Identification of feature autophagy-related genes in patients with acute myocardial infarction based on bioinformatics analyses

Running title: Autophagy-Related Genes in AMI

Yajuan Du*, MD, Enfa Zhao*, MD, Yushun Zhang ✉, MD

Department of Structural Heart Disease, The First Affiliated Hospital of Xi’an Jiaotong University, Xi’an 710061, China

✉ Correspondence to:
Yushun Zhang, PhD
Department of structural heart disease, The First Affiliated Hospital of Xi’an Jiaotong University, Xi’an 710061, China

Email: zys2889@sina.com

Address:
No.277, Yanta West Road, Xi’an 710061, Shaanxi, PR China

Tel: +86 85324128
Fax: +86 85324128

*Yajuan Du and Enfa Zhao contributed equally to this work and should be considered co-first authors.

Disclosure of grants or other funding

None

Ethics statement

Not necessary

Conflict of interest

None
Abstract

Objective: To identify feature autophagy-related genes (ARGs) in patients with acute myocardial infarction (AMI) and further investigate their value in the diagnosis of AMI.

Methods: Gene microarray expression data of AMI peripheral blood samples were downloaded from the GSE66360 dataset. The data were randomly classified into a discovery cohort (21 AMI patients and 22 healthy controls) and a validation cohort (28 AMI patients and 28 healthy controls). Differentially expressed ARGs between patients with AMI and healthy controls in the discovery cohort were identified using a statistical software package. Feature ARGs were screened based on support vector machine-recursive feature elimination (SVM-RFE), and an SVM classifier was constructed. Receiver operating characteristic (ROC) analysis was used to investigate the predictive value of the classifier, which was further verified in an independent external cohort.

Results: A total of seven genes were identified based on SVM-RFE. The SVM classifier had an excellent discrimination ability in both the discovery cohort (area under the curve [AUC] = 0.968) and the validation cohort (AUC = 0.992), which was further confirmed in the GSE48060 dataset (AUC = 0.963). Furthermore, the SVM classifier showed outstanding discrimination between AMI patients with and without recurrent events in the independent external cohort (AUC = 0.992). The identified genes are mainly involved in the cellular response to autophagy, macroautophagy, apoptosis, and the FoxO signaling pathway.

Conclusion: Our study identified feature ARGs and indicated their potential roles in AMI diagnosis to improve our understanding of the molecular mechanism underlying the occurrence of AMI.

Keywords: acute myocardial infarction; autophagy-related genes; support vector machine; recursive feature elimination; diagnosis

Introduction

Acute myocardial infarction (AMI), also known as acute heart attack in a clinical background of acute myocardial ischemia, is induced by the sudden blockade or occlusion of a major coronary artery branch, resulting in ischemia or infarction of cardiomyocytes\(^1\). It is estimated that there were approximately 7.29 million AMIs occurred worldwide in 2015, which contributed to high morbidity and mortality in terms of global health\(^2\). The current diagnostic approaches for the evaluation of AMI generally comprise a physical examination, a surface electrocardiogram (ECG), and blood testing for gold standard cardiac biomarkers, such as cTnI/T and CKMB. However, these biomarkers have been criticized for their lack of sufficient specificity and sensitivity\(^3\). Classic risk
factors, such as smoking, hypertension, high serum cholesterol, hypercholesterolemia, diabetes mellitus, and obesity, fail to fully explain the risks of morbidity and mortality from AMI. Furthermore, numerous patients present an inexplicable etiology of myocardial damage. Previous studies have reported that patients with coronary heart disease often lack common risk factors; in fact, more than 20% of patients have no commonly accepted risk factors and 40% have only one. This evidence indicates that genetic factors may also play an important role in the development of AMI.

Autophagy (meaning “self-eating”), a common phenomenon in eukaryotic cells, is a crucial metabolic process for the degradation of cytoplasmic proteins and organelles into amino acids and fatty acids for energy production and recycling. Previous research has revealed that autophagy in ischemic cardiomyocytes can provide essential energy for cell survival by clearing disordered organelles or aging proteins. At present, more than 30 kinds of autophagy-related genes (ARGs) are believed to be closely involved in autophagy. For example, it has been shown that NR4A2 knockdown aggravates cardiomyocyte apoptosis while NR4A2 overexpression ameliorates it. NR4A2 upregulation is regarded as an adaptive response to ischemia-induced cardiomyocyte apoptosis. However, the role of ARGs in AMI has not been fully elucidated.

In this study, we utilized gene expression microarray data from the Gene Expression Omnibus (GEO) database and performed mRNA profiling on two cohorts of AMI patients. Differentially expressed ARG expression profiles were identified in control and AMI samples. The recursive feature elimination (RFE) algorithm for effectively improving classification accuracy was used to select the risk genes among feature ARGs, which were subsequently applied to construct a SVM classifier for distinguishing between AMI and control samples. The predictive value of the classifier in the diagnosis of AMI in two cohorts was explored using receiver operating characteristic (ROC) analysis and further verified in an independent external cohort.

Materials and methods

Patient samples and ARGs

Two RNA sequencing datasets were downloaded from the GEO database. The GSE66360 dataset, which was conducted in circulating endothelial cells (CECs), consisted of a discovery cohort (21 AMI patients and 22 healthy controls) and a validation cohort (28 AMI patients and 28 healthy controls). The GSE48060 dataset, which was conducted in whole blood samples, contained 31 AMI patients and 21 healthy controls. Among the AMI patients, 5 had recurrent events while the remaining 26 were event-free over a 1.5-year follow-up. All datasets were produced using the Affymetrix Human Genome U133 Plus 2.0 Array. A total of 232 ARGs were obtained by searching the
Human Autophagy Database http://autophagy.lu/) and had been confirmed in previous studies as involved in the autophagy process\textsuperscript{14}. The ARG expression matrix was extracted from the GSE66360 dataset using R statistical software (The R Foundation, Vienna, Austria), and a total of 210 ARG expression values were obtained.

**Differentially expressed ARGs**

R statistical software (version 3.5.1; The R Foundation; https://www.r-project.org/) and the Bioconductor linear models for microarray data (LIMMA) package were used to screen the significance analysis of differentially expressed ARGs between AMI and control samples, as well as background correction and normalization between arrays\textsuperscript{15}. The Bonferroni method was applied to perform multiple test corrections. The threshold for identification of differentially expressed ARGs was set to a $P$-value $< 0.05$ and $|\log_2$fold change (FC)$| \geq 1$.

**Feature gene selection and SVM classifier construction**

SVM is a supervised learning model that aims to classify data points by maximizing the distance of a hyperplane for classification and regression analysis with high accuracy\textsuperscript{16}. The RFE algorithm uses the coefficients of the weight vector to compute the feature ranking score\textsuperscript{13}. RFE used for dimensionality reduction clearly contributes to classification accuracy and yields a better classification performance than many other feature reduction methods\textsuperscript{13}. Therefore, to select smaller subsets of the most relevant features, we proposed a model combining SVM and RFE algorithms to search for best parameters for gene selection among all ARGs. To test the predictive value of the identified ARG, a heat map of the genes was clustered using the pheatmap package in R (The R Foundation; https://cran.rproject.org/web/packages/pheatmap/index.html) for all samples in the discovery cohort and subsequently confirmed in the validation cohort. Finally, the optimal feature genes after recursive feature selection were put into an SVM classifier with a radial basis function (RBF) kernel and five-fold cross-validation to make predictions. Furthermore, the performance of the SVM classifier in discriminating AMI from healthy controls in the discovery and validation cohorts was measured as the area under the curve (AUC). The predictive capacity was further validated in an independent external cohort. Moreover, the SVM classifier was used to test the performance in terms of AMI recurrence and non-recurrence in subsets of the GSE48060 dataset, which consisted of 5 AMI patients with recurrent events and 26 event-free patients.

**Functional enrichment analysis of the identified feature genes**

Gene ontology (GO) functional enrichment analysis is widely used in functional studies of large-scale transcriptomic or genomic data and defines the concepts used to describe gene function, including biological process (BP), cellular component (CC), and molecular function (MF). To characterize the functions and pathways of the identified feature genes, GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment
analyses were performed to identify functional and metabolic pathways using the biological tool. The clusterProfiler package (The R Foundation) was used to explore the biological meaning and pathways behind numerous genes.

**Statistical analysis**

R statistical software (version 3.5.1; The R Foundation; https://www.r-project.org/) was used to perform all the statistical analyses in this study. The feature ARG selection process was performed using the RFE function in the caret package (The R Foundation; https://cran.r-project.org/web/packages/caret/index.html). The SVM classifier from R package e1071 (The R Foundation) was used for classification analysis of the identified feature genes. A five-fold cross-validation was utilized to evaluate the performance of the prediction model. ROC analyses were performed using the ROCR package in R (The R Foundation), and the AUCs were calculated to evaluate the predictive ability of the built model. Hierarchical clustering analysis was conducted for the identified feature genes using the pheatmap package in R (The R Foundation) based on the Euclidean distance. P-value < 0.05 were considered to indicate a statistically significant difference.

**Results**

**Identification of feature ARGs**

After background correction and normalization, ten differentially expressed feature genes were identified among the 210 ARGs. All of the aforementioned feature genes were upregulated (Figure 1). We used RFE analysis to further screen for the optimal feature genes in the discovery cohort. As shown in Figure 2, using RFE, we were able to reduce the features to seven genes (WDFY3, TP53INP2, GABARAPLI, CDKN1A, DDIT3, NAMPT, and FOS) with minimum root mean-square error. Hierarchical clustering analysis was carried out for all samples in the discovery cohort according to the seven identified feature genes. As shown in Figure 3A, hierarchical cluster analysis revealed that feature gene expression patterns varied between the AMI and control groups. Hierarchical clustering analysis in the validation cohort also revealed that the patients were clearly separated into two clusters based on the expression levels of the identified feature genes (Figure 3B). To validate the feature ARG expression levels, the identified genes were selected for validation using the validation cohort. As shown in Figure 4, the expression levels of the feature genes in AMI tissue were significantly higher than those in the control group (P < 0.05). Therefore, the optimal seven feature genes were subsequently utilized to construct the SVM classifier, which was further verified using the independent cohort.

**Diagnostic value of feature ARGs in AMI**
We further explored the diagnostic value of the feature genes in AMI using SVM with the RBF kernel method and five-fold cross-validation. As shown in Figure 5A, the results of cross-validation suggested that the feature gene-based SVM classifier yield an accurate classifying of the discovery cohort between AMI and healthy controls with an AUC of 0.968 (95% confidence interval [CI]: 0.863-0.998), a sensitivity of 90.48% (95% CI: 69.6%-98.8%), and a specificity of 95.45% (95% CI: 77.2%-99.9%). The validation cohort was used to validate the performance of the SVM classifier in predicting AMI. The SVM classifier also demonstrated excellent discriminatory ability in this group with an AUC of 0.992 (95% CI: 0.922-1.00), a sensitivity of 96.43% (95% CI: 81.7%-99.9%; Figure 5B). Furthermore, the strong discrimination power was confirmed in the independent cohort with an AUC of 0.963 (95% CI: 0.870-0.996), a sensitivity of 90.32% (95% CI: 74.2%-98.0%), a specificity of 90.48% (95% CI: 69.6%-98.8%; Figure 5C). In addition, we explored the discrimination power of the classifier in terms of recurrent AMI. Surprisingly, the classifier presented outstanding discrimination ability for recurrent AMI with an AUC of 0.992 (95% CI: 0.874-1.00), a sensitivity of 100% (95% CI: 47.8%-100.0%), and a specificity of 96.15% (95% CI: 80.4%-99.9%; Figure 5D).

Functional analysis of the feature ARGs

To study the functional roles of the identified feature genes, GO enrichment and KEGG pathway analyses were performed. Cellular responses to extracellular/external stimuli, autophagy, and macroautophagy were the most significantly enriched BPs (Figure 6A). In addition, apoptosis, autophagy, and the FoxO signaling pathway were considered to be the most remarkably enriched pathways (Figure 6B).

Discussion

AMI is the primary cause of human mortality and morbidity worldwide. The sensitivity and specificity of the existing biomarkers for AMI remain to be further elucidated. Recently, an association between autophagy and cardiac disorders such as myocardial infarction has been demonstrated. However, numerous autophagy genes in AMI are not fully understood. The role of autophagy in diseases has received increasing attention from researchers. Therefore, to identify the ARGs in AMI, the discovery cohort (21 AMI samples and 22 healthy controls) from the GSE66360 dataset was used to screen differentially expressed genes in patient tissues compared with those in healthy tissues. Using the RFE method, the feature genes in AMI samples were identified, which allowed the AMI samples to be distinguished from the control samples. The identified feature genes were used to build an SVM classifier with an AUC of 0.968 for the patient samples. The discrimination power of the classifier for the
validation cohort and the independent validation cohort was 0.992 and 0.963, respectively. Furthermore, the feature autophagy gene-based classifier successfully distinguished between recurrent AMI events and non-recurrent events with an AUC of 0.992. Therefore, the present study suggests that feature ARGs may be useful markers for identifying patients with AMI. In the original paper, 11 gene signatures (HBEGF, SYTL3, EDN1, NRP2, NFkBIA, VPS8, NRP4A3, SULF1, RNASE1, CCL20, and MGP) derived from gene expression profiling at the transcriptome level in whole blood were identified. In our study, we identified seven feature ARGs in patients with AMI. There was no overlap between the ARGs and 11 gene signatures. The discrimination ability of the SVM classifier was slightly less than that of the original paper in terms of AUC. This is reasonable because the ARGs are parts of all the genes used in the original paper.

Our results demonstrated the potential value of the feature ARGs for the diagnosis of AMI in the clinical setting. GABARAPL1, CDKN1A, DDIT3, NAMPT, FOS, WDFY3, and TP53INP2 were identified as potential genes associated with AMI and recurrence. GABA type A receptor associated protein like 1 (GABARAPL1) genes are located on human chromosome 12p12.3. They are involved in the proliferation, migration, angiogenesis and autophagy of endothelial progenitor cells via regulation of miR-195 expression. GABARAPL1 has also been reported to be involved in autophagy signaling in myocardial infarction-induced muscle atrophy in rats. The CDKN1A gene is a broad-acting cyclin-dependent kinase inhibitor that encodes the p21 protein. In a Langendorff mouse model of hypoxia-reoxygenation myocardial injury, Rev-Erbα gene deletion or antagonist treatment weakened injury at the time of the sleep-to-wake transition by enhancing the expression of the ischaemia-reperfusion injury modulator CDKN1alp21. CDKN1alp21 was regarded as a downstream target of Rev-Erbα in the human myocardium, and Rev-Erbα protected cardiomyocytes from cell death. DDIT3 (DNA damage inducible transcript 3, also known as CHOP) is a key mediator of the endoplasmic reticulum stress response. It is considered to be a vital apoptotic signaling molecule after axonal injury. Furthermore, DDIT3 protein overexpression has been shown to induce apoptosis. Mammalian cell nicotinamide phosphoribosyl transferase (NAMPT) is the rate-limiting enzyme for nicotinamide adenine dinucleotide salvage synthesis in mammals. Increasing evidence suggests that NAMPT is involved in metabolism and the immune response and plays pleiotropic roles in vascular homeostasis by regulating the functions of endothelial cells, vascular smooth muscle cells, endothelial progenitor cells, and perivascular cells. Overexpression of NAMPT has been shown to enhance SIRT1 activity, and protect cells from apoptosis via the activation of SIRT3 and SIRT4. NAMPT can contribute to human vascular smooth muscle cell survival by activating SIRT1 and preventing the accumulation of p53. Therefore, NAMPT may be an effective protective agent in the vasculature. FOS, also known as the
proto-oncogene c-fos, has been demonstrated to be associated with apoptotic cell death. The high expression of c-fos may play a role in the pathogenesis of viral myocarditis, cardiac ischemia-reperfusion, myocardial stunning, and heart failure. For example, c-fos mRNA expression and protein expression rapidly increased at three days after coronary ligation in a rat model; furthermore, the infarct size was diminished in rats treated with c-fos monoclonal antibody compared to that of the control group. This observation indicated that c-fos plays a vital role in myocardial lesions and is likely to be involved in the pathogenesis of AMI. As for WDFY3 and TP53INP2, their roles in AMI have not been well studied and thus should be the focus of further research.

As revealed by the GO and KEGG pathway analyses, the identified ARGs were mainly involved in apoptosis, autophagy, and the FoxO signaling pathway. Autophagy is a physiological process for tissue survival. Apoptosis is a process of programmed cell death that naturally occurs in human body homeostasis and is known to be associated with myocardial ischemia. The balance between autophagy and apoptosis determines cell survival and cardiac function, and the inflammatory response may be protective in the early stages of AMI via stimulation of myocyte autophagy. However, it has also been noted that excessive autophagy may result in detrimental effects of cells. Inflammation is an important regulator of autophagy. For example, TNF-α induce autophagy of cardiomyocytes in vitro and in vivo and contributes to autophagy by rapamycin protection against LPS-mediated myocyte apoptosis. The FoxO signaling pathway has been reported to be involved in hypoxia injury in cardiac cells. FOXO3 serves as a surveillance mechanism to detect and correct autophagy flux disruptions as well as ensure that the cells where autophagy is lacking are subjected to apoptosis. Together, these finding and our results indicate that autophagy and ARGs play an important role in AMI.

However, we also recognize the shortcomings of this study. First, we failed to validate the discriminatory ability of the SVM classifier in the independent patient cohort because of the lack of patient dataset with recurrent events. Second, there was an extremely significant difference in age distribution between AMI patients and control samples; thus, we were unable to adjust for the potential confounding effect in this study. Furthermore, it should be noted that this research was based on bioinformatics analyses. Therefore, further validation from both in vivo and in vitro studies is needed.

In summary, WDFY3, TP53INP2, GABARAPL1, CDKN1A, DDIT3, NAMPT, and FOS are differentially expressed ARGs in AMI patients that may be used as potential biomarkers in the diagnosis of AMI. Identification of these ARGs helps to improve understanding of the molecular mechanism underlying AMI occurrence.
Disclosure of grants or other funding

None

Ethics approval and consent to participate

This study does not contain any work with human participants conducted by any of the authors.

Competing interests

The authors declare that they have no competing interests.

Data statement

The raw data of this study are derived from the GEO data portal (https://www.ncbi.nlm.nih.gov/geo/; Accession number: GSE66360, GSE48060), which is a publicly available database.

Acknowledgement

The authors gratefully acknowledge The Gene Expression Omnibus (GEO) database, which made the data of AMI available.

Authors’ contributions

Yajuan Du and Enfa Zhao is the principle investigator. Yajuan Du and Enfa Zhao conducted statistical analysis and draft the manuscript. Yajuan Du performed data management and bioinformatics analysis. Enfa Zhao and Yushun Zhang edited and revised the manuscript. All authors read and approved the final manuscript.

References

1. Musher DM, Abers MS, Corrales-Medina VF. Acute Infection and Myocardial Infarction. N Engl J Med. 2019;380(2):171-176.
2. Roth GA, Johnson C, Abajobir A, et al. Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015. J Am Coll Cardiol. 2017;70(1):1-25.
3. Park JP, Park MK, Yun JW. Proteomic biomarkers for diagnosis in acute myocardial infarction. Biomarkers. 2011;16(1):1-11.
4. Agewall S, Beltrame JP, Reynolds HR, et al. ESC working group position paper on myocardial infarction with non-obstructive coronary arteries. Eur Heart J. 2017;38(3):143-153.
5. Ajani UA, Ford ES. Has the risk for coronary heart disease changed among U.S. adults? J Am Coll
Cardiol. 2006;48(6):1177-1182.

6. Hozawa A, Folsom AR, Sharrett AR, Chambless LE. Absolute and attributable risks of cardiovascular disease incidence in relation to optimal and borderline risk factors: comparison of African American with white subjects—Atherosclerosis Risk in Communities Study. Arch Intern Med. 2007;167(6):573-579.

7. Cătană C-S, Atanasov AG, Berindan-Neagoe I. Natural products with anti-aging potential: Affected targets and molecular mechanisms. Biotechnol Adv. 2018;36(6):1649-1656.

8. Moe GW, Marín-García J. Role of cell death in the progression of heart failure. Heart Fail Rev. 2016;21(2):157-167.

9. Díaz M, García C, Sebastiani G, de Zegher F, López-Bermejo A, Ibáñez L. Placental and Cord Blood Methylation of Genes Involved in Energy Homeostasis: Association With Fetal Growth and Neonatal Body Composition. Diabetes. 2017;66(3):779-784.

10. Wildenberg ME, Koelink PJ, Diederik K. The ATG16L1 risk allele associated with Crohn’s disease results in a Rac1-dependent defect in dendritic cell migration that is corrected by thiopurines. Mucosal Immunol. 2017;10(2):352-360.

11. Liu H, Liu P, Shi X, Yin D, Zhao J. NR4A2 protects cardiomyocytes against myocardial infarction injury by promoting autophagy. Cell Death Discov. 2018;4:27-27.

12. Matwin S, Sazonova V. Direct comparison between support vector machine and multinomial naive Bayes algorithms for medical abstract classification. J Am Med Inform Assoc. 2012;19(5):917-917.

13. Guyon I, Weston J, Barnhill S, Vapnik V. Gene Selection for Cancer Classification using Support Vector Machines. Machine Learning. 46(1-3):389-422.

14. Moussay E, Kaoma T, Baginska J, et al. The acquisition of resistance to TNFα in breast cancer cells is associated with constitutive activation of autophagy as revealed by a transcriptome analysis using a custom microarray. Autophagy. 2011;7(7):760-770.

15. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol. 2004;3:Article3-Article3.

16. Cherkassky V. The nature of statistical learning theory-. IEEE Trans Neural Netw. 1997;8(6):1564-1564.

17. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284-287.

18. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROC: visualizing classifier performance in R. Bioinformatics. 2005;21(20):3940-3941.

19. Boersma E, Mercado N, Poldermans D, Gardien M, Vos J, Simoons ML. Acute myocardial infarction. Lancet. 2003;361(9360):847-858.

20. Sun T, Li MY, Li PF, Cao JM. MicroRNAs in Cardiac Autophagy: Small Molecules and Big Role. Cells. 2018;7(8).

21. Muse ED, Kramer ER, Wang H, et al. A Whole Blood Molecular Signature for Acute Myocardial Infarction. Scientific reports. 2017;7(1):12268.

22. Mo J, Zhang D, Yang R. MicroRNA-195 regulates proliferation, migration, angiogenesis and autophagy of endothelial progenitor cells by targeting GABARAPL1. Biosci Rep. 2016;36(5):e00396.

23. Jannig PR, Moreira JBN, Bechara LRG, et al. Autophagy signaling in skeletal muscle of infarcted rats. PLoS One. 2014;9(1):e85820-e85820.

24. Warfel NA, El-Deiry WS. p21WAF1 and tumourigenesis: 20 years after. Curr Opin Oncol. 2013;25(1):52-58.

25. Montaigne D, Marechal X, Modine T, et al. Daytime variation of perioperative myocardial injury in cardiac surgery and its prevention by Rev-Erbα antagonism: a single-centre propensity-matched cohort
study and a randomised study. *The Lancet*. 2018;391(10115):59-69.

26. Adachi S, Ito H, Tamamori-Adachi M, et al. Cyclin A/cdk2 activation is involved in hypoxia-induced apoptosis in cardiomyocytes. *Circulation Research*. 2001;88(4):408-414.

27. Kim I, Xu W, Reed JC. Cell death and endoplasmic reticulum stress: disease relevance and therapeutic opportunities. *Nat Rev Drug Discov*. 2008;7(12):1013-1030.

28. Wang P, Li W-L, Liu J-M, Miao C-Y. NAMPT and NAMPT-controlled NAD Metabolism in Vascular Repair. *J Cardiovasc Pharmacol*. 2016;67(6):474-481.

29. Bi T-q, Che X-m. Nampt/PBEF/visfatin and cancer. *Cancer Biol Ther*. 2010;10(2):119-125.

30. van der Veer E, Nong Z, O'Neil C, Urquhart B, Freeman D, Pickering JG. Pre-B-cell colony-enhancing factor regulates NAD+-dependent protein deacetylase activity and promotes vascular smooth muscle cell maturation. *Circulation research*. 2005;97(1):25-34.

31. Zhang S, Zhang M, Goldstein S, et al. The effect of c-fos on acute myocardial infarction and the significance of metoprolol intervention in a rat model. *Cell Biochem Biophys*. 2013;65(2):249-255.

32. Anversa P, Cheng W, Liu Y, Leri A, Redaelli G, Kajstura J. Apoptosis and myocardial infarction. *Basic Res Cardiol*. 1998;93 Suppl 3:8-12.

33. Wang X, Guo Z, Ding Z, Mehta JL. Inflammation, Autophagy, and Apoptosis After Myocardial Infarction. *J Am Heart Assoc*. 2018;7(9):e008024.

34. Bai Y-D, Yang Y-R, Mu X-P, et al. Hydrogen Sulfide Alleviates Acute Myocardial Ischemia Injury by Modulating Autophagy and Inflammation Response under Oxidative Stress. *Oxid Med Cell Longev*. 2018;2018:3402809-3402809.

35. Yuan H, Perry CN, Huang C, et al. LPS-induced autophagy is mediated by oxidative signaling in cardiomyocytes and is associated with cytoprotection. *Am J Physiol Heart Circ Physiol*. 2009;296(2):H470-H479.

36. Fitzwalter BE, Thorburn A. FOXO3 links autophagy to apoptosis. *Autophagy*. 2018;14(8):1467-1468.
Figure legends

Figure 1 Differential expression of feature autophagy-related genes in acute myocardial infarction tissue samples.

Figure 2 A plot of feature gene selection by recursive feature elimination

Figure 3 Hierarchical clustering analysis demonstrates the identified autophagy-related gene expression patterns in the discovery cohort (A) and validation cohort (B) in the GSE66360 dataset.

Figure 4 Validation of the expression levels of identified autophagy-related genes in patients with acute myocardial infarction (AMI) in the validation cohort (28 AMI patients and 28 healthy controls)

Figure 5 Receiver operating characteristic curves of the support vector machine (SVM) classifier for the discovery cohort (A: 21 acute myocardial infarction [AMI] patients and 22 healthy controls), the validation cohort (B: 28 AMI patients and 28 healthy controls), the independent external cohort (C: 31 AMI patients and 21 healthy controls), and recurrent prediction in subsets of the independent external cohort (D: 5 recurrent AMI patients and 22 non-recurrent patients).

Figure 6 Gene ontology and Kyoto Encyclopedia of Genes and Genomes pathway analyses of the identified feature genes.
Gene expression data showing significant differences in expression levels between AMI and Normal groups. The p-values associated with each gene are as follows:

- **CDKN1A**: P = 2.5e-05
- **DDIT3**: P = 8.7e-06
- **FOS**: P = 4.5e-06
- **GABARAPL1**: P = 4.6e-11
- **NAMPT**: P = 5e-06
- **TP53INP2**: P = 3.4e-07
- **WDFY3**: P = 4.8e-05
Specificity
Sensitivity

1.0 0.8 0.6 0.4 0.2 0.0
0.0 0.2 0.4 0.6 0.8 1.0

AUC=0.968 (95%CI=0.863-0.998)
Sensitivity=90.48% (95%CI=69.6%-98.8%)
Specificity=95.45% (95%CI=77.2%-99.9%)

AUC=0.992 (95%CI=0.922-1.00)
Sensitivity=96.43% (95%CI=81.7%-99.9%)
Specificity=96.43% (95%CI=81.7%-99.9%)

AUC=0.963 (95%CI=0.870-0.996)
Sensitivity=90.32% (95%CI=74.2%-98.0%)
Specificity=90.48% (95%CI=69.6%-98.8%)

AUC=0.992 (95%CI=0.874-1.00)
Sensitivity=100% (95%CI=47.8%-100%)
Specificity=96.15% (95%CI=80.4%-99.9%)
