

RESEARCH

Open Access



An exploration into the diagnostic capabilities of microRNAs for myocardial infarction using machine learning

Mehرداد Samadishadlou¹, Reza Rahbarghazi^{2,3}, Kaveh Kavousi⁴ and Farhad Bani^{1,5*}

Abstract

Background MicroRNAs (miRNAs) have shown potential as diagnostic biomarkers for myocardial infarction (MI) due to their early dysregulation and stability in circulation after MI. Moreover, they play a crucial role in regulating adaptive and maladaptive responses in cardiovascular diseases, making them attractive targets for potential biomarkers. However, their potential as novel biomarkers for diagnosing cardiovascular diseases requires systematic evaluation.

Methods This study aimed to identify a miRNA biomarker panel for early-stage MI detection using bioinformatics and machine learning (ML) methods. miRNA expression data were obtained for early-stage MI patients and healthy controls from the Gene Expression Omnibus. Separate datasets were allocated for training and independent testing. Differential expression analysis was performed to identify dysregulated miRNAs in the training set. The least absolute shrinkage and selection operator (LASSO) was applied for feature selection to prioritize relevant miRNAs associated with MI. The selected miRNAs were used to develop ML models including support vector machine, Gradient Boosted, XGBoost, and a hard voting ensemble (HVE).

Results Differential expression analysis discovered 99 dysregulated miRNAs in the training set. LASSO feature selection prioritized 21 miRNAs. Ten miRNAs were identified in both the LASSO subset and independent test set. The HVE model trained with the selected miRNAs achieved an accuracy of 0.86 and AUC of 0.83 on the independent test set.

Conclusions An integrated framework for robust miRNA selection from omics data shows promise for developing accurate diagnostic models for early-stage MI detection. The HVE model demonstrated good performance despite differences between training and test datasets.

Keywords MicroRNA, Machine learning, Myocardial infarction, Bioinformatics, Diagnosis

*Correspondence:

Farhad Bani

bainf@tbzmed.ac.ir; bani.farhad56@gmail.com

¹ Department of Medical Nanotechnology, Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran

² Stem Cell Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

³ Department of Applied Cell Sciences, Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran

⁴ Laboratory of Complex Biological Systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran

⁵ Drug Applied Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

Introduction

Based on released statistics, myocardial infarction (MI) was the leading cause of human mortality in clinical settings over the past two decades and responsible for over 2 million additional deaths each year [1]. The occlusion of coronary arteries and lack of sufficient oxygen and blood support result in the occurrence of ischemic/necrotic changes in cardiomyocytes and MI [2]. In recent decades, significant efforts have been made to identify suitable biomarkers for in-time diagnosis, risk assessment, treatment monitoring, and evaluation of therapeutic protocols in individuals at risk of cardiovascular diseases (CVDs). Of



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

all the introduced diagnostic factors, cardiac troponins, especially troponin I, are reliable, and susceptible signaling molecules for accurate detection of MI individuals using published guidelines [2]. Despite these advantages, the high sensitivity of troponin-based tests has increased the concerns about false positive results, emphasizing the necessity for the development of novel diagnostic approaches with clinical relevance. To enhance the diagnostic value of current laboratory approaches in patients with MI, simultaneous analysis of other biological factors such as microRNAs (miRNAs) has been suggested [3].

The early and accurate diagnosis of MI is crucial for timely therapeutic intervention and improved patient outcomes. While cardiac troponin has been the gold standard biomarker for MI diagnosis, it has limitations in terms of sensitivity and specificity, particularly in the early hours following the onset of symptoms. Emerging research has highlighted the potential of miRNAs, including hsa-miR-1-3p, hsa-miR-19b-3p, hsa-miR-208a, hsa-miR-223-3p, hsa-miR-483-5p, and hsa-miR-499a-5p, as promising biomarkers that could complement and potentially surpass the diagnostic capabilities of cardiac troponin [4]. Moreover, the combined use of miRNAs and classical biomarkers like cardiac troponin can compensate for the limitations of individual biomarkers, leading to enhanced sensitivity and specificity in MI diagnosis [5]. Therefore, as techniques for miRNA detection and quantification continue to evolve, it is anticipated that miRNAs will become an integral part of routine diagnostic approaches in personalized medicine. This could pave the way for more targeted and tailored therapeutic interventions, enabling a more precise and individualized management of MI and potentially other cardiovascular diseases [6].

From structural aspects, miRNAs are short-length non-coding RNA sequences (18 to 22 nucleotides in length), that play a crucial role in regulating gene expression [7]. Studies have elucidated the involvement of miRNAs in the pathogenesis of MI. These biological elements can control various biological processes that are important in the function of the cardiovascular system [8]. For instance, miRNAs are involved in angiogenesis, cardiomyocyte growth, lipid metabolism, plaque formation, and heart rhythm [9]. Of note, both circulating miRNAs and local miRNAs in cardiac tissue exhibit diagnostic potential and prognostic biomarkers for various CVDs such as MI. It is thought that stability and rapid release into the bloodstream following cardiomyocyte injuries highlight the diagnostic and prognostic properties of miRNAs [8].

A recent systematic review article has compared the timing of peak expression and diagnostic accuracy of miRNAs in serum or plasma samples compared to conventional biomarkers used for MI diagnosis [5]. The

results showed that miRNAs such as hsa-miR-1-3p, hsa-miR-19b-3p, hsa-miR-208a, hsa-miR-223-3p, hsa-miR-483-5p, and hsa-miR-499a-5p exhibited earlier peak expression within 4 h after MI with better diagnostic performance compared to cardiac troponins and creatine kinase-MB enzyme. Early peak expression, adequate sensitivity and specificity, and heightened diagnostic accuracy, particularly during the critical initial hours of symptom onset, make miRNA superior elements for MI diagnosis [5]. In another systematic review article, valid miRNAs enlisted in terms of coronary artery disease (CAD) with diagnostic potential for distinguishing acute coronary syndrome (ACS) from stable CAD [8]. Of note, hsa-miR-21, hsa-miR-133, and hsa-miR-499 are valid and promising diagnostics to differentiate ACS versus stable CAD states. Unfortunately, overlapping expression of miRNA families, including hsa-miR-1, hsa-miR-133a/b, hsa-miR-208a/b, and hsa-miR-30, can be evident in both ACS and stable CAD patients. However, other miRNAs appear more specific to either ACS (hsa-miR-499, hsa-miR-1, hsa-miR-133a/b, and hsa-miR-208a/b) or stable CAD (hsa-miR-215, hsa-miR-487a, and hsa-miR-502) conditions [8].

Routinely, serum and plasma samples have also been used as suitable biofluids for miRNA isolation and MI diagnosis [2], but leveraging different blood sources for miRNA biomarker discovery may enable more accurate diagnostic performance and improved understanding of underlying MI mechanisms [10]. In a study conducted by present research group, five peripheral blood mononuclear cells (PBMCs) miRNAs, hsa-miRNA-186-5p, hsa-miRNA-29a-5p, hsa-miRNA-21-3p, hsa-miRNA-296-5p, and hsa-miRNA-197-5p) were introduced as a promising biomarkers for early-stage MI detection [11]. The potential of platelet-derived miRNAs has also been explored. In a small study of select miRNAs, hsa-miR-126 was identified as a possible novel biomarker for ST-segment elevation MI diagnosis [12]. A recent study has highlighted the potential role of miRNAs in modulating platelet function and maturity, which may impact the efficacy of antiplatelet therapy in patients with ST-segment elevation myocardial infarction (STEMI). An observational study comparing 61 STEMI patients with 50 healthy individuals revealed significant differences in the expression of several miRNAs. Notably, STEMI patients exhibited higher levels of hsa-miR-21-5p, hsa-miR-26b-5p, and hsa-miR-223-3p, while showing lower expression of hsa-miR-150-5p, hsa-miR423-5p, and hsa-miR-1180-3p compared to healthy controls. These findings suggest that specific miRNAs, particularly hsa-miR-26b-5p, may play a crucial role in regulating platelet function during acute STEMI, potentially influencing the effectiveness of antiplatelet therapy and opening new avenues for more

accurate diagnosis and personalized treatment strategies in ACS [13]. Also, exosomal miRNAs have highlighted as promising novel biomarkers for acute myocardial infarction (AMI), recently. A study involving 93 individuals (62 AMI patients and 31 healthy controls) identified exosomal miR-4516 (exo-miR-4516) and exo-miR-203 as promising diagnostic indicators for AMI. These miRNAs, along with secretory frizzled-related protein 1 (SFRP1), were found to be significantly elevated in the plasma exosomes of AMI patients compared to healthy controls. These findings suggest that a combination of exo-miR-4516, exo-miR-203, and SFRP1 could serve as a valuable tool for both AMI diagnosis and evaluation of coronary stenosis severity, potentially improving timely diagnosis and treatment strategies for AMI patients [14].

MiRNAs have also been explored for their prognostic value in CVDs. A comprehensive meta-analysis of studies from January 1989 to March 2019 demonstrated the diagnostic and prognostic utility of specific miRNAs in AMI, particularly hsa-miR-1, hsa-miR-133, hsa-miR-208, and hsa-miR-499, which showed strong diagnostic potential. Among these, hsa-miR-208 and hsa-miR-499 emerged as particularly valuable, with hsa-miR-208 exhibiting high sensitivity (0.83, 95% CI: 0.74–0.89) and specificity (0.96, 95% CI: 0.82–0.99), and hsa-miR-499 displaying similarly robust metrics (sensitivity: 0.84, 95% CI: 0.70–0.92; specificity: 0.97, 95% CI: 0.87–0.99). Additionally, hsa-miR-208 was associated with post-AMI mortality, with a hazard ratio of 1.09 (95% CI: 1.01–1.18), further supporting its prognostic value. These findings highlight the potential of miRNAs, particularly hsa-miR-208 and hsa-miR-499, as biomarkers for both the diagnosis and prognosis of AMI, offering promising pathways for enhanced cardiac care and risk assessment [15]. In a prospective nested case–control study spanning a decade, several circulating miRNAs were identified as potential predictors of fatal AMI in healthy individuals. This study, which assessed 179 miRNAs in serum samples from 212 participants, identified 10 miRNAs significantly linked to future AMI risk. A panel of five miRNAs (hsa-miR-106a-5p, hsa-miR-424-5p, hsa-let-7 g-5p, hsa-miR-144-3p, and hsa-miR-660-5p) demonstrated strong predictive power, accurately classifying 77.6% of cases across both sexes [16]. In another study, Gigante et al. analyzed plasma samples from 100 individuals who experienced major adverse coronary events (MACEs) and 100 MACE-free referents over an 11-year follow-up. They identified 55 miRNAs with significant expression differences between the two groups, with hsa-miR-145-3p associated with the highest increase in MACE risk and hsa-miR-720 associated with reduced risk. Interaction analysis revealed 16 miRNA signatures, with hsa-miR-320b playing a central role in modulating MACE risk in combination with

other miRNAs [17]. Another recent study involving 175 AMI patients and 90 healthy controls demonstrated significantly elevated circulating hsa-miR-19a levels in AMI patients compared to the control group. Notably, patients with severe atherosclerosis and > 50% vessel stenosis exhibited substantially higher hsa-miR-19a levels compared to those with normal vessels and no significant stenosis. Hsa-miR-19a levels showed significant positive correlations with established cardiac markers, including CK-MB, cardiac troponin I, and creatinine. These findings suggest that hsa-miR-19a could serve as a sensitive and specific biomarker for early AMI prognosis [18].

Machine learning (ML) methods can automatically construct computational models that capture complex relationships between input variables and outcomes of interest [19]. By utilizing available data and optimizing specified performance metrics, ML models identify multivariate patterns for making predictions on new data. This approach enables the detection of intricate signatures from combinations of multiple biomarkers that may be difficult to discern using traditional techniques [20]. Additionally, breakthroughs in omics technologies have enabled the production of biological datasets, encompassing thousands of observations across hundreds to thousands of samples. This has opening doors for ML to extract valuable biological insights from these intricate omics data landscapes [21]. Consequently, ML presents groundbreaking methods for integrating and analyzing different types of omics data, which can lead to the discovery of revolutionary biomarkers. These biomarkers have the potential to revolutionize disease prediction, patient stratification, and the development of innovative therapies. In summary, ML leverages existing omics data to uncover multidimensional patterns as the basis for making accurate predictions and revealing new biomarkers with clinical utility [22].

Here, the diagnostic potential of miRNAs was investigated for early-stage diagnosis of MI. To achieve this, a set of miRNA biomarkers was selected from microarray data using bioinformatics techniques. These miRNAs were used to train ML models, whose performance was evaluated on an independent test set. The integration of bioinformatics and ML for analyzing omics data represents a promising approach for discovering accurate and robust miRNA biomarkers for MI monitoring. More cohesive studies leveraging these techniques could yield universal miRNA signatures, paving the way for a new generation of molecular diagnostics for MI.

Methods

Data collection and preprocessing

The miRNA expression data were obtained from the Gene Expression Omnibus (GEO) database (<https://>

www.ncbi.nlm.nih.gov/geo/) under accession numbers GSE61741 [23] and GSE29532 [24]. GSE61741 was generated using the GPL9040 platform, which specifically profiles miRNA expression. This dataset contains acute MI samples and was used as the training set. GSE29532 was produced on the GPL5175 platform, which is not miRNA-specific but contains probes for some miRNAs based on standard platform annotations. GSE29532 includes samples from ST-elevation MI patients alongside healthy controls and served as the independent test set. For both datasets, whole blood of case samples was collected at hospital admission for MI before any intervention, while control samples were from individuals without any known disease.

After downloading, datasets were checked for proper normalization. Expression values were log₂ transformed if not already. For both datasets, expression profiles for early-stage MI and healthy controls were extracted for further analyses. Details on the datasets are provided in Table 1. Bioinformatics analyses including preprocessing, differential expression analysis, functional and pathway enrichment analyses, and miRNA selection were conducted using R, ver. 4.3.2, and RStudio. All plots and graphics of these sections were created using the ggplot2 R package.

Differential expression analysis

To evaluate the overall dispersion between the control and MI groups in the training set, principal component analysis (PCA) was performed and the top two principal components were extracted and plotted. Differentially expressed miRNAs (DEMs) between controls and MI were then identified using the R limma package. DEMs were defined as having $|\log_2 \text{fold change} (\log_2\text{FC})| > 1$ and adjusted p-value < 0.05 compared to healthy controls. Subsequently, the volcano plot was plotted based on log₂ fold change (log₂FC) and adjusted p-values.

Functional enrichment analysis

To investigate the potential biological functions and pathways associated with the DEMs, enrichment analysis was performed using mirPATH v.4.0 [25]. For this purpose, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis and Gene Ontology (GO) functional annotation were used for different analyses. GO

analysis examined three domains as follows; biological process (BP), cellular component (CC), and molecular function (MF). Statistically significant enrichment was achieved with p-values less than 0.05. In mirPATH-based enrichment analysis, TargetScan software v8.0 was used to predict DEMs target genes. In all other cases, default settings in miRPath v.4.0 were applied. An additional GeneRatio metric was calculated for each enriched using the below formula:

$$\text{GeneRatio} = \frac{\text{Predicted target genes in term}}{\text{Total genes in term}}$$

For each GO domain, the terms involving the highest number of DEMs were selected and then sorted by GeneRatio. The top 5 terms by GeneRatio in each GO domain were plotted. The same procedure of selecting terms with the most DEMs and sorting by GeneRatio was applied to KEGG pathway analysis. The top 20 KEGG terms were selected and plotted.

miRNA selection

To reduce the number of miRNAs for model development, the least absolute shrinkage and selection operator (LASSO) method [26] was applied to DEMs using the glmnet R package. The LASSO minimizes the residual sum of squares subject to the sum of the absolute coefficients being less than a constant. This constraint shrinks some coefficients to zero, effectively selecting a subset of miRNAs. LASSO was implemented with 10-fold cross-validation (CV) using default parameters in glmnet including Logistic Regression as the classifier. LASSO 10-fold CV coefficient profiles were plotted for DEMs. The 10-fold CV was used to determine the optimal lambda value that minimized the mean CV error. The final subset of miRNAs was selected based on the coefficients obtained from fitting the LASSO model using the optimal lambda across the entire dataset. Subsequently, miRNAs with non-zero coefficients were extracted from the model. This approach was used to identify miRNAs with potential diagnostic value. A heatmap was generated to visualize the correlation between the expression of the LASSO-selected miRNAs and the samples.

Machine learning analysis

ML analysis was implemented in Python v3.9 using Pandas and Scikit-Learn packages. The Scikit-Optimize package was utilized for hyper-parameter tuning. On the training set (GSE61741), a 0.7:0.3 split into training and validation sets was used with performance metrics reported as the average over 5-fold cross-validation. The best-performing model on the validation set was then trained on all training data and evaluated on the independent test set (GSE29532). To develop an ML model

Table 1 Detailed information about train and test datasets

GEO accession	Platform	MI samples	Healthy samples	Category
GSE61741	GPL9040	62	94	Train
GSE29532	GPL5175	8	6	Test

on the training set and evaluate its performance on the independent test set, the miRNAs must be present in both datasets. Therefore, only the subset of miRNAs that were identified by LASSO feature selection in the training set and were also present in the test set were used as features for model training. This procedure ensured the model could be applied to the test set for performance assessment. Three models were selected for evaluation: support vector machine (SVM), Gradient Boosting (GB), and XGBoost (XGB). SVM was chosen for its strong performance with low feature numbers and small sample sizes [27]. Ensemble methods GB and XGB can achieve high accuracy in complex datasets [28]. All models were trained, hyper-tuned, and tested using expression values for the 10 miRNAs that overlapped between the LASSO-selected miRNAs and miRNAs present in test datasets. The evaluation metric optimized during tuning was selected based on the initial evaluation of the models using their preset hyper-parameters. This ensured tuning focused on improving the most critical metric needing optimization for each model. The Scikit-Optimize package was utilized for hyperparameter tuning of the machine learning models. Specifically, the `gp_minimize` algorithm implementing Bayesian optimization was used. For the GB classifier, the loss function (deviance and exponential), number of estimators (10-120), Maximum depth of the individual regression estimators (`max_depth`, 2-5) and learning rate (0.0001-0.1) were tuned hyperparameters. The XGB classifier was tuned over the booster type (`gbtree` and `dart`), Maximum depth of a tree (`max_depth`, 2-10), learning rate (`eta`, 0.01-0.99), and number of estimators (10-200). For the SVM classifier, tuning was performed over the kernel type (linear, rbf, and poly), regularization parameter (`C`, 0.1-100),

kernel coefficient (`gamma`, 0.0001-1), and degree (for polynomial kernel only, 2-5). To leverage the strengths of the SVM, GB, and XGB models, a hard voting ensemble (HVE) was constructed. The HVE aggregates predictions by taking the output of each classifier for each sample and selecting the majority vote as the final prediction. The performance of the HVE was then compared to that of each individual model.

Results

Differential expression analysis

PCA was performed using all miRNA expression data from the training set. The PCA plot in Fig. 1A shows that the MI and healthy control samples are intermixed with no clear separation.

Differential expression analysis identified 99 DEMs in MI samples compared to healthy controls, with 67 up-regulated and 32 down-regulated at thresholds of $|\log_2FC| > 1$ and adjusted p-value < 0.05 . The volcano plot illustrating the DEMs was then generated utilizing the \log_2FC and $-\log_{10}$ of adjusted p-value (Fig. 1B). The full list of DEMs is presented in Table 2.

Functional enrichment analysis

Functional analysis of the DEMs was performed using GO and KEGG pathway enrichment. GO analysis revealed the DEMs were enriched in “growth factor activity” and “enzyme binding” for Molecular Function; “transcription factor complex” for Cellular Component; and “positive regulation of transcription, DNA-templated” for Biological Process. The top 5 significant terms in each GO category are shown in Fig. 2A. KEGG pathway analysis indicated the DEMs were involved with the “regulation of actin cytoskeleton”, “calcium signaling pathway”,

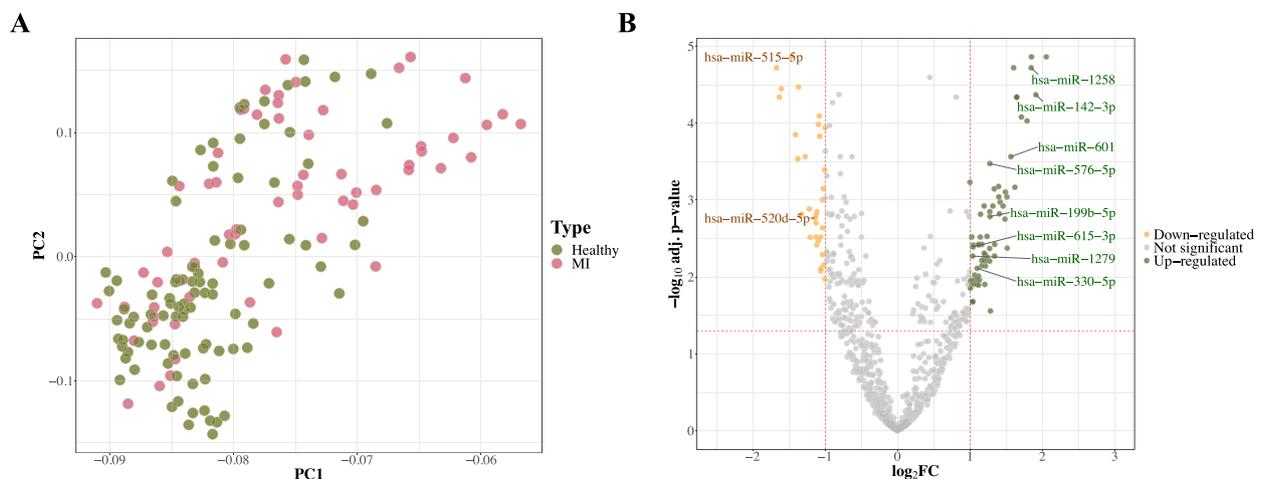


Fig. 1 Data processing and identification of DEMs. **(A)** Principal component analysis plots for the training data. **(B)** Volcano plot of DEMs

Table 2 Complete list of up- and down-regulated miRNAs in the training dataset

Up-regulated	Down-regulated
hsa-miR-375-3p, hsa-miR-29c-5p, hsa-miR-132-5p, hsa-miR-1258, hsa-miR-142-3p, hsa-miR-204-5p, hsa-miR-520c-3p, hsa-miR-1468-5p, hsa-miR-302b-3p, hsa-miR-601, hsa-miR-576-5p, hsa-miR-598-3p, hsa-miR-154-5p, hsa-miR-609, hsa-miR-302d-3p, hsa-miR-508-3p, hsa-miR-135b-5p, hsa-miR-942-5p, hsa-miR-522-3p, hsa-miR-373-3p, hsa-miR-488-3p, hsa-miR-133a-3p, hsa-miR-126-5p, hsa-miR-483-3p, hsa-miR-1262, hsa-miR-199b-5p, hsa-miR-190b-5p, hsa-miR-520f-3p, hsa-miR-1290, hsa-miR-130b-5p, hsa-miR-26a-2-3p, hsa-miR-29b-1-5p, hsa-miR-578, hsa-miR-892b, hsa-miR-1825, hsa-miR-615-3p, hsa-miR-1282, hsa-miR-450a-5p, hsa-let-7b-3p, hsa-miR-1299, hsa-miR-1279, hsa-miR-98-5p, hsa-miR-556-3p, hsa-miR-26b-3p, hsa-miR-876-3p, hsa-miR-382-5p, hsa-miR-1238-3p, hsa-miR-644a, hsa-miR-196a-5p, hsa-miR-330-5p, hsa-miR-34a-5p, hsa-miR-23a-5p, hsa-miR-509-3p, hsa-miR-1246, hsa-miR-629-5p, hsa-miR-515-3p, hsa-let-7f-2-3p, hsa-miR-454-5p, hsa-miR-340-3p, hsa-miR-133b, hsa-miR-302a-3p, hsa-miR-335-3p, hsa-miR-581, hsa-miR-613, hsa-miR-376a-5p, hsa-miR-455-5p, hsa-miR-9-5p, hsa-miR-520d-3p	hsa-miR-515-5p, hsa-miR-31-3p, hsa-miR-488-5p, hsa-miR-200a-3p, hsa-miR-1283, hsa-miR-489-3p, hsa-miR-455-3p, hsa-miR-1245a, hsa-miR-646, hsa-miR-519b-5p, hsa-miR-155-3p, hsa-miR-1291, hsa-miR-31-5p, hsa-miR-518a-5p, hsa-miR-545-5p, hsa-miR-589-5p, hsa-miR-21-3p, hsa-miR-136-3p, hsa-miR-520d-5p, hsa-miR-564, hsa-miR-933, hsa-miR-96-3p, hsa-miR-491-3p, hsa-miR-518d-5p, hsa-miR-568, hsa-miR-450b-5p, hsa-miR-10b-3p, hsa-miR-556-5p, hsa-miR-518d-3p, hsa-miR-889-3p, hsa-miR-298

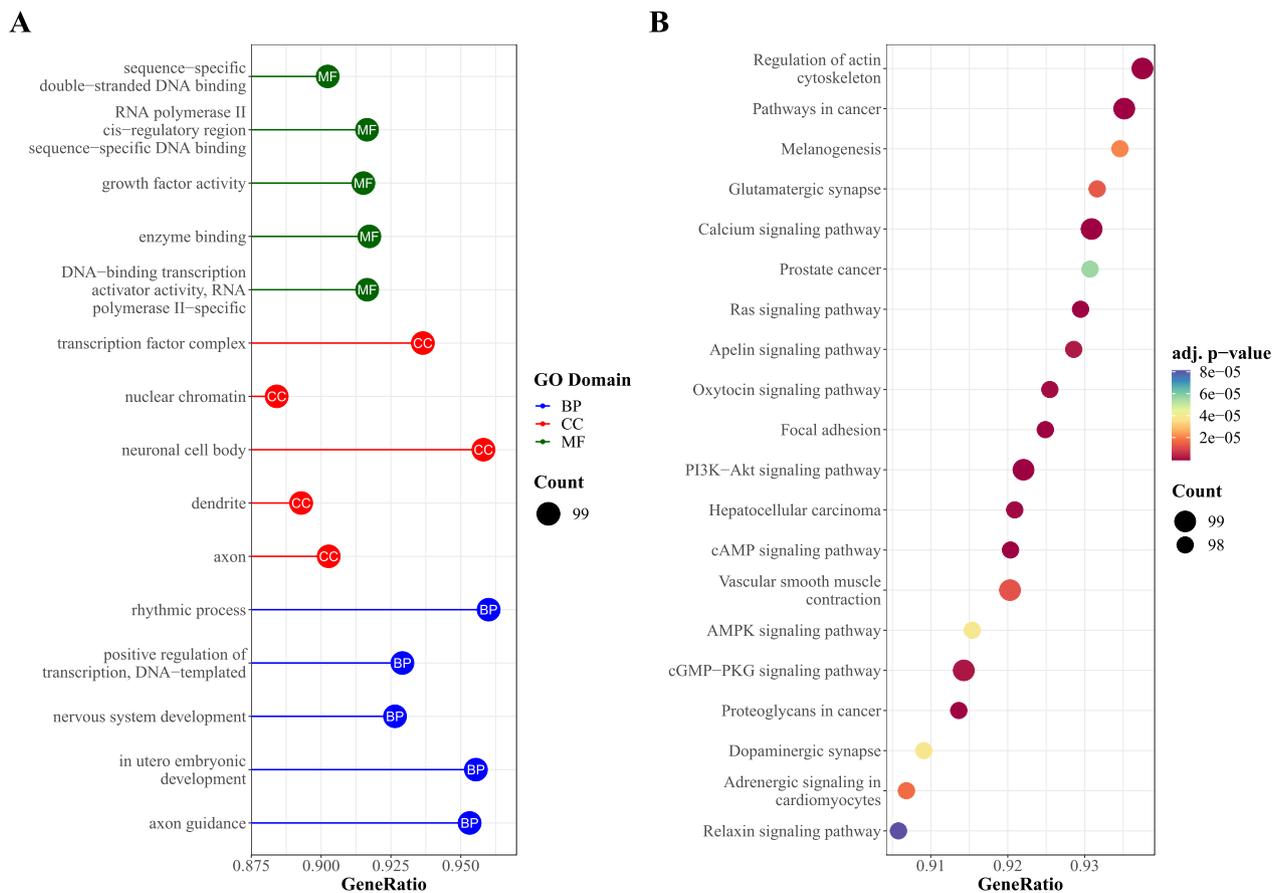


Fig. 2 Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment result in DEMs. **A** Biological process terms (BP), Cellular component terms (CC), and Molecular function terms (MF). **B** KEGG analysis

“Ras signaling pathway”, “PI3K-Akt signaling pathway”, “cAMP signaling pathway”, “AMPK signaling pathway”, and “cGMP-PKG signaling pathway.” The top 20 enriched KEGG pathways are presented in Fig. 2B.

miRNA selection

The LASSO method was applied to the 99 DEMs identified previously. LASSO coefficient profiles presented in Fig. 3A. LASSO coefficient profiles for 10-fold CV presented in Fig. 3A. The 10-fold CV LASSO analysis yielded an optimal lambda value that minimized the mean cross-validation error. This optimal lambda was subsequently employed to fit a LASSO model to the complete dataset. Based on the resulting model, 21 miRNAs with non-zero coefficients were identified. A heatmap depicting the correlation between expression of these 21 miRNAs across samples is shown in Fig. 3B.

Machine learning analysis

The test set expression data were generated using the GPL5175 platform, which contains 229 probes annotated as miRNAs based on the standard platform annotations. To utilize this dataset as an independent test set, the selected miRNAs during training must also be present in

the test data. Of the 21 LASSO-selected miRNAs from the training set, 10 were also detected in the GPL5175 platform. The miRNAs common to both the training and independent test datasets were hsa-miR-601, hsa-miR-375-3p, hsa-miR-34a-5p, hsa-miR-330-5p, hsa-miR-29c-5p, hsa-miR-200a-3p, hsa-miR-199b-5p, hsa-miR-142-3p, hsa-miR-133a-3p, and hsa-miR-132-5p (Fig. 4A). These ten miRNAs, identified through LASSO feature selection on the training set and present in the test set, were utilized as features for developing and evaluating ML models to distinguish MI from healthy samples. The expression profiles of these 10 selected miRNAs are shown in Fig. 4B and their log2FC and adjusted p-values reported in Table 3. Also, the receiver operating characteristic (ROC) curve of each individual miRNA for separating MI from healthy samples on the training set using a simple logistic regression model is presented in Fig. 5A.

The performance of the SVM, GB, and XGB models using default hyper-parameters on the validation set is reported in Table 4. The accuracy, area under the ROC curve (AUC-ROC), sensitivity, and specificity achieved were 0.76, 0.82, 1.0, and 0.0 for SVM; 0.72, 0.78, 1.0, and 0.33 for GB; and 0.75, 0.76, 0.62, and 0.50 for XGB.

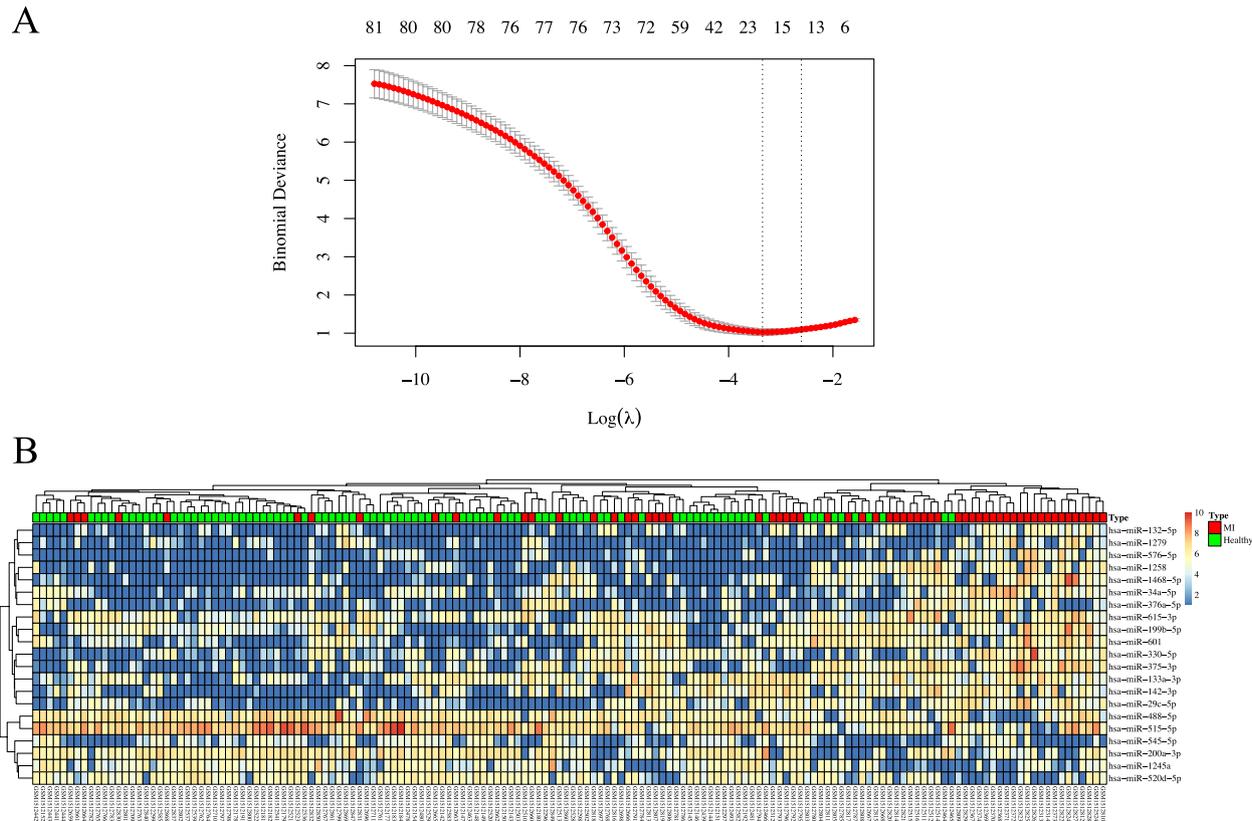


Fig. 3 Selection of the most important miRNAs. **A** LASSO coefficient profiles for DEMs and **B** Expression heatmap of 21 LASSO-selected miRNAs

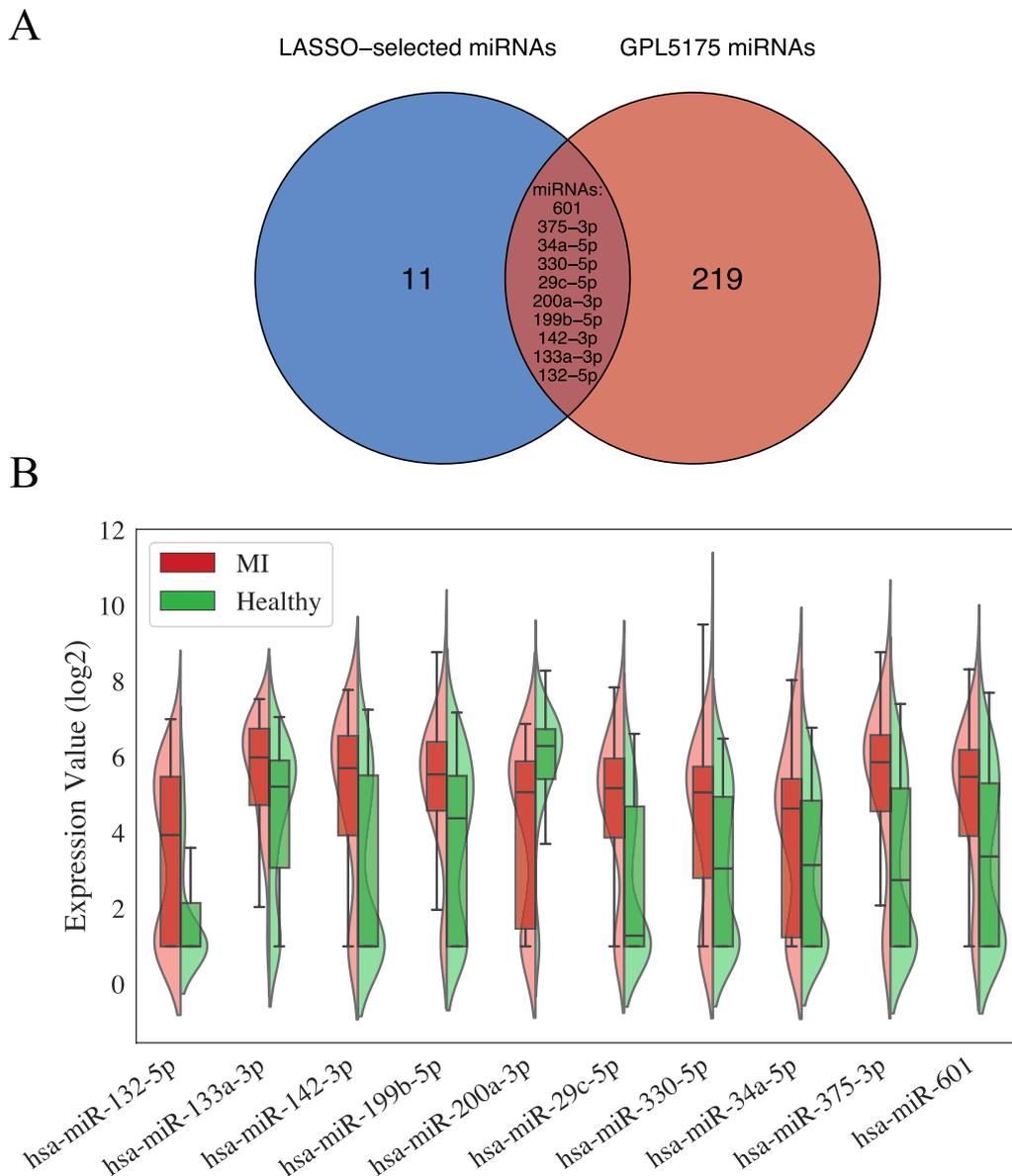


Fig. 4 **A** The intersection of LASSO-selected miRNAs and miRNAs was selected on test dataset. **B** Expression profile of selected miRNAs for training machine learning models

ROC curves are shown in Fig. 5B. All models had low specificity, which was identified as the metric requiring optimization.

Hyper-parameter tuning was performed to maximize the recall average. This approach improved model performance on the validation set such that accuracy and AUC-ROC reached 0.79 and 0.85 for the SVM model; 0.77 and 0.78 for GB; and 0.81 and 0.82 for XGB after tuning. Moreover, the performance of the optimized models on the validation set showed improved specificity across all models, with specificity reaching 0.82 for SVM, Gradient

Boosted, and XGB models. The sensitivity changed to 0.74 for SVM, 0.68 for Gradient Boosted, and 0.79 for XGB. The HVE model was created by integrating the optimized SVM, GB, and XGB models. This ensemble classifier was trained on the training data and its performance was evaluated on the validation set. The HVE maintained overall performance in comparison to the individual models, with an accuracy of 0.79, AUC-ROC of 0.82, sensitivity of 0.68, and specificity of 0.86 on the validation data. The ROC curves for the tuned individual models and HVE are shown in Fig. 5C. The performance

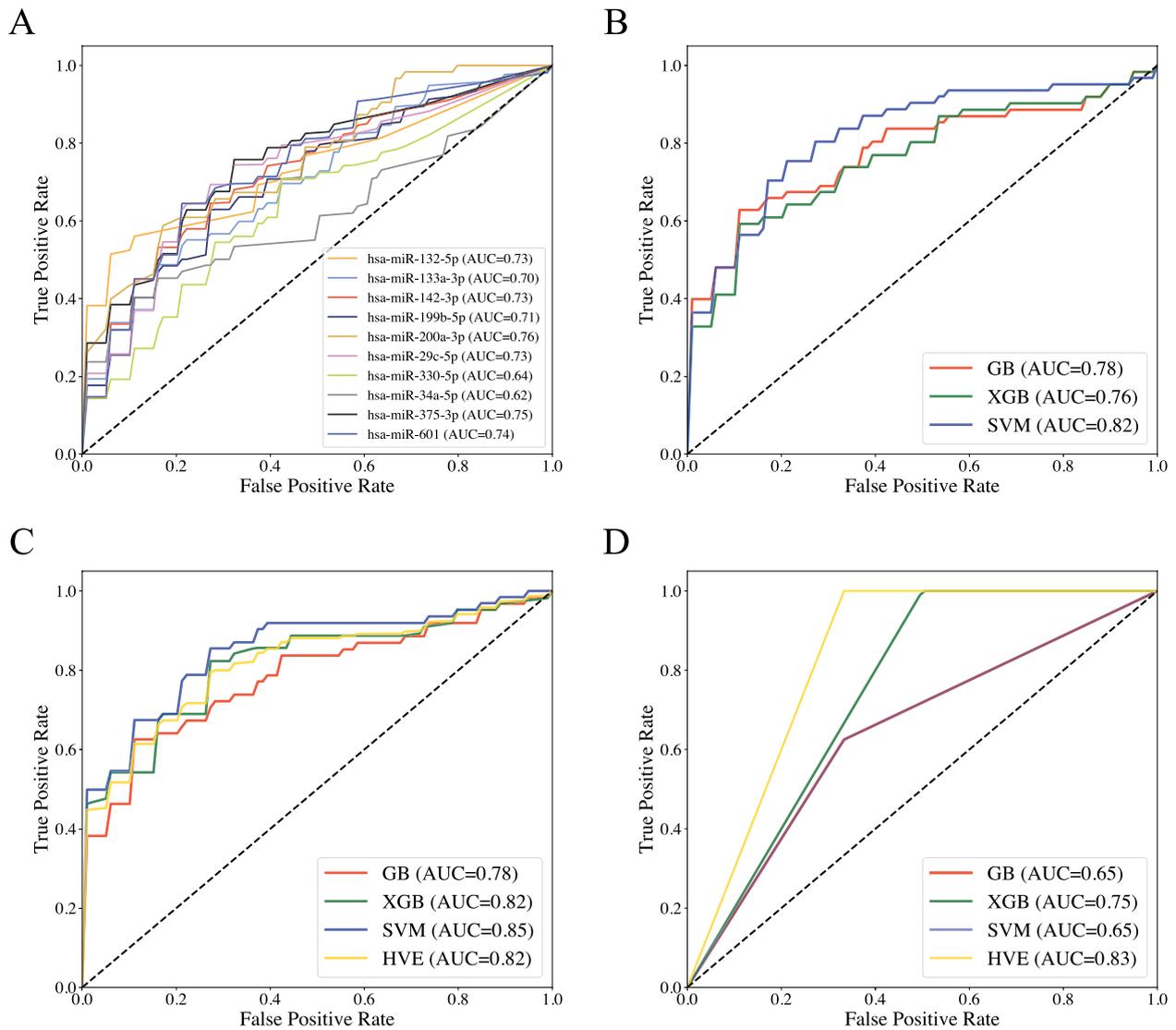


Fig. 5 ROC curve for **A** individual miRNAs on the validation set. **B** Gradient Boosted (GB), XGBoost (XGB), and support vector machine (SVM) with preset hyper-parameters on the validation set. **C** Hyper-tuned GB, XGB, and SVM and a hard voting ensemble (HVE) made of the three hyper-tuned models on the validation set. **D** GB, XGB, SVM, and HVE on test set

Table 3 Adjusted p-value and log2FC for 10 selected miRNAs

	adj. p-value	log2FC
hsa-miR-601	0.0002724	1.565302
hsa-miR-375-3p	0.0000137	2.055180
hsa-miR-34a-5p	0.0094190	1.065633
hsa-miR-330-5p	0.0076999	1.100359
hsa-miR-29c-5p	0.0000137	1.850096
hsa-miR-200a-3p	0.0000355	-1.604994
hsa-miR-199b-5p	0.0016343	1.276995
hsa-miR-142-3p	0.0000426	1.911637
hsa-miR-133a-3p	0.0011926	1.196438
hsa-miR-132-5p	0.0000190	1.603354

Table 4 Accuracy, AUC-ROC, sensitivity, and specificity for machine learning models with preset and hyper-tuned hyper-parameters on the validation set

	Accuracy	AUC-ROC	Sensitivity	Specificity
<i>Preset Models</i>				
SVM	0.76	0.82	1.00	0.00
GB	0.72	0.78	1.00	0.33
XGB	0.75	0.76	0.62	0.50
<i>Hyper-tuned Models</i>				
SVM	0.79	0.85	0.74	0.82
GB	0.77	0.78	0.68	0.82
XGB	0.81	0.82	0.79	0.82
HVE	0.79	0.82	0.68	0.86

metrics for all models on the validation set are reported in Table 4.

The optimized SVM, GB, and XGB models were trained on the combined training and validation sets and evaluated on the independent test set. Test accuracy, AUC-ROC, sensitivity, and specificity achieved were 0.64, 0.65, 0.62, and 0.67 for both SVM and Gradient Boosted models and 0.79, 0.75, 1.0, and 0.5 for XGB. A HVE of models attained improved test performance with an accuracy of 0.86 and AUC-ROC of 0.83. The HVE maintained maximal sensitivity and specificity among individual models which were 1.0 and 0.67 respectively. Prediction scores on the test set for each model, along with the HVE, are reported in Table 5. The ROC curves for all models on the test set are shown in Fig. 5D.

Discussion

The occurrence of MI in different societies leads to high-rate human mortalities with socioeconomic outcomes [1]. The early-stage diagnosis and proper treatment can reduce mortality rates in clinical settings [2, 3]. Molecular investigations have revealed that miRNA levels are changed during various biological processes [29]. Changes in miRNA expression are known to significantly contribute to MI progression and subsequent cardiac tissue damage [30]. Following MI, the occurrence of pathological changes can alter the expression of specific miRNAs associated with cardiomyocyte function [31]. According to regulatory roles of miRNAs, even small changes in their transcription can potentially be detected before alterations in mRNA and protein levels [3, 5]. These characteristics make miRNAs a potentially useful and valid early-stage diagnostic tool to identify both minor and major host cell damage. While some studies have explored serum or plasma miRNAs as indicators of cardiac tissue injury [5, 8, 32], further in-depth characterization and elucidation are still needed for the miRNA landscape across all relevant blood components. Numerous studies have reported altered miRNA profiles related to MI in specific blood components such as platelets, and

PBMCs [11, 33], but there is limited research analyzing miRNA expression changes in whole blood samples. By conducting a whole blood analysis, our study provides a more integrated view of systemic miRNA responses, encompassing signals from various cell types involved in MI pathogenesis and progression. Here, we utilized bioinformatics approaches to identify a whole blood cell miRNA biomarker panel for early-stage MI detection and developed ML models using this panel.

Initial PCA exhibited no clear separation between MI and healthy samples. To overcome this challenge, a bioinformatics pipeline was implemented for biomarker selection and model development. DEMs included 99 miRNAs with $|\log_2FC| > 1$ and adjusted p-value < 0.05 (Table 2), and further enrichment analysis confirmed their close association with MI-related pathways. Calcium homeostasis maintains cardiomyocyte excitation-contraction coupling. Therefore, disruption in the calcium signaling pathway may affect heart function, leading to the development of various heart disease [34]. The regulation of the actin cytoskeleton pathway impacts the migration and proliferation of vascular smooth muscle and cardiomyocytes in the healing process after MI. This viewpoint influences the development of cardiovascular systems and healing post-MI [35]. The Ras family of GTPases, including H-Ras and K-Ras isoforms, are critical for developing cardiac hypertrophy and heart failure through regulating cardiomyocyte size and proliferation. Thus, alterations in Ras signaling could potentially impact MI progression [36]. The PI3K-Akt pathway significantly regulates cardiomyocyte size/survival, angiogenic processes that repair ischemic damage, and inflammatory responses following MI [37]. Chronic activation of the cAMP pathway in cardiomyocytes is known to compensate for losing cells post-MI by modifying responses to extracellular stimuli [38]. AMPK signaling regulates multiple facets of MI outcomes, including cell apoptosis/survival, cardiac fibrosis, contractility, and gene expression. Manipulating this pathway suppresses cardiomyocyte death in heart failure [39, 40]. Finally, the cGMP-PKG pathway's role in platelet inhibition prevents clot occlusion of vessels post-MI [41]. Given that the samples analyzed were derived from whole blood, it is expected that the enriched pathways correspond to various blood components, beyond just cardiomyocyte-specific processes. Notably, the cGMP-PKG and platelet inhibition signaling noted indicates the involvement of circulating platelets. Additionally, the enrichment of inflammation and survival-related pathways points to signaling and gene expression changes in immune cell populations found in whole blood.

The LASSO method further narrowed the miRNA panel to 21 key miRNAs, of which 10 were also detected

Table 5 Accuracy, AUC-ROC, sensitivity, and specificity for machine learning models with preset and hyper-tuned hyper-parameters on test set

	Accuracy	AUC-ROC	Sensitivity	Specificity
<i>Hyper-tuned Models</i>				
SVM	0.64	0.65	0.62	0.67
GB	0.64	0.65	0.62	0.67
XGB	0.79	0.75	1.00	0.50
HVE	0.86	0.83	1.00	0.67

in the test data. These 10 miRNAs (hsa-miR-132-5p, hsa-miR-133a-3p, hsa-miR-142-3p, hsa-miR-199b-5p, hsa-miR-200a-3p, hsa-miR-29c-5p, hsa-miR-330-5p, hsa-miR-34a-5p, hsa-miR-375-3p, and hsa-miR-601) were used as features for training ML models. Notably, 9 of the 10 were up-regulated in MI samples compared to healthy controls, with only hsa-miR-200a-3p being down-regulated.

According to the literature, nine out of ten miRNAs presented in this study play pivotal roles in the pathophysiology of MI, and being involved in various pathological processes such as myocardial hypertrophy, fibrosis, apoptosis, angiogenesis, calcium regulation, neuroendocrine activation, and oxidative stress. Notably, miR-199b has been demonstrated to exacerbate pathological remodeling in a murine model of MI, while its pharmacological silencing exhibited therapeutic effects, proposed to arise from the interplay between the calcineurin A (CnA)/nuclear factor of activated T-cells (NFAT) and Notch signaling pathways [42]. In contrast, miR-200a exhibited protective effects in neonatal mouse ventricular cardiomyocytes subjected to oxidative stress by restoring cell viability, inhibiting apoptosis, and suppressing pro-inflammatory factors through directly targeting and inhibiting Keap1 and β -catenin, respectively [43].

In addition, miR-29 family has garnered significant attention due to its multifaceted regulatory functions in cardiovascular diseases, particularly its influence on myocardial fibrosis, fibrocyte activation, and extracellular matrix production. Both circulating and tissue/cellular levels of miR-29 are associated with the occurrence and progression of these diseases, highlighting its potential as a diagnostic biomarker and therapeutic target [44]. Additionally, miR-375-3p plays a crucial role in promoting cardiac fibrosis following ischemia-reperfusion injury, a common occurrence in MI, by accelerating ferroptosis of cardiomyocytes through the regulation of glutathione peroxidase 4 (GPX4). Treating rats or cell models with miR-375-3p antagonist and the ferroptosis inhibitor Ferrostatin-1 alleviated ischemia-reperfusion-induced cardiac fibrosis, suggesting its potential as a therapeutic target [45].

Along with the above-mentioned biomarkers, miR-132-5p is a master regulator of various pathological processes in ischemic or non-ischemic heart failure, such as myocardial hypertrophy, fibrosis, apoptosis, angiogenesis, calcium homeostasis, neuroendocrine activation, and oxidative stress. Preliminary data have suggested that antisense oligonucleotide targeting miR-132 may be a potential therapeutic approach for these conditions [46]. Moreover, miR-133a-3p has been shown to ameliorate myocardial ischemic injury and is

associated with platelet reactivity and functionality in acute coronary syndrome. Overexpression of miR-133a downregulates TGF- β 1 and connective tissue growth factor expression, potentially reducing myocardial collagen deposition, inhibiting myocardial fibrosis, and improving cardiac function after MI [47].

Also, miR-330-5p inhibits NLRP3 inflammasome-mediated myocardial ischemia-reperfusion injury by directly targeting TIM3, an inhibitor of the NLRP3 inflammasome pathway. Inhibition of miR-330-5p aggravates myocardial ischemia-reperfusion injury, suggesting its potential as a therapeutic target for MI [48]. miR-34a-5p is implicated in the regulation of apoptosis and cardiac dysfunction in MI, with its upregulation in hypoxia-induced cardiomyocytes contributing to cardiomyocyte apoptosis and impaired cardiac function by targeting genes involved in various signaling pathways related to MI pathophysiology. Inhibition of miR-34a-5p has been shown to improve cardiac function and reduce cardiac injury, indicating its potential as a therapeutic target. Moreover, elevated plasma levels of miR-34a-5p have been observed in patients with MI, highlighting its diagnostic potential as a biomarker for early-stage MI [49].

Furthermore, miR-142-3p has been found to be associated with platelet reactivity and functionality in acute coronary syndrome and has been suggested as a potential biomarker in modeling the risk of acute coronary syndrome [50]. Notably, hsa-miR-601 is the only miRNA mentioned that has not been studied for its role in MI.

Comparing individual miRNA models to using all miRNAs shows SVM and GB had higher AUC-ROC than the best single miRNA, while XGB matched the top miRNA, using default hyper-parameters. All models suffered from poor specificity initially. After tuning to optimize the recall average, validation accuracy rose for all models. AUC-ROC increased for SVM and XGB, remaining equal for GB. Sensitivity fell for SVM and GB but improved for XGB. As intended, specificity increased across the board. The HVE achieved the highest specificity on the validation set while having an average of other metrics. As expected, performance declined on the independent test set, with accuracy, AUC-ROC, and sensitivity dropping for all individual models but XGB sensitivity. However, the HVE increased accuracy, AUC-ROC, and sensitivity over individual models, despite a decrease in specificity. Through integrating the optimized models' predictions, the HVE attained equal or enhanced performance metrics compared to the best individual model across all evaluation criteria. Specifically, it retained the highest accuracy and AUC-ROC values among models, while improving specificity balance compared to the more

skewed individual models, maximizing both sensitivity and specificity.

The observed performance decline on the independent test set is attributable to platform differences between the training and test datasets, despite similar sampling protocols. Specifically, miRNA expression was profiled on completely different microarray platforms. In addition to differing miRNA profiling platforms used for the training and test sets, having only 10 of the 21 LASSO-selected miRNAs present in the test data likely impacted model performance. However, the HVE classifier still achieved satisfactory prediction accuracy in the face of these challenges.

Conclusion

miRNAs are valid signaling biomarkers for monitoring the occurrence and progression of CVDs, especially MI. Finding and introducing more specific miRNA types can help clinicians distinguish the non-CVD pathologies. While platform mismatches impacted models' performance, our HVE leveraged model complements to maintain acceptable accuracy. With a more unified profiling and sampling protocol, it is conceivable that models with lower complexity could reach even higher performance. Given the proven effectiveness of whole-blood miRNAs as biomarkers for myocardial infarction diagnosis, developing tailored profiling platforms that focus on known biologically relevant miRNAs and integrate crucial complementary information such as sex and age could facilitate the production of more consistent and comparable data. Analyzing such integrated datasets could identify robust miRNA signatures and improved ML models to enhance diagnosis. Additionally, it may provide better elucidation of mechanisms in MI pathogenesis and consistent therapeutic targets for intervention.

Abbreviations

MI	Myocardial infarction
CVD	Cardiovascular disease
miRNA	MicroRNA
CAD	Coronary artery disease
ACS	Acute coronary syndrome
PBMC	Peripheral blood mononuclear cell
STEMI	ST-segment elevation myocardial infarction
AMI	Acute myocardial infarction
MACE	Major adverse coronary events
ML	Machine learning
GEO	Gene expression omnibus
PCA	Principal component analysis
DEM	Differentially expressed microRNA
log ₂ FC	Log ₂ fold change
KEGG	Kyoto encyclopedia of genes and genomes
GO	Gene ontology
BP	Biological process
CC	Cellular component
MF	Molecular function
LASSO	Least absolute shrinkage and selection operator
CV	Cross-validation
SVM	Support vector machine

GB	Gradient boosting
XGB	XGBoost
ROC	Receiver operating characteristic
AUC-ROC	Area under the receiver operating characteristic curve
HVE	Hard voting ensemble

Author contributions

F.B. and K.K. conceived the idea and coordinated the project. M.S. researched, collected the data, performed the analyzes, assembled the results, and drafted the manuscript. F.B. and R.R. edited and revised the manuscript. F.B. is author responsible for contact and ensures communication. All authors read the content of final manuscript.

Funding

This is a report of result from Ph.D. thesis registered in Tabriz University of Medical Sciences with the Number 66372. This work was extracted from Mehrdad Samadishadlou's thesis titled "Developing and manufacturing of a paper-based Nanobiosensor in order to diagnosing myocardial infarction using a set of blood microRNAs".

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>), reference numbers GSE61741 and GSE29532. The codes associated with this article are accessible from the GitHub repository (https://github.com/MehrdadSamadishadlou/MI_miRNA_signature.git). All data generated or analysed during this study are included in this published article.

Declarations

Ethics approval and consent to participate

The study was approved by the research ethics committee of Tabriz University of Medical Sciences (approval ID: IR.TBZMED.VCR.REC.1399.388, date of approval: 2021/1/11).

Consent for publication

All authors gave consent for the publication of the article

Competing interests

The authors declare that they have no Conflict of interest.

Received: 11 June 2024 Accepted: 7 October 2024

Published online: 10 December 2024

References

- Nowbar AN, Gitto M, Howard JP, Francis DP, Al-Lamee R. Mortality from ischemic heart disease: Analysis of data from the World Health Organization and coronary artery disease risk factors From NCD Risk Factor Collaboration. *Circ Cardiovasc Qual Outcomes*. 2019;12(6): e005375. <https://doi.org/10.1161/CIRCOUTCOMES.118.005375>.
- Thygesen K, Alpert JS, Jaffe AS, Chaitman BR, Bax JJ, Morrow DA, et al. Fourth Universal Definition of Myocardial Infarction (2018). *Circulation*. 2018;138(20):e618–51. <https://doi.org/10.1161/CIR.0000000000000617>.
- Schulte C, Barwari T, Joshi A, Zeller T, Mayr M. Noncoding RNAs versus Protein Biomarkers in Cardiovascular Disease. *Trends Mol Med*. 2020;26(6):583–96. <https://doi.org/10.1016/j.molmed.2020.02.001>.
- Cheng C, Wang Q, You W, Chen M, Xia J. MiRNAs as biomarkers of myocardial infarction: a meta-analysis. *PLoS ONE*. 2014;9(2):88566.
- Wang B, Li Y, Hao X, Yang J, Han X, Li H, et al. Comparison of the Clinical Value of miRNAs and Conventional Biomarkers in AMI: A Systematic Review. *Front Genet*. 2021. <https://doi.org/10.3389/fgene.2021.668324>.
- Condrat CE, Thompson DC, Barbu MG, Bugnar OL, Boboc A, Cretoiu D, et al. miRNAs as Biomarkers in Disease: Latest Findings Regarding Their Role in Diagnosis and Prognosis. *Cells*;9(2):276. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute. <https://doi.org/10.3390/cells9020276>.

7. Schulte C, Karakas M, Zeller T. microRNAs in cardiovascular disease - clinical application. *Clin Chem Lab Med*. 2017;55(5):687–704. <https://doi.org/10.1515/cclm-2016-0576>.
8. Kaur A, Mackin ST, Schlosser K, Wong FL, Elharram M, Delles C, et al. Systematic review of microRNA biomarkers in acute coronary syndrome and stable coronary artery disease. *Cardiovasc Res*. 2020;116(6):1113–24. <https://doi.org/10.1093/cvr/cvz302>.
9. Tanase DM, Gosav EM, Ouatu A, Badescu MC, Dima N, Ganceanu-Rusu AR, et al. Current Knowledge of MicroRNAs (miRNAs) in Acute Coronary Syndrome (ACS): ST-Elevation Myocardial Infarction (STEMI). *Life*. 2021;11(10):1057. <https://doi.org/10.3390/life11101057>.
10. Dutttagupta R, Jiang R, Gollub J, Getts RC, Jones KW. Impact of Cellular miRNAs on Circulating miRNA Biomarker Signatures. *PLoS ONE*. 2011;6(6):e20769. <https://doi.org/10.1371/journal.pone.0020769>.
11. Samadishdlou M, Rahbarghazi R, Piryaei Z, Esmaili M, Avci CB, Bani F, et al. Unlocking the potential of microRNAs: machine learning identifies key biomarkers for myocardial infarction diagnosis. *Cardiovasc Diabetol*. 2023;22(1):247. <https://doi.org/10.1186/s12933-023-01957-7>.
12. Li S, Guo LZ, Kim MH, Han JY, Serebruany V. Platelet microRNA for predicting acute myocardial infarction. *J Thromb Thrombolysis*. 2017;44(4):556–64. <https://doi.org/10.1007/s12399-017-1537-6>.
13. Pedersen OB, Hvas AM, Pasalic L, Kristensen SD, Grove EL, Nissen PH. Platelet Function and Maturity and Related microRNA Expression in Whole Blood in Patients with ST-Segment Elevation Myocardial Infarction. *Thromb Haemost*. 2024;124(03):192–202. <https://doi.org/10.1055/s-0043-1776305>.
14. Liu P, Wang S, Li K, Yang Y, Man Y, Du F, et al. Exosomal microRNA-4516, microRNA-203 and SFRP1 are potential biomarkers of acute myocardial infarction Corrigendum in /10.3892/mmr.2024.13289. *Molecular Medicine Reports*. 2023 Jun;27(6):1–12. Publisher: Spandidos Publications. <https://doi.org/10.3892/mmr.2023.13010>.
15. Lee GK, Hsieh YP, Hsu SW, Lan SJ. Exploring diagnostic and prognostic predictive values of microRNAs for acute myocardial infarction. *Medicine*. 2021;100(29):e26627. <https://doi.org/10.1097/MD.00000000000026627>.
16. Bye A, Røsjø H, Nauman J, Silva GJJ, Follstad T, Omland T, et al. Circulating microRNAs predict future fatal myocardial infarction in healthy individuals - The HUNT study. *J Mol Cell Cardiol*. 2016;97:162–8. <https://doi.org/10.1016/j.jmcc.2016.05.009>.
17. Gigante B, Papa L, Bye A, Kunderfranco P, Viviani C, Roncarati R, et al. MicroRNA signatures predict early major coronary events in middle-aged men and women. *Cell Death & Disease*. 2020 Jan;11(1):1–3. Publisher: Nature Publishing Group. <https://doi.org/10.1038/s41419-020-2291-9>.
18. Mansouri F, Seyed Mohammadzad MH. Molecular miR-19a in Acute Myocardial Infarction: Novel Potential Indicators of Prognosis and Early Diagnosis. *Asian Pacific Journal of Cancer Prevention*. 2020, 21(4), 975–982. <https://doi.org/10.31557/APJCP.2020.21.4.975>
19. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. *Science*. 2015;349(6245):255–60. <https://doi.org/10.1126/science.aaa8415>.
20. Xu C, Jackson SA. Machine learning and complex biological data. *Genome Biol*. 2019;20(1):76. <https://doi.org/10.1186/s13059-019-1689-0>.
21. Joyce AR, Palsson B. The model organism as a system: integrating omics' data sets. *Nat Rev Mol Cell Biol*. 2006;7(3):198–210. <https://doi.org/10.1038/nrm1857>.
22. Ng S, Masarone S, Watson D, Barnes MR. The benefits and pitfalls of machine learning for biomarker discovery. *Cell Tissue Res*. 2023;394(1):17–31. <https://doi.org/10.1007/s00441-023-03816-z>.
23. Keller A, Leidinger P, Vogel B, Backes C, ElSharawy A, Galata V, et al. miRNAs can be generally associated with human pathologies as exemplified for miR-144*. *BMC Med*. 2014;12(1):224. <https://doi.org/10.1186/s12916-014-0224-0>.
24. Silbiger VN, Luchessi AD, Hirata RDC, Lima-Neto LG, Cavichioli D, Caracado A, et al. Novel genes detected by transcriptional profiling from whole-blood cells in patients with early onset of acute coronary syndrome. *Clin Chim Acta*. 2013;421:184–90. <https://doi.org/10.1016/j.cca.2013.03.011>.
25. Tastsoglou S, Skoufos G, Miliotis M, Karagkouni D, Koutsoukos I, Karavangeli A, et al. DIANA-miRPath v4.0: expanding target-based miRNA functional analysis in cell-type and tissue contexts. *Nucleic Acids Res*. 2023;51:W154–9. <https://doi.org/10.1093/nar/gkad431>.
26. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *J R Stat Soc Ser B Methodol*. 1996;58(1):267–88. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>.
27. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. *PLoS ONE*. 2019;14(11):e0224365. <https://doi.org/10.1371/journal.pone.0224365>.
28. Mienye ID, Sun Y. A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects. *IEEE Access*. 2022;10:99129–49. <https://doi.org/10.1109/ACCESS.2022.3207287>.
29. Kalayinia S, Arjmand F, Maleki M, Malakootian M, Singh CP. MicroRNAs: roles in cardiovascular development and disease. *Cardiovasc Pathol*. 2021;50:107296. <https://doi.org/10.1016/j.carpath.2020.107296>.
30. Lagerbauer B, Engelhardt S. MicroRNAs as therapeutic targets in cardiovascular disease. *J Clin Investig*. 2022. <https://doi.org/10.1172/JCI159179>.
31. Khan AA, Gupta V, Mahapatra NR. Key regulatory miRNAs in lipid homeostasis: Implications for cardiometabolic diseases and development of novel therapeutics. *Drug Discov Today*. 2022;27(8):2170–80. <https://doi.org/10.1016/j.drudis.2022.05.003>.
32. Guan R, Zeng K, Zhang B, Gao M, Li J, Jiang H, et al. Plasma Exosome miRNAs Profile in Patients With ST-Segment Elevation Myocardial Infarction. *Front Cardiovasc Med*. 2022. <https://doi.org/10.3389/fcvm.2022.848812>.
33. Gidlöf O, van der Brug M, Öhman J, Gilje P, Olde B, Wahlestedt C, et al. Platelets activated during myocardial infarction release functional miRNA, which can be taken up by endothelial cells and regulate ICAM1 expression. *Blood*. 2013;121(19):3908–17. <https://doi.org/10.1182/blood-2012-10-461798>.
34. Wang R, Wang M, He S, Sun G, Sun X. Targeting Calcium Homeostasis in Myocardial Ischemia/Reperfusion Injury: An Overview of Regulatory Mechanisms and Therapeutic Reagents. *Front Pharmacol*. 2020. <https://doi.org/10.3389/fphar.2020.00872>.
35. Tang DD, Gerlach BD. The roles and regulation of the actin cytoskeleton, intermediate filaments and microtubules in smooth muscle cell migration. *Respir Res*. 2017;18(1):54. <https://doi.org/10.1186/s12931-017-0544-7>.
36. Ramos-Kuri M, Meka SH, Salamanca-Buentello F, Hajjar RJ, Lipskaia L, Chemaly ER. Molecules linked to Ras signaling as therapeutic targets in cardiac pathologies. *Biol Res*. 2021;54(1):23. <https://doi.org/10.1186/s40659-021-00342-6>.
37. Ghafouri-Fard S, Khanbabapour Sasi A, Hussen BM, Shoori H, Siddiq A, Taheri M, et al. Interplay between PI3K/AKT pathway and heart disorders. *Mol Biol Rep*. 2022;49(10):9767–81. <https://doi.org/10.1007/s11033-022-07468-0>.
38. Colombe AS, Pidoux G. Cardiac cAMP-PKA Signaling Compartmentalization in Myocardial Infarction. *Cells*. 2021;10(4):922. <https://doi.org/10.3390/cells10040922>.
39. Wang L, Yu F. SCD leads to the development and progression of acute myocardial infarction through the AMPK signaling pathway. *BMC Cardiovasc Disord*. 2021;21(1):197. <https://doi.org/10.1186/s12872-021-02011-8>.
40. Liu C, Guo X, Zhou Y, Wang H. AMPK Signaling Pathway: A Potential Strategy for the Treatment of Heart Failure with Chinese Medicine. *J Inflamm Res*. 2023;16:5451–64. <https://doi.org/10.2147/JIR.S441597>.
41. Gambaryan S. The Role of NO/sGC/cGMP/PKG Signaling Pathway in Regulation of Platelet Function. *Cells*. 2022;11(22):3704. <https://doi.org/10.3390/cells11223704>.
42. Duygu B, Poels EM, Juni R, Bitsch N, Ottaviani L, Olieslagers S, de Windt LJ, da Costa Martins PA. miR-199b-5p is a regulator of left ventricular remodeling following myocardial infarction. *Non-coding RNA Res*. 2017;2(1):18–26.
43. Ma Y, Pan C, Tang X, Zhang M, Shi H, Wang T, et al. MicroRNA-200a represses myocardial infarction-related cell death and inflammation by targeting the Keap1/Nrf2 and Beta-catenin pathways. *Hellenic J Cardiol*. 2021;62(2):139–48.
44. Liu MN, Luo G, Gao WJ, Yang SJ, Zhou H. miR-29 family: A potential therapeutic target for cardiovascular disease. *Pharmacol Res*. 2021;166:105510. <https://doi.org/10.1016/j.phrs.2021.105510>.
45. Zhuang Y, Yang D, Shi S, Wang L, Yu M, Meng X, et al. MiR-375-3p Promotes Cardiac Fibrosis by Regulating the Ferroptosis Mediated by GPX4. *Computat Intell Neurosci*;2022. <https://doi.org/10.1155/2022/9629158>.
46. Xu K, Chen C, Wu Y, Wu M, Lin L. Advances in miR-132-Based Biomarker and Therapeutic Potential in the Cardiovascular System. *Front Pharmacol*.

47. Wang W, Zheng H. Myocardial Infarction: The Protective Role of MiRNAs in Myocardium Pathology. *Front Cardiovasc Med*;8. <https://doi.org/10.3389/fcvm.2021.631817>.
48. Harada M, Okuzaki D, Yamauchi A, Ishikawa S, Nomura Y, Nishimura A, et al. Circulating miR-20b-5p and miR-330-3p are novel biomarkers for progression of atrial fibrillation: Intracardiac/extracardiac plasma sample analysis by small RNA sequencing. *PLOS ONE*;18(4):e0283942. Publisher: Public Library of Science. <https://doi.org/10.1371/journal.pone.0283942>.
49. Hua CC, Liu XM, Liang LR, Wang LF, Zhong JC. Targeting the microRNA-34a as a Novel Therapeutic Strategy for Cardiovascular Diseases. *Frontiers in Cardiovascular Medicine*;8. Publisher: Frontiers. <https://doi.org/10.3389/fcvm.2021.784044>.
50. Szelenberger R, Karbownik MS, Kacprzak M, Maciak K, Bijak M, Zielińska M, et al. Screening Analysis of Platelet miRNA Profile Revealed miR-142-3p as a Potential Biomarker in Modeling the Risk of Acute Coronary Syndrome. *Cells*;10(12):3526. Number: 12 Publisher: Multidisciplinary Digital Publishing Institute. <https://doi.org/10.3390/cells10123526>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.