Comprehensive bioinformatics analytics and *in vivo* validation reveal SLC31A1 as an emerging diagnostic biomarker for acute myocardial infarction

Shujing Zhou[1,2,*], Longbin Wang[1,*], Xufeng Huang[1,2,*], Ting Wang[1], Yidan Tang[2], Ying Liu[3], Ming Xu[1]

[1]Department of Clinical Veterinary Medicine, Huazhong Agricultural University, Wuhan, China
[2]Faculty of Medicine, University of Debrecen, Debrecen, Hungary
[3]Department of Cardiology, Sixth Medical Center, PLA General Hospital, Beijing, China
*Equal contribution

Correspondence to: Ying Liu, Ming Xu; **email:** liuying@medmail.com.cn, mingxu110@mail.hzau.edu.cn

ABSTRACT

Background: Globally, Acute Myocardial Infarction (AMI) is a common cause of heart failure (HF), which has been a leading cause of mortality resulting from non-communicable diseases. On the other hand, increasing evidence suggests that the role of energy production within the mitochondria strongly links to the development and progression of heart diseases, while Cuproptosis, a newly identified cell death mechanism, has not yet been comprehensively analyzed from the aspect of cardiovascular medicine.

Materials and Methods: 8 transcriptome profiles curated from the GEO database were integrated, from which a diagnostic model based on the Stacking algorithm was established. The efficacy of the model was evaluated in a multifaced manner (i.e., by Precision-Recall curve, Receiver Operative Characteristic curve, etc.). We also sequenced our animal models at the bulk RNA level and conducted qPCR and immunohistochemical staining, with which we further validated the expression of the key contributor gene to the model. Finally, we explored the immune implications of the key contributor gene.

Results: A merged machine learning model containing 4 Cuproptosis-related genes (i.e., PDHB, CDKN2A, GLS, and SLC31A1) for robust AMI diagnosis was developed, in which SLC31A1 served as the key contributor. Through *in vivo* modeling, we validated the aberrant overexpression of SLC31A1 in AMI. Besides, further transcriptome analysis revealed that its high expression was correlated with significant potential immunological implications in the infiltration of many immune cell types, especially monocyte.

Conclusions: We constructed an AMI diagnostic model based on Cuproptosis-related genes and validated the key contributor gene in animal modeling. We also analyzed the effects on the immune system for its overexpression in AMI.

INTRODUCTION

Acute Myocardial Infarction (AMI) is one of the leading causes of death in developed countries. Globally, the incidence of this disease approaches 3 million people, posing significant challenges to global healthcare. Its etiology involves reduced coronary artery blood flow, leading to insufficient oxygen supply and subsequent myocardial ischemia. The decrease in coronary artery blood flow is multifactorial, potentially caused by the rupture of atherosclerotic plaques that lead to thrombus formation, coronary artery embolism,

abuse of cocaine that causes cardiac ischemia, coronary artery dissection, and coronary vasospasm. AMI is characterized by its sudden onset, rapid progression, and grave prognosis. Despite the extensive efforts in pharmaceutical and surgical interventions, there has been limited improvement in the incidence and mortality rates associated with AMI over the past decades [1]. Within this context, emphasizing the urgency of AMI diagnosis cannot be overstated, as timely identification is crucial for patient survival.

Currently, cardiac troponin (encoded by the TNNI3 gene) and creatine kinase MB isoenzyme (encoded by the CKM gene) serve as the gold standard biomarkers for AMI diagnosis, although their specificity and sensitivity are less satisfactory [2, 3]. Furthermore, classic risk factors such as smoking, obesity, and hypertension play pivotal roles in prevention and clinical management, although they are insufficient for immediate diagnosis [4]. As such, the pursuit of more precise diagnostic methods and a comprehensive understanding of the complex pathogenesis and risk factors are crucial for enhancing AMI management and outcomes.

Fortunately, thanks to the rapid development in genetic engineering, microarray technology was advanced and applied in both clinical investigations and academic research. Some more new biomarkers with the potential to outperform the aforementioned were proposed in recent years. For example, Zhang et al. found that ARG1 might play a key role in the pathogenesis of AMI, which could be a biomarker of AMI and provide a reference for a more in-depth study [5]. Chen et al. identified TBX21 and PRF1 as novel diagnostic biomarkers and potential modulatory targets through comprehensive bioinformatic analytics [6]. Furthermore, the activation of the immune system after the occurrence of AMI is more and more aware by scientists and clinicians these days. Immune cell types including M2 macrophages, mast cells, and eosinophils have been proven to possess certain impacts on patients after AMI, providing new insights into the immune mechanisms of AMI pathogenesis [7]. Meanwhile, the introduction of artificial intelligence (AI) into the field of bioinformatics analytics improves the robustness of *in-silico* methods significantly. By combining these techniques, the discovery of even more reliable biomarkers can be expected in the foreseeable future [8].

Abnormal copper metabolism has been proven to be linked with heart ischemia for a long time [9–11]. Cuproptosis, a novel cell death mechanism induced by an intracellular imbalance of copper ions may also have an important role in this regard as copper has been proven to coordinate a variety of cellular biological processes such as lipolysis, cell proliferation, autophagy, and neural activity [12].

Inspired by the aforementioned, the present study aimed to establish a novel diagnostic model for early AMI detection based on the Cuproptosis-related gene set. In total, the AI for AMI diagnosis was constructed on the basis of 4 contributor genes, including PDHB, CDKN2A, GLS, and SLC31A1. Then, we performed bulk RNA sequencing with our *in vivo* models to assemble our own small cohort. By comparing with the public datasets, we determined the central role of SLC31A1. Furthermore, we validated the expression of SLC31A1 by a series of *in vivo* assays (i.e., quantitative real-time PCR, immunohistochemical staining, and their quantitative analyses by the ImageJ software) and explored the potential immunological implications of SLC31A1 expression.

Figure 1 demonstrates the general design of the present study.

MATERIALS AND METHODS

Data curation and processing

In total, 8 datasets (i.e., GSE29111, GSE60993, GSE109048, GSE29532, GSE19339, GSE48060, GSE66360, GSE97320) were curated from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) [13–20]. A total of 318 samples were therefore enrolled in the present study. The normalization and calibration were done through the "Normalize Between Arrays" function of the R package, "limma". As for the genes set involved in the present study, a total of 13 Cuproptosis-related genes, including 7 pro-Cuproptosis genes (i.e., FDX1, LIAS, LIPT1, DLD, DLAT, PDHA1, and PDHB), 3 anti-Cuproptosis genes (i.e., MTF1, GLS, and CDKN2A), and 3 copper transporter-encoding genes (i.e., ATP7A, ATP7B, and SLC31A1) [21, 22]. Of note that some genes (i.e., LIPT1, MTF1, ATP7A, and ATP7B) were not present in all the datasets and thus were not considered in the statistics. The aforementioned analyses were done online through the integrative tool set "Xsmart plateform" (https://www.xsmartanalysis.com/) by different packages embedded in the R studio software and Python (especially the Python package "sklearn"). If not specifically mentioned, the statistic test used in the analytics is the Wilcoxon rank sum test. Notably, within some figures, *, **, and *** may occur, indicative of a P-value less than 0.05, 0.01, and 0.001, respectively.

Selection of contributor genes

To avoid linearity problems, we used not only popular linear algorithms such as LASSO but also non-linear

ones like SVM-RFE, XGBoost-RFE, and Boruta algorithms [23–25]. Each algorithm gives a certain list of genes that can serve as contributor genes for the construction of diagnostic AI, but we intended to limit the number of genes involved to make it more clinically applicable. Therefore, a Venn diagram was drawn to screen the overlapping genes out. Those genes were thought as most ideal contributor genes for modeling in the next step.

Mainstream machine learning algorithms used in the present study

According to the "no free lunch" theorem, if one machine learning algorithm outperforms the others on a specific assessment, it should sacrifice certain points on the other assessment measurements [26]. Therefore, to overcome this disadvantage, we first exhaustively tried 9 mainstream stream algorithms, including XGBoost, Logistic, Random Forest, AdaBoost, GNB, LightGBM, MLP, KNN, and SVM, to select the most suitable algorithms for this task, and then merge them by stacking method. The concept of Stacking involves first training base learners on the original data. These base learners each generate outputs based on the original data. These outputs from the multiple models are then combined to form new data, which is subsequently fed into a second-level model for fitting. This usually results in a more accurate output ultimately. In the present study, the training set and validation set contain a sample ratio of 7:3. In the modulation of stacking, the sample allocation ratio of the training set:validation set: test set is 7:2:1. All samples involved in the machine learning process were randomized first, and then allocated.

Decision curve analysis (DCA)

Most of the time, in the case of clinical questions, false positives and false negatives are inevitable. To address
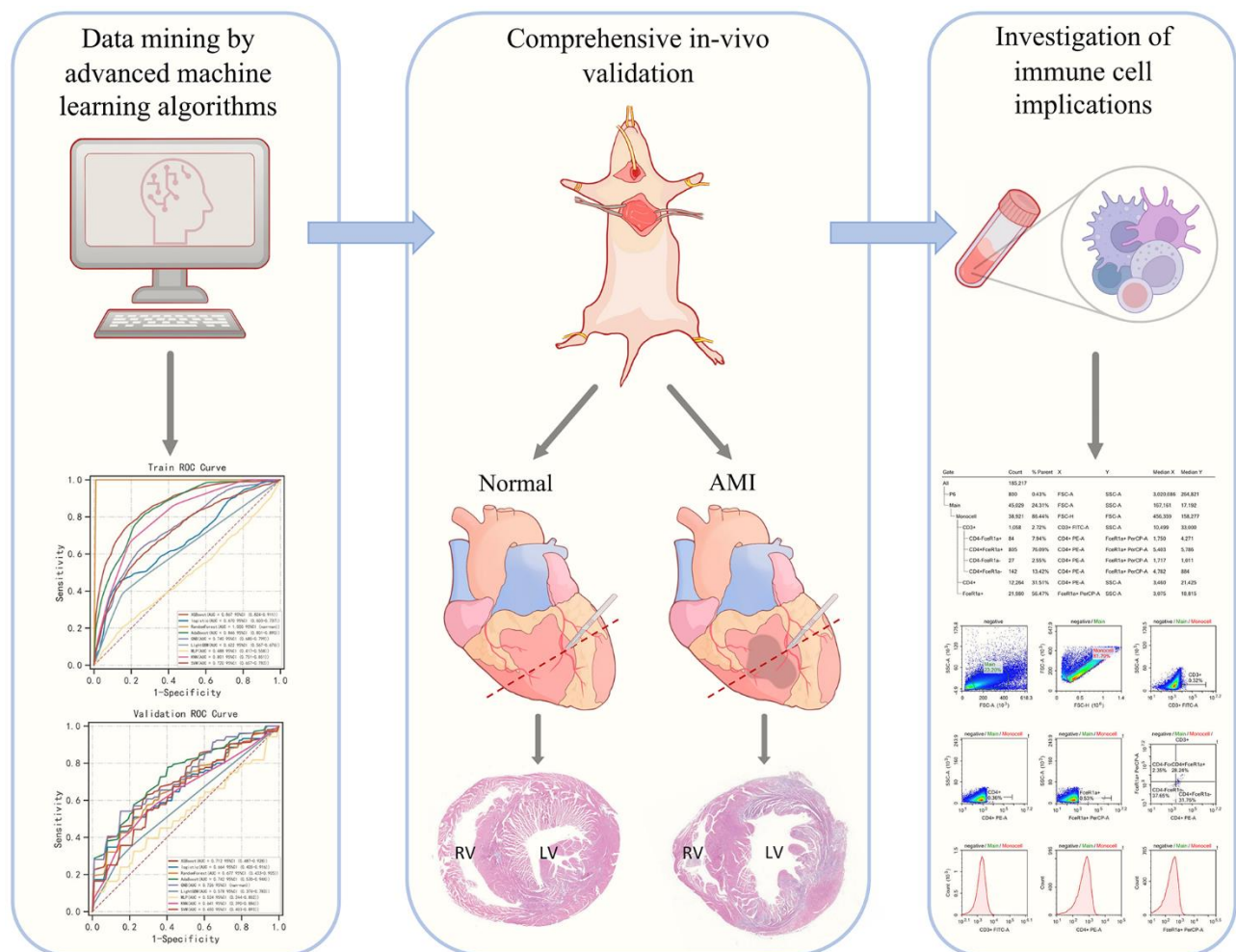

Figure 1. Graphical abstract of the present study.

this issue, DCA, which is a popular method to compare the efficacy of different predictive models under such circumstances, was used, so that the clinical benefits are maximized [27, 28].

Calibration of predictions

To further assess the deviation between the predictive results and the reality, a calibration curve in which the predictive results were plotted against the observed reality within a randomly divided subset from the whole merged dataset. The ratio of the cases contained by this subset to the whole merged dataset is 3:10. The closer the curve is situated to the 450 ideal dash line, the more accurate prediction is given.

Analysis of the immunological microenvironment

CIBERSORT (https://cibersort.stanford.edu/) was used to assess the abundance of various infiltrating immune cells [29, 30]. Overall, 22 immune cell types were quantified. Correlation analysis between the immune cell types and SLC31A1 was done by the Spearman method. The visualization was achieved by the R package "ggplot2".

Construction of *in vivo* model

All experiments were performed on male C57BL/6 J mice (6-8 weeks of age, n=6/group). Animals were housed in groups in an environment with a 12/ 12 h day/night cycle and free access to water and food.

The mice were anesthetized with isoflurane first and then opened the chest quickly. Then, the left anterior descending coronary artery was ligated together with the great cardiac vein. Immediately after the coronary artery is lapped, the heart is placed back into the chest cavity, where it is gently pressed to expel air from the chest cavity while sutures from the skin are closed. The mice were evaluated by observing their status after surgery. In the healthy control group (Sham group), the thoracic cavity was opened without ligation of the left anterior descending branch of the coronary artery. The observation indexes were the vitality and normal life of mice, and the changes in electrocardiogram before and after the operation. Finally, the mice were sacrificed by neck amputation.

RNA sequencing and data analysis

Whole-genome gene expression analysis was performed using the heart tissues of mice from AMI and Sham groups (n=3/group) at 24h. The total RNA was extracted using Trizol (Vazyme), and cDNA samples were sequenced using a sequencing system (Novaseq 6000; Illumina). The reference Mus musculus genome and gene information were downloaded from the National Center for Biotechnology Information database. Raw reads were filtered to produce high-quality clean data. All the subsequent analyses were performed with clean data. The expression matrixes of selected genes involved in the figures were organized as tables in Supplementary Table 1.

Quantitative real-time PCR (qPCR)

Total RNA was isolated from tissues or cells using Trizol (Vazyme), and RNA concentration and purity were measured using spectrophotometry. RNA was reverse transcribed using the PrimeScript RT reagent Kit (Vazyme) under the manufacturer's instructions. Quantitative PCR was performed using LightCycler 96 (Roche) and SYBR Mastermix (Vazyme) in accordance with the manufacturer's instructions. The fold difference in gene expression was 6 calculated using the 2-$\Delta\Delta$ Ct method and is presented relative to Gapdh mRNA. All reactions were performed in triplicate, and specificity was monitored using melting curve analysis.

The PCR primer used in the present study was designed for the SLC31A1 gene, shown as follows:

Forward 5'-3': GGAGAAATGGCTGGAGCTTTT
Reverse 5'-3': CGGGCTATCTTGAGTCCTTCA

Histological examination

For histological analysis, hearts were fixed overnight in 4% paraformaldehyde (pH 7.4), embedded in paraffin, and serially sectioned at 5-μm thickness. The sections were stained with Hematoxylin and Eosin (H&E) for routine histological examination with a light microscope. To measure collagen deposits, select sections were stained with Masson's trichrome (MT) reagent. For each mouse, 3 random sections were quantified using ImageJ software (National Institutes of Health). Details can be checked in Supplementary Figure 1.

Immunohistochemistry

Serial sections were deparaffinized and blocked with phosphate-buffered saline (PBS) containing 5% (v/v) normal goat serum and 1% (w/v) BSA; the sections were then incubated with anti-GLS1 (Cat. No. bs-10341R, Bioss) antibody and anti-SLC31a1 (Cat. No. bs-10773R, Bioss) overnight, followed by incubation with a secondary antibody for 1 hour at room temperature. The relative intensity of protein staining was analyzed in five random sections, chosen 40× fields for each sample and quantified using ImageJ software

(National Institutes of Health). Detailed slides can be checked in Supplementary Figure 2.

Statistical analysis used in *in vivo* experiments

The data were analyzed and graphed using GraphPad Prism 9.4.1 software and are shown as mean ± SD. The Shapiro–Wilk-test was used to detect the normal distribution. Student's t-test or one-way ANOVA followed by Tukey's post-hoc test was used for statistical analysis as appropriate. For the Kaplan–Meier survival plots, statistical significance was measured by the log-rank (Mantel–Cox) test. A P-value < 0.05 was deemed statistically significant. All experiments were repeated independently 3 times.

RESULTS

Data normalization and the selection of contributor genes for machine learning

The expression matrixes of the 8 GEO datasets were normalized with a baseline correction and then merged. To see if the samples were homogenized, we visualized them in the form of UMAP. As shown in the figure, the samples involved in the present study were well-mixed with one another, indicating a good homogeneity (Figure 2A, 2B).

Then, we went on with contributor gene selection. Through the SVM-RFE algorithm, 7 genes (i.e., FDX1, DLD, PDHA1, PDHB, GLS, CDKN2A, and SLC31A1) were thought to be the most suitable for the following AI construction (Figure 2C), while the results of XGBoost-RFE algorithm recommended LIAS, PDHA1, PDHB, GLS, CDKN2A, and SLC31A1 (Figure 2D). In the Boruta algorithm, FDX1, DLAT, LIAS, PDHA1, PDHB, GLS, CDKN2A, and SLC31A1 were filtered out and ranked with specific weights of importance (Figure 2E). Besides the aforementioned non-linear algorithms, the regularly used linear algorithm in the field of bioinformatics, LASSO, was also employed in the present study, through which we found FDX1, LIAS, DLD, PDHB, GLS, CDKN2A, and SLC31A1 were the strongest candidates (Figure 2F). To sum up, we used a Venn diagram to intersect the overlapping genes (Figure 2G), so that the contributor genes for AI construction could be determined. As a result, we decided to use PDHB, CDKN2A, GLS, and SLC31A1 in the rest of the study.

Selecting the best algorithms

We exhaustively went through 9 mainstream machine learning algorithms (i.e., XGBoost, Logistic, Random Forest, AdaBoost, GNB, LightGBM, MLP, KNN, and SVM) to identify the ideal machine learning algorithms for this task. Consequently, we found that regarding the Precision-Recall rate (i.e., the so-called "PR" annotated in the figure), XGBoost, AdaBoost, and GBN were the most superior algorithms (Figure 3A, 3B), which was further confirmed by the value of Area under Curve (AUC) on the Receiver Operative Characteristic (ROC) curve (Figure 3C, 3D). Although the AdaBoost algorithm possessed a little bit higher bias in the calibration curve (i.e., 0.182) than that of the Random Forest algorithm (i.e., 0.116), and was located comparatively lower in the plot of DCA, comprehensively considering, we believed XGBoost, AdaBoost, and GBN were the most ideal options for further AI modulation (Figure 3E, 3F).

The stacking of machine learning algorithms outperformed the gold standard biomarkers in AMI

As previously described, we found that XGBoost, AdaBoost, and GBN were seemingly the best options for AI modulation as they possessed relatively high PR-AUC and ROC-AUC values. While in terms of the stacking method, usually 2 layers of classifying logic are set, and the difference between the 2 layers is positively associated with the final predictive outcome. Therefore, besides placing XGBoost, AdaBoost, and GBN into the first layer, we used MLP which performed most distinctly from these algorithms as a second layer. Subsequently, we found the merged version behaved satisfying predictions with a ROC-AUC value over 0.7 in not only the training set but also the validation set and even the test set respectively, showcasing its outstanding calculation power (Figure 4A–4C). Meanwhile, we examined the individual predictive ability of each contributor gene, finding that when using these genes as diagnostic biomarkers solely, the outcomes were fall out of our expectations as PDHB (Figure 4E), SLC31A1 (Figure 4F), and GLS (Figure 4G) possessed a ROC-AUC value of around 0.6 respectively and CDKN2A (Figure 4D) had only reached 0.54. Gold standard biomarkers (i.e., TNNI3 and CKM) on the other hand were not playing well either, as they were scored with a ROC-AUC value of 0.62 and 0.59, respectively (Figure 4H, 4I). Overall, by combing the 4 contributor genes as a genetic signature, with the proper assistance of AI, the early diagnosis of AMI would likely reach a new level.

***In vivo* validation of SLC31A1 expression**

We modeled the mice with surgical techniques described in the methodology section and validated them from histological slides with Masson staining. Representative images were exhibited within the figure.
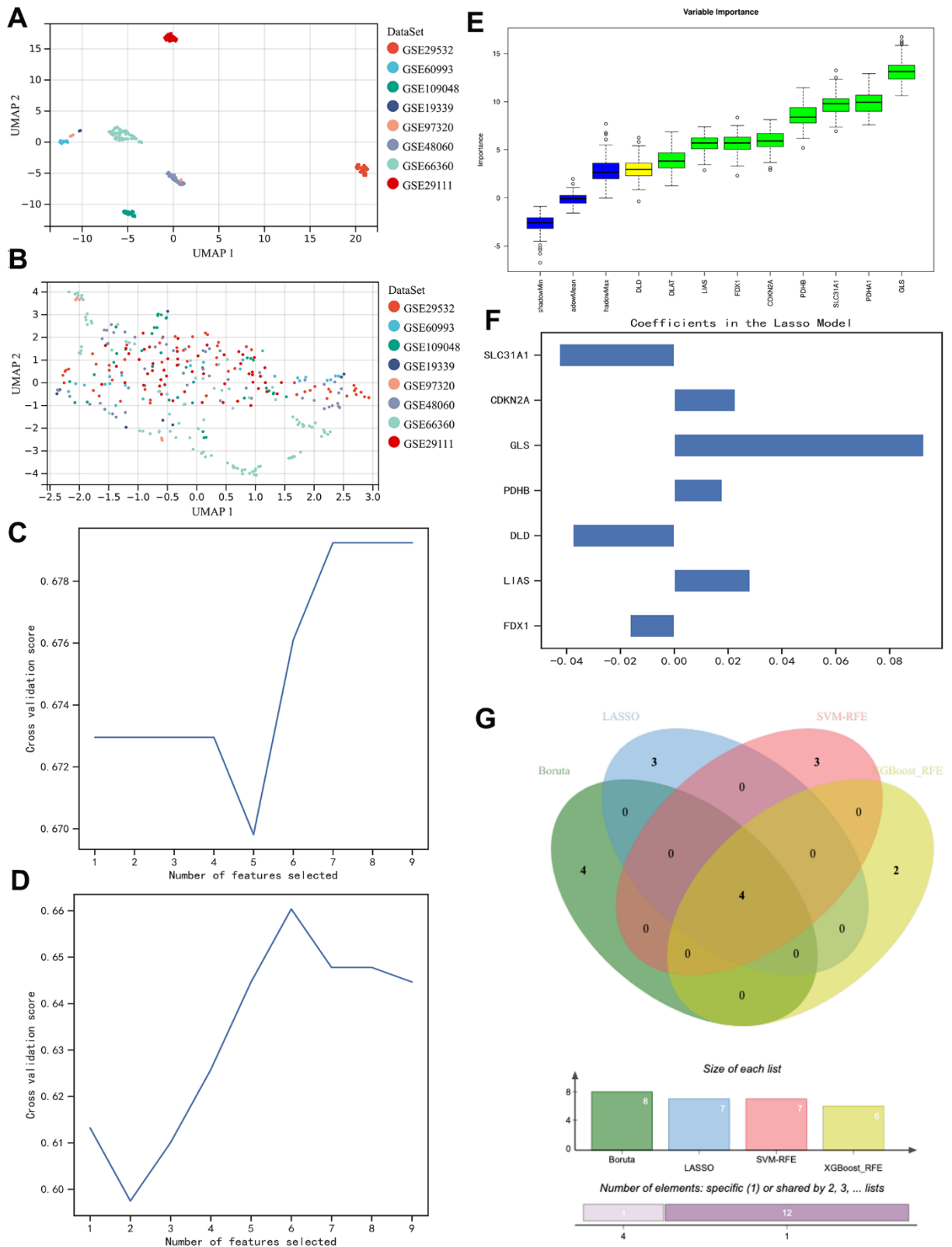
Figure 2. Homogenization of different samples from various datasets and selection of the contributor genes for the construction of AI diagnostic predictor. (**A**, **B**) UMAP plot visualizes the sample distribution. As shown, originally the samples were fairly separated (**A**), but they were very well homogenous after normalization (**B**). (**C**–**F**) Determination of suitable contributor genes by SVM-RFE, XGBoost-RFE, Boruta, and LASSO, respectively. (**G**) Venn diagram demonstrated the overlapping genes. 4 genes (i.e., PDHB, CDKN2A, GLS, and SLC31A1) were finally selected.
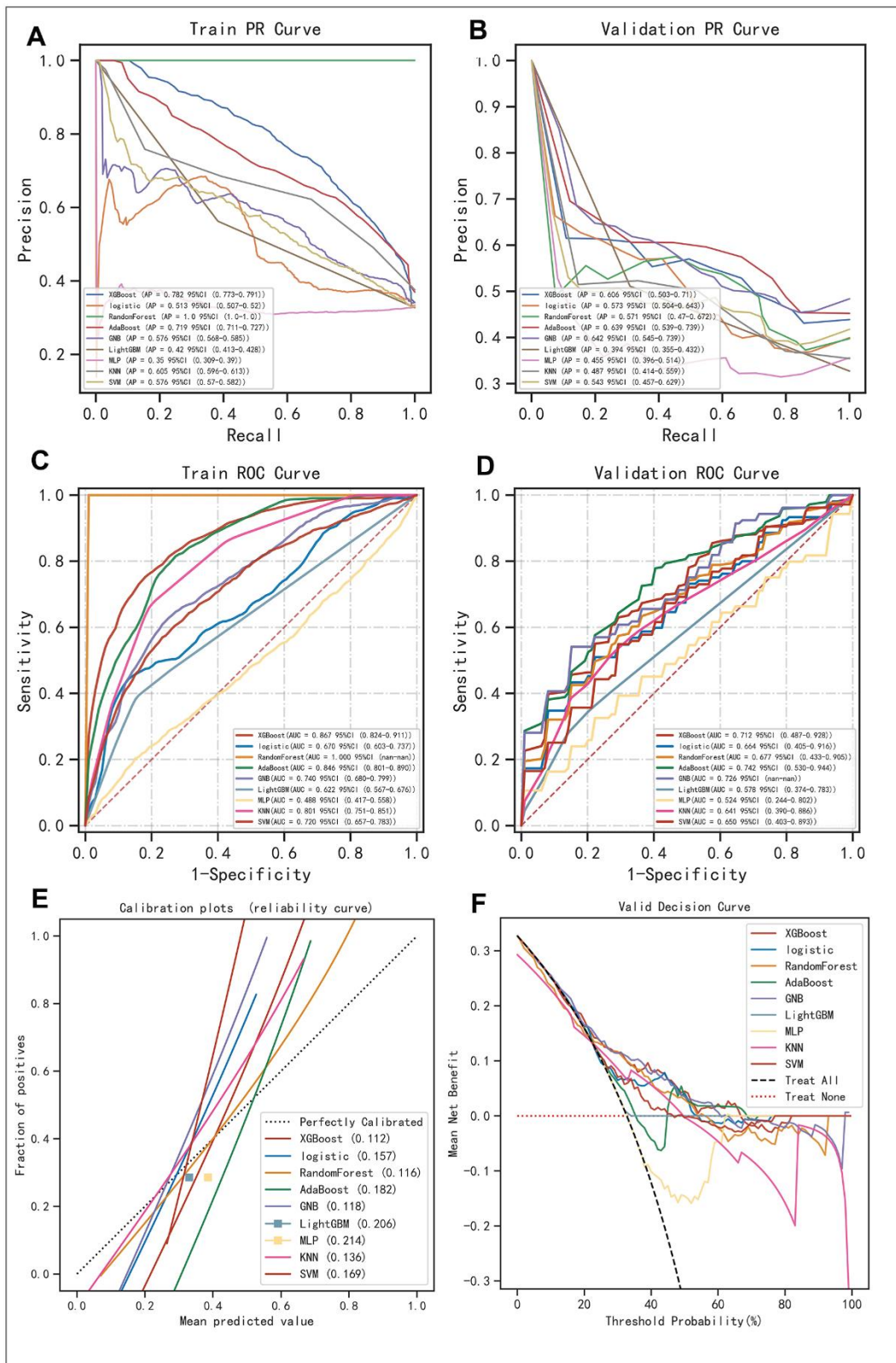
Figure 3. Evaluation of the robustness of the 9 mainstream machine algorithms to identify the best ones for AI construction. (**A**, **B**) Precision-Recall (PR) curve in the training set and validation set, respectively. (**C**, **D**) Receiver Operative Characteristic (ROC) curve performed in the training set and validation set, respectively. (**E**) Calibration curve demonstrating the bias between predictive values and realistic values for the machine learning algorithms involved in the present study. (**F**) DCA for the machine learning algorithms involved in the present study.

In a nutshell, we compared the AMI group and the Sham group macroscopically (Figure 5A) and under magnification (Figure 5B), both demonstrating significant differences regarding the color of staining. In the AMI group, within the region of infarction, the myocardium was in much deeper purplish, indicating a significant pathological status, while in the Sham group, normal myocardium could be clearly observed in a healthy red color.

In the GEO datasets, we tested the expression of the contributor genes in the manner of violin plots. It was found that except for the CDKN2A gene, the rest genes had very high statistical significance regarding differential expression (Figure 5C). However, it was found that in our own bulk sequencing cohort, merely the SLC31A1 gene was granted statistical significance, and the other genes possessed P-values over 0.05, although meanwhile, the mean differences were
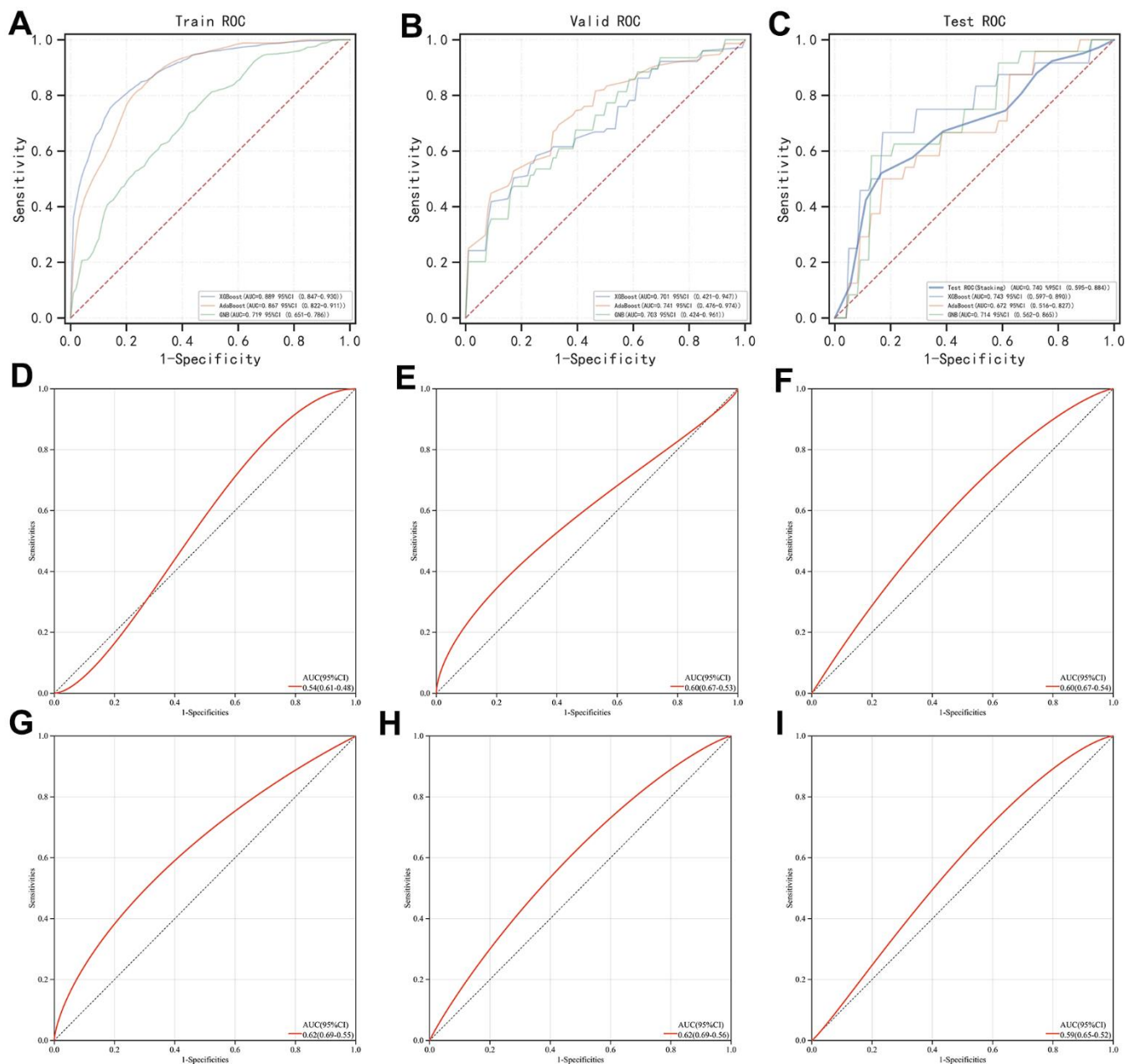
Figure 4. Evaluation of the stacking-based AI predictor for AMI diagnosis and comparison of its efficacy with individual genes involved (i.e., PDHB, CDKN2A, GLS, and SLC31A1) and gold standard biomarkers (i.e., TNNI3 and CKM). (**A**–**C**) ROC curve of the stacking-based AI predictor in the training set, validation set, and test set, respectively. (**D**–**G**) ROC curve of PDHB, CDKN2A, GLS, and SLC31A1, respectively. (**H**, **I**) ROC curve of TNNI3 and CKM, respectively.
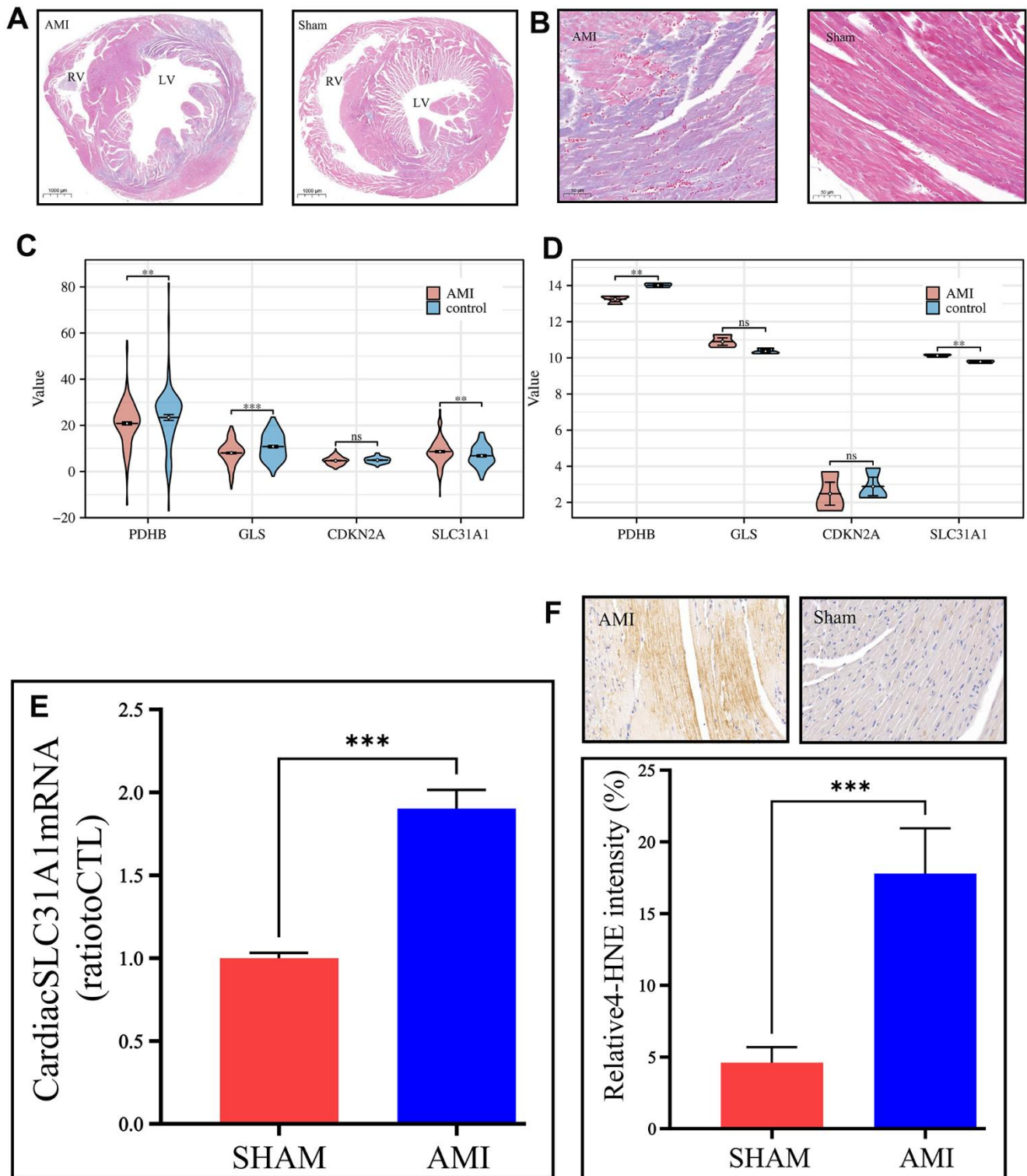
Figure 5. *In vivo* validation of the expression of the contributor genes, with a focus on the SLC31A1 gene. (**A**) Masson staining slides on a macroscopic view, demonstrating the coronal section of the mice's hearts. The left panel is the AMI sample, and the right panel is the Sham sample. RV: right ventricle, LV: left ventricle. (**B**) Masson staining slides of the coronary heart section of the AMI group and Sham group under magnification. The purplish color indicates hypoxia, thus the area of AMI. The deeper the color, the more severe infarction. (**C**, **D**) Expression analysis of the contributor genes in the form of violin plots in the merged GEO dataset and our own bulk RNA sequencing cohort, respectively. (**E**) qPCR results of the SLC31A1 gene expression. (**F**) Upper panel: immunohistochemical staining of the SLC31A1 protein in the AMI group and Sham group. The more brownish color, the more abundant the SLC31A1 protein. Lower panel: quantitative analysis of the immunohistochemical staining for the SLC31A1 protein in the AMI group and Sham group in the manner of bar plot.

apparent (Figure 5D). As such, we continued the *in vivo* studies focusing on the SLC31A1 gene. Through the qPCR, the expression of the SLC31A1 gene was determined at the mRNA level in the AMI group and Sham groups. It was found that the SLC31A1 gene was expressed at a relatively higher level in the AMI group (Figure 5E), which supported our computational results. At the protein level, we immunohistochemically stained the samples to visualize the abundance of SLC31A1 protein (Figure 5F). Subsequently, we found the SLC31A1 protein was much more abundant in the AMI group than that in the Sham group.

SLC31A1 expression was potentially positively correlated with monocyte infiltration

We first obtained the abundance of various infiltrating immune cells in AMI using the CIBERSORT algorithm. There are significant differences in multiple immune cell types in AMI in comparison with the healthy control, including plasma cells, CD4+ memory T cells, Gamma-Delta T cells, activated NK cells, monocytes, macrophage M2, activated dendritic cells, mast cells, eosinophils, and neutrophils (Figure 6A). Such results were cross-validated with previous studies which
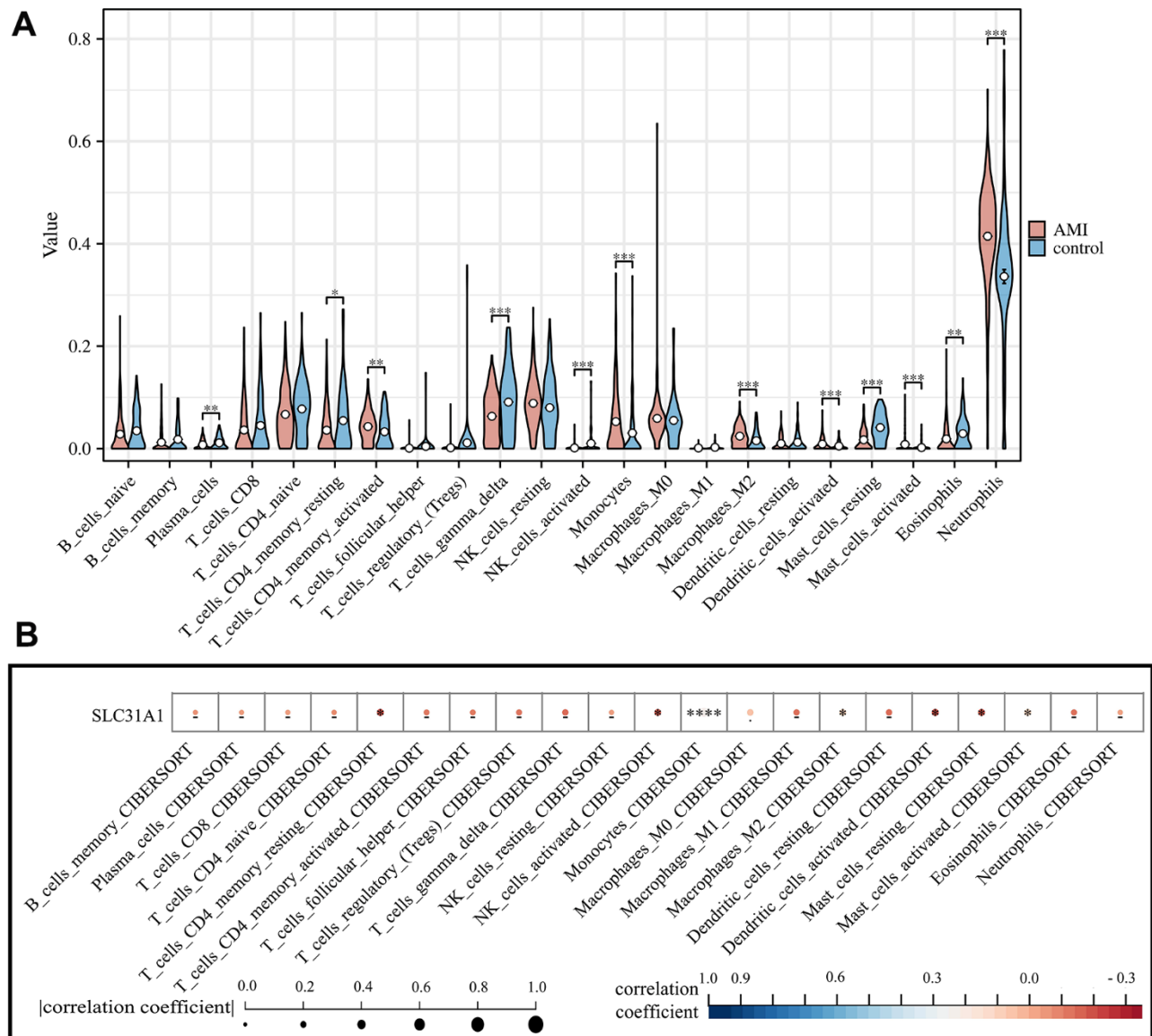

Figure 6. Assessment of the infiltration of various immune cell types in both AMI and healthy control and their correlation with SLC31A1 expression. (**A**) Violin plot demonstrating the comparison of abundance of infiltrating immune cell types in the AMI group and healthy control. (**B**) Spearman correlation analysis between the SLC31A1 expression and diverse infiltrating immune cell types.

pointed out that the infarction status of the myocardium might trigger a diverse immune cell activation in which there was a sophisticated interplay [31–33]. We further analyzed the association between the SLC31A1 expression and cellular immunity. We found a significant association between SLC31A1 and most of the immune cell types, especially monocytes, with a synchronizing trend in Figure 6A, hinting at the potentially central role of SLC31A1 among the contributor genes in AMI from the immunological aspect (Figure 6B).

DISCUSSION

Cardiovascular diseases as one of the most serious diseases worldwide, especially AMI, have been affecting millions of patients yearly and can recur in more than half of the population [34, 35]. Regardless of the rapid advancements in the healthcare industry, there have been limited improvements in morbidity and mortality from AMI over the past few decades, particularly in young ages [36]. For a long time, cardiac troponin and creatine kinase-MB isoenzyme are viewed as gold standard biomarkers for AMI diagnosis [37–39], but their real effectiveness and sensitivity are questionable. In the present study, their predictive performance in our merged GEO dataset was indeed less satisfying. Identifying new biomarkers for accurate and robust AMI diagnosis remains in demand.

In 2022, Tsvetkov and colleagues discovered Cuproptosis, a novel cellular demise pathway contingent upon mitochondrial respiration [12]. It stands distinct from apoptosis by its unique reliance on copper and mitochondria interaction, meanwhile, unlike ferroptosis (also a result of imbalanced ion hemostasis), which involves iron-dependent lipid peroxidation, Cuproptosis emphasizes copper homeostasis disruption, unveiling a novel avenue in understanding cellular demise with implications beyond classical apoptotic pathways. Subsequently, extensive research has delved into its implications within the realm of oncology [40–42], although there remains a conspicuous dearth of studies investigating its pertinence to non-neoplastic conditions.

On the other hand, the wide use of high-throughput sequencing technique generated a wealth of big data that is being analyzed again and again to further explore the pathomechanisms of various diseases. However, mining valuable knowledge from such overwhelming amounts of data is a challenging and vital task in modern medicine. In this regard, machine learning, a successful methodology to extract valuable knowledge under this background, has been applied in order to identify novel therapeutic targets and optimize current

treatment strategies, making early detection of the disease of interest and predictions of the corresponding prognosis, and so forth more precise and efficient.

Our study is the first to employ up to nine mainstream machine learning algorithms, including XGBoost, Logistic, Random Forest, AdaBoost, GNB, LightGBM, MLP, KNN, and SVM, as well as stacking methods to create a powerful AI for AMI diagnosis. By using these methods, we identified XGBoost, AdaBoost, GBN, and MLP as the most suitable algorithms for the task. Using the stacking method, we integrated XGBoost, AdaBoost, and GBN into the first layer of the AI's logic chain, while MLP served as the second layer. This resulted in a highly robust AI with superior predictive power compared to any of the single machine learning algorithms, the single contributor genes to the AI, and even the gold standard biomarkers (i.e., TNNI3 and CKM). In regard to the evaluation of all the aforementioned, we mainly made it based on the AUC value. In general, a model that has an AUC of 0.5 does no better than random guessing, and an AUC value over 0.7 is usually seen as a good value. By building the highly accurate model for AMI diagnosis, we hope that with the help of our model, someday physicians can estimate reliably the risk of AMI for patients under suspicion, by simply testing the expression of the contributor genes of the model.

Among the genes contributing to AMI (PDHB, CDKN2A, GLS, and SLC31A1), SLC31A1 was found to be of particular importance. In fact, SLC31A1 is a gene that encodes for the copper transporter protein 1 (CTR1) which plays a crucial role in copper homeostasis in cells. CTR1 is responsible for the uptake of copper from the extracellular environment and its transport into cells. The protein is particularly important in cells that require copper for their function, such as those involved in oxidative phosphorylation, iron metabolism, and neurotransmitter synthesis. From this end, it is somehow reasonable that SLC31A1 may impact significantly in AMI. To validate our hypothesis, we carried out qPCR, western blot, and immunohistochemical staining in *in vivo* models. The results showed that SLC31A1 was aberrantly overexpressed in cases of AMI, underscoring its potential as a diagnostic biomarker. This finding supported the previous works done by Zheng et al. and Wang et al. in which variations in the SLC31A1 gene have been associated with an increased risk of AMI and decreased CTR1 expression has been observed in rat hearts after myocardial infarction [43, 44].

Accumulated evidence had suggested that immune cells might play a critical role in the development and progression of AMI [45–48]. In particular, the infiltration of monocytes in the myocardium has been implicated in

the development of ventricular remodeling and the potentially resulting heart failure [49–51]. Therefore, we were also interested in the correlation between SLC31A1 expression and various immune cell types. It turned out that there was a potentially positive association between SLC31A1 expression and the infiltration of many immune cell populations, especially monocyte. In fact, monocytes have been proven to be the versatile cells of the innate immune system, indispensable in the initial inflammatory response to injuries and subsequent wound healing processes in many tissues, including the heart [52]. As such, these findings highlighted the possible central role of SLC31A1 expression in AMI's immune landscape and its probability in improving heart tissue recovery in the post-AMI scenario.

CONCLUSIONS

In summary, the present study established a novel diagnostic model for early AMI detection based on the Cuproptosis-related gene set, identifying the central role of SLC31A1, and validated the aberrant overexpression of SLC31A1 in *in vivo* assays, exploring its potential immunological implications, sharing new perspective toward alternative AMI biomarker development.

AUTHOR CONTRIBUTIONS

Conceptualization: Shujing Zhou and Longbin Wang. Data curation: Shujing Zhou, Longbin Wang, Ting Wang, and Yidan Tang. Formal analysis: Shujing Zhou and Xufeng Huang. Investigation: Longbin Wang, Ting Wang, Yidan Tang, and Ming Xu. Visualization: Shujing Zhou and Xufeng Huang. Software: Shujing Zhou and Xufeng Huang. Writing – original draft: Shujing Zhou, Longbin Wang, Ting Wang, and Yidan Tang. Writing - review and edition: Ming Xu. All authors agreed to submit the present version.

ACKNOWLEDGMENTS

We want to express our deep gratitude to the public databases, including GEO, GeneCards, OMIM, HPA, CTD, and more, for providing open-accessible and high-quality research resources. We also sincerely thank the Fundamental Research Funds for the Central Universities of China for their financial support of our present study.

CONFLICTS OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in the present study.

REFERENCES

1. Kapur NK, Thayer KL, Zweck E. Cardiogenic Shock in the Setting of Acute Myocardial Infarction. Methodist Debakey Cardiovasc J. 2020; 16:16–21.
https://doi.org/10.14797/mdcj-16-1-16
PMID:32280413

2. de Winter RJ, Koster RW, Sturk A, Sanders GT. Value of myoglobin, troponin T, and CK-MBmass in ruling out an acute myocardial infarction in the emergency room. Circulation. 1995; 92:3401–7.
https://doi.org/10.1161/01.cir.92.12.3401
PMID:8521560

3. 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.
https://doi.org/10.3389/fcvm.2020.586871
PMID:33195475

4. Law MR, Watt HC, Wald NJ. The underlying risk of death after myocardial infarction in the absence of treatment. Arch Intern Med. 2002; 162:2405–10.
https://doi.org/10.1001/archinte.162.21.2405
PMID:12437397

5. Zhang R, Ji Z, Qu Y, Yang M, Su Y, Zuo W, Zhao Q, Ma G, Li Y. Clinical value of ARG1 in acute myocardial infarction patients: Bioinformatics-based approach. Biomed Pharmacother. 2020; 121:109590.
https://doi.org/10.1016/j.biopha.2019.109590
PMID:31733574

6. Chen J, Yu L, Zhang S, Chen X. Network Analysis-Based Approach for Exploring the Potential Diagnostic Biomarkers of Acute Myocardial Infarction. Front Physiol. 2016; 7:615.
https://doi.org/10.3389/fphys.2016.00615
PMID:28018242

7. Xu JY, Xiong YY, Lu XT, Yang YJ. Regulation of Type 2 Immunity in Myocardial Infarction. Front Immunol. 2019; 10:62.

https://doi.org/10.3389/fimmu.2019.00062
PMID:30761134

8. Zou Y, Xie J, Zheng S, Liu W, Tang Y, Tian W, Deng X, Wu L, Zhang Y, Wong CW, Tan D, Liu Q, Xie X. Leveraging diverse cell-death patterns to predict the prognosis and drug sensitivity of triple-negative breast cancer patients after surgery. Int J Surg. 2022; 107:106936.
https://doi.org/10.1016/j.ijsu.2022.106936
PMID:36341760

9. Culotta V. Cell biology of copper. J Biol Inorg Chem. 2010; 15:1–2.
https://doi.org/10.1007/s00775-009-0601-x
PMID:19882181

10. Liu Y, Miao J. An Emerging Role of Defective Copper Metabolism in Heart Disease. Nutrients. 2022; 14:700.
https://doi.org/10.3390/nu14030700 PMID:35277059

11. Li K, Li C, Xiao Y, Wang T, James Kang Y. The loss of copper is associated with the increase in copper metabolism MURR domain 1 in ischemic hearts of mice. Exp Biol Med (Maywood). 2018; 243:780–5.
https://doi.org/10.1177/1535370218773055
PMID:29763364

12. Tsvetkov P, Coy S, Petrova B, Dreishpoon M, Verma A, Abdusamad M, Rossen J, Joesch-Cohen L, Humeidi R, Spangler RD, Eaton JK, Frenkel E, Kocak M, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. Science. 2022; 375:1254–61.
https://doi.org/10.1126/science.abf0529
PMID:35298263

13. Zhang N, Zhou B, Tu S. Identification of an 11 immune-related gene signature as the novel biomarker for acute myocardial infarction diagnosis. Genes Immun. 2022; 23:209–17.
https://doi.org/10.1038/s41435-022-00183-7
PMID:36182975

14. Park HJ, Noh JH, Eun JW, Koh YS, Seo SM, Park WS, Lee JY, Chang K, Seung KB, Kim PJ, Nam SW. Assessment and diagnostic relevance of novel serum biomarkers for early decision of ST-elevation myocardial infarction. Oncotarget. 2015; 6:12970–83.
https://doi.org/10.18632/oncotarget.4001
PMID:26025919

15. Gobbi G, Carubbi C, Tagliazucchi GM, Masselli E, Mirandola P, Pigazzani F, Crocamo A, Notarangelo MF, Suma S, Paraboschi E, Maglietta G, Nagalla S, Pozzi G, et al. Sighting acute myocardial infarction through platelet gene expression. Sci Rep. 2019; 9:19574.
https://doi.org/10.1038/s41598-019-56047-0
PMID:31863085

16. Silbiger VN, Luchessi AD, Hirata RD, Lima-Neto LG, Cavichioli D, Carracedo A, Brión M, Dopazo J, García-García F, dos Santos ES, Ramos RF, Sampaio MF, Armaganijan D, 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
PMID:23535507

17. Suresh R, Li X, Chiriac A, Goel K, Terzic A, Perez-Terzic C, Nelson TJ. Transcriptome from circulating cells suggests dysregulated pathways associated with long-term recurrent events following first-time myocardial infarction. J Mol Cell Cardiol. 2014; 74:13–21.
https://doi.org/10.1016/j.yjmcc.2014.04.017
PMID:24801707

18. Muse ED, Kramer ER, Wang H, Barrett P, Parviz F, Novotny MA, Lasken RS, Jatkoe TA, Oliveira G, Peng H, Lu J, Connelly MC, Schilling K, et al. A Whole Blood Molecular Signature for Acute Myocardial Infarction. Sci Rep. 2017; 7:12268.
https://doi.org/10.1038/s41598-017-12166-0
PMID:28947747

19. Shao G. Integrated RNA gene expression analysis identified potential immune-related biomarkers and RNA regulatory pathways of acute myocardial infarction. PLoS One. 2022; 17:e0264362.
https://doi.org/10.1371/journal.pone.0264362
PMID:35231061

20. Wu Y, Jiang T, Hua J, Xiong Z, Chen H, Li L, Peng J, Xiong W. Integrated Bioinformatics-Based Analysis of Hub Genes and the Mechanism of Immune Infiltration Associated With Acute Myocardial Infarction. Front Cardiovasc Med. 2022; 9:831605.
https://doi.org/10.3389/fcvm.2022.831605
PMID:35463752

21. Lv H, Liu X, Zeng X, Liu Y, Zhang C, Zhang Q, Xu J. Comprehensive Analysis of Cuproptosis-Related Genes in Immune Infiltration and Prognosis in Melanoma. Front Pharmacol. 2022; 13:930041.
https://doi.org/10.3389/fphar.2022.930041
PMID:35837286

22. Liu H. Pan-cancer profiles of the cuproptosis gene set. Am J Cancer Res. 2022; 12:4074–81.
https://doi.org/10.21203/rs.3.rs-1716214/v1
PMID:36119826

23. Tibshirani, R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological). 1996; 58:267–88.
https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

24. Sanz H, Valim C, Vegas E, Oller JM, Reverter F. SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. BMC Bioinformatics. 2018; 19:432.

https://doi.org/10.1186/s12859-018-2451-4
PMID:30453885

25. Kursa MB, Rudnicki WR. "Feature Selection with the Boruta Package." Journal of Statistical Software. 2010; 36: 1–13.
https://doi.org/10.18637/jss.v036.i11

26. Sammut, C., Webb, G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. 2011.

27. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006; 26:565–74.
https://doi.org/10.1177/0272989X06295361
PMID:17099194

28. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. Diagn Progn Res. 2019; 3:18.
https://doi.org/10.1186/s41512-019-0064-7
PMID:31592444

29. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol. 2018; 1711:243–59.
https://doi.org/10.1007/978-1-4939-7493-1_12
PMID:29344893

30. Craven KE, Gökmen-Polar Y, Badve SS. CIBERSORT analysis of TCGA and METABRIC identifies subgroups with better outcomes in triple negative breast cancer. Sci Rep. 2021; 11:4691.
https://doi.org/10.1038/s41598-021-83913-7
PMID:33633150

31. Anzai A, Ko S, Fukuda K. Immune and Inflammatory Networks in Myocardial Infarction: Current Research and Its Potential Implications for the Clinic. Int J Mol Sci. 2022; 23:5214.
https://doi.org/10.3390/ijms23095214 PMID:35563605

32. Kologrivova I, Shtatolkina M, Suslova T, Ryabov V. Cells of the Immune System in Cardiac Remodeling: Main Players in Resolution of Inflammation and Repair After Myocardial Infarction. Front Immunol. 2021; 12:664457.
https://doi.org/10.3389/fimmu.2021.664457
PMID:33868315

33. Frangogiannis NG. The immune system and the remodeling infarcted heart: cell biological insights and therapeutic opportunities. J Cardiovasc Pharmacol. 2014; 63:185–95.
https://doi.org/10.1097/FJC.0000000000000003
PMID:24072174

34. Mangion K, Gao H, Husmeier D, Luo X, Berry C. Advances in computational modelling for personalised medicine after myocardial infarction. Heart. 2018; 104:550–7.

https://doi.org/10.1136/heartjnl-2017-311449
PMID:29127185

35. Nichols M, Townsend N, Scarborough P, Rayner M. Cardiovascular disease in Europe 2014: epidemiological update. Eur Heart J. 2014; 35:2929.
https://doi.org/10.1093/eurheartj/ehu378
PMID:25381246

36. Gulati R, Behfar A, Narula J, Kanwar A, Lerman A, Cooper L, Singh M. Acute Myocardial Infarction in Young Individuals. Mayo Clin Proc. 2020; 95:136–56.
https://doi.org/10.1016/j.mayocp.2019.05.001
PMID:31902409

37. Puleo PR, Guadagno PA, Roberts R, Scheel MV, Marian AJ, Churchill D, Perryman MB. Early diagnosis of acute myocardial infarction based on assay for subforms of creatine kinase-MB. Circulation. 1990; 82:759–64.
https://doi.org/10.1161/01.cir.82.3.759
PMID:2203554

38. Pervaiz S, Anderson FP, Lohmann TP, Lawson CJ, Feng YJ, Waskiewicz D, Contois JH, Wu AH. Comparative analysis of cardiac troponin I and creatine kinase-MB as markers of acute myocardial infarction. Clin Cardiol. 1997; 20:269–71.
https://doi.org/10.1002/clc.4960200316
PMID:9068915

39. Robinson DJ, Christenson RH. Creatine kinase and its CK-MB isoenzyme: the conventional marker for the diagnosis of acute myocardial infarction. J Emerg Med. 1999; 17:95–104.
https://doi.org/10.1016/s0736-4679(98)00129-2
PMID:9950395

40. Xiong C, Ling H, Hao Q, Zhou X. Cuproptosis: p53-regulated metabolic cell death? Cell Death Differ. 2023; 30:876–84.
https://doi.org/10.1038/s41418-023-01125-0
PMID:36755067

41. Chen L, Min J, Wang F. Copper homeostasis and cuproptosis in health and disease. Signal Transduct Target Ther. 2022; 7:378.
https://doi.org/10.1038/s41392-022-01229-y
PMID:36414625

42. Tong X, Tang R, Xiao M, Xu J, Wang W, Zhang B, Liu J, Yu X, Shi S. Targeting cell death pathways for cancer therapy: recent developments in necroptosis, pyroptosis, ferroptosis, and cuproptosis research. J Hematol Oncol. 2022; 15:174.
https://doi.org/10.1186/s13045-022-01392-3
PMID:36482419

43. Zheng, Y., Li, Y., Zhang, Y., Zhang, X., Wang, H., Sun, Z., Liu, J. Association between SLC31A1 gene polymorphisms and the risk of acute myocardial

infarction. Journal of the American Heart Association. 2016; 5:e004257.

44. Wang, Y., Lv, Q., Deng, X., Liu, B., Liu, Y., Shi, D., Xia, J. Copper transporter 1 expression in rat myocardium following myocardial infarction. Journal of Cardiovascular Pharmacology and Therapeutics, 2016; 21:301–7.

45. Libby P. Inflammation in atherosclerosis. Nature. 2002; 420:868–74.
https://doi.org/10.1038/nature01323 PMID:12490960

46. Frangogiannis NG. The immune system and cardiac repair. Pharmacol Res. 2008; 58:88–111.
https://doi.org/10.1016/j.phrs.2008.06.007
PMID:18620057

47. Katusic ZS. Superoxide anion and endothelial regulation of arterial tone. Free Radic Biol Med. 1996; 20:443–8.
https://doi.org/10.1016/0891-5849(96)02116-8
PMID:8720916

48. Liuzzo G, Biasucci LM, Gallimore JR, Grillo RL, Rebuzzi AG, Pepys MB, Maseri A. The prognostic value of C-reactive protein and serum amyloid a protein in severe unstable angina. N Engl J Med. 1994; 331:417–24.
https://doi.org/10.1056/NEJM199408183310701
PMID:7880233

49. Ma Y, Yabluchanskiy A, Iyer RP, Cannon PL, Flynn ER, Jung M, Henry J, Cates CA, Deleon-Pennell KY, Lindsey ML. Temporal neutrophil polarization following myocardial infarction. Cardiovasc Res. 2016; 110:51–61.
https://doi.org/10.1093/cvr/cvw024
PMID:26825554

50. Frangogiannis NG. Regulation of the inflammatory response in cardiac repair. Circ Res. 2012; 110:159–73.
https://doi.org/10.1161/CIRCRESAHA.111.243162
PMID:22223212

51. Dewald O, Zymek P, Winkelmann K, Koerting A, Ren G, Abou-Khamis T, Michael LH, Rollins BJ, Entman ML, Frangogiannis NG. CCL2/Monocyte Chemoattractant Protein-1 regulates inflammatory responses critical to healing myocardial infarcts. Circ Res. 2005; 96:881–9.
https://doi.org/10.1161/01.RES.0000163017.13772.3a
PMID:15774854

52. Peet C, Ivetic A, Bromage DI, Shah AM. Cardiac monocytes and macrophages after myocardial infarction. Cardiovasc Res. 2020; 116:1101–12.
https://doi.org/10.1093/cvr/cvz336
PMID:31841135

Supplementary Figures
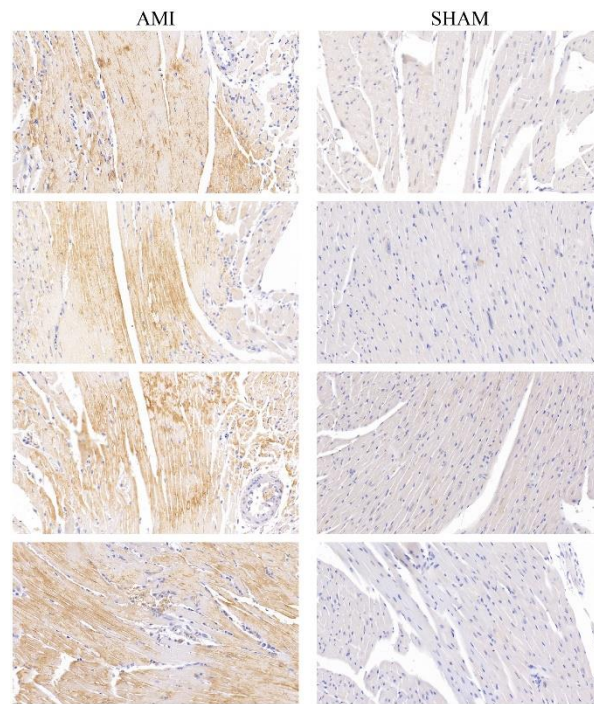
Supplementary Figure 1. Supplementary images of Masson staining slides.
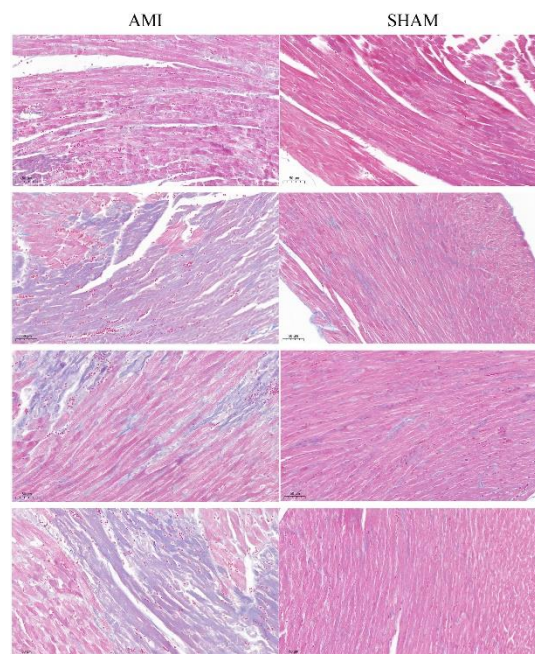
Supplementary Figure 2. IHC slides.

Supplementary Table

Please browse Full Text version to see the data of Supplementary Table 1.

Supplementary Table 1. Detailed expression of the contributor genes in our own bulk sequencing cohort.