

RESEARCH

Open Access



# Identification of novel biomarkers based on lipid metabolism-related molecular subtypes for moderately severe and severe acute pancreatitis

Jifeng Liu<sup>1†</sup>, Lei Zhong<sup>1†</sup>, Yunshu Zhang<sup>1†</sup>, Jingyuan Ma<sup>1</sup>, Tong Xie<sup>1</sup>, Xu Chen<sup>1\*</sup>, Biao Zhang<sup>1\*</sup> and Dong Shang<sup>1\*</sup>

## Abstract

**Background** Acute pancreatitis (AP) is an unpredictable and potentially fatal disorder. A derailed or unbalanced immune response may be the root of the disease's severe course. Disorders of lipid metabolism are highly correlated with the occurrence and severity of AP. We aimed to characterize the contribution and immunological characteristics of lipid metabolism-related genes (LMRGs) in non-mild acute pancreatitis (NMAP) and identify a robust subtype and biomarker for NMAP.

**Methods** The expression mode of LMRGs and immune characteristics in NMAP were examined. Then LMRG-derived subtypes were identified using consensus clustering. The weighted gene co-expression network analysis (WGCNA) was utilized to determine hub genes and perform functional enrichment analyses. Multiple machine learning methods were used to build the diagnostic model for NMAP patients. To validate the predictive effectiveness, nomograms, receiver operating characteristic (ROC), calibration, and decision curve analysis (DCA) were used. Using gene set variation analysis (GSVA) and single-cell analysis to study the biological roles of model genes.

**Results** Dysregulated LMRGs and immunological responses were identified between NMAP and normal individuals. NMAP individuals were divided into two LMRG-related subtypes with significant differences in biological function. The cluster-specific genes are primarily engaged in the regulation of defense response, T cell activation, and positive regulation of cytokine production. Moreover, we constructed a two-gene prediction model with good performance. The expression of CARD16 and MSGT1 was significantly increased in NMAP samples and positively correlated with neutrophil and mast cell infiltration. GSVA results showed that they are mainly upregulated in the T cell receptor complex, immunoglobulin complex circulating, and some immune-related routes. Single-cell analysis indicated that CARD16

<sup>†</sup>Jifeng Liu, Lei Zhong and Yunshu Zhang contributed equally to this work and share the first authorship.

\*Correspondence:

Xu Chen

1169777928@qq.com

Biao Zhang

2962674630@qq.com

Dong Shang

shangdong@dmu.edu.cn

Full list of author information is available at the end of the article



was mainly distributed in mixed immune cells and macrophages, and MGST1 was mainly distributed in exocrine glandular cells.

**Conclusions** This study presents a novel approach to categorizing NMAP into different clusters based on LMRGs and developing a reliable two-gene biomarker for NMAP.

**Keywords** Acute pancreatitis, Lipid metabolism, Machine learning, Biomarkers, Molecular clusters

## Introduction

Acute pancreatitis (AP) is an inflammatory disease of the pancreas whose course depends on severity. According to the 2012 revision of the Atlanta classification, AP could be classified as mild, moderately severe, or severe [1]. Moderately severe acute pancreatitis (MSAP) is characterized by temporary organ failure, local complications, or an aggravation of a co-morbid condition. Severe acute pancreatitis (SAP) refers to continued organ failure that lasts more than 48 h [1]. Non-mild acute pancreatitis (NMAP), including MSAP and SAP, occurs in 20% of AP patients [2]. Although treatment methods such as early enteral feeding, fluid resuscitation, and organ support therapy are often used in clinical settings, the mortality rate of NMAP can reach 35%, which is much greater than MAP [2]. A derailed or unbalanced immune response may be the root of the disease's severe course. Stimulation of acinar cells initiates a local immune response, and cells of the innate immune system from the bone marrow, such as monocytes and neutrophils, are recruited to areas of damaged tissue in the pancreas [3]. The immediate response of the immune system to local cellular injury influences the progression of AP. Intervention targeting the immune response mechanisms involved is one of the most promising therapeutic options for AP. Therefore, there is a need to study the immune characteristics in AP and find potentially modifiable risk factors and novel biomarkers for the early identification of high-risk AP individuals.

Hypertriglyceridemia has long been thought of as the third most frequent cause of AP [4]. With more studies being done on dyslipidemia during AP, the connection between hypertriglyceridemia and AP is becoming more and more distinct [4]. Acute renal damage, acute respiratory distress syndrome, and a longer duration of hospital stay are all independently linked to higher blood triglyceride levels during AP, according to a large number of recent clinical investigations [5–7]. The study found that AP patients with hyperlipidemia were at greater risk for more severe pancreatitis [8]. At the same time, it was discovered that strict control of triglyceride concentration after acute pancreatitis presentation decreased the probability of recurrence [8]. In addition, it has been found that hypertriglyceridemia during AP, regardless of etiology, is linked to continuous organ failure, which is a leading

decisive factor in AP mortality [9–11]. As a result, it is promising to study the molecular subtypes and develop new diagnostic biomarkers for NMAP patients based on lipid metabolism-related genes (LMRGs).

First, we performed a systematic analysis of the LMRGs and immunological characteristics between normal and NMAP individuals. Then, based on DE-LMRGs, NMAP individuals were split into two lipid metabolism-related subgroups. The differences in immune cell infiltration and key pathways were studied further between the two subgroups. NMAP-specific and cluster-specific genes were identified using weighted gene co-expression network analysis (WGCNA), and the biological roles and pathways of gene enrichment were investigated. A two-gene diagnostic model consisting of CARD16 and MGST1 was discovered by applying the least absolute shrinkage and selection operator (LASSO), random forest (RF), and support vector machine recursive feature elimination (SVM-RFE). The receiver operating characteristic (ROC) curves, nomogram, decision curve analysis (DCA), and calibration curves were used to verify the diagnostic model's effectiveness. At last, we explored the biological characteristics, single-cell maps, and immunofluorescence of these two model genes and constructed their ceRNA regulatory networks. These findings would improve our comprehension of the lipid metabolism mechanism in AP and offer fresh ideas for the early detection of NMAP.

## Method

### Data preparation

The GSE194331 dataset, including 87 AP patients and 32 normal samples, was obtained from the Gene Expression Omnibus (GEO) database, and the clinical information on the samples can be obtained from previous studies [12]. After excluding the MAP samples, 30 NMAP patients (20 MSAP and 10 SAP) and 32 normal individuals were included in this research. Considering the number of genes and their relevance score, as well as previous studies, we selected 1004 LMRGs (relevance score > 10) from the GeneCards database for this analysis [13, 14]. With adjusted  $P < 0.05$  and  $|\log_{2}FC| > 2$  as the criterion, differentially expressed genes (DEGs) between normal and NMAP samples were found by the “limma” program [15].

### Evaluating the immune cell infiltration

We looked at the differences in the immune microenvironment between people with NMAP and people who are healthy using the CIBERSORT algorithm, which can figure out the percentages of infiltration of 22 different types of immune cells. Individuals with a  $p$ -value  $< 0.05$  represented that the evaluation of the proportions of 22 immune cell subsets generated by CIBERSORT was accurate. These samples were then utilized for further assessing the variations in immune infiltration between different molecular subtypes [16].

### Cluster analysis

The “ConsensusClusterPlus” package was applied to cluster NMAP samples using the expressed levels of the DE-LMRGs [17]. The reliability of clustering results was demonstrated by PCA. The biological characteristics of different subtypes were further compared.

### Identifying specific genes in key modules

We used the “WGCNA” package in R to find modules and genes that are linked to NMAP and lipid metabolism-related clusters [18]. Using the optimal soft power, the weighted neighbor matrices were built and converted to a topological overlap matrix (TOM). The minimum module size was set to 100, and then the TOM dissimilarity metric was used to construct the modules. Genes with gene significance (GS)  $> 0.5$  and module membership (MM)  $> 0.8$  were considered specific genes [19].

### Functional enrichment analysis

The STRING database was utilized to examine protein interactions using a composite score  $> 0.15$  as a condition to determine the validity of such interactions [20]. Functional enrichment analysis was conducted using the Metascape database [21].

### Identifying the best model genes for NMAP

LASSO, RF, and SVM-RFE were performed to filter genes. LASSO was performed by the R package “glmnet” to avoid overfitting and by 10-fold cross-validation to tune the optimal penalty parameter  $\lambda$  [22, 23]. RF was performed using the R package “Random Forest” (ntree = 500) [24]. The feature importance was determined by the mean decrease Gini index calculated by RF, and genes with relative importance  $> 1.5$  were determined as characteristic genes. Meanwhile, SVM-RFE was employed using the R package “e1071” [25]. SVM-RFE uses the principle of structural risk minimization, which aims to optimize the learning performance by minimizing the empirical error. Then, the hub genes for

the following studies were chosen from the intersection of the three subsets. The “pROC” package was used to calculate the ROC curves to evaluate the model’s predictability [26]. The correlation of the model genes CARD16 and MGST1 in peripheral blood was assessed by the GTEx dataset in GEPIA, an online tool for bioinformatics based on TCGA and GTEx [27].

### Establishment of a nomogram

A nomogram integrating model genes was constructed by the “rms” R package [28]. Meanwhile, the calibration and DCA curves were used to evaluate the predictive accuracy of the nomogram [29].

### Analysis of the model genes

Gene set variation analysis (GSVA) was conducted for model genes by the “GSVA” package [30]. The association between model genes and immunochemicals was evaluated based on CIBERSORT results. Meanwhile, using the Human Protein Atlas (HPA: <https://www.proteinatlas.org/>), immunofluorescence was used to demonstrate cellular localization. The latest version of the HPA database has integrated large datasets with single-cell type level information, and the single-cell type data is integrated into the gene search results page, making it possible for visitors to access all the information from a gene-specific perspective [31]. Therefore, we also examined the single-cell type atlases of the model genes in the pancreatic tissues using the HPA platform.

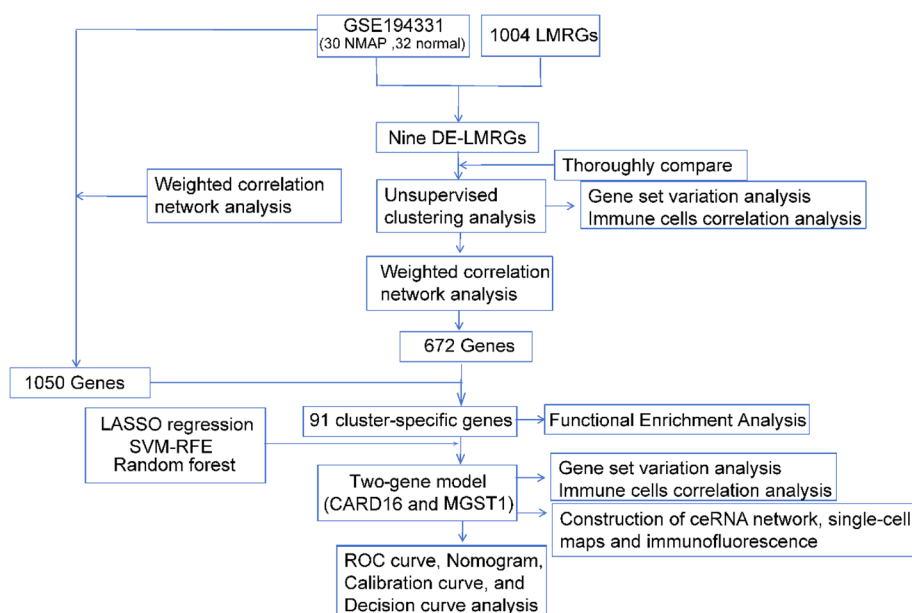
### Construction of ceRNA network

To discover the ceRNA network that may be regulated by model genes. The TargetScan, miRDB, and miRanda databases were utilized to forecast miRNA-mRNA pairs [32]. Only genes that were simultaneously included in all three databases were considered potential mRNA targets for additional investigation. The spongeScan database was utilized to forecast miRNA-lncRNA pairs [33]. Cytoscape was finally used to visualize the ceRNA network.

## Result

### Dysregulation of LMRGs and immune characteristics in NMAP

The current study was illustrated in the flow chart (Fig. 1). Figure 2A illustrates the DEGs between NMAP and normal individuals. We took the intersection of DEGs and LMRGs, and nine LMRGs were found to have differential expression (Fig. 2B). Among them, the expression levels of XIST and ALOX15 were lower, whereas CES1, RETN, HP, IL10, CYP19A1, PPARG, and ARG1 gene expression levels were higher in NMAP than in normal samples (Fig. 2C). Correlation analysis showed mostly positive



**Fig. 1** The research's analytical workflow in detail

correlations between DE-LMRGs, except for ALOX15 and XIST (Fig. 2D). Figure 2E shows the location of the DE-LMRGs in chromosomes.

We employed the CIBERSORT algorithm to assess the infiltration proportions of various immune cell subpopulations within each sample and to explore differences in immune infiltration between NMAP and normal samples. It was found that NMAP patients presented higher infiltration levels of T cells CD4 memory activated, neutrophils, and mast cells resting. The number of T cells CD8, T cells CD4 memory resting, B cells naive, and NK cells resting was higher in normal individuals (Fig. 2F and G). Correlation analyses showed that except for CES1 and XIST, the remaining seven DE-LMRGs were significantly positively or negatively correlated with T cells gamma delta and T cells CD4 memory resting (Fig. 2H).

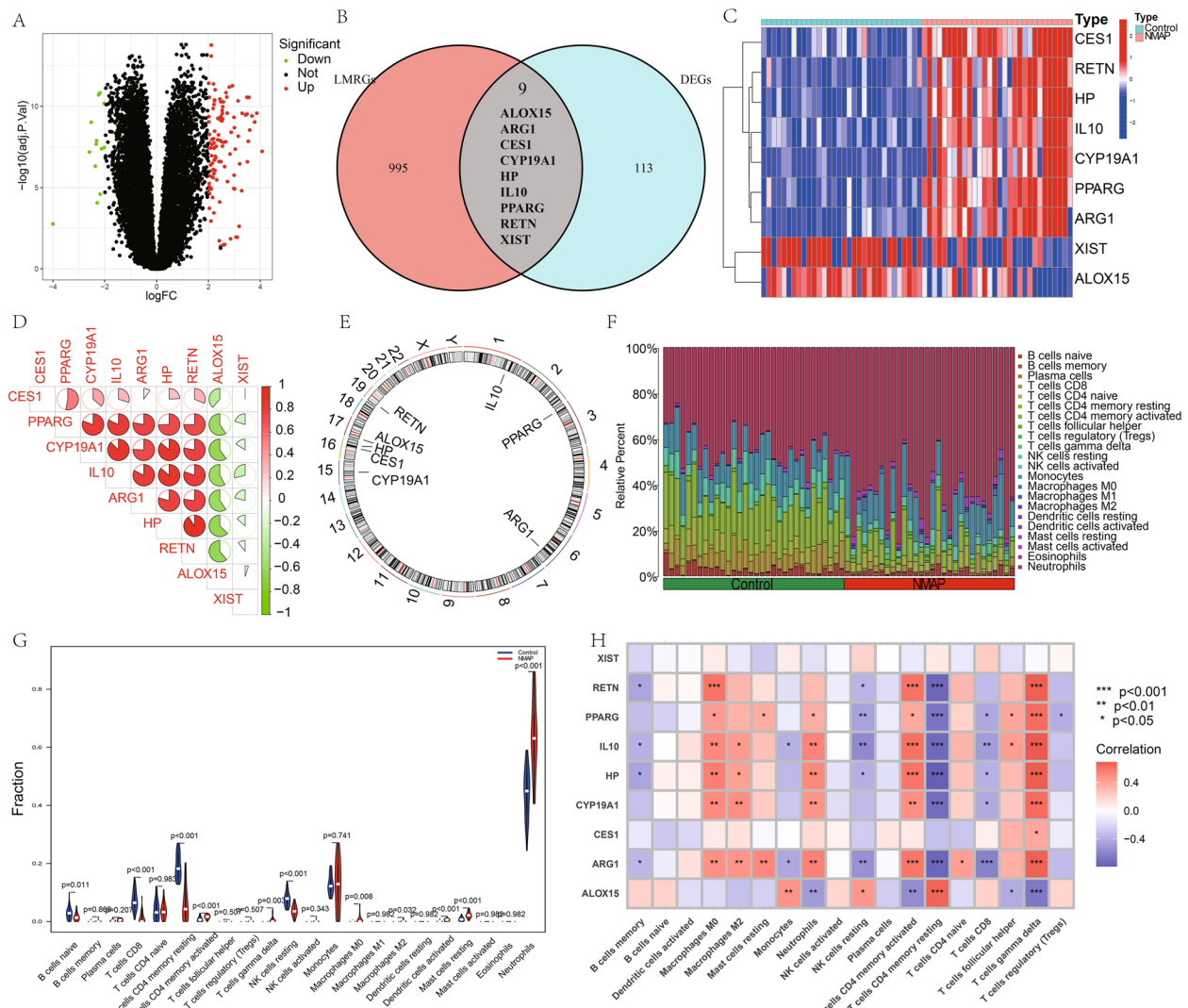
#### Identification of LMRG-related clusters in NMAP

To further understand the LMRG-related expression patterns in NMAP, we classified the NMAP data based on nine DE-LMRG expression profiles using a consensus clustering technique. When the NMAP individuals were split into two clusters, the results were the most stable (Fig. 3A). A significant difference across the two groups was identified by PCA (Fig. 3B). We looked at the gene expression variations of 9 LMRGs between C1 and C2 to investigate the molecular disparities among clusters. PPARG, HP, CYP19A1, ARG1, IL10, RETN, and CES1 were all strongly elevated in the C2 cluster, whereas ALOX15 was substantially decreased (Fig. 3C).

The differences in immune cell infiltration between the two clusters were then compared. We found that the C1 group had a significantly higher abundance of T cells CD4 memory resting and NK cells resting, whereas the C2 group had a significantly higher abundance of T cells CD4 memory activated, T cells gamma delta, and macrophages (Fig. 3D and E). Additionally, we used GSVA to look into possible distinctions in biological function among these two clusters. The findings showed that C1 was primarily involved in the regulation of endoplasmic reticulum tubular network organization, regulation of plasma cell differentiation, and protein export, while C2 was primarily enhanced in antigen processing and presentation, regulation of histone modification, and methylation (Fig. 3F and G).

#### Gene module screening and co-expression network development

WGCNA was first utilized to discover the important modules related to NMAP. The scale-free R2 parameter was set to 0.9 to identify co-expressed modules (Fig. 4A). The most relevant module was the blue one (Fig. 4B). For further investigation, the hub genes in the blue module were picked (Fig. 4C). Given the important role of lipid metabolism in AP and the significant biological differences between LMRG-related molecule clusters, WGCNA was also used to assess the critical modules that had strong connections to clusters associated with LMRG (Fig. 4D). The LMRG-related clusters were most



**Fig. 2** The comparison between NMAP and normal samples. **A** Volcano plot of the DEGs. **B** Venn diagram of the LMRGs and the DEGs. **C** Heatmap showing the expression levels of 9 DE-LMRGs in C1 and C2. **D** The correlation between nine DE-LMRGs. **E** Location of the DE-LMRGs in chromosomes. **F, G** Comparison of 22 immune cell infiltration levels between NMAP and normal samples. **H** The relationship between nine DE-LMRGs and immune cells

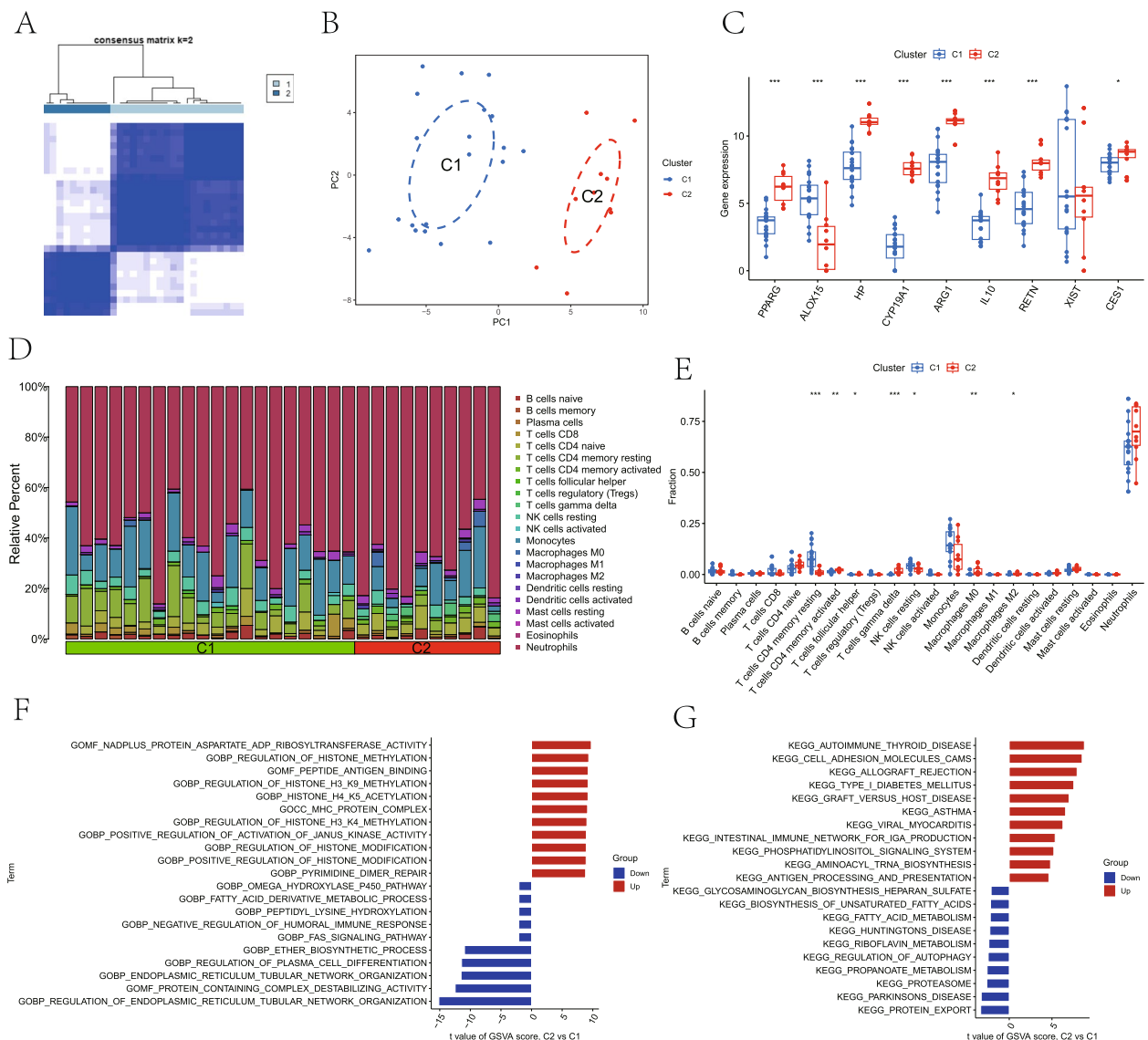
connected to the turquoise module (Fig. 4E). The crucial genes for the turquoise module were also picked (Fig. 4F).

### Functional enrichment analysis

When the module-associated genes of LMRG-related clusters were intersected with the module-associated genes of NMAP, a total of 91 cluster-specific genes were found (Fig. 5A). The PPI analysis showed that most genes were closely linked, except for POR, ROPN1L, NDUFAF1, C20orf24, and PNPLA1 (Fig. 5B). These genes were considerably abundant in the regulation of the defense response and T cell activation, according to the Metascape (Fig. 5C).

### Construction of the model

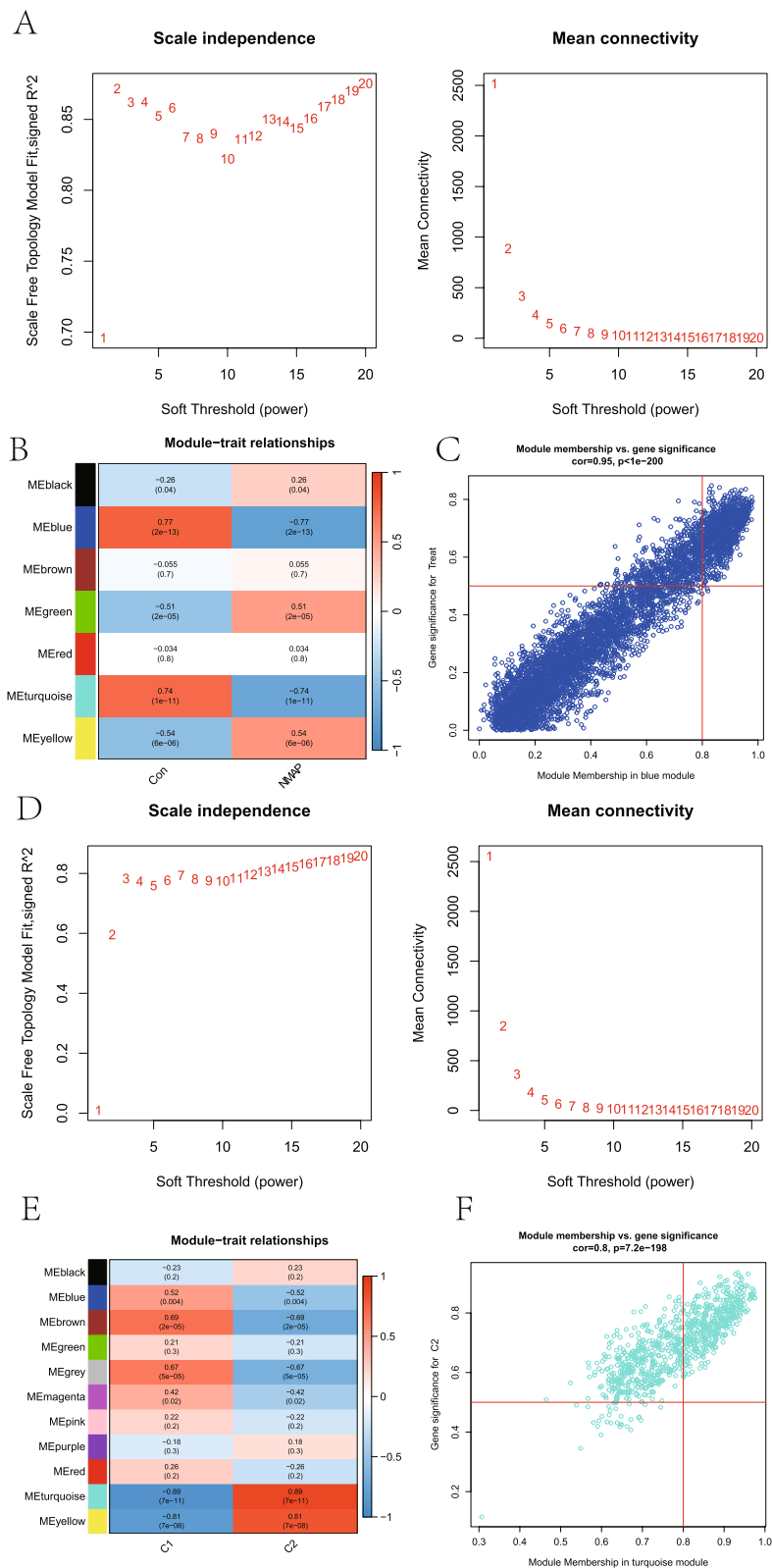
According to the 91 cluster-specific genes, three machine learning models were created. Six genes have been discovered to be possible biomarkers for diagnosis using the LASSO regression technique (Fig. 6A and B). Eighteen genes were extracted from these genes as candidate biomarkers by the SVM-RFE algorithm (Fig. 6C and D). Regarding the RF algorithm, six genes with scores for importance greater than 1.5 were included in the ensuing analysis (Fig. 6E and F). Finally, the two genes (CARD16 and MGST1) were then overlaid by a Venn diagram and used as powerful diagnostic biomarkers (Fig. 6G).



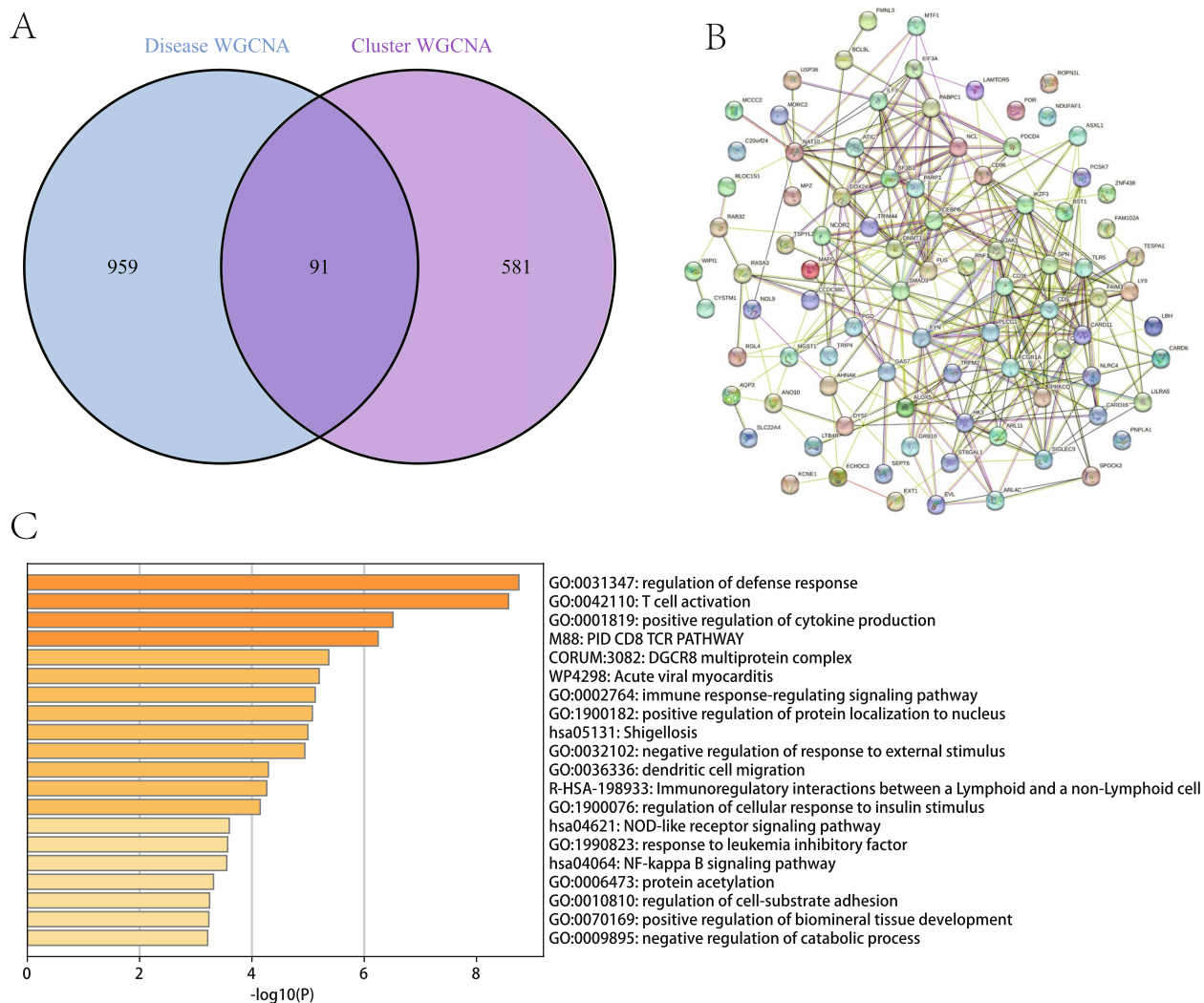
**Fig. 3** Two LMRG-related clusters in NMAP. **A** NMAP patients are divided into two subgroups. **B** PCA can clearly distinguish two subgroups. **C** The boxplot demonstrates the difference in expression of nine DE-LMRGs between the two clusters. **D, E** Comparison of 22 immune cell infiltration levels between the two clusters. **F** Differences in function enrichment between C1 and C2 by GSEA. **G** Differences in the KEGG pathway between C1 and C2 using GSEA. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$

(See figure on next page.)

**Fig. 4** Construction of the co-expressed network. **A** The scale-free fit index and the mean connectivity for various soft-thresholding powers in the first WGCNA. **B** Heatmap showing the correlation between different modules and NMAP traits. The color of the module indicates the correlation between the corresponding module and the trait. **C** Scatter plot of blue module genes. Genes with  $GS > 0.5$  and  $MM > 0.8$  were considered specific genes. **D** The scale-free fit index and the mean connectivity for various soft-thresholding powers in the second WGCNA. **E** Heatmap showing the correlation between different modules and clusters. The color of the module indicates the correlation between the corresponding module and the trait. **F** Scatter plot of turquoise module genes. Genes with  $GS > 0.5$  and  $MM > 0.8$  were considered specific genes



**Fig. 4** (See legend on previous page.)



**Fig. 5** Functional enrichment analysis of hub genes. **A** Taking the intersection of the specific genes from the two cluster analyses yielded 91 hub genes. **B** PPI analysis reveals correlations between hub genes. **C** Functional enrichment analysis of hub genes using the Metascape database

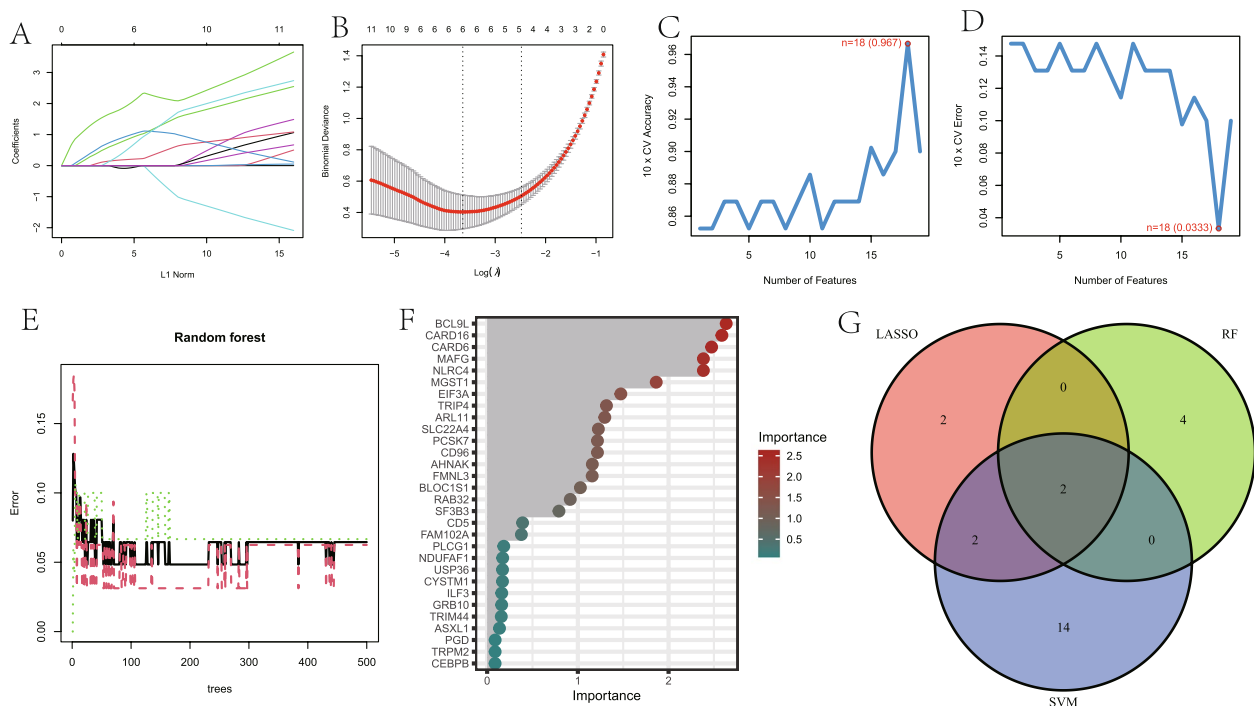
### Evaluation of the model

Figure 7A illustrates the expression levels of CARD16 and MGST1 in each sample. Boxplots revealed that NMAP samples had substantially greater levels of CARD16 and MGST1 expression than normal samples (Fig. 7B and C). In addition, we analyzed the correlation between CARD16 and MGST1 in normal samples of whole blood using the GTEx dataset in GEPIA. It was found that there was a significant positive correlation between CARD16 and MGST1 (Fig. 7D), suggesting an interaction between the two genes. According to the ROC curves, it can be found that CARD16 and MGST1 both have good predictive value (Fig. 7E). More importantly, the two-gene prediction model demonstrated

more favorable performance with an AUC value of 0.994 (Fig. 7F). To further confirm the model's accuracy, we compared it with the previously published AP diagnostic biomarkers. We found that the AUC of CARD16, MGST1, and our two-gene model was significantly higher than that of existing diagnostic biomarkers (Fig. 7G and H) [34, 35].

In addition, we also created a nomogram to estimate the potential hazards of NMAP individuals (Fig. 7I). The calibration curve demonstrated that there was little variation between the projected and reality risk for NMAP (Fig. 7J). In addition, the DCA curve demonstrated that the overall clinical benefit was greater than in the case in which either all or none of the tests were used for diagnosis (Fig. 7K).





**Fig. 6** Identifying the best model genes for NMAP by multiple machine learning methods. **A, B** The LASSO regression selected 6 genes based on minimum lambda values (**C, D**) Identify 18 genes based on SVM-RFE (**E, F**) RF ranked the importance of all genes to get 6 genes with scores for importance greater than 1.5. **G** The Venn diagram exhibiting the intersection of three machine learning models

### Biological functions of CARD16 and MGST1 in NMAP

To investigate the potential functional and molecular mechanisms of CARD16 and MGST1 in NMAP, GSEA was performed. It was found that they both were mainly upregulated in the T cell receptor complex, immunoglobulin complex circulating, as well as some immune-related routes, and downregulated in the negative regulation of chronic inflammatory response and oxidative phosphorylation (Fig. 8A-D). Then, the relationship between two model genes and immunochemicals was further explored. The results revealed that they largely linked negatively with resting NK cells, T cells CD4 memory resting, and T cells CD8 and positively with neutrophils, mast cells resting, and T cells gamma delta (Fig. 8E and F). These results suggest that CARD16 and MGST1 may promote NMAP progression by mediating these inflammatory and immune-related pathways and cells.

### The ceRNA networks, single-cell maps, and immunofluorescence analysis

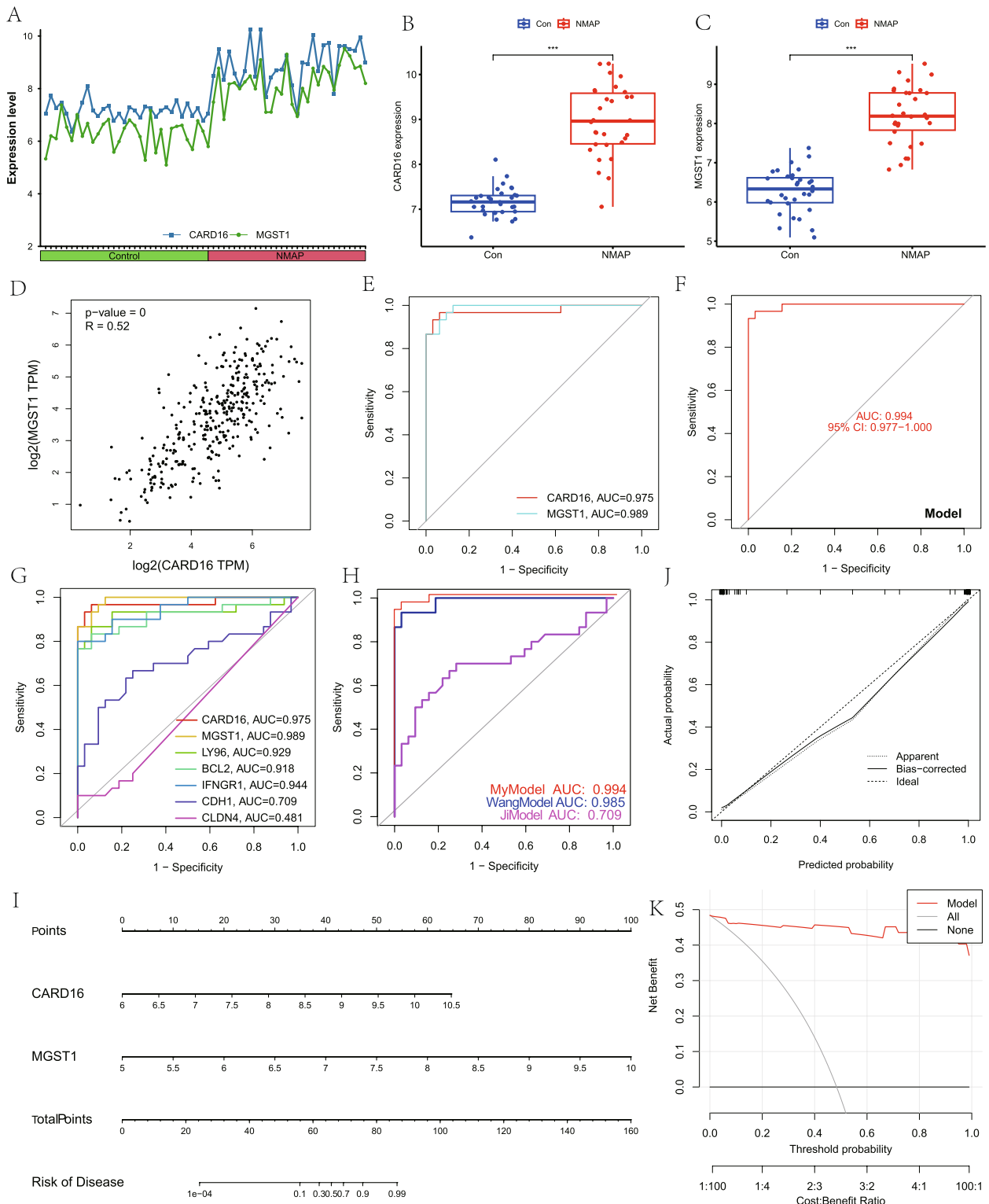
It is generally accepted that miRNAs bind to mRNAs and then cause gene silencing and decrease the expression of genes. However, by binding to miRNA response regions, its upstream molecule, lncRNA, can influence miRNA function and boost the expression of genes. The interactions between RNAs are known as ceRNA networks.

Using multiple public databases, CARD16- and MGST1-based ceRNA networks were constructed separately. In the end, we identified 5 objective miRNAs and 30 objective lncRNAs of CARD16 (Fig. 9A) and 4 objective miRNAs and 14 objective lncRNAs of MGST1 (Fig. 9B). The network revealed mechanisms that regulate model genes at the transcriptional level.

Additionally, we analyze the single-cell type for CARD16 and MGST1 using the HPA database. It was found that CARD16 was mainly distributed in mixed immune cells and macrophages, and MGST1 was mainly distributed in exocrine glandular cells (Fig. 9C and D). Then, we explored the cellular localization of CARD16 and MGST1. CARD16 was detected in mitochondria, and MGST1 was detected in mitochondria and endoplasmic reticulum (Fig. 9E and F).

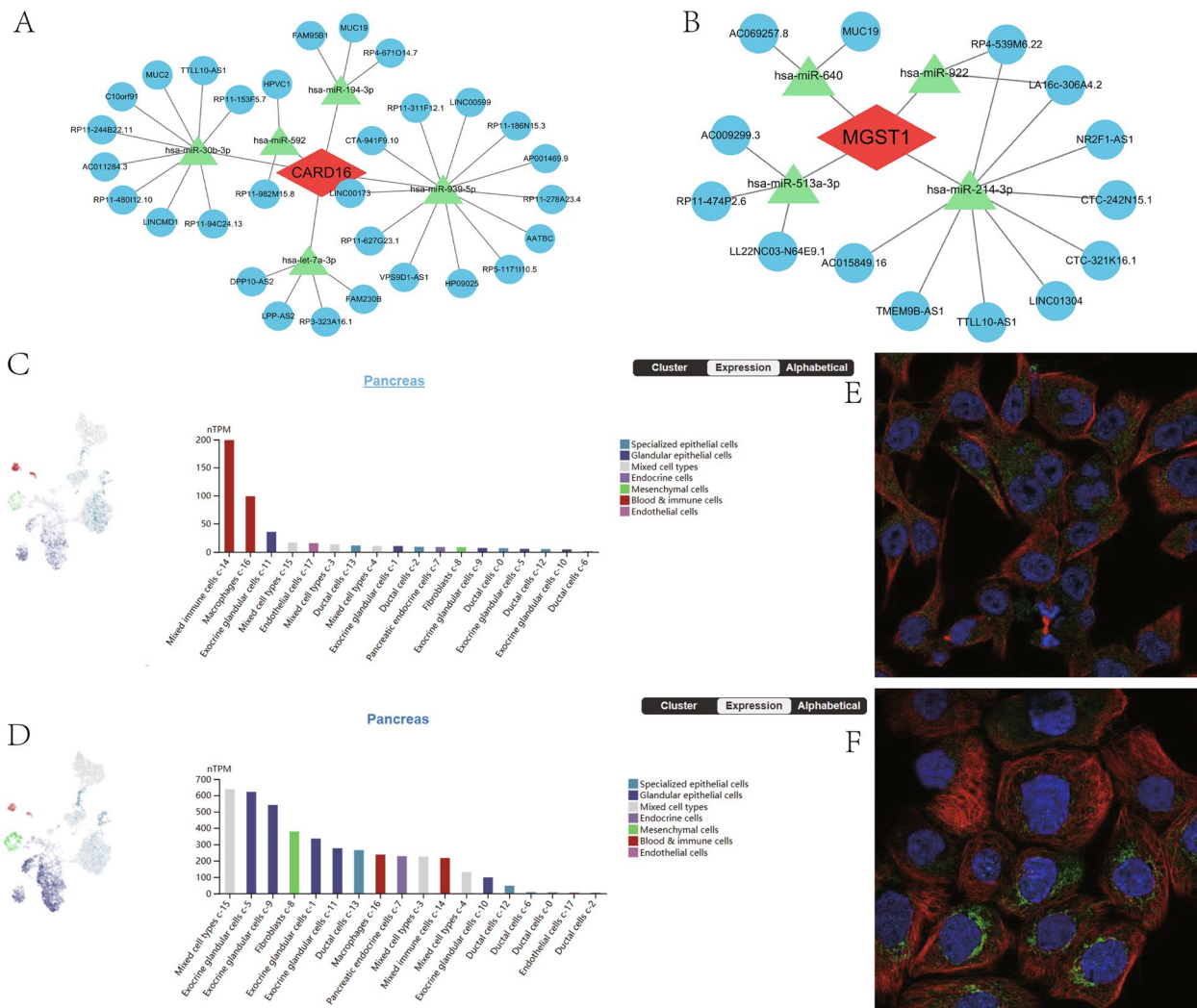
### Discussion

The degree of seriousness of AP, a condition characterized by inflammation, varies greatly, from mild manifestations with a low death rate to serious ones with a high death rate [36]. Therefore, early identification of NMAP is needed to identify individuals who are at risk for organ dysfunction or complications and which patients might benefit from earlier, more intensive treatment. Lipid metabolism has an important impact on AP and is closely



**Fig. 7** Diagnostic effect of the two-gene model on NMAP. **A** Expression of CARD16 and MSGT1 in different individuals. **B, C** Box plots show the expression difference in CARD16 and MSGT1 between NMAP and normal samples. **D** Correlation analysis of CARD16 and MSGT1. **E** ROC curves for the evaluation of the CARD16's and MSGT1's prediction ability for NMAP. **F** ROC curves for the evaluation of the two-gene model's prediction abilities for NMAP. **G, H** ROC curves of our biomarkers and existing AP diagnostic biomarkers. **I** The nomogram consisting of CARD16 and MSGT1 for forecasting NMAP. **J, K** Calibration curve and DCA curve to measure the prediction capability of the model. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$





**Fig. 9** The ceRNA networks, single-cell maps, and immunofluorescence of CARD16 and MSGT1. **A, B** The ceRNA networks of CARD16 and MSGT1. **C, D** The single-cell type atlases of CARD16 and MSGT1 in the pancreatic tissues. **E, F** The immunofluorescence of CARD16 and MSGT1

regulation of histone modification and methylation, while C1 was mainly involved in the regulation of endoplasmic reticulum tubular network organization, regulation of plasma cell differentiation, and protein export. Therefore, risk stratification based on LMRG is a potential method for determining prognosis and managing people with NMAP.

Although molecular typing is critical for the functional mining of LMRGs, it is difficult to accurately predict risk scores for specific individuals. To solve this problem, a model consisting of two genes with outstanding performance in predicting NMAP was developed using WGCNA and machine learning techniques. In addition, a nomogram for the diagnosis of NMAP was created using CARD16 and MSGT1. The ROC, DCA, and calibration

curves demonstrated the strong predictive value of the model. Meanwhile, we compared it with the previously published AP diagnostic biomarkers. Wang et al. identified three immunogenic cell death-related genes (LY96, BCL2, and IFNGR1) using the same dataset as ours, and Ji et al. identified two genes (CDH1 and CLDN4) that may serve as diagnostic biomarkers for AP using mouse samples. Compared with them, our two-gene model has a significant advantage in the discovery of NMAP.

CARD16, also known as Caspase Recruitment Domain Family Member 16, is a small molecule consisting of 97 amino acids and characterised by a solitary CARD motif. The motif exhibits a sequence identity of 92% with the prodomain of caspase-1 [42]. The CARD16 protein functions as a decoy to hinder

the binding of RIP2 to caspase-1 by means of CARD-CARD interaction in the context of the inflammatory response. Consequently, this leads to the suppression of caspase-1 activity and the initiation of NF- $\kappa$ B activation [43]. Furthermore, it has been found that CARD16 plays a role in the stemness maintenance of glioma stem cells and the inhibition of cAMP-induced differentiation [44]. Microsomal glutathione S-transferase 1 (MGST1) is a member of the MAPEG family of membrane proteins. It is a homotrimeric protein with three glutathione binding sites that is found in large amounts in the outer mitochondrial membranes and the endoplasmic reticulum. MGST1 was shown to be an important mediator of inflammation with glutathione S-transferase and peroxidase activities [45]. Emerging data suggest mitochondrial MGST1 may contribute to apoptotic cell death under chemical or oxidative stress [46]. Furthermore, MGST1 exerts inhibitory effects on the process of ferroptosis in cancer cells by interacting with ALOX5, or its lipid peroxidase activities [47]. MGST1 is considered a tumor marker, and its overexpression has been linked to a poor prognosis in several malignancies [48–50], and proposed as a potential therapeutic target for pancreatic cancer [51]. Additionally, it could serve as a biomarker to predict treatment efficacy and assess cancer risk [52, 53]. MGST1 polymorphisms may also influence an individual's susceptibility to specific cancers [54]. However, no studies have reported the roles of CARD16 and MGST1 in AP, and their potential mechanisms in the development of AP still need to be further explored.

For a deeper investigation of the potential functions and molecular mechanisms of CARD16 and MGST1 of NMAP, GSVA was performed. CARD16 and MGST1 were found to be mainly upregulated in the T cell receptor complex, immunoglobulin complex circulating, as well as some immune-related routes, and downregulated in the negative regulation of chronic inflammatory response and oxidative phosphorylation. Further immune cell correlation analysis identified that the model genes were closely linked to neutrophils and mast cells. Numerous studies have shown a close correlation between neutrophils and the seriousness of AP and that they are crucial to the system-wide inflammatory reaction and other damage to organs brought on by AP. An important factor in determining pancreatic damage and inflammation is increased and persistent neutrophil activation [40]. Mast cells play a significant role in activating local and systemic inflammatory responses in the early stages of the disease [55] and are intimately related to the severity and

unfavorable outcomes of AP [41]. These results suggest that CARD16 and MGST1 may promote the activation and recruitment of neutrophils and mast cells through the regulation of the above pathways and functions, thereby promoting pancreatic injury and systemic inflammation.

### Study strengths and limitations

As far as we know, this is the first bioinformatics research to thoroughly examine LMRGs in AP. Given the commonalities of high mortality and concomitant organ dysfunction in MSAP and SAP, both MSAP and SAP samples were included in the present study to enable early diagnosis of NMAP. We innovatively used LMRGs as the basis for grouping and further identified model genes by WGCNA and multiple machine learning methods. However, it is important to recognize that our model validation is still not robust enough, and further basic experimental and clinical studies are still needed. Specifically, blood samples from NMAP and normal samples will be collected for metabolomics analysis to verify the presence of abnormal lipid metabolism in NMAP. The dysregulation of model gene expression in NMAP will be verified by PCR and other methods, and the changes in the levels of AP-related markers will be observed after the knockdown of target genes in AP cell models. Meanwhile, the expression of the target gene in human mast cell lines and monocyte cell lines needed further investigation to clarify its role in immune cells. In addition, the sequencing of clinical samples is needed to validate the predictive ability of our model in a clinical cohort.

### Conclusion

This study presents a novel approach to categorizing NMAP into different clusters based on LMRGs and developing a reliable two-gene diagnostic biomarker for NMAP patients. Importantly, this study may provide a theoretical basis for future studies of AP lipid metabolism.

### Acknowledgements

None.

### Authors' contributions

The study's design, data analysis, and paper preparation benefited from the work of Jifeng Liu, Lei Zhong, and Yunshu Zhang. Jingyuan Ma and Tong Xie participated in the manuscript writing. Xu Chen, Biao Zhang, and Dong Shang developed the research conceptualization and revised the manuscript.

### Funding

This work was financially supported by grants from the National Natural Science Foundation of China (No. 82374248).

**Availability of data and materials**

The datasets analyzed in this work may be found in the GEO databases. Additionally, any raw data and analytic technologies can be requested by directly contacting the author if the request is reasonable.

**Declarations****Ethics approval and to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare no competing interests.

**Author details**

<sup>1</sup>Department of General Surgery, The First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China.

Received: 27 July 2023 Accepted: 14 November 2023

Published online: 02 January 2024

**References**

- Banks PA, Bollen TL, Dervenis C, Gooszen HG, Johnson CD, Sarr MG, Tsio-  
tos GG, Vege SS, Acute Pancreatitis Classification Working G. Classifica-  
tion of acute pancreatitis–2012: revision of the Atlanta classification and  
definitions by international consensus. *Gut*. 2013;62:102–11.
- Yang DJ, Lu HM, Liu Y, Li M, Hu WM, Zhou ZG. Development and valida-  
tion of a prediction model for moderately severe and severe acute  
pancreatitis in pregnancy. *World J Gastroenterol*. 2022;28:1588–600.
- Glaubitiz J, Asgarbeik S, Lange R, Mazloum H, Elsheikh H, Weiss FU, Sandler  
M. Immune response mechanisms in acute and chronic pancreatitis:  
strategies for therapeutic intervention. *Front Immunol*. 2023;14:1279539.
- Yang AL, McNabb-Baltar J. Hypertriglyceridemia and acute pancreatitis.  
*Pancreatol*. 2020;20:795–800.
- Yu B, He W-H, Lu N. Risk factors for acute kidney injury in acute pancrea-  
titis: a 7-year retrospective analysis of patients in a large tertiary hospital.  
*Pancreas*. 2020;49:1057–62.
- Tariq H, Gaduputi V, Peralta R, Abbas N, Nayudu SK, Thet P, Zaw T, Hui  
S, Chilimuri S. Serum triglyceride level: a predictor of complications  
and outcomes in acute pancreatitis? *Can J Gastroenterol Hepatol*.  
2016;2016:8198047.
- Yang X, He J, Ma S, Wang T, Zhu Q, Cao F, Li Y, Yang C, Chen C, Lu G, et al.  
The role of comorbid hypertriglyceridemia and abdominal obesity in  
the severity of acute pancreatitis: a retrospective study. *Lipids Health Dis*.  
2021;20:171.
- Adiamah A, Psaltis E, Crook M, Lobo DN. A systematic review of the epide-  
miology, pathophysiology and current management of hyperlipidaemic  
pancreatitis. *Clin Nutr*. 2018;37:1810–22.
- Pascual I, Sanahuja A, García N, Vázquez P, Moreno O, Tosca J, Peña A,  
Garayoa A, Lluch P, Mora F. Association of elevated serum triglyceride  
levels with a more severe course of acute pancreatitis: cohort analysis of  
1457 patients. *Pancreatol*. 2019;19:623–9.
- Nawaz H, Koutroumpakis E, Easler J, Slivka A, Whitcomb DC, Singh VP,  
Yadav D, Papachristou GI. Elevated serum triglycerides are independently  
associated with persistent organ failure in acute pancreatitis. *Am J Gas-  
troenterol*. 2015;110:1497–503.
- Schepers NJ, Bakker OJ, Besselink MG, Ahmed Ali U, Bollen TL, Gooszen  
HG, van Santvoort HC, Bruno MJ. Impact of characteristics of organ  
failure and infected necrosis on mortality in necrotising pancreatitis. *Gut*.  
2019;68:1044–51.
- Nesvaderani M, Dhillon BK, Chew T, Tang B, Baghela A, Hancock RE, Eslick  
GD, Cox M. Gene expression profiling: identification of novel pathways  
and potential biomarkers in severe acute pancreatitis. *J Am Coll Surg*.  
2022;234:803–15.
- Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, Nativ N,  
Bahir I, Doniger T, Krug H, et al. GeneCards Version 3: the human gene  
integrator. *Database (Oxford)*. 2010;2010:baq020.
- Liu J, Li Y, Ma J, Wan X, Zhao M, Zhang Y, Shang D. Identification and  
immunological characterization of lipid metabolism-related molecular  
clusters in nonalcoholic fatty liver disease. *Lipids Health Dis*. 2023;22:124.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma pow-  
ers differential expression analyses for RNA-sequencing and microarray  
studies. *Nucleic Acids Res*. 2015;43:e47.
- Miao Y, Liu J, Liu X, Yuan Q, Li H, Zhang Y, Zhan Y, Feng X. Machine learn-  
ing identification of cuproptosis and necroptosis-associated molecular  
subtypes to aid in prognosis assessment and immunotherapy response  
prediction in low-grade glioma. *Front Genet*. 2022;13:951239.
- Ren J, Yuan Q, Liu J, Zhong L, Li H, Wu G, Chen F, Tang Q. Identifying  
the role of transient receptor potential channels (TRPs) in kidney renal  
clear cell carcinoma and their potential therapeutic significances using  
genomic and transcriptome analyses. *BMC Med Genomics*. 2022;15:156.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation  
network analysis. *BMC Bioinformatics*. 2008;9:559.
- Zhang S, Niu Q, Tong L, Liu S, Wang P, Xu H, Li B, Zhang H. Identification  
of the susceptible genes and mechanism underlying the comorbid  
presence of coronary artery disease and rheumatoid arthritis: a network  
modularization analysis. *BMC Genomics*. 2023;24:411.
- Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R,  
Gable AL, Fang T, Doncheva NT, Pyysalo S, et al. The STRING database  
in 2023: protein-protein association networks and functional enrich-  
ment analyses for any sequenced genome of interest. *Nucleic Acids Res*.  
2023;51:D638–46.
- Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O,  
Benner C, Chanda SK. Metascape provides a biologist-oriented resource  
for the analysis of systems-level datasets. *Nat Commun*. 2019;10:1523.
- Bai J, Huang JH, Price CPE, Schauer JM, Suh LA, Harmon R, Conley DB,  
Welch KC, Kern RC, Shintani-Smith S, et al. Prognostic factors for polyp  
recurrence in chronic rhinosinusitis with nasal polyps. *J Allergy Clin  
Immunol*. 2022;150:352–361 e357.
- Blanco JL, Porto-Pazos AB, Pazos A, Fernandez-Lozano C. Prediction of  
high anti-angiogenic activity peptides in silico using a generalized linear  
model and feature selection. *Sci Rep*. 2018;8:15688.
- Rigatti SJ. Random forest. *J Insur Med*. 2017;47:31–9.
- Yang L, Pan X, Zhang Y, Zhao D, Wang L, Yuan G, Zhou C, Li T, Li W. Bioin-  
formatics analysis to screen for genes related to myocardial infarction.  
*Front Genet*. 2022;13:990888.
- Qin Q, Gu Z, Li F, Pan Y, Zhang T, Fang Y, Zhang L. A Diagnostic model for  
Alzheimer's disease based on blood levels of autophagy-related genes.  
*Front Aging Neurosci*. 2022;14:881890.
- Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer  
and normal gene expression profiling and interactive analyses. *Nucleic  
Acids Res*. 2017;45:W98–102.
- Liu J, Yuan Q, Ren J, Li Y, Zhang Y, Shang D. Single-cell sequencing and  
bulk RNA sequencing reveal a cell differentiation-related multigene panel  
to predict the prognosis and immunotherapy response of hepatocellular  
carcinoma. *Chin Med J (Engl)*. 2023;136:485–7.
- Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vick-  
ers AJ, Roobol MJ, Steyerberg EW. Reporting and interpreting decision  
curve analysis: a guide for investigators. *Eur Urol*. 2018;74:796–804.
- Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for  
microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.
- Digre A, Lindskog C. The human protein atlas-Integrated omics for single  
cell mapping of the human proteome. *Protein Sci*. 2023;32:e4562.
- Li Y, Li R, Cheng D, Fu X, Fu L, Peng S. The potential of CircRNA1002 as a  
biomarker in hepatitis B virus-related hepatocellular carcinoma. *PeerJ*.  
2022;10:e13640.
- Furio-Tari P, Tarazona S, Gabaldon T, Enright AJ, Conesa A. spongeScan:  
A web for detecting microRNA binding elements in lncRNA sequences.  
*Nucleic Acids Res*. 2016;44:W176–180.
- Wang Z, Liu J, Wang Y, Guo H, Li F, Cao Y, Zhao L, Chen H. Identification  
of key biomarkers associated with immunogenic cell death and their  
regulatory mechanisms in severe acute pancreatitis based on WGCNA  
and machine learning. *Int J Mol Sci*. 2023;24:3033.

35. Ji R, Chen Y, Chen W, Wang Y, Gong F, Huang S, Xie R, Zhong M, Liu Z, Yang Z, et al. Identification of significant genes and pathways in acute pancreatitis via bioinformatical analysis. *Dig Dis Sci*. 2021;66:3045–53.
36. Mederos MA, Reber HA, Gargis MD. Acute pancreatitis: a review. *JAMA*. 2021;325:382–90.
37. Peng C, Li Z, Yu X. The role of pancreatic infiltrating innate immune cells in acute pancreatitis. *Int J Med Sci*. 2021;18:534–45.
38. Stojanovic B, Jovanovic IP, Stojanovic MD, Jovanovic M, Vekic B, Milosevic B, Cvetkovic A, Spasic M, Stojanovic BS. The emerging roles of the adaptive immune response in acute pancreatitis. *Cells*. 2023;12:1495.
39. Zhou X, Jin S, Pan J, Lin Q, Yang S, Ambe PC, Basharat Z, Zimmer V, Wang W, Hong W. Damage associated molecular patterns and neutrophil extracellular traps in acute pancreatitis. *Front Cell Infect Microbiol*. 2022;12:927193.
40. Wan J, Ren Y, Yang X, Li X, Xia L, Lu N. The role of neutrophils and neutrophil extracellular traps in acute pancreatitis. *Front Cell Dev Biol*. 2020;8:565758.
41. Braganza JM. Mast cell: pivotal player in lethal acute pancreatitis. *QJM*. 2000;93:469–76.
42. Lee SH, Stehlik C, Reed JC. Cop, a caspase recruitment domain-containing protein and inhibitor of caspase-1 activation processing. *J Biol Chem*. 2001;276:34495–500.
43. Bouchier-Hayes L, Martin SJ. CARD games in apoptosis and immunity. *EMBO Rep*. 2002;3:616–21.
44. Chen Z, Zhong Y, Chen J, Sun S, Liu W, Han Y, Liu X, Guo C, Li D, Hu W, et al. Disruption of  $\beta$ -catenin-mediated negative feedback reinforces cAMP-induced neuronal differentiation in glioma stem cells. *Cell Death Dis*. 2022;13:493.
45. Schaffert CS. Role of MGST1 in reactive intermediate-induced injury. *World J Gastroenterol*. 2011;17:2552–7.
46. Zhang J, Ye Z-W, Morgenstern R, Townsend DM, Tew KD. Microsomal glutathione transferase 1 in cancer and the regulation of ferroptosis. *Adv Cancer Res*. 2023;160:107–32.
47. Kuang F, Liu J, Xie Y, Tang D, Kang R. MGST1 is a redox-sensitive repressor of ferroptosis in pancreatic cancer cells. *Cell Chem Biol*. 2021;28:765–775. e765.
48. Li Y, Xu X, Wang X, Zhang C, Hu A, Li Y. MGST1 expression is associated with poor prognosis, enhancing the Wnt/ $\beta$ -Catenin pathway via regulating AKT and inhibiting ferroptosis in gastric cancer. *ACS Omega*. 2023;8:23683–94.
49. Yan J, Ye G, Shao Y. High expression of the ferroptosis-associated MGST1 gene in relation to poor outcome and maladjusted immune cell infiltration in uterine corpus endometrial carcinoma. *J Clin Lab Anal*. 2022;36:e24317.
50. Yu L, Gao Z, Li Z, Liu P, Gao Y, Liang G. Identification of ferroptosis-related molecular subtypes and a methylation-related ferroptosis gene prognostic signature in cervical squamous cell carcinoma. *J Cancer Res Clin Oncol*. 2023;149:14673–89.
51. Kuang F, Liu J, Xie Y, Tang D, Kang R. MGST1 is a redox-sensitive repressor of ferroptosis in pancreatic cancer cells. *Cell Chem Biol*. 2021;28:765–775. e765.
52. Rahmat JN, Esuvaranathan K, Mahendran R. Bacillus Calmette-Guérin induces rapid gene expression changes in human bladder cancer cell lines that may modulate its survival. *Oncol Lett*. 2018;15:9231–41.
53. Messinger JE, Dai J, Stanland LJ, Price AM, Luftig MA. Identification of host biomarkers of Epstein-Barr virus latency IIb and latency III. *mBio*. 2019;10:e01006-19.
54. Zhang H, Liao L-H, Liu S-M, Lau K-W, Lai AK-C, Zhang J-H, Wang Q, Chen X-Q, Wei W, Liu H, et al. Microsomal glutathione S-transferase gene polymorphisms and colorectal cancer risk in a Han Chinese population. *Int J Colorectal Dis*. 2007;22:1185–94.
55. Lopez-Font I, Gea-Sorli S, de Madaria E, Gutierrez LM, Perez-Mateo M, Closa D. Pancreatic and pulmonary mast cells activation during experimental acute pancreatitis. *World J Gastroenterol*. 2010;16:3411–7.

## Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

