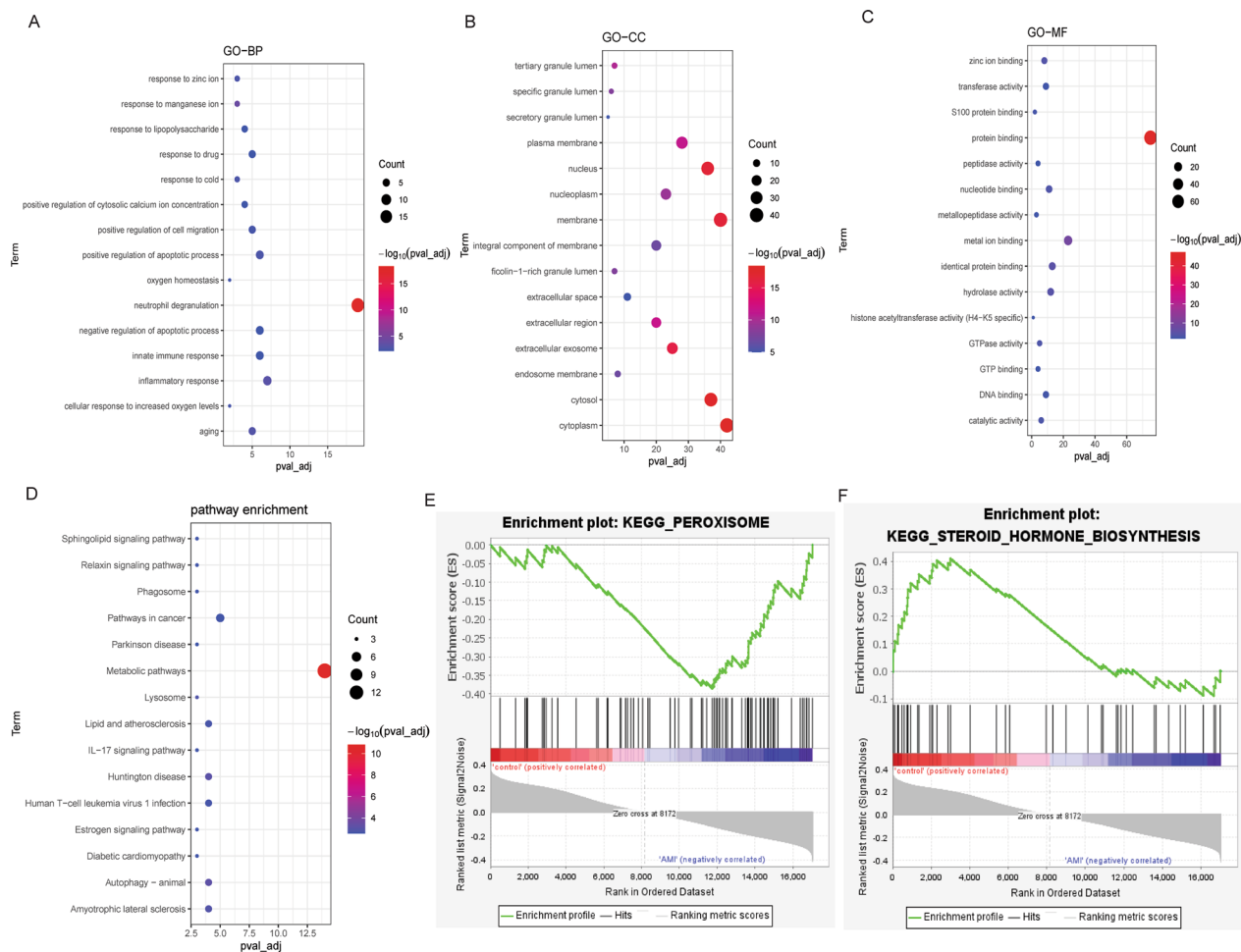
Fig. 1 Functional enrichment analysis of DEmRNAs. **A** Top 15 significantly enriched Gene Ontology (GO): biological process (BP) terms; **B** Top 15 significantly enriched GO: cell composition (CC) terms; **C** Top 15 significantly enriched GO: molecular function (MF) terms; **D** Top 15 signaling pathways of Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. **E** Gene Set Enrichment Analysis (GSEA) analysis showed that steroid hormone biosynthesis signaling pathway was down-regulated in acute myocardial infarction (AMI); **F** GSEA analysis showed that peroxisome signaling pathway was down-regulated in AMI

tenfold cross-validation was used to obtain AUC. The AUC of SVM, RF and DT were 0.719, 0.706, and 0.603, respectively (Additional file 2: Fig. 2G–I). These results imply that the classification models based on 5 key DEm-RNA biomarkers have sufficient diagnostic accuracy.

Diagnostic and expression of key DEmRNA biomarkers

ROC analysis was performed for 5 key DEmRNA bio-markers to evaluate their diagnostic value. The AUC of 5 key DEmRNA biomarkers were all greater than 0.7 (Fig. 4A–E). The result showed that these 5 key DEm-RNA biomarkers may be considered as the potential diagnostic biomarkers in AMI. However, in GSE66360 dataset, compared with SVM, RF and DT classification models, the AUC of 5 key DEmRNA biomarkers was lower than that of all models. This further implies that

classification models may have higher potential diag-nostic value compared with individual key biomarkers. Furthermore, ANPEP, REPS2, TUBB2A and ZNF281 were up-regulated in AMI (Fig. 4F–I) and ARHGEF3 was down-regulated (Fig. 4J) in AMI compared with normal controls. Subsequently, mRNA-miRNA sub-networks containing only key DEmRNA biomarkers were screened from the mRNA-miRNA network. Include 5 key DEm-RNA biomarkers and 8 DEmiRNAs in the sub-network (Fig. 5). Subsequently, the expression of 5 key DEm-RNA biomarkers was verified in the GSE60993 dataset. The results showed that the expression trend of 5 key DEmRNA biomarkers was consistent with that in the GSE66360 dataset (Additional file 3: Fig. 3).
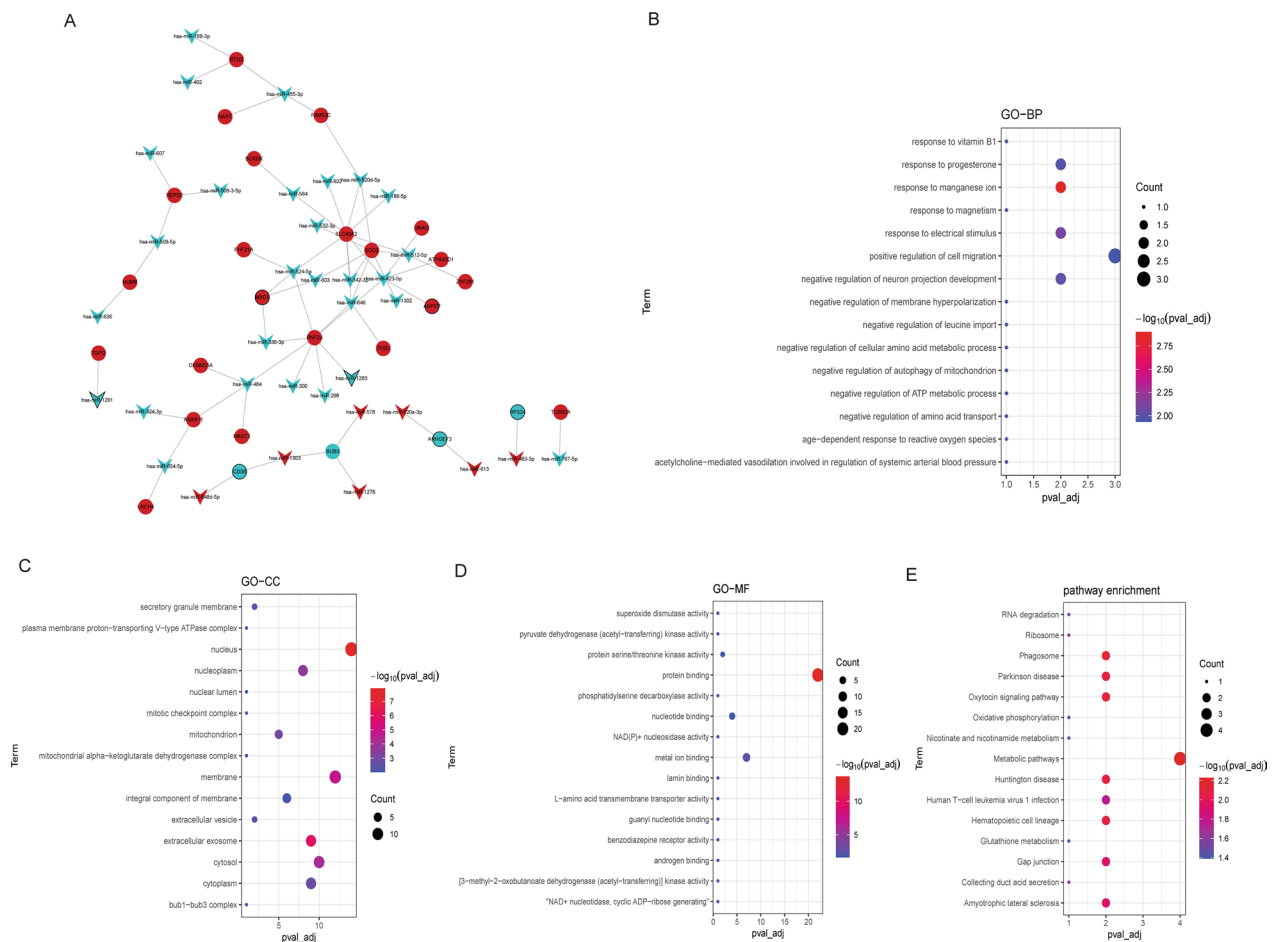
Fig. 2 Construction of miRNA-mRNA network and functional enrichment of target DEmRNAs. **A** MiRNAs-mRNA network. V-shape, circle, blue and red represent miRNA, mRNA, down-regulation, and up-regulation, respectively. The thick black borders are the up-regulated or down-regulated DEmiRNAs or DEmRNAs in top10. **B** Top 15 significantly enriched Gene Ontology (GO): biological process (BP) terms; **C** Top 15 significantly enriched GO: cell composition (CC) terms; **D** Top 15 significantly enriched GO: molecular function (MF) terms **E** Top 15 signaling pathways of Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis

Correlation analysis between immune cells and key DEmRNA biomarkers

The distribution of 12 immune cell types was significantly different between normal controls and AMI by rank-sum test (Fig. 6A). Pearson correlation analysis showed that 5 key DEmRNA biomarkers were correlated with most of the differential infiltrating immune cells (Fig. 6B). The results demonstrated that REPS2 was significantly positively correlated with neutrophils. Moreover, ANPEP was significantly positively correlated with neutrophils and monocytes, and negatively correlated with T cells CD4 memory resting. This implies that key biomarkers play important roles in the immune regulation of AMI.

Discussion

AMI is one of the main diseases threatening human life. Early diagnosis and treatment can reduce mortality and improve the prognosis of AMI. In order to identify potentially effective biomarkers, we analyzed AMI data based on GEO database. In this study, 92 DEmRNAs and 272 DEmiRNAs were identified. Functional analysis showed that DEmRNAs were enriched in various biological processes, suggesting the complexity of AMI pathogenesis. Then, 5 key DEmRNA biomarkers (ANPEP, REPS2, TUBB2A, ZNF281 and ARHGEF3) were obtained by RF analysis of 26 target DEmRNAs based on the scale data of 17,044 mRNAs, and classification diagnosis models of RF, DT and SVM were constructed. The AUC in the ROC curve of SVM, RF and DT were 0.922, 0.962, and 0.880, respectively. Moreover, the RF model has the highest accuracy (0.859). These results suggest that RF model
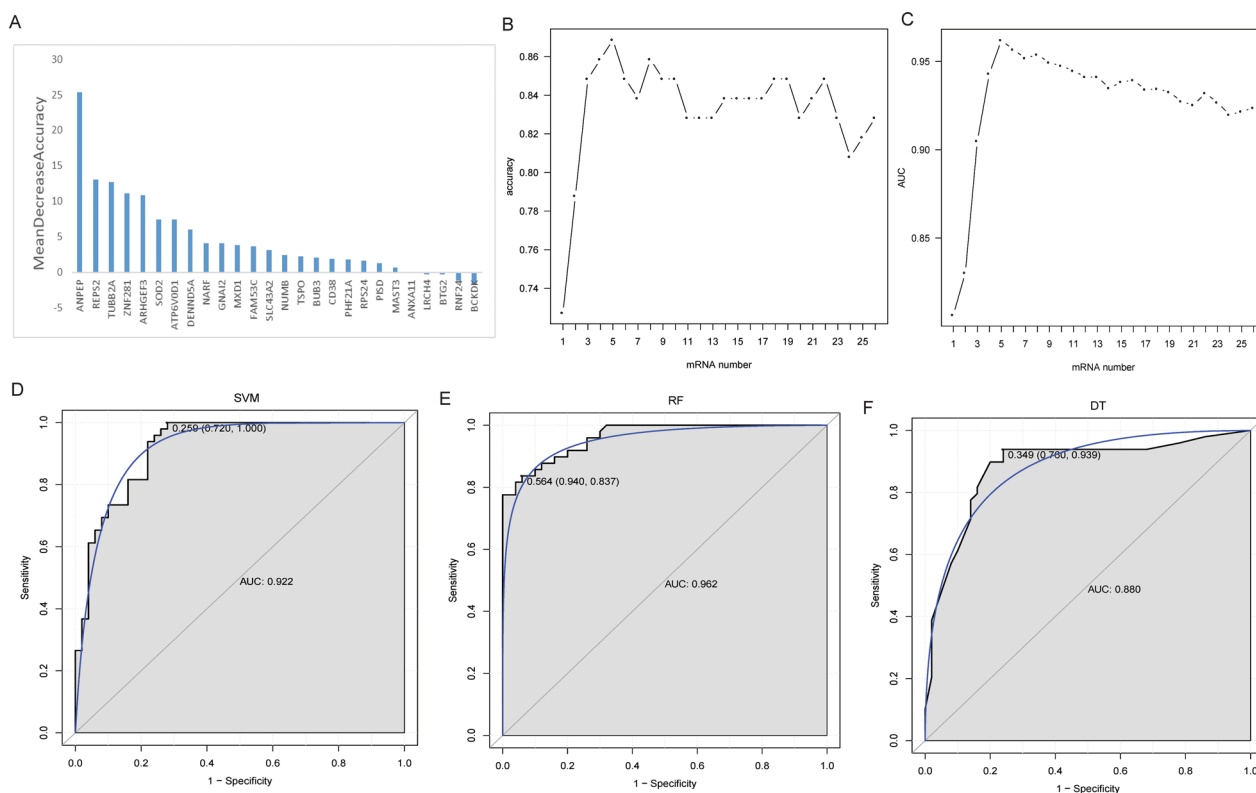
Kang *et al. BMC Cardiovascular Disorders*   (2023) 23:2

Page 7 of 12

Fig. 3 Identification of key DEmRNA biomarkers and construction of classification models. **A** The importance ranking map of 26 target DEmRNAs; **B** Trend chart of accuracy with the increase of DEmRNA quantity; **C** Trend chart of area under curve (AUC) with the increase of DEmRNA quantity; **D** Receiver operating characteristic (ROC) curves of the support vector machine (SVM) classifier; **E** ROC curves of the random forests (RF) classifier; **F** ROC curves of the decision tree (DT) classifier

Table 2 Tenfold cross-validation results of each model

Classifier	Accuracy	Sensitivity	Specificity	Area under curve (AUC)
Support vector machine (SVM)	0.818	0.796	0.840	0.922
Random forest (RF)	0.859	0.898	0.820	0.962
Decision tree (DT)	0.818	0.776	0.860	0.880

has high diagnostic value and may contribute to the early diagnosis of AMI. To further investigate the potential diagnostic value of 5 key DEmRNA biomarkers, we performed ROC analyses. The AUC of 5 key DEmRNA biomarkers were all greater than 0.7. The result showed that 5 key DEmRNA biomarkers may be considered as potential diagnostic biomarkers in AMI. However, compared with SVM, RF and DT classification models, the AUC of 5 key DEmRNA biomarkers was lower than that of all models. This further implies that classification models may have higher potential diagnostic value compared with individual key biomarkers. In addition, we also

found 5 key biomarkers associated with immune cells. This implies that key biomarkers play important roles in the immune regulation of AMI.

Alanyl aminopeptidase, membrane (ANPEP, also known as CD13 and APN) is essential for inflammatory transport and infarct healing after permanent coronary artery occlusion [21]. ANPEP can regulate the repair of atherosclerotic vascular injury [22]. The disorder of ANPEP is also related to the pathogenesis of hypertension [23]. To our knowledge, this is the first study to show that ANPEP is differentially expressed in AMI and may be a potential diagnostic biomarker for AMI. KEGG analysis
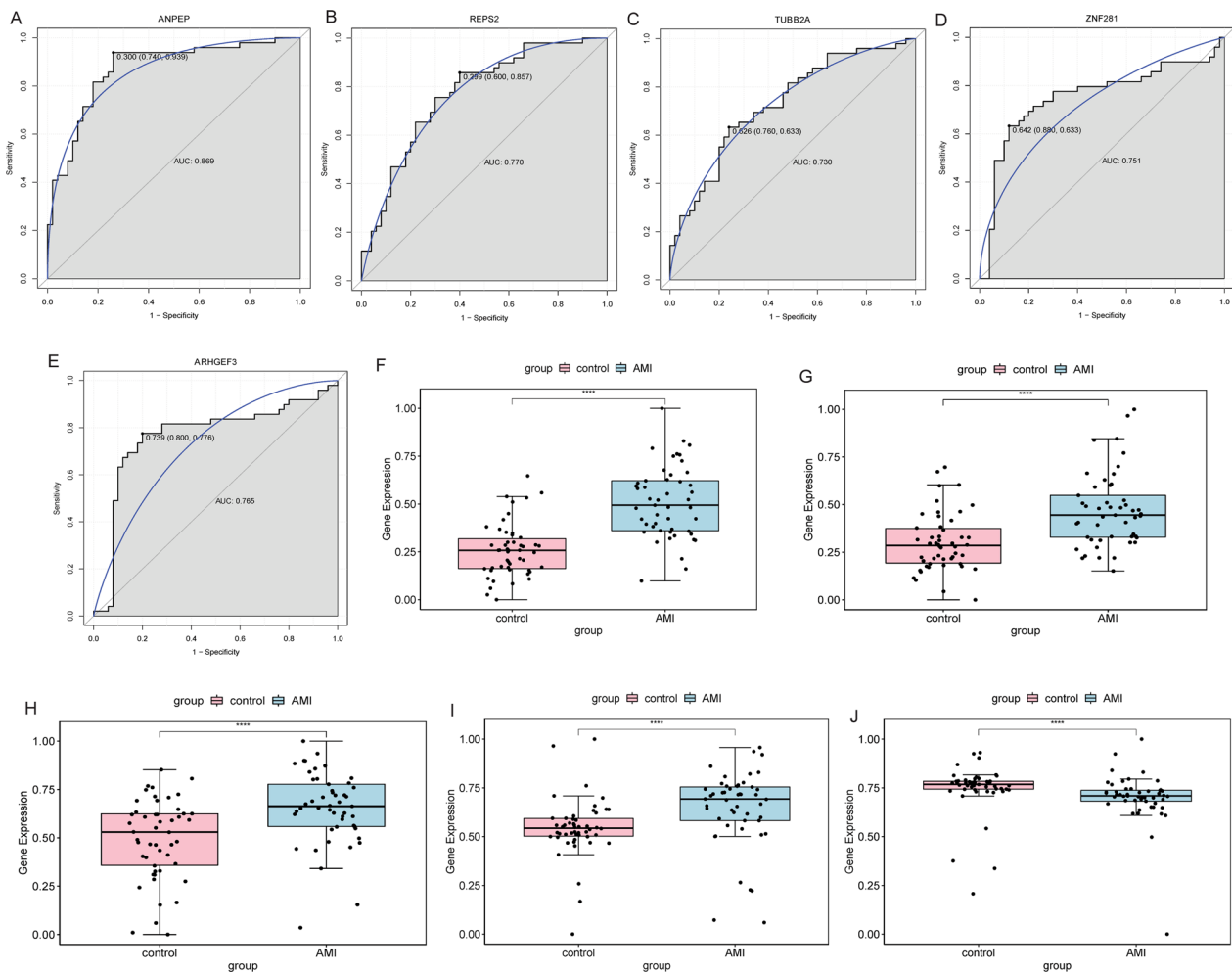
Fig. 4 Receiver operating characteristic (ROC) diagnostic analysis and expression analysis of key DEmRNA biomarkers in GSE66360 dataset. **A** ROC diagnostic analysis of the ANPEP; **B** ROC diagnostic analysis of the REPS2; **C** ROC diagnostic analysis of the TUBB2A; **D** ROC diagnostic analysis of the ZNF281; **E** ROC diagnostic analysis of the ARHGEF3; **F** Expression analysis of ANPEP; **G** Expression analysis of REPS2; **H** Expression analysis of TUBB2A; **I** Expression analysis of ZNF281; **J** Expression analysis of ARHGEF3

found that metabolic pathways, glutathione metabolism, hematopoietic cell lineage, and renin-angiotensin system were significantly enriched signaling pathways (FDR < 0.05), and involved multiple DEmRNAs including ANPEP. Patients with AMI have abnormal metabolic pathways [24, 25]. Abnormal glutathione levels in AMI are associated with oxidative stress mechanisms [26]. Hematopoietic stem cells have potential regenerative capacity in AMI xenotransplantation [27]. Wang S et al. also found that hematopoietic cell lineage has potential regulatory roles in AMI [28]. Renin-angiotensin system plays an important role in the mediation of AMI [29]. Exosomes prevent AMI in rats by modulating the renin-angiotensin system [30]. Therefore, we speculate that ANPEP may play an important role in the ongoing case mechanism of AMI by regulating metabolic pathways,

glutathione metabolism, hematopoietic cell lineage, and renin-angiotensin system with other mRNAs. In addition, we also predicted that ANPEP (up-regulated) and hsa-miR-423-5p (down-regulated) have negative targeting relationship based on the miRWalk database. Previous studies have found that hsa-miR-423-5p is abnormally expressed in AMI and is associated with the mortality of cardiogenic shock [31–33]. Based on this study, we speculated that ANPEP may be a potential target gene for hsa-miR-423-5p to play a regulatory role in AMI. Remarkably, we also found that ANPEP was significantly associated with neutrophils, monocytes and T cells CD4 memory resting. Neutrophils [34], monocytes [35] and T cells CD4 memory resting [36] play important roles in the pathophysiology of AMI. This further
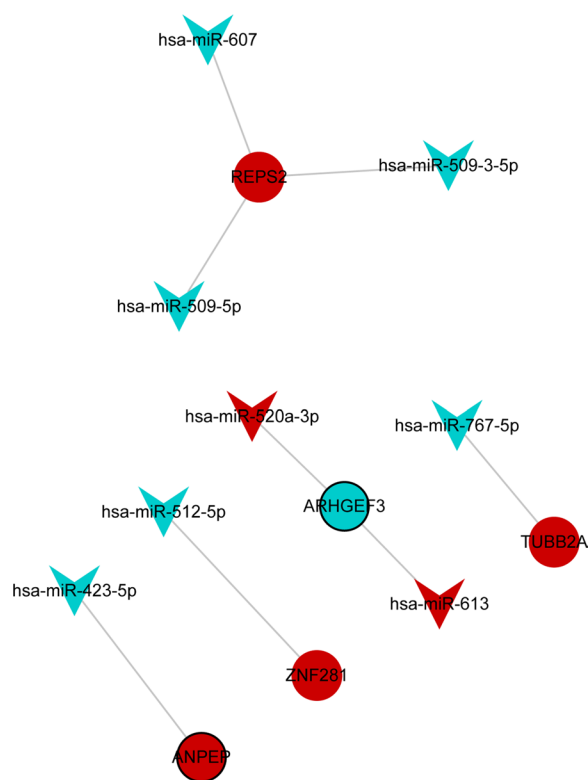
Fig. 5 MiRNA-key DEmRNA biomarkers sub-network. V-shape, circle, blue and red represent miRNA, mRNA, down-regulation, and up-regulation, respectively. The thick black borders are the up-regulated or down-regulated DEmiRNAs or DEmRNAs in top10

suggests that ANPEP may mediate the progression of AMI by regulating immune cells.

Tubulin beta 2A class IIa (TUBB2A) protein is a major component of microtubule structure [37]. Previous studies have found that TUBB2A is up-regulated in acute myocardial infarction, which is consistent with our results [38]. KEGG analysis found that Parkinson disease, gap junction and phagosomes were significantly enriched signaling pathways (FDR < 0.05), and involved multiple DEmRNAs including TUBB2A. Studies have found that Parkinson's patients are more likely to suffer from AMI [39]. Enhancement of gap junction function during AMI improves healing and reduces susceptibility to late ventricular arrhythmias [40]. Phagosomes are involved in tissue remodeling, clearing apoptotic cells and limiting the spread of intracellular pathogens [41]. Phagosomes are closely related to autophagy [42]. Autophagosomes are present in surviving cardiomyocytes in the chronic phase of MI to inhibit apoptosis and mitigate harmful effects, which may alleviate cardiac dysfunction and enhance remodeling [43]. Therefore, we speculate that TUBB2A may play an important role in the ongoing case mechanism of AMI by regulating Parkinson disease, gap junction and phagosomes with other mRNAs. In addition, we also predicted that TUBB2A (up-regulated) and hsa-miR-767-5p (down-regulated) have negative targeting relationship based on the miRWalk database. However, there is no evidence for the role of miR-767-5p in AMI. These results provide a direction for further exploring the molecular mechanism of TUBB2A in AMI.

Previous studies on RALBP1 associated Eps domain containing 2 (REPS2, also known as POB1) mostly focus on cancer, such as breast cancer [44], esophageal squamous cell carcinoma [45] and prostate cancer [46]. Transcription Factor zinc finger protein 281 (ZNF281) play a regulatory role in intestinal fibrosis [47]. ZNF281 acts at a nexus of cardiac and inflammatory gene programs, which exert influences on fibroblast and cardiac reprogramming [48]. In this study, rho guanine nucleotide exchange factor 3 (ARHGEF3, also known as XPLN) was the only down-regulated DEmRNA biomarker among 5 key DEmRNA biomarkers. A study has shown that ARHGEF3 plays a regulatory role in pulmonary fibrosis through mTORC2 [49]. Modulation of ARHGEF3 gene expression plays a role in human megakaryocytes and platelet function [50]. To our knowledge, this is the first study to show a potential regulatory role of REPS2, ZNF281 and ARHGEF3 in AMI progression. This provides a new theoretical basis for studying the molecular mechanism of AMI. In addition, we also predicted that REPS2 (up-regulated) and hsa-miR-509–3-5p (down-regulated), hsa-miR-509-5p (down-regulated) and hsa-miR-607 (down-regulated) have negative targeting relationships, ZNF281 (up-regulated) and hsa-miR-512-5p (down-regulated) have negative targeting relationships, and ARHGEF3 (down-regulated) and hsa-miR-520a-3p (up-regulated) and hsa-miR-613 (up-regulated) have negative targeting relationships based on the miRWalk database. These results indicate that REPS2, ZNF281 and ARHGEF3 are regulated by other molecules in AMI. Remarkably, REPS2 was significantly positively correlated with neutrophils. This suggests that REPS2 may play a role in the dysregulation of immune regulation in AMI. The identification of REPS2, ZNF281 and ARHGEF3 and related miRNAs provides potential research directions for further exploring the molecular mechanism of AMI.

GSEA analysis found that peroxisome was up-regulated in AMI and steroid hormone biosynthesis was down-regulated in AMI compared to normal controls. The peroxisome proliferator-activated receptor-gamma plays a role in promoting cardiac healing after AMI [51]. Peroxisome proliferator-activated receptor down-regulates the expression of pro-inflammatory molecules in MI [52]. Sex hormone is an important regulator of acute inflammatory response after cardiac injury [53]. Hormone therapy with estrogen can reduce the risk of AMI [54].
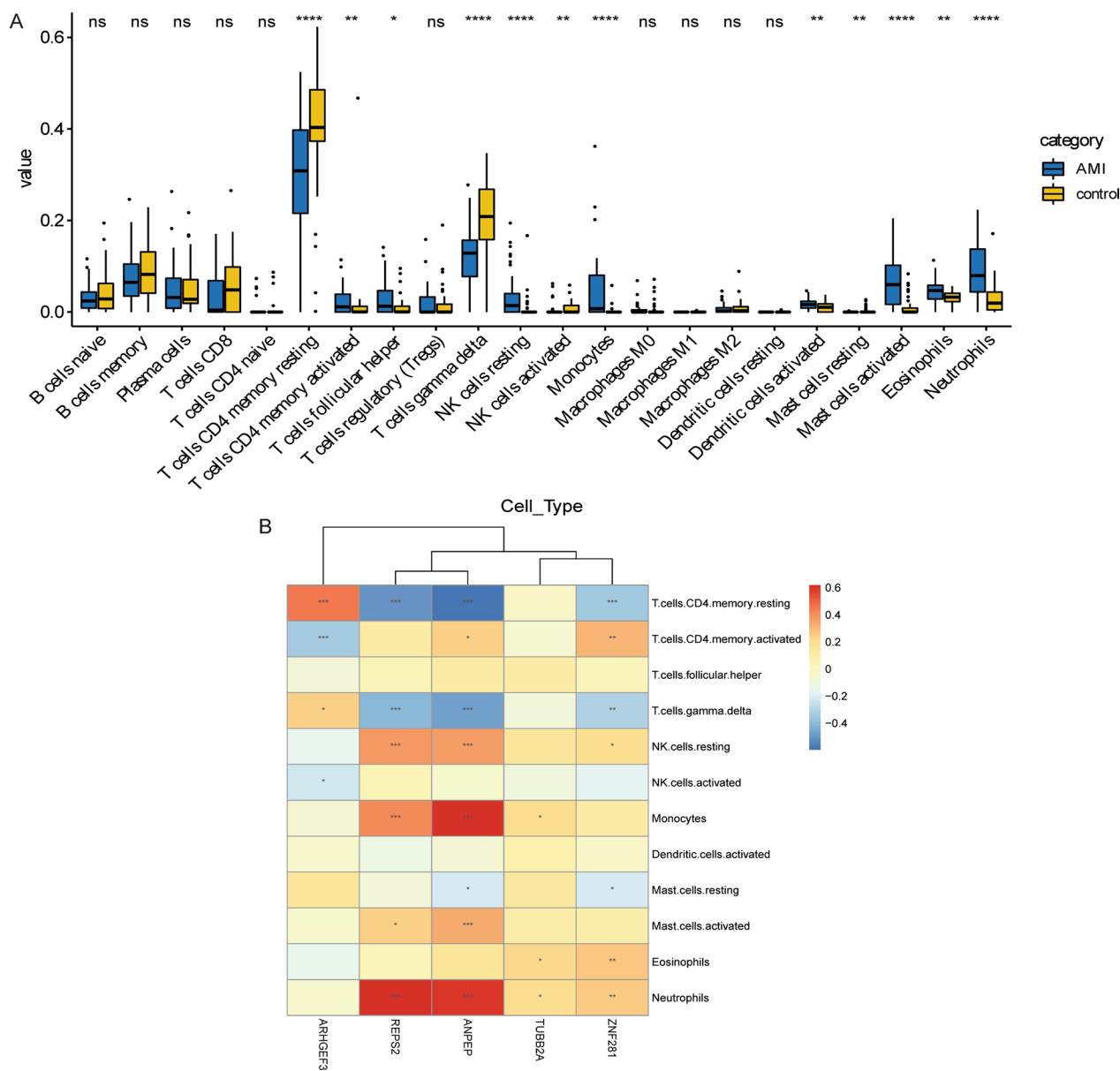
Kang *et al. BMC Cardiovascular Disorders*   (2023) 23:2

Page 10 of 12

Fig. 6 Immune correlation analyses. **A** Distribution of immune cells between normal control and AMI groups; **B** Pearson correlation coefficient analysis the correlation between immune cells and 5 key DEmRNA biomarkers. The darker the color, the stronger the correlation. Red and blue represent positive and negative correlations, respectively. * represent $P < 0.05$, ** represent $P < 0.01$, *** represent $P < 0.001$, **** represent $P < 0.0001$; ns represent no statistical significance

Therefore, further exploration of the molecular mechanisms of peroxisome and steroid hormone biosynthesis signaling pathways are conducive to understanding the molecular regulatory mechanism of AMI.

However, this study has a certain degree of limitations. All data in this study came from public databases, lacking clinical sample validation. Therefore, clinical samples need to be collected for further research at a later stage. Furthermore, the specific roles of the identified key mRNA biomarkers and related biological pathways and miRNAs in AMI remain unclear. Therefore, further studies in vitro are required to understand the molecular mechanism of AMI.

In summary, 5 key DEmRNA biomarkers were obtained by RF analysis of 26 target DEmRNAs based on the scale data of 17,044 mRNAs, and classification diagnosis models of RF, DT and SVM were constructed. The AUC in the ROC curve of SVM, RF and DT were 0.922, 0.962, and 0.880, respectively. Moreover, the

Kang *et al. BMC Cardiovascular Disorders*   (2023) 23:2

Page 11 of 12

RF model also has the highest accuracy. These results indicate that the RF model has high diagnostic value and may contribute to the early diagnosis of AMI. In addition, we found 5 key biomarkers associated with immune cells. Identification of new molecular biomarkers provides potential research directions for exploring the molecular mechanism of AMI. Furthermore, it is important to explore new diagnostic genetic biomarkers for the diagnosis and treatment of AMI.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12872-022-02999-7.

Additional file 1: Fig. S1 Volcano plot of DEmRNAs **A** and Volcano plot of DEmiRNAs **B**. DEmRNAs, differentially expressed mRNAs; DEmiRNAs, differentially expressed miRNAs.

Additional file 2: Fig. S2 Validation of classification models constructed by 5 key DEmRNA biomarkers. **A**: Receiver operating characteristic (ROC) curves of the support vector machine (SVM) classifier in test1; **B**: ROC curves of the random forests (RF) classifier in test1; **C**: ROC curves of the decision tree (DT) classifier in test1; **D**: ROC curves of the SVM classifier in test2; **E**: ROC curves of the RF classifier in test2; **F**: ROC curves of the DT classifier in test2; **G**: ROC curves of the SVM classifier in total data (GSE66360, GSE48060, GSE34198 and GSE97320); **H**: ROC curves of the RF classifier in total data; **I**: ROC curves of the DT classifier in total data.

Additional file 3: Fig. S3 Expression analysis of key DEmRNA biomarkers in GSE60993 dataset. **A**: Expression analysis of ANPEP; **B**: Expression analysis of REPS2; **C**: Expression analysis of TUBB2A; **D**: Expression analysis of ZNF281; **E**: Expression analysis of ARHGEF3. * represent $P < 0.05$, ** represent $P < 0.01$; ns represent no statistical significance.

Additional file 4: Table S1 Top 10 up-regulated and down-regulated DEmRNAs.

Additional file 5: Table S2 Top 10 up-regulated and down-regulated DEmiRNAs.

Additional file 6: Table S3 hold-out validation results of each model.

Author contributions
Conception and design: LK and QZ; Administrative support: KJ and YW; Provide materials and samples: XY and HC; Data collection and collation: LY; Data analysis and interpretation: LK and QZ. All authors have made important contributions to data analysis, drafting the article or revising the article. All authors read and approved by the final manuscript.

Availability of data and materials
The database analysed during the current study are available in the GEO database, persistent accessible web link to database is https://www.ncbi.nlm.nih.gov/geo/. Accession numbers of the datasets used in the current study are GSE66360, GSE48060, GSE34198, GSE97320 and GSE31568 in Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo). All data generated or analyzed during this study are included in this published article.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Conflict of interest
The authors declare that they have no conflict of interest.

Competing interests
The authors declare no competing interests.

References
1. Wu X, Reboll MR, Korf-Klingebiel M, Wollert KC. Angiogenesis after acute myocardial infarction. Cardiovasc Res. 2021;117(5):1257–73.
2. Gulati R, Behfar A, Narula J, Kanwar A, Lerman A, Cooper L, et al. Acute myocardial infarction in young individuals. Mayo Clin Proc. 2020;95(1):136–56.
3. Liu W, Shen J. Pyroptosis inhibition improves the symptom of acute myocardial infarction. Cell Death Disease. 2021;12(10):852.
4. Song Z, Chen L, Pang S, Yan B. Molecular genetic study on GATA5 gene promoter in acute myocardial infarction. PLoS ONE. 2021;16(3): e0248203.
5. Kuhn TC, Knobel J, Burkert-Rettenmaier S, Li X, Meyer IS, Jungmann A, et al. Secretome analysis of cardiomyocytes identifies PCSK6 (proprotein convertase subtilisin/Kexin type 6) as a novel player in cardiac remodeling after myocardial infarction. Circulation. 2020;141(20):1628–44.
6. Zhang L, Zhu H, Teng X, Sheng X, Yu B. Modulation of miR-382-5p reduces apoptosis of myocardial cells after acute myocardial infarction. Autoimmunity. 2021;54(4):195–203.
7. Wu CG, Huang C. MicroRNA-147 inhibits myocardial inflammation and apoptosis following myocardial infarction via targeting HIPK2. Eur Rev Med Pharmacol Sci. 2020;24(11):6279–87.
8. Wang Y, Dembowsky K, Chevalier E, Stüve P, Korf-Klingebiel M, Lochner M, et al. C-X-C motif chemokine receptor 4 blockade promotes tissue repair after myocardial infarction by enhancing regulatory T cell mobilization and immune-regulatory function. Circulation. 2019;139(15):1798–812.
9. Mo F, Luo Y, Yan Y, Li J, Lai S, Wu W. Are activated B cells involved in the process of myocardial fibrosis after acute myocardial infarction? An in vivo experiment. 2021;21(1):5.
10. Deng F, Huang J, Yuan X, Cheng C. Performance and efficiency of machine learning algorithms for analyzing rectangular biomedical data. Lab Invest. 2021;101(4):430–41.
11. Du H, Feng L. Clinical influencing factors of acute myocardial infarction based on improved machine learning. J Healthcare Eng. 2021;2021:5569039.
12. Lee HC, Park JS, Choe JC, Ahn JH, Lee HW, Oh JH, et al. Prediction of 1-year mortality from acute myocardial infarction using machine learning. Am J Cardiol. 2020;133:23–31.
13. Wang S, Li J. Application of machine learning to predict the occurrence of arrhythmia after acute myocardial infarction. BMC Med Inform Decision Making. 2021;21(1):301.
14. Yifan C, Jianfeng S, Jun P. Development and validation of a random forest diagnostic model of acute myocardial infarction based on ferroptosis-related genes in circulating endothelial cells. Front Cardiovasc Med. 2021;8: 663509.
15. Edgar R, Domrachev M, Lash AE. gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30(1):207–10.
16. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30.
17. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. Protein Sci Publ Protein Soc. 2019;28(11):1947–51.
18. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. Nucleic Acids Res. 2021;49(D1):D545–51.
19. Šimundić AM. Measures of diagnostic accuracy: basic definitions. Ejifcc. 2009;19(4):203–11.

20. Dittrich T, Marsch S, Egli A, Rüegg S, De Marchis GM, Tschudin-Sutter S, et al. Predictors of infectious meningitis or encephalitis: the yield of cerebrospinal fluid in a cross-sectional study. BMC Infect Dis. 2020;20(1):304.
21. Pereira FE, Cronin C, Ghosh M, Zhou SY, Agosto M, Subramani J, et al. CD13 is essential for inflammatory trafficking and infarct healing following permanent coronary artery occlusion in mice. Cardiovasc Res. 2013;100(1):74–83.
22. Du H, Yang L, Zhang H, Zhang X, Shao H. LncRNA TUG1 silencing enhances proliferation and migration of ox-LDL-treated human umbilical vein endothelial cells and promotes atherosclerotic vascular injury repairing via the Runx2/ANPEP axis. Int J Cardiol. 2021;338:204–14.
23. Danziger RS. Aminopeptidase N in arterial hypertension. Heart Fail Rev. 2008;13(3):293–8.
24. Zuurbier CJ, Bertrand L, Beauloye CR, Andreadou I, Ruiz-Meana M, Jespersen NR, et al. Cardiac metabolism as a driver and therapeutic target of myocardial infarction. J Cell Molecul Med. 2020;24(11):5937–54.
25. Vignoli A, Tenori L. Differential network analysis reveals metabolic determinants associated with mortality in acute myocardial infarction patients and suggests potential mechanisms underlying different clinical scores used to predict death. J Proteome Res. 2020;19(2):949–61.
26. Iqbal MP, Ishaq M, Mehboobali N. Increased levels of erythrocyte glutathione in acute myocardial infarction: an antioxidant defence. JPMA J Pakistan Med Assoc. 2004;54(5):254–8.
27. Sondergaard CS, Hess DA, Maxwell DJ, Weinheimer C, Rosová I, Creer MH, et al. Human cord blood progenitors with high aldehyde dehydrogenase activity improve vascular density in a model of acute myocardial infarction. J Transl Med. 2010;8:24.
28. Wang S, Wang E, Chen Q, Yang Y, Xu L, Zhang X, et al. Uncovering potential lncRNAs and mRNAs in the progression from acute myocardial infarction to myocardial fibrosis to heart failure. Front Cardiovas Med. 2021;8:664044.
29. Lee HS, Kang J. Renin-angiotensin system blockade in acute myocardial infarction: is there a winner? Korean Circ J. 2020;50(11):995–7.
30. Xiao M, Zeng W, Wang J, Yao F, Peng Z, Liu G, et al. Exosomes protect against acute myocardial infarction in rats by regulating the renin-angiotensin system. Stem Cells Develop. 2021;30(12):622–31.
31. Nabiałek E, Wańha W, Kula D, Jadczyk T, Krajewska M, Kowalówka A, et al. Circulating microRNAs (miR-423-5p, miR-208a and miR-1) in acute myocardial infarction and stable coronary heart disease. Minerva Cardioangiol. 2013;61(6):627–37.
32. Rizzacasa B, Morini E, Mango R, Vancheri C, Budassi S, Massaro G, et al. MiR-423 is differentially expressed in patients with stable and unstable coronary artery disease: a pilot study. PLoS ONE. 2019;14(5):e0216363.
33. Jäntti T, Segersvärd H, Tolppanen H, Tarvasmäki T, Lassus J, Devaux Y, et al. Circulating levels of microRNA 423–5p are associated with 90 day mortality in cardiogenic shock. ESC heart failure. 2019;6(1):98–102.
34. Carbone F, Nencioni A, Mach F, Vuilleumier N, Montecucco F. Pathophysiological role of neutrophils in acute myocardial infarction. Thromb Haemost. 2013;110(3):501–14.
35. Peet C, Ivetic A, Bromage DI, Shah AM. Cardiac monocytes and macrophages after myocardial infarction. Cardiovasc Res. 2020;116(6):1101–12.
36. Zhao E, Xie H, Zhang Y. Predicting diagnostic gene biomarkers associated with immune infiltration in patients with acute myocardial infarction. Front Cardiovasc Med. 2020;7: 586871.
37. Leandro-García LJ, Leskelä S, Jara C, Gréen H, Avall-Lundqvist E, Wheeler HE, et al. Regulatory polymorphisms in *β*-tubulin IIa are associated with paclitaxel-induced peripheral neuropathy. Clinic Cancer Res Off J Am Assoc Cancer Res. 2012;18(16):4441–8.
38. Zhang T, Zhao LL, Cao X, Qi LC, Wei GQ, Liu JY, et al. Bioinformatics analysis of time series gene expression in left ventricle (LV) with acute myocardial infarction (AMI). Gene. 2014;543(2):259–67.
39. Sheen SH, Hong JB, Kim H, Kim J, Han IB, Sohn S. The relationship between Parkinson's disease and acute myocardial infarction in korea : a nationwide longitudinal cohort study. J Korean Neurosurg Soc. 2022;65(4):507.
40. Ng FS, Kalindjian JM, Cooper SA, Chowdhury RA, Patel PM, Dupont E, et al. Enhancement of gap junction function during acute myocardial infarction modifies healing and reduces late ventricular arrhythmia susceptibility. JACC Clinic Electrophysiol. 2016;2(5):574–82.
41. Garin J, Diez R, Kieffer S, Dermine JF, Duclos S, Gagnon E, et al. The phagosome proteome: insight into phagosome functions. J Cell Biol. 2001;152(1):165–80.
42. Peña-Ramos O, Chiao L, Liu X, Yu X, Yao T, He H, et al. Autophagosomes fuse to phagosomes and facilitate the degradation of apoptotic cells in Caenorhabditis elegans. Elife. 2022. https://doi.org/10.7554/eLife.72466.
43. Lavandero S, Troncoso R, Rothermel BA, Martinet W, Sadoshima J, Hill JA. Cardiovascular autophagy: concepts, controversies, and perspectives. Autophagy. 2013;9(10):1455–66.
44. Doolan P, Clynes M, Kennedy S, Mehta JP, Germano S, Ehrhardt C, et al. TMEM25, REPS2 and Meis 1: favourable prognostic and predictive biomarkers for breast cancer. Tumour Biol J Int Soc Oncodevelop Biol Med. 2009;30(4):200–9.
45. Zhang H, Duan CJ, Zhang H, Cheng YD, Zhang CF. Expression and clinical significance of REPS2 in human esophageal squamous cell carcinoma. Asian Pacific J Cancer Prevent APJCP. 2013;14(5):2851–7.
46. Oosterhoff JK, Kühne LC, Grootegoed JA, Blok LJ. EGF signalling in prostate cancer cell lines is inhibited by a high expression level of the endocytosis protein REPS2. Int J Cancer. 2005;113(4):561–7.
47. Pierdomenico M, Palone F, Cesi V, Vitali R, Mancuso AB, Cucchiara S, et al. Transcription factor ZNF281: a novel player in intestinal inflammation and fibrosis. Front Immunol. 2018;9:2907.
48. Zhou H, Morales MG, Hashimoto H, Dickson ME, Song K, Ye W, et al. ZNF281 enhances cardiac reprogramming by modulating cardiac and inflammatory gene expression. Genes Dev. 2017;31(17):1770–83.
49. Kamio K, Azuma A, Usuki J, Matsuda K, Inomata M, Nishijima N, et al. XPLN is modulated by HDAC inhibitors and negatively regulates SPARC expression by targeting mTORC2 in human lung fibroblasts. Pulm Pharmacol Ther. 2017;44:61–9.
50. Zou S, Teixeira AM, Kostadima M, Astle WJ, Radhakrishnan A, Simon LM, et al. SNP in human ARHGEF3 promoter is associated with DNase hypersensitivity, transcript level and platelet function, and Arhgef3 KO mice have increased mean platelet volume. PLoS ONE. 2017;12(5):e0178095.
51. Tokutome M, Matoba T, Nakano Y, Okahara A, Fujiwara M, Koga JI, et al. Peroxisome proliferator-activated receptor-gamma targeting nanomedicine promotes cardiac healing after acute myocardial infarction by skewing monocyte/macrophage polarization in preclinical animal models. Cardiovasc Res. 2019;115(2):419–31.
52. Ibarra-Lara Mde L, Sánchez-Aguilar M, Soria E, Torres-Narváez JC, Del Valle-Mondragón L, Cervantes-Pérez LG, et al. Peroxisome proliferator-activated receptors (PPAR) downregulate the expression of pro-inflammatory molecules in an experimental model of myocardial infarction. Can J Physiol Pharmacol. 2016;94(6):634–42.
53. Meldrum DR, Wang M, Tsai BM, Kher A, Pitcher JM, Brown JW, et al. Intracellular signaling mechanisms of sex hormones in acute myocardial inflammation and injury. Front Biosci. 2005;10:1835–67.
54. Falkeborn M, Persson I, Adami HO, Bergström R, Eaker E, Lithell H, et al. The risk of acute myocardial infarction after oestrogen and oestrogen-progestogen replacement. Br J Obstet Gynaecol. 1992;99(10):821–8.

Publisher's Note