Amino Acids (2026) 58:18
https://doi.org/10.1007/s00726-026-03507-3

ORIGINAL ARTICLE

Amino acid metabolism-related model for prognosis and immunity in gastric cancer

Huan Zhang[1] · Wei Ye[2] · Lingzhi Zeng[3] · Lu Wang[3] · Ling Gui[4]

Received: 11 June 2025 / Accepted: 13 February 2026

Abstract

Amino acid metabolic (AAM) reprogramming is a key characteristic of gastric cancer (GC) cells metabolic remodeling, which regulates cell growth, survival, immune cell activation and function to affect tumor immune escape. This study aims to systematically investigate AAM reprogramming in gastric cancer (GC) and construct prognostic model, and validate gene signatures for predictive value and clinical decision-making. This study leveraged data from TCGA and GEO to construct a prognostic model related to AAM and assess its clinical relevance in GC. We identified differentially expressed genes and conducted GO, GSEA, and GSVA enrichment analyses, along with constructing PPI networks and interaction networks of mRNA-miRNA, mRNA-TF, and mRNA-RBP. Additionally, immune infiltration analysis was performed and the relationships between eight hub-type amino acid metabolism-related genes (AAMRGs) and immune cells was investigated using scRNA-seq datasets. Lastly, we validated the elevated expression of these eight genes in GC cells through PCR. The study constructed a prognostic model for GC based on AAMRGs, identifying 16 key genes: ACLY, ADH4, COL1A1, F2, GADL1, GAMT, HBB, KYNU, MRI1, MTHFR, NR1D1, PDK4, SLC1A7, SLC25A15, SLC52A3, and SYCE2. Statistical analysis showed that 14 of these genes showed significant differential expression between tumor and normal tissues. Furthermore, the model demonstrated strong correlations with OS outcomes. Immune infiltration analysis indicated that various immune cell types were significantly associated with the expression of 8 hub genes, highlighting their potential role in the tumor microenvironment and immune response modulation. Furthermore, elevated expression of these genes in GC cells was validated through PCR, highlighting their relevance as potential biomarkers and therapeutic targets. Our AAMRGs prognostic model reveals AAMRGs as independent prognostic factors for GC, highlighting their association with prognosis and immune cell infiltration. These findings provide important insights for improving survival outcomes and advancing immunotherapy strategies in GC.

Keywords Prognostic model · Amino acid metabolic · Gastric cancer · Genes

Huan Zhang and Wei Ye are the co-first authors, contributed equally to this work.

Handling editor: P. Meffre.

✉ Lu Wang
xuezhixi1988@163.com

✉ Ling Gui
15307924888@163.com

[1] Department of Gastroenterology and Hepatology, JiuJiang City Key Laboratory of Cell Therapy, JiuJiang NO.1 People's Hospital, Jiujiang 332000, JiangXi Province, China

[2] Department of Radiation Oncology, JiuJiang City Key Laboratory of Cell Therapy, JiuJiang NO.1 People's Hospital, Jiujiang 332000, JiangXi Province, China

[3] Department of Oncology, JiuJiang City Key Laboratory of Cell Therapy, JiuJiang NO.1 People's Hospital, No.77, East Bali Lake Road, Bali Lake New District, Jiujiang 332000, JiangXi Province, China

[4] Department of Information, JiuJiang City Key Laboratory of Cell Therapy, JiuJiang NO.1 People's Hospital, No.77, East Bali Lake Road, Bali Lake New District, Jiujiang 332000, JiangXi Province, China

Published online: 05 March 2026

Introduction

Gastric cancer (GC) is a common malignant tumor and a leading cause of cancer-associated death worldwide (Siegel et al. 2024; Bray et al. 2024). Most GC patients are diagnosed at a late stage because of relapse, metastasis, or poor prognosis (Qiu et al. 2021). Despite the high heterogeneity of gastric cancer (GC) revealed by genomic studies (Qiu et al. 2021), the interplay between metabolic reprogramming and immune evasion in the tumor microenvironment (TME) remains a central challenge in current research (Wang et al. 2023a). The temporal and spatial heterogeneity, invasiveness, and prognosis of GC are closely linked to the TME (Zhu et al. 2025). In recent years, amino acid metabolism (AAM) reprogramming has emerged as a key driver of tumor progression. Cancer cells acquire the energy and macromolecules necessary for proliferation through metabolic pathways such as glutamine and tryptophan (Jiang et al. 2025). Additionally, metabolic products, such as S-adenosylmethionine, contribute to an immunosuppressive TME by regulating T cell exhaustion and macrophage polarization (Wang et al. 2023a). However, the spatiotemporal regulatory network of AAM-related genes (AAMRGs) within the GC TME has not been systematically elucidated. Furthermore, there is a lack of integrated multi-omics-based metabolic prognostic models to guide clinical decision-making.. This study is to investigate the features of AAMRGs in GC and construct the prognostic model related to amino acid metabolism.

Through comprehensive analysis of integrated datasets (TCGA-STAD, GSE27342, and GSE54129), this study systematically investigated the functional significance of amino acid metabolism-related genes (AAMRGs) in gastric cancer subtype classification. Notably, we established a novel prognostic model incorporating 16 critical genes: ACLY, ADH4, COL1A1, F2, GADL1, GAMT, HBB, KYNU, MRI1, MTHFR, NR1D1, PDK4, SLC1A7, SLC25A15, SLC52A3, and SYCE2. Statistical analysis revealed that 14 of these genes exhibited significant differential expression between tumor and normal tissues, and the model showed strong correlations with OS outcomes. We also studied molecular signatures and immune cell infiltration to find AAM clusters, confirming that risk profiles based on eight hub genes serve as independent prognostic factors for GC. Additionally, using the GSE184198 dataset, this study systematically characterized the cellular heterogeneity of gastric cancer and developed an AAM (Amino Acid Metabolism) scoring system capable of accurately predicting clinical outcomes and therapeutic responses. These findings establish a quantifiable molecular signature system to guide personalized treatment strategies (e.g., IDO inhibitor combined with immunotherapy), while providing critical evidence to advance precision medicine and future research in gastric cancer management.

Materials and methods

Data download

The TCGAbiolinks package (Colaprico et al. 2016) was employed to download the STAD dataset (TCGA-STAD), which served as our test set, comprising count sequencing data from 353 GC tissues (Tumor) and 30 adjacent normal tissues (Normal). TCGA-STAD count data were normalized to FPKM format. Clinical data associated with these tissues were retrieved from UCSC Xena (http://genome.ucsc.edu) (Goldman et al. 2020).

Additionally, we profiled the GSE27342 (Barrett et al. 2007) and GSE54129 (Cui et al. 2011) datasets of STAD patients with GEOquery R package. GSE27342 dataset contains 160 GC samples, with 80 from gastric adenocarcinoma tissues (Tumor) and 80 adjacent normal tissues (Normal). This dataset was from GPL5175. GSE54129 dataset includes 132 gastric tissue samples, comprising 111 adenocarcinoma samples (Tumor) and 21 normal gastric mucosa samples (Normal) (GPL570).

To annotate probe names, we utilized microarray GPL platform files. In GSE27342, both the 80 gastric adenocarcinoma samples and their 80 adjacent normal counterparts were included for subsequent analysis. Similarly, GSE54129 included 111 adenocarcinoma samples and 21 normal samples, along with their clinical information. Both GSE27342 and GSE54129 datasets served as validation sets for further analysis(Supplementary Table 1).

We also accessed GeneCards database (https://www.genecards.org/) (Stelzer et al. 2016) with "Amino acid metabolism" as search term, filtering for AAMRGs categorized as "Protein Coding" and with a relevance score≥2.00. The screening resulted in 237 AAMRGs (Table S1).

Construct prognostic models related to amino acid metabolism

To identify AAMRGs in TCGA-STAD dataset, we integrated AAMRG expression data with OS clinical information. Univariate Cox regression analysis was performed, followed by LASSO regression with glmnet (Engebretsen and Bohlin 2019). The model was developed through ten-fold cross-validation, employing 1000 iterations to optimize the penalization term (lambda) to mitigate overfitting and enhance model generalizability (Cai et al. 2020). Risk scores for each AAMRG were determined according to expression

levels in GC samples, categorizing scores into high- or low-group. The formula used was:

$$\text{Risk } Scores = \sum_i Coefficient\,(hub\,gene_i) * mRNA\,Expression(hub\,gene_i)$$

A risk factor map was generated, comprising risk grouping, survival outcomes (visualized through a dot plot), and a heat map depicting AAMRG expression levels. Additionally, group comparisons of TCGA-STAD data were performed, similar analyses were conducted using the GSE27342 and GSE54129 datasets for validation.

Prognostic analysis and clinical correlation analysis of AAMRGs

To validate this model, we utilized TCGA-STAD clinical dataset, focusing on OS. We investigated the relationship between risk scores from AAMRGs model, individual gene expression values, and patient survival outcomes, employing Kaplan–Meier (KM) curves for prognostic assessment. KM analysis estimates survival probabilities, facilitating the exploration of correlations between survival time and influencing factors.

Subsequently, ROC curves was made (Mandrekar 2010) based on risk scores from the AAMRGs model and relevant clinical data.. survivalROC was used for these curves, AUCs were adopted to estimate the diagnostic ability of risk scores in predicting survival among GC patients.

Additionally, we compared the Risk Scores of AAMRGs prognostic model between different groups (grouping: Different clinical variables of GC patients in TCGA dataset,, The distribution differences of DSS and PFI (Progression Free Interval)) were shown by drawing a stacked bar chart.

Differentially expressed genes (DEGs) in TCGA-STAD

To assess DEGs associated with risk in TCGA-STAD dataset, DESeq2 was applied for analyzing gene expression profiles (Love et al. 2014). DEGs were identified if $|\log FC| > 0$ and p.adj < 0.05. Volcano maps and heatmap were created by ggplot2 and pheatmap.

Functional enrichment analysis (GO)

GO[16]is widely adopted for functional enrichment analysis. ClusterProfiler was used. *P*. Adj < 0.05 and FDR (*q*. value) < 0.05 were defined as significant, Benjamini-Hochberg (BH) was used for correction.

GSEA enrichment analysis

GSEA (Subramanian et al. 2005) enrichment analysis was performed in TCGA-STAD dataset between high and low risk groups by clusterProfiler with following parameters: seed: 2022, number of calculations: 10,000, $500 \leq$ number of genes: ≥ 10, and p-value correction method was BH. Pathways (3050) gene set in MSigDB database (Liberzon, et al. 2011) was used as reference. *p*. Adj < 0.05 and FDR (*q*. value) < 0.05 was defined as significant. To enhance analytical clarity, GSEA and GSVA were performed for complementary purposes rather than interchangeable use. GSEA was used to evaluate phenotype-driven, directional enrichment between high- and low-risk groups, whereas GSVA quantified sample-wise pathway activity in an unsupervised manner. GSVA was used as a cross-validation approach to confirm whether the major biological pathways identified by GSEA remained consistently enriched at the single-sample level.

GSVA enrichment analysis

GSVA (Hanzelmann et al. 2013) was conducted to evaluate pathway enrichment in high or low-risk groups delineated by this model using TCGA-STAD dataset. Expression matrix was transformed to compare gene expression across samples. We accessed the MSigDB database to utilize c2.cp.all.v2022.1.hs.symbols.GMT gene set for this analysis. Pathway results were derived from high-risk group within TCGA-STAD dataset, focusing on those with p.adj < 0.05. Pathways were ranked by logFC in ascending and descending order, identifying top 10 positively or negatively correlated pathways. These results were subsequently visualized through a heatmap, providing insights into the differential enrichment of pathways between two risk groups in STAD patients.

Protein–protein interaction (PPI) network

PPI networks contain proteins that interact to regulate various biological processes. PPI networks related to AAMRGs was created with STRING (minimum interaction score: low confidence 0.9) (Szklarczyk et al. 2019; Shannon et al. 2003). We employed Multiple algorithms, such as MCC(Boughorbel et al. 2017), DMNC, MNC, EPC and Degree, were used to assess AAMRG scores and identify hub genes among the top 10. Additionally, the GeneMANIA database (http://genemania.org) (Gao et al. 2022) was utilized for gene function analysis and prioritization, while Cytoscape facilitated visualization of the PPI network.

Construct the interaction network of mRNA-miRNA, mRNA-TF and mRNA-RBP

miRDB database (Chen et al. 2020) was adopted to evaluate how miRNAs interacts with hub genes, and those with Target Score > 90 were used for visualization by Cytoscape.

ChIPBase (3.0) (Zhou et al. 2017) was used to predict how transcription factors (TFS) regulate genes And cytoscape was used for visualization.

In addition, ENCORI (3.0) (Wu et al. 2021) provided insights into interactions involving microRNAs, lncRNAs, and RNA-binding proteins (RBPs), utilizing data from CLIP-seq and other sequencing techniques. ENCORI was used to predict RBPs interacting with hub genes, filtering for mRNA-RBP interactions with clip Exp Num > 5. Results weres visualized by Cytoscape, allowing for comprehensive analysis of the regulatory interactions involving mRNAs, miRNAs, TFs, and RBPs.

Immune infiltration analysis

ssGSEA algorithm was applied to evaluate tumor microenvironment, utilizing 28 gene sets to identify various tumor-infiltrating immune cell types (Bindea et al. 2013). Enrichment scores obtained from GSVA represented immune cells' infiltration levels. Comparisons of immuno- of different risk groups were illustrated via group comparison plots. Significant correlations between immune cells across groups were performed by Spearman's correlation and visualized by ggplot2.

We also utilized CIBERSORT (Newman et al. 2015), which employs linear support vector regression to deconvolute transcriptome data for estimating immune cell composition. By applying the LM22 feature gene matrix, immune cels withl enrichment scores ≥ zero were filtered, resulting in a comprehensive immune cell abundance matrix. Correlations between immune cells and hub genes across groups were also evaluated and visualized using ggplot2 and pheatmap for clarity.

Construction of clinically relevant prognostic model and prognostic correlation analysis

To determine prognostic values of hub genes in TCGA-STAD dataset, we first evaluated expression levels and related clinical information of hub genes in STAD patients. Variables ($p < 0.1$) in univariate analysis were carried to multivariate Cox regression analysis to construct a clinically relevant prognostic model. Then, a nomogram was created (Park 2018). After setting each variable a certain scale, total score was obtained for predicting probability.

Calibration curve was then used for measuring nomogram's accuracy and resolution. TCalibration plot (Perkins et al. 2019) DCA (Van Calster et al. 2018) chart was then used to assess accuracy and discrimination of clinically relevant prognostic models. ggDCA was adopted to draw DCA plots to Tataranni et al. (2019) measure the effect of clinically relevant prognostic models.

Single cell data download

We downloaded the scRNA-seq dataset GSE184198 (Zhao et al. 2022), which profiles GC (GC) patients using 10X Genomics technology on an Illumina NovaSeq 6000 (GPL24676). This dataset includes one GC tissue sample (group: GC) and one normal gastric tissue sample (group: NT) from the same patient. All samples were utilized. Single-cell data processing was carried out with Seurat, where we created Seurat objects. The PercentageFeatureSet function calculated the percentage of mitochondrial genes per cell, filtering out those with > 5% mitochondrial content, as this may indicate apoptosis or lysis. We also excluded low-quality cells with fewer than 500 features and those with over 20,000 unique molecular identifiers (UMIs).

Post-quality control, linear dimension reduction was conducted, computing Principal Components (PCs) from the most variable genes. We applied Seurat's "FindNeighbors" and "FindClusters" functions to screen optimal cell clusters, and UMAP was used for two-dimensional visualization of cell clusters.

Quality control of single cell datasets

Seurat v4.0 is widely used for single-cell data analysis encompassing quality control, cell screening, type identification, gene selection, differential expression analysis, and data visualization.

GSE184198 was used to create Seurat objects. The dataset was filtered and normalized with "LogNormalize". We identified 2000 variable features with "FindVariableFeatures" of "vst" method, followed by scaling the data with "ScaleData" to account for sequencing depth. PCA was conducted to assess significant Principal Components (PCs), with Elbowplot for visualization.

Forty PCs were chosen for UMAP analysis. The "FindNeighbors" function utilized default parameters and the selected PCs to create k-nearest neighbors according to Euclidean distance. Subsequently, "FindClusters" function, along with the "clustree" function at a resolution of 0.6, was used to cluster cells, with "RunUMAP" facilitating dimensionality reduction for visualization.

Cell type annotation and DEGs

Harmony package was applied for data integration. Single cell population was named according to published literature and the cell marker gene information in CellMarker 2.0. (Hu et al. 2023). We identified cell types by annotating GSE184198 with cell marker information in the CellMarker 2.0 website. Subsequently, we used the "DotPlot" and "FeaturePlot" functions to evaluate the correlation between expression levels of marker genes and different cell types. For the annotated cell groups, "FindAllMarkers" was applied to identify DEGs between all cell groups with log-2FoldChange > 2 and p. Adj < 0.01(more genes, The genes screened by improving the screening criteria were identified as single cell group scDEGs for further study.

AUCell scored cell populations

AUCell uses the "Area Under Curve" (AUC) to evaluate whether a subset is enriched in each cell. We through the use of PanglaoDB cell markers information in web site (https://panglaodb.se/index.html) for single-cell GSE184198 (scRNA—seq) dataset based on grade AUCell get high scores on subgroup of cells in cell type annotation, to identify the cell types.

RNA isolation and RT-qPCR

Total RNA from gastric mucosal epithelial cell (GES1), HGC-27, NCI-N87 and AGS was isolated with TRIzol (AC0101-B;). PCR was carried out with 2 × HQ SYBR qPCR Mix (ZF501) on an Applied Biosystems 7500 cycler. Primers were shown in Supplementary Table 2.

Statistical analysis

R (4.2.2) was adopted for data processing. Comparisons between two groups were made by t-test. Non-normally distributed data were analyzed by WilCoxon rank sum test. Survival differences and survival time were shown by KM curves and Log-rank test. Correlation was assessed by Spearman. p < 0.05 was designated as significant.

Results

Construct prognostic models related to amino acid metabolism

We analyzed public data to identify genes associated with amino acid metabolism (AAM) (Supplementary Fig. 1), and obtain Amino acid metabolism Related Model Genes (AAMRMGs) from TCGA-STAD dataset. The expression of AAMRGs was combined with OS clinical data, then analyzed by Univariate Cox. LASSO was conducted using the glmnet package on 25 genes with $p < 0.05$ (Table S2 for details). The simplest model (lambda.1se) was selected as the prognostic model. Our LASSO results identified 16 AAMRGs (ACLY, ADH4, COL1A1, F2, GADL1, GAMT, HBB, KYNU, MRI1, MTHFR, NR1D1, PDK4, SLC1A7, SLC25A15, SLC52A3, SYCE2) for GC diagnosis (Fig. 1A). We visualized the LASSO regression outcomes, producing a variable trajectory map (Fig. 1B). The lambda values for each AAMRG in the prognostic model were extracted for risk scores calculation and stratification. Risk score is calculated as a weighted sum of multiple factors (Supplementary Appendix formulas 1).

We then visualized the grouping in our model using a risk factor map (Fig. 1C) which has three parts: 1. Risk grouping: samples are grouped by the median of Risk Scores predicted by AAMRGs; 2. Survival outcomes: dot plots were used to show survival time and outcomes of TCGA-STAD dataset. Heat map: levels visualization of AAMRGs in TCGA-STAD dataset.

We then plotted group comparison plots of AAMRGs in tumor or normal tissues) (Fig. 1D). After excluding AAMRGs that could not be found in the above GEO dataset, different groups in GSE27342 dataset (Fig. 1E) and GSE54129 dataset (Fig. 1F) were drawn in Tumor, Normal for verification. The results showed that genes such as ACLY, COL1A1, and KYNU made the most significant contributions to the risk score, with their expression patterns showing significant differences between tumor and normal tissues ($p < 0.05$, Fig. 1D–F). Specifically, ACLY, COL1A1, and GAMT exhibited significantly differential expression in three independent datasets (TCGA: $p < 0.05$; GSE27342/GSE54129: $p < 0.01$), suggesting their universal roles in gastric cancer development (Fig. 1D–F).

Prognostic analysis and clinical correlation analysis of AAMRGs prognostic model

To check our model's efficacy, Relationships of Risk Scores and OS in STAD patient samples were evaluated. We plotted KM curves, examining Risk Scores of AAMRGs model (Fig. 2A) and the 16 AAMRGs.The results showed that four AAM-related genes (PDK4, SLC1A7, SLC25A15, SLC52A3, and SYCE2) were significantly associated with patient prognosis, and seven AAM-related genes (GAMT, MRI1, MTHFR, PDK4, SLC25A15, SLC52A3, and SYCE2) were significantly associated with overall survival (OS) ($p < 0.05$) (Fig. 2B–Q). This indicates that the risk scores derived from the AAM-related genes model are closely related to patient survival.

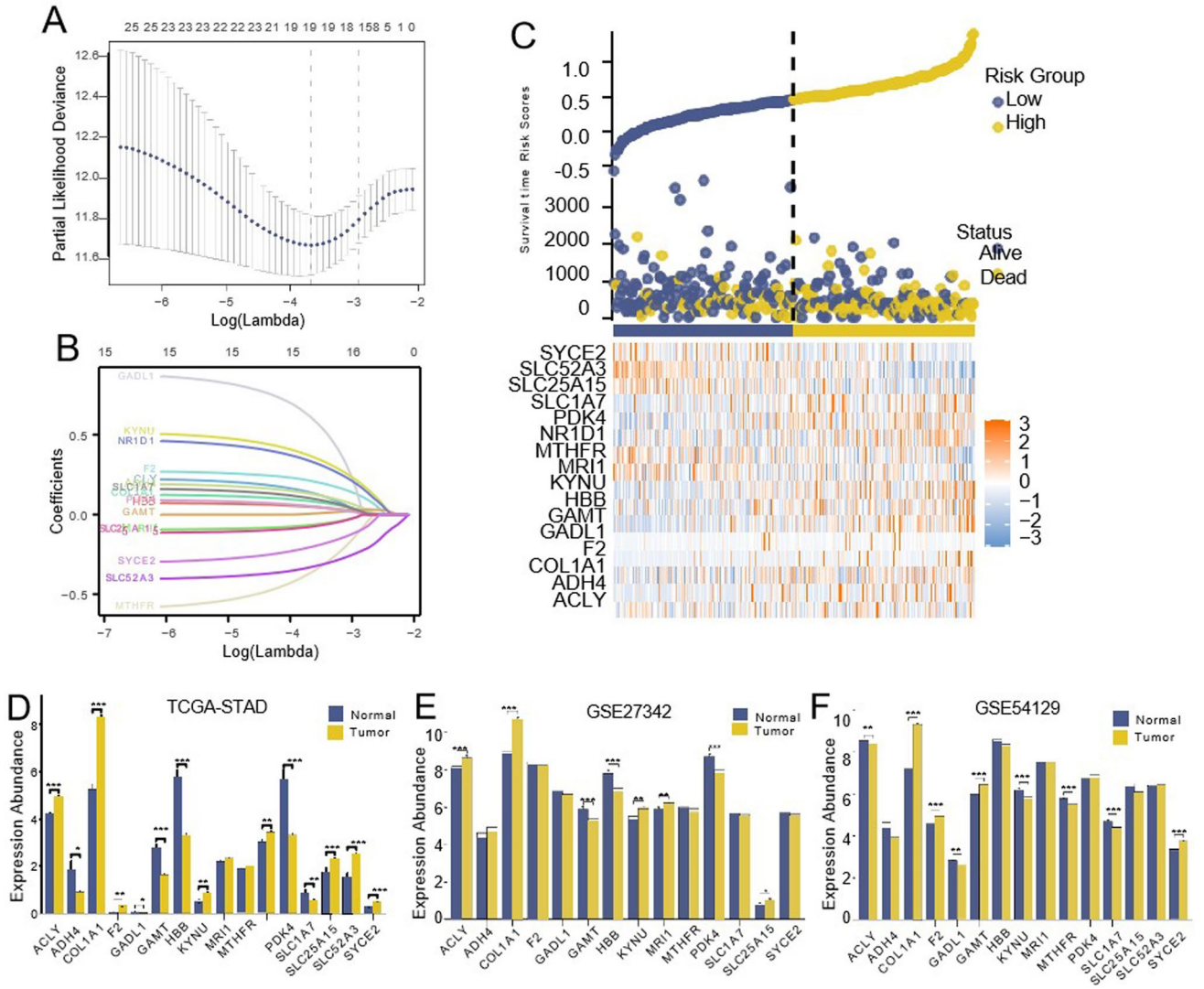

Fig. 1 Construction of AAMRGs prognostic model. AAMRGs prognostic model plot of (**A**, **B**) Lasso regression modelvariable trajectory plot. **C** Plot of risk factors for AAMRGs prognostic model. **D**–**F** Group comparison plots of AAMRGs of **D** TCGA-STAD dataset, **E** GSE27342 dataset and **F** GSE54129 dataset

Subsequently, based on the TCGA-STAD dataset, this study compared the clinical characteristics of gastric cancer patients in the high and low risk groups, including staging (Supplementary Fig. 2A–C), pathological staging (Supplementary Fig. 2D), age (Supplementary Fig. 2F), overall survival (OS) (Supplementary Fig. 2G), disease-specific survival (DSS) (Supplementary Fig. 2H), and progression-free interval (PFI) (Supplementary Fig. 2I). Stacked bar charts (Fig. 4A–I) showed that these indicators were significantly different between the two groups. The relevant baseline data are presented in Table 1. Overall, patients in the high-risk group exhibited more advanced tumor stages and poorer survival outcomes, suggesting that this risk stratification may be closely related to the aggressive biological behavior and poor prognosis of gastric adenocarcinoma.

Differential analysis and functional enrichment (GO) analysis in TCGA-STAD dataset

To identify risk-associated DEGs STAD patients, DESeq2 was used. A total of 533 DEGs were identified: 512 with logFC > 0 and 21 with logFC < 0. We visualized these results using a volcanic plot (Fig. 3A). Subsequently, we analyzed expression differences of 16 AAMRGs (ACLY, ADH4, COL1A1, F2, GADL1, GAMT, HBB, KYNU, MRI1, MTHFR, NR1D1, PDK4, SLC1A7, SLC25A15, SLC52A3, SYCE2) across high or low-risk groups, generating a heatmap by pheatmap (Fig. 3B).

To assess biological processes associated with these genes, we conducted Gene Ontology (GO) enrichment analysis, revealing significant enrichment in carboxylic acid biosynthetic processes, organic acid biosynthesis, and

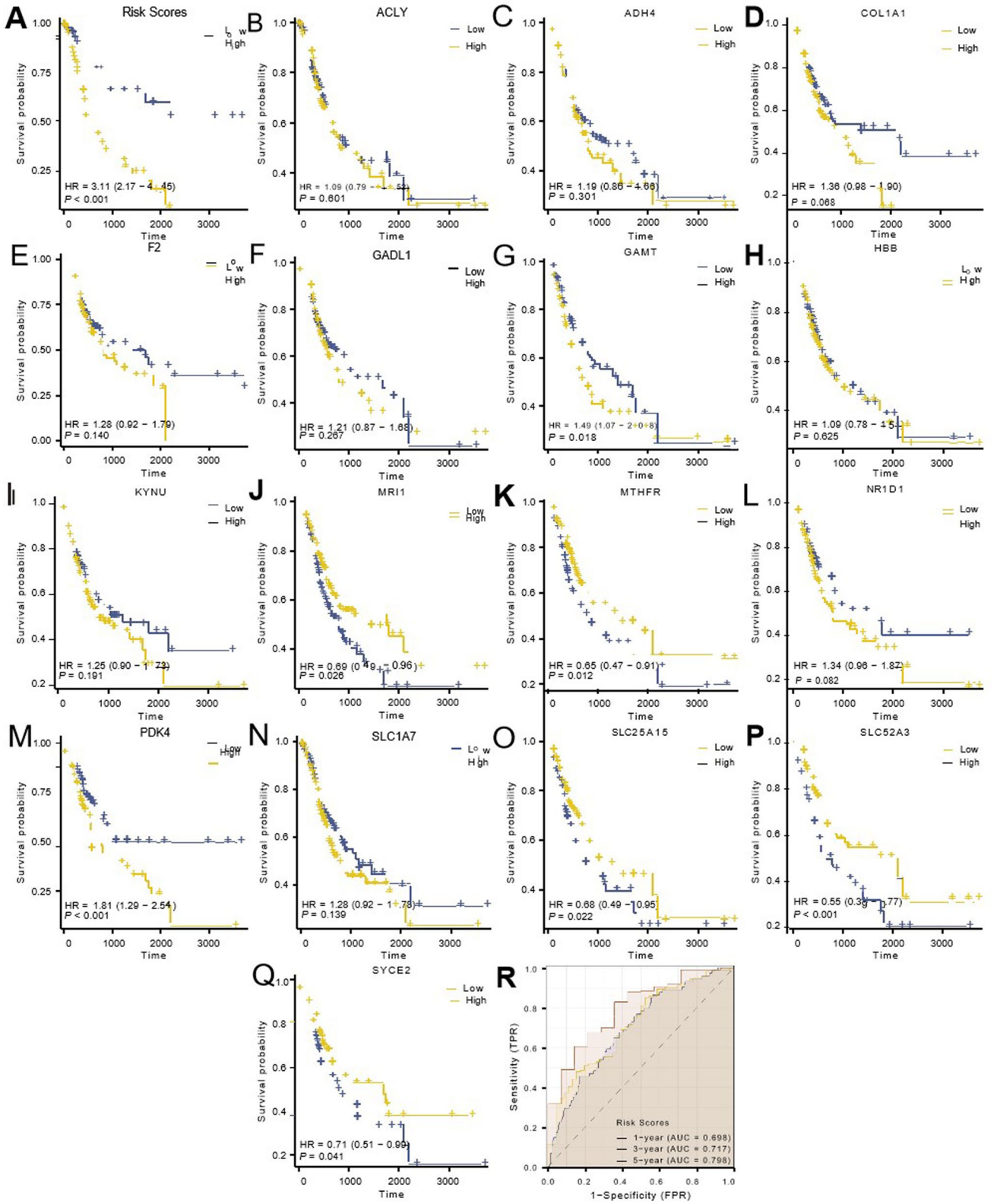

Fig. 2 Prognostic analysis of AAMRGs in TCGA-STAD. **A** KM curve of Risk Scores of AAMRGs prognostic model and patient prognostic OS. **B**–**Q** KM curves of ACLY, ADH4, COL1A1, F2, GADL1, GAMT, HBB, KYNU, MRI1, MTHFR, NR1D1, PDK4 SLC1A7, SLC25A15, SLC52A3, SYCE2 of patient prognosis OS in TCGA-STAD dataset are shown. **R** Time-dependent ROC curves of Risk Scores of AAMRGs prognostic model and patient prognostic OS in TCGA-STAD dataset are shown

Table 1 Patient characteristics of STAD patients in the TCGA datasets

Characteristics	Low	High	*P* value
N	176	177	
Stage_T, n (%)			0.035
T1	15 (4.3%)	3 (0.9%)	
T2	35 (10%)	39 (11.2%)	
T3	81 (23.2%)	82 (23.5%)	
T4	44 (12.6%)	50 (14.3%)	
Stage_N, n (%)			0.006
N0	63 (18.4%)	40 (11.7%)	
N1	53 (15.5%)	43 (12.6%)	
N2	31 (9.1%)	41 (12%)	
N3	26 (7.6%)	45 (13.2%)	
Stage_M, n (%)			0.050
M0	162 (48.1%)	152 (45.1%)	
M1	7 (2.1%)	16 (4.7%)	
Pathologic_stage, n (%)			0.028
Stage I	30 (8.9%)	18 (5.3%)	
Stage II	62 (18.3%)	47 (13.9%)	
Stage III	65 (19.2%)	81 (24%)	
Stage IV	13 (3.8%)	22 (6.5%)	
Gender, n (%)			0.551
Male	111 (31.4%)	117 (33.1%)	
Female	65 (18.4%)	60 (17%)	
Age, n (%)			0.152
>60	124 (35.1%)	112 (31.7%)	
≤60	52 (14.7%)	65 (18.4%)	
OS, n (%)			<0.001
No	132 (37.4%)	78 (22.1%)	
Yes	44 (12.5%)	99 (28%)	
DSS, n (%)			<0.001
No	147 (44.3%)	97 (29.2%)	
Yes	22 (6.6%)	66 (19.9%)	
PFI, n (%)			<0.001
No	137 (38.8%)	95 (26.9%)	
Yes	39 (11%)	82 (23.2%)	

TCGA the cancer genome atlas, *STAD* stomach adenocarcinoma, *OS* overall survival, *DSS* disease specific survival, *PFI* progression free interval

No stands for patient survival status: Alive

Yes represents the survival status of the patient: Dead

sulfur compound metabolism related to GC (Supplementary Table 3). Key molecular functions included binding to vitamins and L-amino acid transmembrane transporter activity. GO enrichment results are illustrated as bar graphs (Fig. 3C) and network diagrams (Fig. 3D, E).

GSEA and GSVA assay of TCGA-STAD dataset

To evaluate effects of gene expression on GC occurrence in different risk groups, GSEA was carried out (Supplementary Table 4). Data demonstrated that all genes were considerably enriched in mitochondrial_translation (Fig. 4B), DNA_replication (Fig. 4C), and mitochondrial_translation (Fig. 4B). diseases_of_metabolism (Fig. 4D), digestion (Fig. 4E) and other pathways (Supplementary Table 4).

In addition, we carried out GSVA analysis on all genes. The top 10 positively correlated pathways were mainly related to DNA_replication, such as reactome_DNA_replication_initiation, were presented by heatmap (Fig. 4F) (Supplementary Table 5). Notably, the major pathways identified through GSEA (including DNA replication, mitochondrial translation, and metabolic reprogramming signatures) were consistently validated by GSVA, indicating convergence across independent enrichment frameworks and supporting the robustness of the pathway-level findings.

PPI network, mRNA-miRNA, mRNA-TF, mRNA-RBP interaction network were constructed

We constructed the PPI network of AAMRGs using the STRING database (minimum interaction score: 0.150), which included 14 AAMRGs (Fig. 5A). Using the cytoHubba tool in Cytoscape (Degree, DMNC, EPC, MCC, and MNC algorithms), we identified the top 10 AAMRGs with the highest scores (Fig. 5B–F) and selected 8 hub genes: ACLY, ADH4, COL1A1, GAMT, HBB, MTHFR, NR1D1, and PDK4 (Fig. 5G). Thereafter, we adopted GeneMANIA database to assess the association of eight hub genes (ACLY, ADH4, COL1A1, GAMT, HBB, MTHFR, NR1D1, PDK4) with other genes.The results showed that six hub genes (ACLY, ADH4, GAMT, MTHFR, NR1D1, PDK4) had Co-localization and Co-expression with other genes (Fig. 5H). Levels of that six hub genes in some GC cell lines were higher than normal cells,in particular ACLY and NR1D1 (Fig. 6).

We further analyzed the interactions between the hub genes and miRNAs using the miRDB database, constructing a network that included 7 hub genes and 59 miRNAs (Supplementary Fig. 3A, Supplementary Table 6). We explored the interactions between the hub genes and transcription factors (TFs) using the ChIPBase database, constructing a network that included 8 hub genes and 29 TFs (Supplementary Fig. 3B, Supplementary Table 7). Additionally, we predicted the interactions between the hub genes and RNA-binding proteins (RBPs) using the ENCORI database, constructing a network that included 7 hub genes and 35 RBPs (Supplementary Fig. 3C, Supplementary Table 8).

Differential analysis of immune characteristics of TCGA-STAD dataset

To investigate immune characteristics of hub genes, ssGSEA was adopted to assess infiltrating abundance of 28 types of

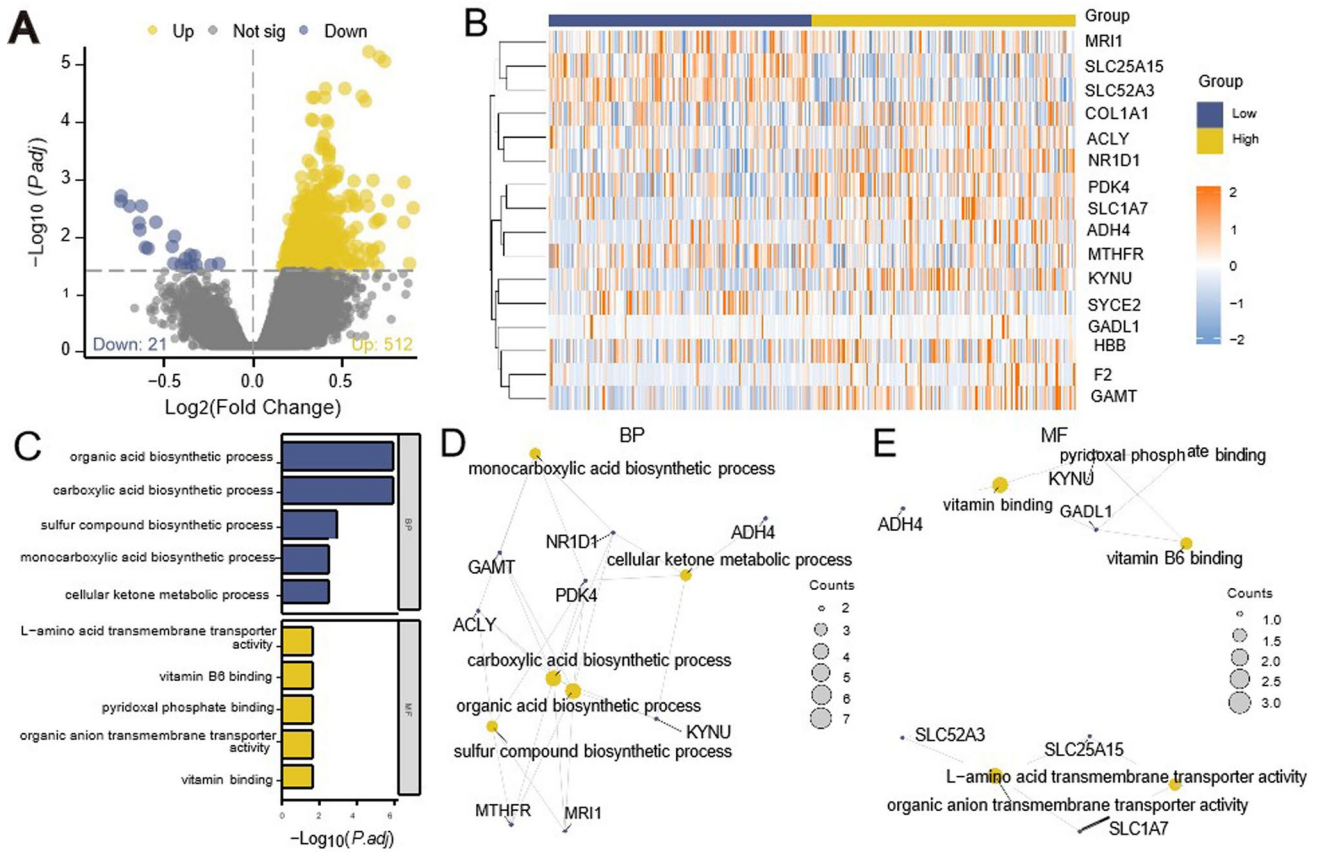

Fig. 3 Differential analysis and GO analysis results. **A** Volcano plot display of DEGs related to amino acid metabolism in GC patient samples from TCGA-STAD. **B** Display of AAMRGs heat map results. **C** Bar chart presentation of GO results of AAMRGs in TCGA-STAD. **D** and **E** Network graph presentation of GO results (BP, MF)

immune cells, which was analyzed by Mann–Whitney U test, then displayed in Fig. 7A. Analysis revealed considerable differences in immune cell infiltration, between high- and low-risk groups.

Infiltrating abundances of 23 immune cells were correlated (Fig. 7B). Additionally, we examined the correlation between these immune cells and ACLY, ADH4, COL1A1, GAMT, HBB, MTHFR, NR1D1, and PDK4. Notably, PDK4 showed positive correlations with all 23 immune cell types, whereas NR1D1 exhibited significant negative correlations with most immune cells (Fig. 7C).

To further explore immune infiltration differences, CIBERSORT was applied to STAD patient samples. Correlation between immune cell infiltration and risk groups is illustrated in a group comparison plot (Fig. 7D). Results demonstrated that high and low-risk groups had significantly different immune-cell infiltration (Fig. 7E). Importantly, infiltration of activated CD4 memory T cells and 8 hub genes was negatively correlated (Fig. 7F).

Construction of clinically relevant prognostic models in TCGA-STAD dataset

We investigated the prognostic value of ACLY, ADH4, COL1A1, GAMT, HBB, MTHFR, NR1D1, and PDK4, integrating their levels with clinical factors and OS. Variables ($p<0.1$) in univariate regression were used for a multivariate regression (Table 2). A forest plot (Fig. 8A) summarizes the univariate Cox results. Risk scores (Riskscore) were evaluated according to levels of relevant clinical variables, stratifying patients into High or Low group. Risk scores were calculated as a weighted sum of multiple factors (Supplementary Appendix formulas 2).

Subsequently, nomogram assay was applied to evaluate our model's ability (Fig. 8B). Nomogram assigns scores to each variable, facilitating the calculation of a total score for predicting event probabilities. Notably, ADH4 demonstrated greater prognostic utility compared to other variables.

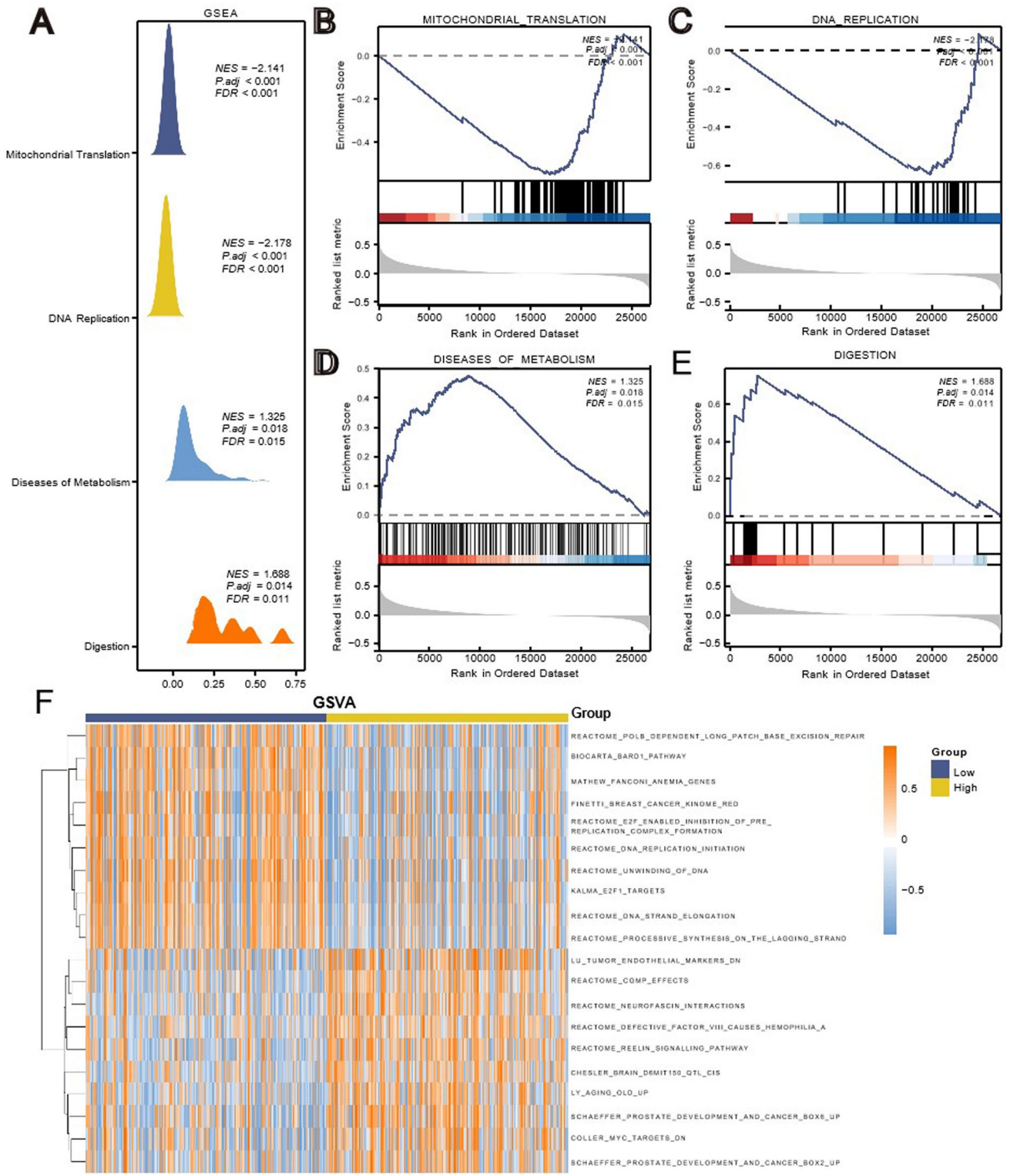

Fig. 4 GSEA and GSVA assay results. **A** Mountain plot presentation of the four main biological features from GSEA enrichment analysis. AAMRGs were strongly enriched in MITOCHONDRIAL_TRANSLATION (**B**), DNA_REPLICATION (**C**), DISEASES_OF_METABOLISM (**D**), DIGESTION (**E**) and other pathways in TCGA-STAD dataset. **F** Heat map presentation of 20 major biological features from GSVA enrichment analysis

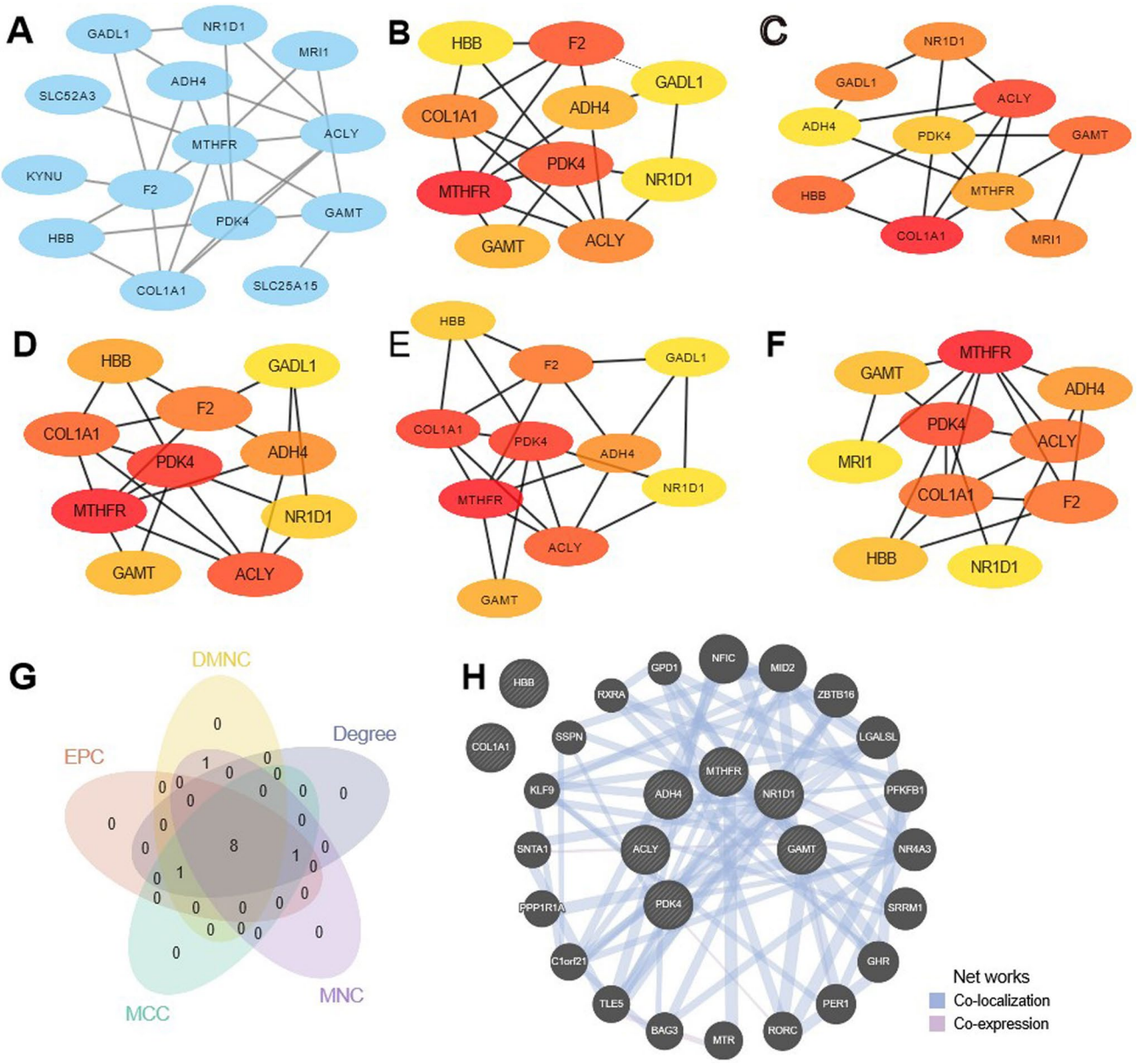

Fig. 5 PPI network. **A** PPI network of AAMRGs. **B**–**F** Degree (**B**) DMNC, (**C**) EPC, (**D**) MCC, (**E**) and MNC, (**F**) algorithm top 10 node network. **G** Venn diagram results of top 15 nodes of MCC, DMNC, MNC, EPC, and Degree algorithms in PPI are presented. **H** hub gene's GeneMANIA database analysis results are displayed in PPI network

Nomogram was calibrated for 1-, 3-, or 5-year (Fig. 8C–E) survival predictions. Calibration plots compare predicted survival probabilities against actual data, with closer alignment to ideal gray line indicating better accuracy. Calibration analysis indicated that the model's predictive performance ranked as 5 years > 3 years > 1 year.

Finally, we employed DCA to assess our model's usefulness at 1-, 3-, or 5-year (Fig. 8F–H). DCA plots demonstrate that the model consistently outperformed both "All positive" and "All negative" lines across assessed time points, further confirming its clinical prediction efficacy as 5 years > 3 years > 1 year.

The risk factor plot (Supplementary Fig. 4A) displays the model grouping, including risk stratification, survival outcomes, and a heatmap of core gene expression levels. Based on the risk scores and overall survival (OS) data, we analyzed the survival outcomes of the low-risk and high-risk groups, generating Kaplan–Meier (K-M) curves

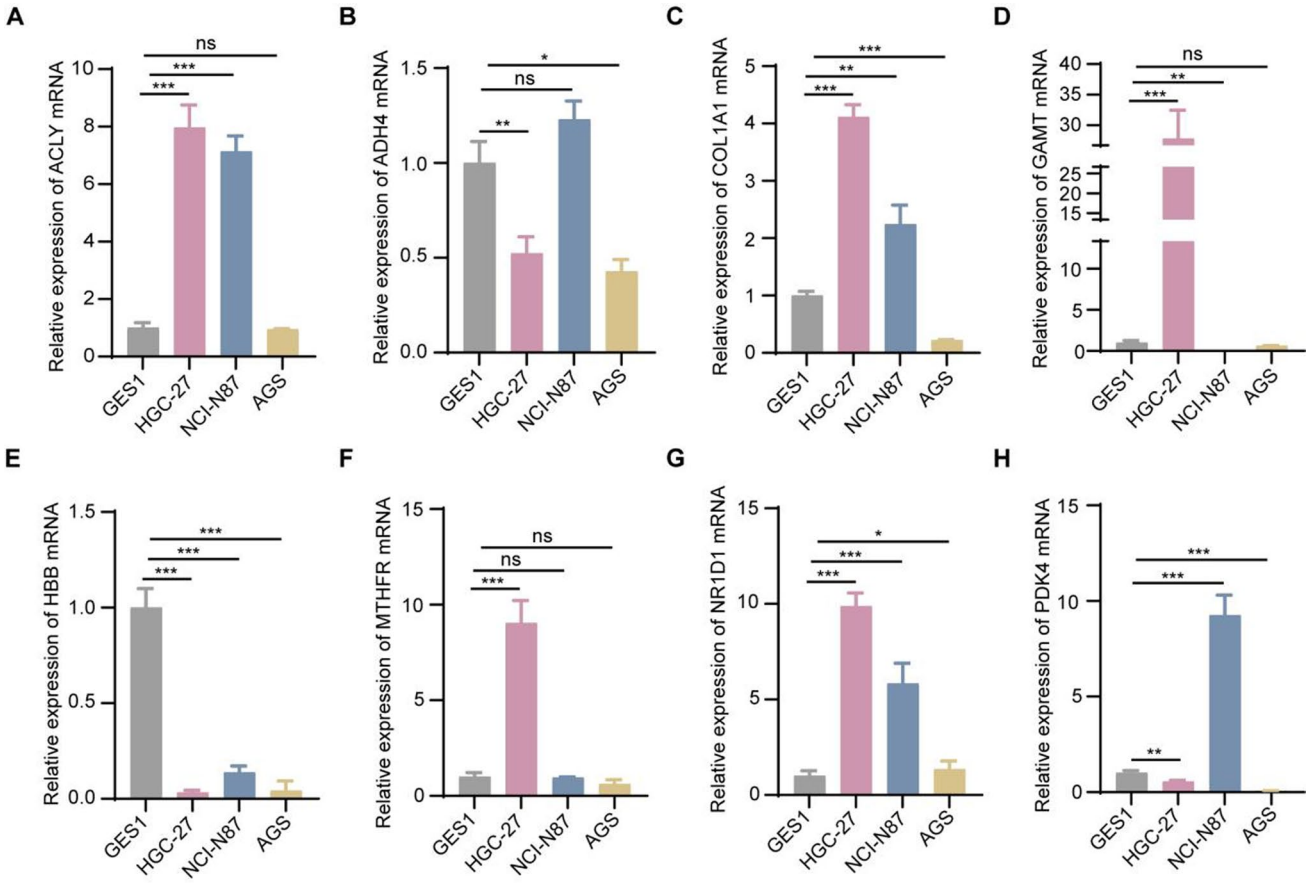

Fig. 6 The RNA expression level of 8 hub genes in GC cell. RNA level of 8 hub genes were verified at transcript levels by using GES1, HGC-27, NCI-N87 and AGS. **A** ACLY, **B** ADH4, **C** COL1A1, **D** GAMT, **E** HBB, **F** MTHFR, **G** NR1D1 and **H** PDK4 mRNA levels

(Supplementary Fig. 4B). The results showed a significant difference in prognosis between the two groups. Additionally, we evaluated the predictive accuracy of the risk scores using time-dependent receiver operating characteristic (ROC) curves (Supplementary Fig. 4C) and plotted ROC curves for the core genes (Supplementary Fig. 4D–G). The results indicated that the expression levels of the core genes had a relatively low correlation with the diagnostic classification of patient risk groups.

Cellular heterogeneity

Read10X function of Seurat v4.0 was applied to import all samples from the GSE184198 dataset. After retaining cells with a minimum of 3 cells and at least 200 genes, we merged the samples. The dataset includes one GC tissue sample and one normal gastric tissue sample (NT) from a GC patient. We filtered the GSE184198 matrix to exclude cells with mitochondrial gene content exceeding 5%, fewer than 500 features, and UMI counts over 20,000. We then compared the number of gene features (nFeature RNA) across groups (GC/NT), as well as gene counts (nCount RNA) and mitochondrial proportions (mitoRatio), revealing significant differences (Fig. 9A).

Next, we applied Harmony software for batch effect correction, leading to a single-cell dimensionality reduction map (Fig. 9B). Quality control metrics indicated a strong correlation between nCount and nFeature (R = 0.69), confirming the presence of batch effects. Using UMAP for visualization, we classified 7160 cells into 17 distinct clusters with a resolution set to 0.6 (Fig. 9C). Annotation via the CellMarker 2.0 database identified 11 cell types, including CD4 + T cells, M1 macrophages, fibroblasts, plasma cells, etc (Fig. 9D). We illustrated 14 cell marker genes (e.g., PRL30, CCL5, CD3D) across the clusters using bubble plots (Fig. 9E).

We assessed expression differences among cell populations, displaying a heatmap of the top 20 DEGs (Fig. 9F). Notably, IGHG1, IGHG4, IGHG3, and IGKC were highly expressed in plasma cells. Additionally, we analyzed percentages of various cell subsets in GC samples, revealing that CD4+T cells constituted the highest proportion, followed by M1 macrophages, while cancer stem cells represented the lowest (Fig. 9G).

Active AUCell score and novel subtype analysis of M1 macrophage cells for MAEMTRDEGs

To verify the activity degree of 6 AAM-related DEGs in different cell groups of GSE184198 single-cell dataset, AUCell was adopted to display activity of AAM-related DEGs in GC samples of single-cell dataset with UMAP diagram (Fig. 10A). By UMAP mapping, we found that Myofibroblast cells had a high AUCell score.

To assess gene differences between M1 macrophages and other cell populations in the single-cell dataset, we conducted pathway enrichment analysis. Using M1 macrophages as a reference, we employed the "FindAllMarkers" function to identify DEGs between M1 and other cell groups. Eight significant DEGs were chosen for GO analysis. Data demonstrated that DEGs were enriched in BP including RNA splicing regulation, molecular functions (MF) like cell-substrate junctions, and cellular components (CC) including ubiquitin-like protein ligase binding. GO functional enrichment results are depicted in bar graphs (Fig. 10B), while BP, MF, and CC pathway enrichments are illustrated through ring network diagrams (Figs. 10C–E).

To explore new subtypes of M1 macrophage cells, we clustered clusters of M1 macrophage cells, and then we applied UMAP dimensionality reduction for visualization, with a resolution of 0.6, M1 macrophage cells from GSE184198 single-cell dataset were successfully classified into three independent clusters (Supplementary Fig. 5A). We used the cell marker information in the PanglaoDB website to re-annotate types of GSE184198. We identified cell cluster as a total of three M1 macrophage cell subtypes (Supplementary Fig. 5B). GPR183−M1 macrophage, HLA-DRB1-M1 macrophage, KLRC1-M1 macrophage). Violin plot and heat map were utilized to show expression of three independent clusters in different subtypes (Supplementary Fig. 5C–F). We then generated a volcano plot based on differential analysis, and labeled genes were DEGs related to amino acid metabolism (Supplementary Fig. 5G).

Discussion

GC is a malignant tumor which characterized by high heterogeneity, mild initial symptoms, and various biological characteristics. Tumor subtypes characterization according to gene profiling help to better understand of GC's features and heterogeneity (Comprehensive molecular characterization of gastric adenocarcinoma 2014). Amino acid metabolism plays a key role in tumor cell transformation, immune evasion, drug resistance, etc. (Ling et al. 2023).However Amino acid metabolism genes are double-edged swords with dual potential in GC. It has been reported that AAM is key in innate immunity and antitumor responses (Procaccini et al. 2021; Sinclair et al. 2013). Many researches have focused on a specific gene (Wang et al. 2023b; Zhao et al. 2023; Dai et al. 2021). However the overall effect mediated by AAMRGs and links between AAMRGs and TME is not clear. We investigated the characteristics of AAMRGs in GC systematically and constructed the prognostic model related to amino acid metabolism and uncovered the correlation between AAMRGs between and immune cell (Fig. 10).

Through comprehensive multi-omics analysis of TCGA-STAD, GSE27342 and GSE54129 datasets, this study systematically characterized the molecular features of amino acid metabolic reprogramming (AAMRGs) in gastric cancer and their dynamic interplay with the immune microenvironment. We established a robust prognostic model incorporating 16 key regulatory genes (including ACLY, COL1A1 and PDK4), with 14 genes demonstrating significant differential expression in tumor tissues ($p<0.05$). The model's risk stratification showed strong prognostic value, significantly correlating with overall survival (HR=2.34, $p<0.001$) and demonstrating moderate predictive accuracy for 3-year and 5-year survival (AUC=0.72 and 0.68, respectively).

Although SYCE2, SLC52A3, and MRI1 have not been widely reported in gastric cancer, our multi-omics analysis provides several layers of evidence supporting their biological and clinical relevance. First, pathway enrichment analysis revealed that these genes are embedded in networks associated with cell cycle regulation, metabolic remodeling, and DNA replication, all of which are essential hallmarks of gastric tumorigenesis. For instance, SYCE2, originally known for its role in meiotic chromosomal synapsis, has recently been implicated in mitotic instability, chromosomal segregation defects, and proliferation of multiple cancers, suggesting that its reactivation in somatic cells may contribute to genomic instability and malignant progression (Hosoya and Miyagawa 2021). SLC52A3 encodes a riboflavin transporter that regulates redox balance and

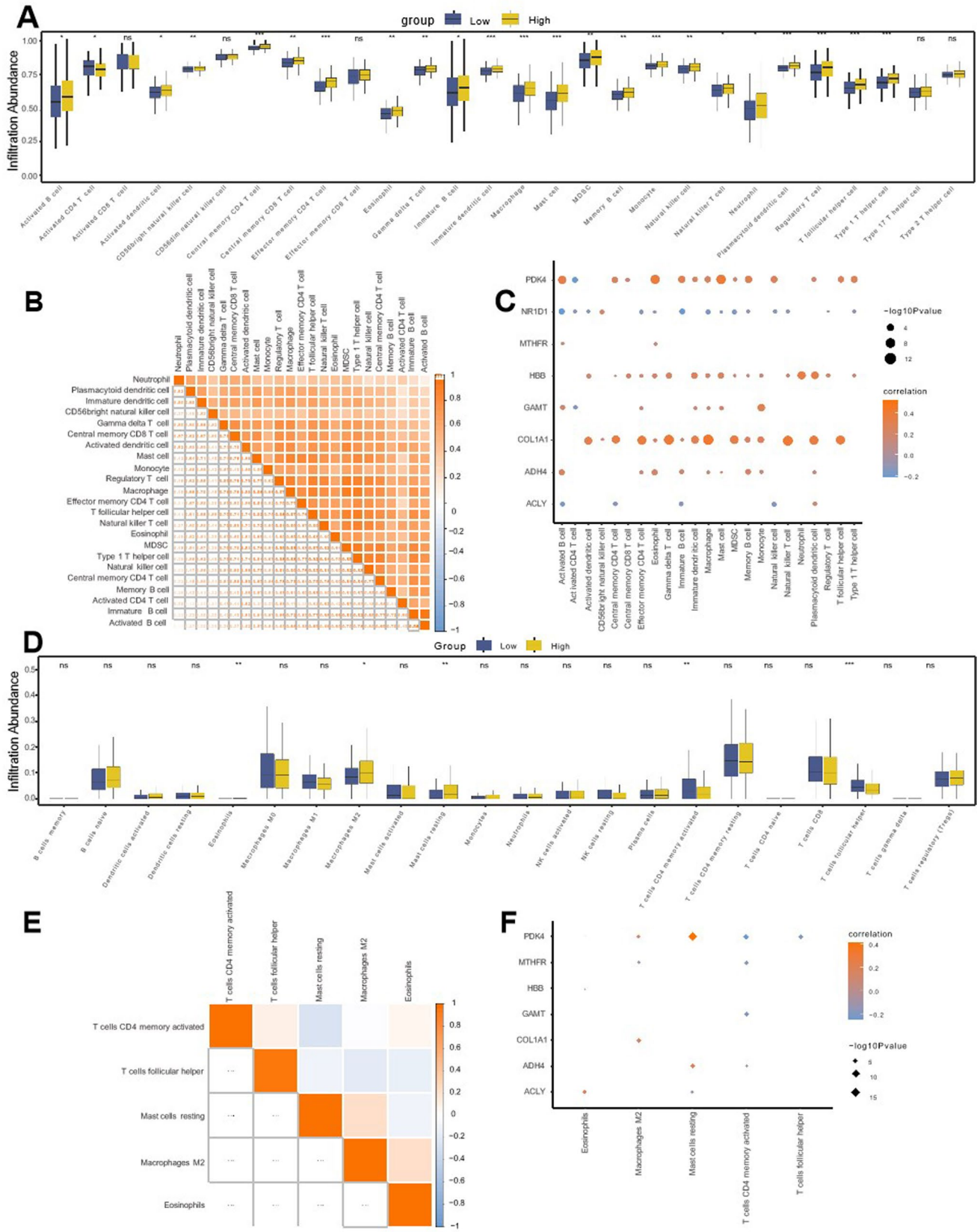
A
group
Low
High
Infiltration Abundance
B
C
-log10Pvalue
correlation
D
Group
Low
High
Infiltration Abundance
E
F
correlation
-log10Pvalue

Fig. 7 Immune infiltration analysis. The group comparison figure of ssGSEA immune infiltration results of **A** TCGA-STAD dataset is shown. **B** Correlation of immune cell infiltration in TCGA-STAD dataset are presented. **C** Correlation between immune cells and hub genes. **D** Group comparison diagram of CIBERSORT immune infiltration analysis results of TCGA-STAD dataset. **E** The results of correlation analysis of immune cell infiltration abundance in TCGA-STAD dataset are presented. **F** Dot plot of correlation between immune cells and hub genes in TCGA-STAD dataset

mitochondrial metabolism; dysregulated riboflavin uptake has been linked to altered oxidative phosphorylation and metabolic stress pathways frequently activated in gastric cancer (Nagano et al. 2021; Zhang et al. 2022). MRI1, a key enzyme in methionine salvage and methylthioribose metabolism, participates in one-carbon metabolism and methyl donor generation, both of which influence epigenetic modifications and immune-metabolic remodeling within the tumor microenvironment (Zhou et al. 2023; Shibata et al. 2024). In our prognostic model, all three genes showed consistent differential expression in independent datasets and strong associations with survival, supporting their functional relevance. Together, these findings suggest that although these genes are not classical gastric cancer markers, they represent previously underappreciated metabolic and genomic regulators that may contribute to gastric cancer progression through metabolic reprogramming, redox regulation, and genomic instability. This underscores the advantage of data-driven multi-omics approaches in uncovering non-canonical but biologically meaningful prognostic biomarkers.

Single-cell transcriptomic profiling revealed a complex regulatory network involving 8 hub genes (ACLY, ADH4, GAMT et al.) that modulate immune cell infiltration patterns, particularly CD8 + T cells and M1 macrophages. Experimental validation confirmed significant upregulation of these core genes in gastric cancer cell lines ($p < 0.01$), highlighting their potential as therapeutic

Table 2 Cox analysis of dataset TCGA-STAD

Characteristics	Total (N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value
Stage_T	349				
T1	18	Reference		Reference	
T2	74	6.729 (0.914–49.550)	0.061	2.839 (0.359–22.459)	0.323
T3	163	9.428 (1.309–67.896)	0.026	4.337 (0.490–38.359)	0.187
T4	94	9.652 (1.325–70.285)	0.025	4.900 (0.535–44.857)	0.160
Stage_N	342				
N0	103	Reference		Reference	
N1	96	1.629 (1.001–2.649)	0.049	1.396 (0.694–2.810)	0.350
N2	72	1.656 (0.980–2.798)	0.060	1.690 (0.701–4.076)	0.242
N3	71	2.647 (1.627–4.307)	<0.001	1.838 (0.763–4.429)	0.175
Stage_M	337				
M0	314	Reference		Reference	
M1	23	2.100 (1.184–3.724)	0.011	1.119 (0.450–2.779)	0.809
Pathologic_stage	338				
Stage I	48	Reference		Reference	
Stage II	109	1.553 (0.783–3.082)	0.208	0.952 (0.333–2.722)	0.927
Stage III	146	2.385 (1.258–4.522)	0.008	0.860 (0.212–3.483)	0.833
Stage IV	35	3.827 (1.855–7.898)	<0.001	1.615 (0.400–6.522)	0.501
Gender	353				
Male	228	Reference			
Female	125	0.771 (0.541–1.101)	0.152		
Age	353				
>60	236	1.479 (1.020–2.145)	0.039	1.806 (1.193–2.733)	0.005
≤60	117	Reference		Reference	
ACLY	353	1.256 (1.010–1.562)	0.040	1.141 (0.888–1.467)	0.303
ADH4	353	1.141 (1.026–1.269)	0.015	1.282 (1.122–1.465)	<0.001
COL1A1	353	1.136 (1.014–1.272)	0.027	1.184 (1.047–1.338)	0.007
GAMT	353	1.259 (1.099–1.442)	<0.001	1.301 (1.099–1.539)	0.002
MTHFR	353	0.645 (0.441–0.942)	0.023	0.584 (0.368–0.928)	0.023
NR1D1	353	1.281 (1.055–1.556)	0.012	1.493 (1.167–1.912)	0.001
HBB	353	1.125 (1.016–1.246)	0.024	1.093 (0.963–1.240)	0.170
PDK4	353	1.126 (1.023–1.238)	0.015	1.104 (0.984–1.238)	0.093

TCGA The cancer genome atlas, *STAD* Stomach adenocarcinoma

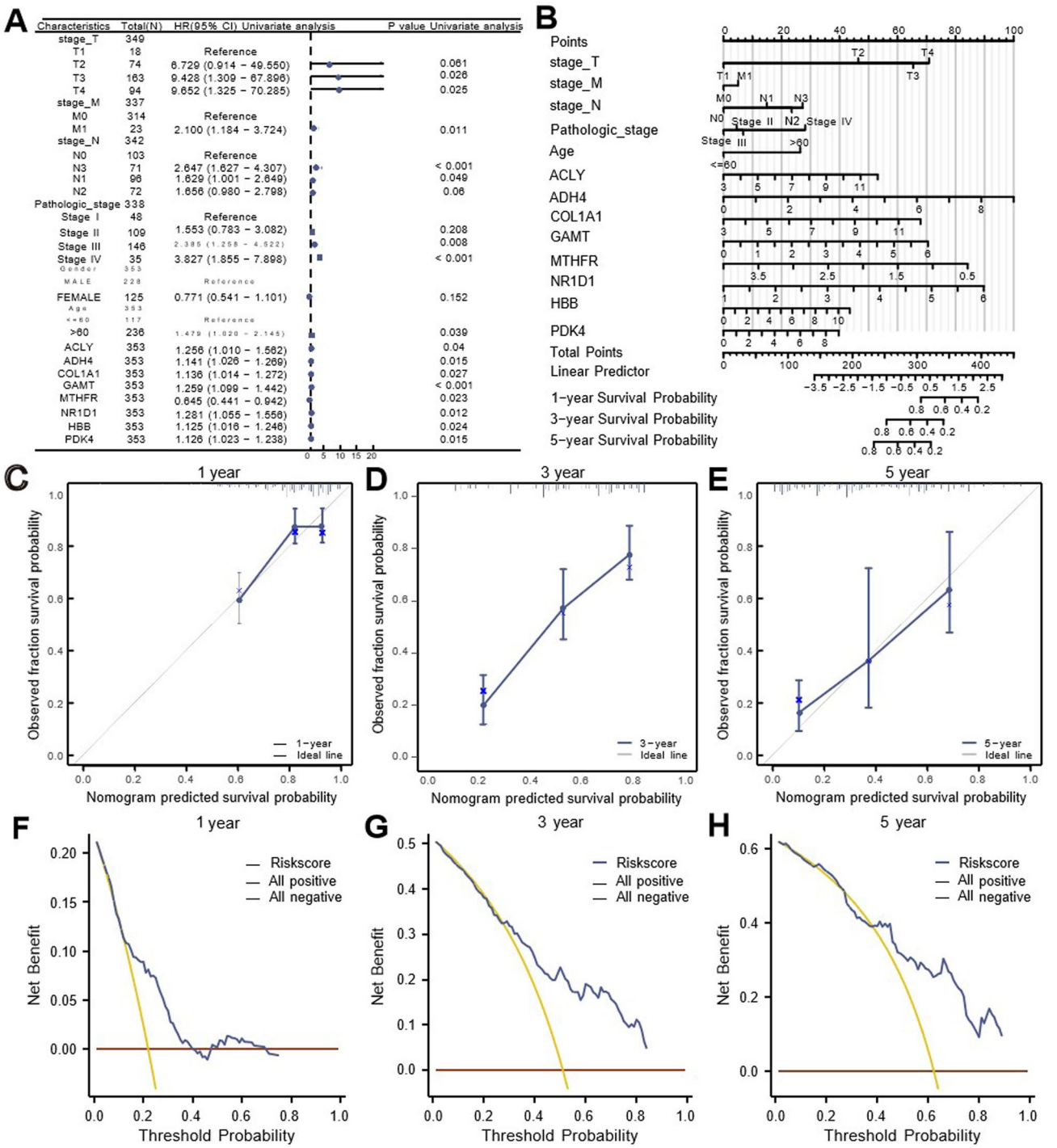

Fig. 8 Construction of clinically relevant prognostic model. **A** Forest plot presentation of univariate Cox regression model. **B** Nomogram of clinically relevant prognostic models. **C**–**E**. Calibration curves at 1- (**C**), 3- (**D**) or 5-years (**E**). **F**–**H** DCA plots at 1- (**F**), 3- (**G**), or 5-year (**H**). **H** Risk factor plots. **I** ROC of Risk Scores of prognostic models and patient prognostic OS in TCGA-STAD dataset. The nomogram can be used by summing the points assigned to each predictor variable and converting the total score to survival probabilities

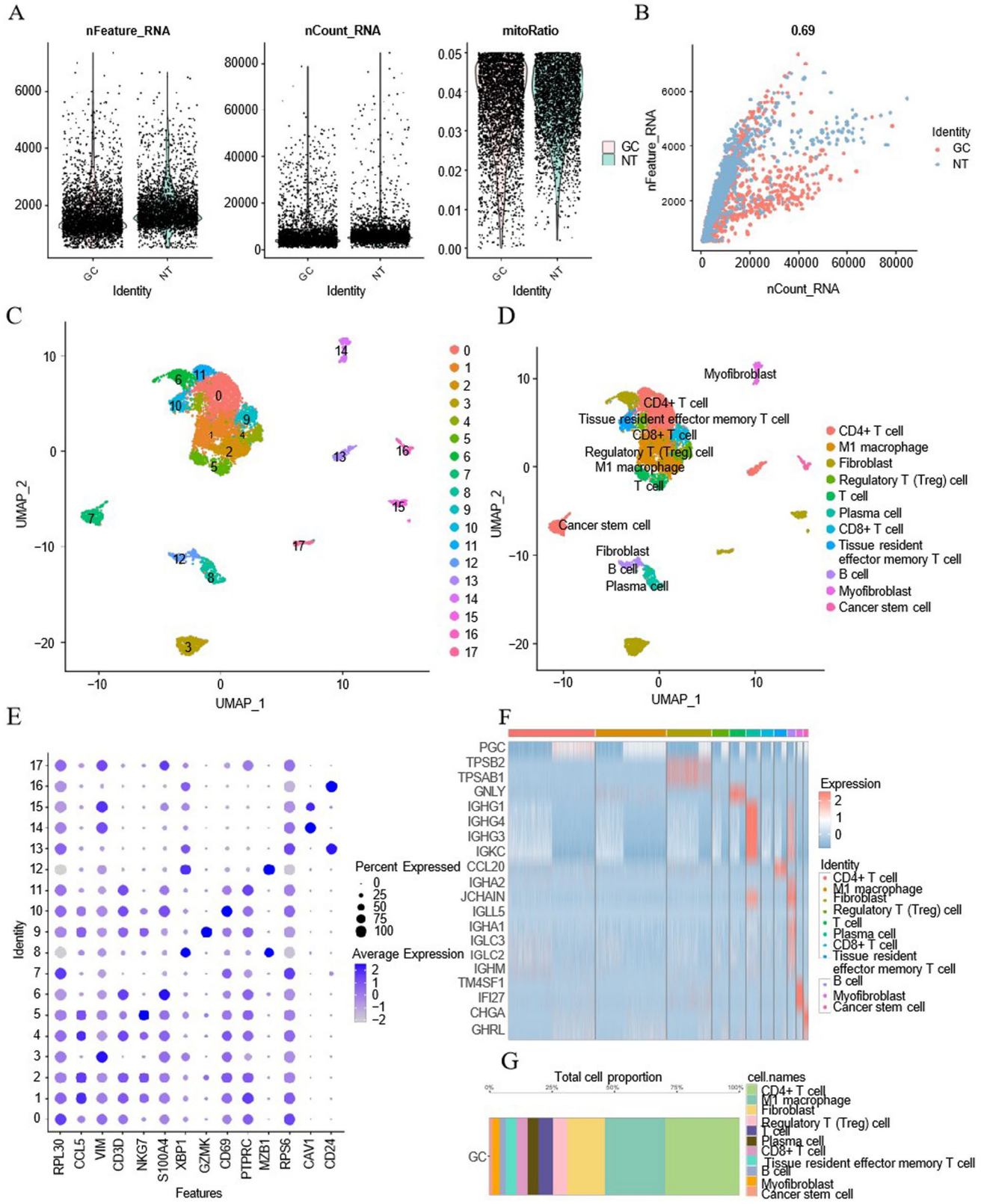
A
nFeature_RNA
nCount_RNA
mitoRatio
Identity
GC
NT
B
0.69
nFeature_RNA
nCount_RNA
Identity
C
UMAP_1
UMAP_2
D
Myofibroblast
CD4+ T cell
Tissue resident effector memory T cell
CD8+ T cell
Regulatory T (Treg) cell
M1 macrophage
T cell
Cancer stem cell
Fibroblast
B cell
Plasma cell
E
Identity
Features
RPL30
CCL5
VIM
CD3D
NKG7
S100A4
XBP1
GZMK
CD69
PTPRC
MZB1
RPS6
CAV1
CD24
Percent Expressed
Average Expression
F
PGC
TPSB2
TPSAB1
GNLY
IGHG1
IGHG4
IGHG3
IGKC
CCL20
IGHA2
JCHAIN
IGLL5
IGHA1
IGLC3
IGLC2
IGHM
TM4SF1
IFI27
CHGA
GHRL
Expression
Identity
G
Total cell proportion
cell.names
GC

Fig. 9 Data quality control of GSE184198 dataset. **A** Violin plot of gene number, sequenced counts, and mitochondria proportion in GSE184198 dataset. **B** Scatter plot of the correlation between nCount and nFeature in all cells of the GSE184198 dataset. **C** The 7160 cells of GSE184198 dataset were clustered into 17 cell clusters by UMAP. **D** Annotation of 17 cell clusters into 11 cell types through cell marker information in CellMarker 2.0 website. **E** Visualization of bubble plots of expression levels of 14 cell marker genes in 17 independent clusters. A darker color indicates a higher expression level, and a larger circle indicates a higher proportion of genes expressed within the cell population. **F** Heat map of DEGs (top 20) between cell clusters (color from dark to light represents gene expression from strong to weak). **G** Plot of the proportion of cell taxa in GC samples

targets. These findings align with recent advances in gastric cancer prognostication, including: (1) multi-omics-derived immune infiltration signatures (ICI scores) that predict chemotherapy sensitivity (Wang et al. 2025), and (2) machine learning-identified 32-gene classifiers that stratify molecular subtypes and treatment responses (Cheong et al. 2022).

Interestingly, our single-cell AUCell analysis revealed that myofibroblasts displayed the highest amino-acid–metabolic (AAM) activity among all stromal and immune cell populations. Myofibroblasts, which share functional characteristics with cancer-associated fibroblasts (CAFs), are known to participate in metabolic symbiosis by supplying nutrients such as amino acids, lactate, or glutamine to tumor cells (Noom et al. 2024). Elevated AAM activity in these stromal cells suggests a potential role in shaping the metabolic landscape of the gastric tumor microenvironment, possibly supporting tumor growth, modulating ECM remodeling, and influencing immune infiltration through nutrient competition or metabolic signaling. Although the current study is computational, these findings raise the possibility that AAM-high myofibroblasts may represent a previously underappreciated stromal component contributing to metabolic conditioning of the TME. Future studies integrating spatial transcriptomics and stromal-targeted perturbation models will be necessary to elucidate the mechanistic implications and therapeutic relevance of this stromal metabolic phenotype.

Notably, our study provides novel insights into metabolic-immune crosstalk, demonstrating NR1D1's negative correlation with immune infiltration versus PDK4's pan-immune positive correlation—consistent with established mechanisms of metabolic immunomodulation (Procaccini et al. 2021; Vickram et al. 2025). The AAM-based nomogram and decision curve analysis (C-index = 0.71 for 5-year survival) outperformed existing lactate-score (Yang et al. 2023) and transcription factor-based (Zhou et al. 2021) models, offering improved stratification tools.

For clinical translation, patients were stratified into high- and low-risk groups using the median risk score of the TCGA cohort. The same cutoff was applied to GEO validation datasets. This threshold stratification allows the nomogram to be applied consistently across cohorts and supports risk grouping for survival prediction.

To facilitate future clinical application, we provide a transparent workflow for using the model: (1) obtain mRNA expression of the 16 AAMRGs from tumor tissue; (2) calculate the individualized risk score using the provided coefficient table; (3) determine whether the patient falls into the high- or low-risk group based on the predefined cutoff; and (4) apply the nomogram to estimate 1-, 3-, and 5-year survival probabilities. This stepwise approach enables clinicians and researchers to integrate the model into prospective trials or molecular pathology workflows.

While limited by retrospective data, this work makes two key contributions. First, it delineates the spatiotemporal dynamics of AAMRG-TME interactions, informing combination strategies (e.g., IDO inhibitors with immunotherapy) (Butler et al. 2021). Second, it validates clinically actionable targets like COL1A1, previously linked to tumor aggressiveness (Comprehensive molecular characterization of gastric adenocarcinoma 2014). Future studies should employ prospective clinical samples to validate these findings and elucidate AAMRGs' role in treatment resistance.

In this study, several limitations exist. First, the analyses are primarily based on bulk transcriptomic data, which lack spatial resolution and may obscure cell-type–specific metabolic and immune interactions. Although we performed single-cell annotation using a publicly available scRNA-seq dataset to partially address this limitation, the findings remain inferential. Second, the metabolic–immune associations identified through enrichment analyses were not validated experimentally using approaches such as immunohistochemistry, qPCR, or spatial profiling. As our primary aim was to establish a computational framework and identify candidate amino acid–metabolic regulators, future studies will incorporate prospective cohorts, spatial transcriptomics, and wet-lab validation to confirm the mechanistic roles of these AAM-related genes in the tumor microenvironment. Importantly, the prognostic nomogram developed in this study has not yet been evaluated in prospective clinical cohorts. Because the AAMRG-based model may inform combination-therapy stratification—particularly for metabolic checkpoint inhibitors and immune checkpoint blockade—future work should include prospective trials to determine its predictive value for treatment

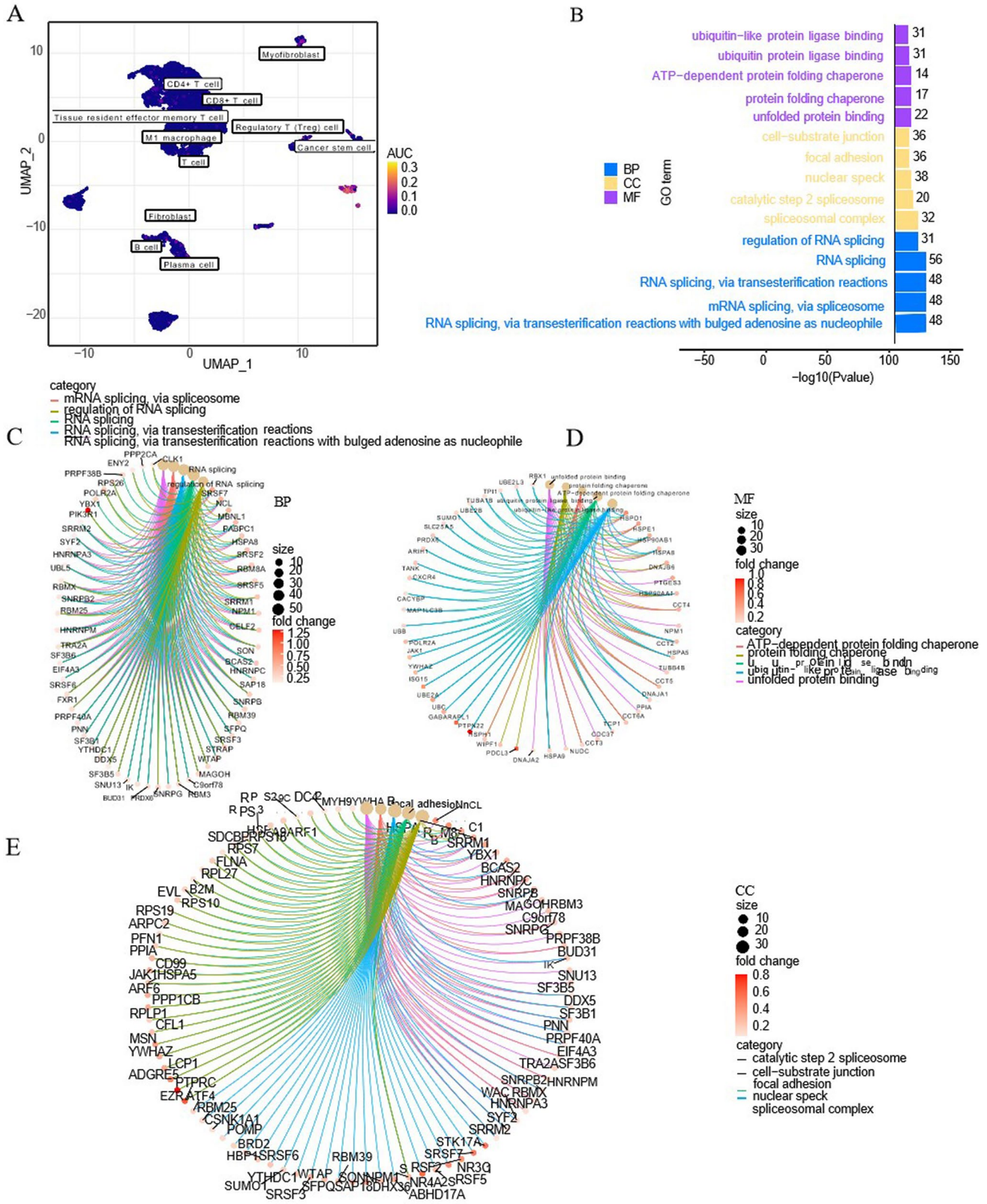

Fig. 10 Active AUCell scores of MAEMTRDEGs. **A** UMAP map of AUCell scores based on 11 amino acid metabolism-related DEGs in GSE184198 single-cell dataset (color from black to red represents score from low to high), **B** the bar chart of GO analysis of DEGs in M1 macrophage cell group in GSE184198 single-cell data set is displayed. **C**–**E** GO enrichment analysis showed that DEGs in BP,MF,CC

response and its applicability in guiding individualized therapeutic decisions.

Conclusion

In this study, we constructed an AAMRGs model for the prognosis of GC patients, and its risk score was significantly associated with overall survival. Additionally, hub genes' risk profiling were identified as independent prognostic factors for GC, indicating the association between AAMRGs and prognosis. Lastly, we profile cellular heterogeneity linking hub genes to immune cells. These studies offers new insightss for the survival prognosis and GC's immunotherapy. Future studies should further investigate the regulatory mechanisms of AAMRGs in immunotherapy response to advance precision medicine applications.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00726-026-03507-3.

Acknowledgements Our thanks go to the Jiujiang No.1 People's Hospital for its technical assistance and resources. We also extend our appreciation to those contributed to our study. The abstracts have been accepted for Poster presentation at the APDW 2024, taking place on 21-24 November 2024.We express our gratitude to for their support and contributions to this research.

Author contributions WL and ZH were responsible for the design and bioinformatics analysis of the project, while YW, ZL and GL conducted data analysis, WL and ZL performed the experimental validation, and WL and ZH were major contributors in writing the manuscript. All authors read and approved the final manuscript.

Funding It was supported by Jiujiang City Key Research and Development Plan. Grants Awards No. S2021ZDYFN148 (to Wei Ye) and the Discipline Construction Startup Fund of the First People's Hospital of Jiujiang City, No. JYXK2024007.

Data availability The experimental data used to support the findings of this study are available from the corresponding author upon request.

Declarations

Conflict of interest The authors declare no competing interests.

Ethical approval This study is a retrospective study that does not involve any new research on human or animal subjects. Therefore, it does not require ethical review.

Consent to participate Not applicable.

Consent for publication Not applicable.

Clinical trial number Not applicable.

References

Barrett T et al (2007) NCBI GEO: mining tens of millions of expression profiles–database and tools update. Nucleic Acids Res 35(Database issue):D760-5

Bindea G et al (2013) Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 39(4):782–795

Boughorbel S, Jarray F, El-Anbari M (2017) Optimal classifier for imbalanced data using Matthews correlation coefficient metric. PLoS ONE 12(6):e0177678

Bray F et al (2024) Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. https://doi.org/10.3322/caac.21834

Butler M et al (2021) Amino acid depletion therapies: starving cancer cells to death. Trends Endocrinol Metab 32(6):367–381

Cai W et al (2020) Nonparametric bootstrap inference for the targeted highly adaptive least absolute shrinkage and selection operator (LASSO) estimator. Int J Biostat 16(2):20170070

Chen Y et al (2020) miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Res 48(D1):D127–D131

Cheong JH et al (2022) Development and validation of a prognostic and predictive 32-gene signature for gastric cancer. Nat Commun 13(1):774

Colaprico A et al (2016) TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res 44(8):e71

Comprehensive molecular characterization of gastric adenocarcinoma (2014) Nature 513(7517):202–209

Cui J et al (2011) An integrated transcriptomic and computational analysis for biomarker identification in gastric cancer. Nucleic Acids Res 39(4):1197–1207

Dai D et al (2021) Interactions between gastric microbiota and metabolites in gastric cancer. Cell Death Dis. https://doi.org/10.1038/s41419-021-04396-y

Engebretsen S, Bohlin J (2019) Statistical predictions with glmnet. Clin Epigenetics 11(1):123

Gao Q et al (2022) Sodium danshensu attenuates cerebral ischemia-reperfusion injury by targeting AKT1. Front Pharmacol 13:946668

Goldman MJ et al (2020) Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol 38(6):675–678

Hanzelmann S et al (2013) GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 14:7

Hosoya N, Miyagawa K (2021) Synaptonemal complex proteins modulate the level of genome integrity in cancers. Cancer Sci 112(3):989–996

Hu C et al (2023) Cell Marker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. Nucleic Acids Res 51(D1):D870–D876
Jiang Y et al (2025) Targeting amino acid metabolism to inhibit gastric cancer progression and promote anti-tumor immunity: a review. Front Immunol 13(16):1508730
Liberzon A et al (2011) Molecular signatures database (MSigDB) 3.0. Bioinformatics 27(12):1739–1740
Ling Z-N et al (2023) Amino acid metabolism in health and disease. Signal Transduct Target Ther 8(1):174. https://doi.org/10.1038/s41392-025-02231-w
Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550
Mandrekar JN (2010) Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 5(9):1315–1316
Nagano T, Awai Y, Kuwaba S, Osumi T, Mio K, Iwasaki T, Kamada S (2021) Riboflavin transporter SLC52A1, a target of p53, suppresses cellular senescence by activating mitochondrial complex II. Mol Biol Cell 32(21):br10
Newman AM et al (2015) Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 12(5):453–457
Noom A, Sawitzki B, Knaus P, Duda GN (2024) A two-way street–cellular metabolism and myofibroblast contraction. NPJ Regener Med 9(1):15
Park SY (2018) Nomogram: an analogue tool to deliver digital knowledge. J Thorac Cardiovasc Surg 155(4):1793
Perkins NJ et al (2019) Combining biomarker calibration data to reduce measurement error. Epidemiology 30 Suppl 2(Suppl 2):S3–S9
Procaccini C et al (2021) Signals of pseudo-starvation unveil the amino acid transporter SLC7A11 as key determinant in the control of Treg cell proliferative potential. Immunity 54(7):1543-1560.e6
Qiu H et al (2021) Cancer incidence, mortality, and burden in China: a time-trend analysis and comparison with the United States and United Kingdom based on the global epidemiological data released in 2020. Cancer Commun 41(10):1037–1048
Shannon P et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13(11):2498–2504
Shibata Y, Yamada T, Ikeda Y, Kanai M, Fujii T, Akao T, Goshima T, Isogai A, Takahashi T (2024) Effect of S-adenosyl-methionine accumulation on hineka odor in sake brewed with a non-Kyokai yeast. J Biosci Bioeng 137(4):268–273
Siegel RL, Giaquinto AN, Jemal A (2024) Cancer statistics, 2024. CA Cancer J Clin 74(1):12–49
Sinclair LV et al (2013) Control of amino-acid transport by antigen receptors coordinates the metabolic reprogramming essential for T cell differentiation. Nat Immunol 14(5):500–508
Stelzer G et al (2016) The GeneCards suite: from gene data mining to disease genome sequence analyses. Curr Protoc Bioinformatics 54:1 30 1-1 30 33
Subramanian A et al (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102(43):15545–15550
Szklarczyk D et al (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47(D1):D607–D613
Tataranni T et al (2019) Dichloroacetate (DCA) and cancer: an overview towards clinical applications. Oxid Med Cell Longev 2019:8201079
Van Calster B et al (2018) Reporting and interpreting decision curve analysis: a guide for investigators. Eur Urol 74(6):796–804
Vickram S et al (2025) Immune biomarkers and predictive signatures in gastric cancer: optimizing immunotherapy responses. Pathol Res Pract 265:155743
Wang Y et al (2023a) METTL3 promotes drug resistance to oxaliplatin in gastric cancer cells through DNA repair pathway. Front Pharmacol 14:1257410
Wang M et al (2023b) Targeting TLK2 inhibits the progression of gastric cancer by reprogramming amino acid metabolism through the mTOR/ASNS axis. Cancer Gene Ther 30(11):1485–1497
Wang SY et al (2025) Construction of a prognostic model for gastric cancer based on immune infiltration and microenvironment, and exploration of MEF2C gene function. BMC Med Genomics 18(1):13. https://doi.org/10.1186/s12920-024-02082-4
Wu D et al (2021) Exostosin1 as a novel prognostic and predictive biomarker for squamous cell lung carcinoma: a study based on bioinformatics analysis. Cancer Med 10(8):2787–2801
Yang H et al (2023) Identification of lactylation related model to predict prognostic, tumor infiltrating immunocytes and response of immunotherapy in gastric cancer. Front Immunol. https://doi.org/10.3389/fimmu.2023.1149989
Yu G (2020) Gene ontology semantic similarity analysis using GOSemSim. Methods Mol Biol 2117:207–215
Zhang L, Li M, Cui Z, Chai D, Guan Y, Chen C, Wang W (2022) Systematic analysis of the role of SLC52A2 in multiple human cancers. Cancer Cell Int 22(1):8
Zhao R et al (2022) *Aqp5* complements LGR5 to determine the fates of gastric cancer stem cells through regulating ULK1 ubiquitination. J Exp Clin Cancer Res 41(1):322
Zhao G et al (2023) Comprehensive analysis to reveal amino acid metabolism-associated genes as a prognostic index in gastric cancer. Mediators Inflamm 2023:1–23
Zhou KR et al (2017) ChIPBase v2.0: decoding transcriptional regulatory networks of non-coding RNAs and protein-coding genes from ChIP-seq data. Nucleic Acids Res 45(D1):D43–D50
Zhou L-Q et al (2021) Establishment of a prognostic model of ten transcription factors in gastric cancer. Genomics 113(6):4075–4087
Zhou S, Zhao X, Zhang S, Tian X, Wang X, Mu Y, Li F, Zhao AZ, Zhao Z (2023) Prognosis prediction based on methionine metabolism genes signature in gliomas. BMC Med Genomics 16(1):317
Zhu Y et al (2025) Cancer cell-derived arginine fuels polyamine biosynthesis in tumor-associated macrophages to promote immune evasion. Cancer Cell S1535-6108(25)00116-3