Article

Discovery Medicine 2026; 38(205): 553–573
https://doi.org/10.24976/Discov.Med.202638205.46

PDK4 and RGMA as Biomarkers for Gastric Cancer Diagnosis and Prognostication: Transcriptomic and Mendelian Randomization Analyses

Mengwei Tu[1,†], Mengxue Wang[1,†], Zhuo Zhen[1], Xie Liu[1], Jianbo Zhang[1,*]

[1]Department of Gastrointestinal Surgery, The Second Affiliated Hospital of Chongqing Medical University, 400010 Chongqing, China
*Correspondence: zhangjianbo@hospital.cqmu.edu.cn (Jianbo Zhang)
[†]These authors contributed equally.

Submitted: 10 November 2025 Revised: 17 December 2025 Accepted: 30 December 2025 Published: 20 February 2026

Background: Gastric cancer has a poor prognosis and remains a major public health challenge. Therefore, identifying reliable biomarkers is essential to enhance early detection and improve patient survival.

Methods: We integrated gene expression data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) to identify markedly dysregulated genes, and performed Mendelian randomization analyses using MR-Base to evaluate potential causal relationships. Colocalization analysis was subsequently conducted to refine candidate loci. Functional enrichment analyses, including Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA), were performed to explore the underlying biological processes, followed by construction of a regulatory network and development of a prognostic model. We further established and validated a risk model for predicting overall survival. Single-cell RNA sequencing data (GSE163558) were analyzed to characterize gene expression across specific cell types. Finally, immunohistochemistry (IHC) was used to verify protein-level differences between gastric cancer tissues and adjacent normal tissues.

Results: Pyruvate dehydrogenase kinase 4 (*PDK4*) and repulsive guidance molecule A (*RGMA*) were identified as biomarkers. These genes play active roles in key biological processes, including cellular transformation and angiogenesis. By integrating *PDK4* and *RGMA* expression with clinical parameters, we constructed a prognostic model that accurately predicted 3- and 5-year survival outcomes in the TCGA-STAD cohort. Single-cell analysis further revealed cell–type–specific expression patterns of these genes. In addition, immunohistochemical assays demonstrated higher protein levels in tumor tissues compared with adjacent normal tissues.

Conclusions: This comprehensive study suggests that *PDK4* and *RGMA* are promising biomarkers that may facilitate the early detection of gastric cancer and improve prognostic assessment. These findings provide new insights into the disease's pathogenesis and may further inform future clinical decision-making.

Keywords: gastric cancer; *PDK4*; *RGMA*; bioinformatics analysis; biomarkers

Introduction

Gastric cancer is an aggressive malignancy arising from the gastric mucosa. According to the International Agency for Research on Cancer (IARC), it is the fifth most commonly diagnosed cancer and the fifth leading cause of cancer-related mortality worldwide. In 2022, more than 968,000 new cases were reported, and approximately 660,000 deaths were attributed to the disease, with the highest incidence observed in East Asia and Eastern Europe [1]. In China, gastric cancer remains a substantial public health burden and continues to be one of the major causes of cancer-related deaths, with a considerable proportion of patients presenting with advanced-stage disease at diagnosis [2]. Management of advanced gastric cancer requires a multimodal approach, primarily involving surgery, chemotherapy, and radiotherapy [3]. In recent years, advances in targeted therapy and immunotherapy have improved the treatment outcomes by inhibiting tumor progression [4–6]. However, metastatic spread and chemoresistance remain major clinical challenges, resulting in suboptimal therapeutic efficacy for patients with advanced disease [7,8].

Mendelian Randomization (MR) is a robust analytical approach that helps minimize random errors, processing bias, and sample bias. By leveraging genetic variants as instrumental variables, MR effectively mitigates confounding and reverse causation, thereby reducing the impact of unmeasured factors on study outcomes [9]. Moreover, MR facilitates the construction of causal networks linking environmental and genetic risk factors to complex diseases and offers important insights for drug repurposing and therapeutic target identification [10].

Through transcriptomic analysis, we identified several candidate genes that may be closely associated with the development of gastric cancer, as they participate in metabolic reprogramming, immune regulation, and epithelial–mesenchymal transition. In this study, we focused primarily on their intrinsic expression levels. Given their biological relevance, the expression patterns of these genes are likely to exhibit causal relationships with gastric cancer risk. We hypothesized that genetic factors regulating the expression of these genes may influence individual susceptibility to gastric cancer. Therefore, our study objectives were as follows: (1) to identify candidate genes by integrating data from the The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases; (2) to assess causal associations using a two-sample Mendelian randomization approach; (3) to perform colocalization analysis, pathway enrichment analysis, regulatory network construction, and single-cell validation; (4) to characterize the functions of robust causal genes that may hold potential for future diagnostic and prognostic applications.

To investigate the pathways associated with gene expression changes, we performed differential expression and functional enrichment analysis on gastric cancer transcriptomic data from the TCGA database. Using expression quantitative trait loci (eQTL) and genome-wide association studies (GWAS) datasets, the coloc method was applied to identify the key genes of interest, pyruvate dehydrogenase kinase 4 (PDK4), and repulsive guidance molecule A (RGMA). To further elucidate the biological roles of these genes in gastric cancer, we conducted Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA) analysis. We also explored the regulatory mechanisms underlying their expression by constructing transcription factor regulatory networks and miRNA interaction networks. Additionally, by integrating gene expression profiles with clinical variables, we developed a prognostic nomogram based on regression modeling to predict clinical outcomes in gastric cancer. To further validate the significance of these key genes, we analyzed single-cell sequencing data from the GEO database, and identified the cell–type–specific expression patterns of *PDK4* and *RGMA*, as well as their associations with immune and metabolic pathways.

We applied multiple bioinformatics approaches to investigate the key mechanisms involved in the development of gastric cancer, including gene expression patterns, miRNA-mediated regulation and transcription factor activity. The findings were validated through a combination of analytical and experimental methods. We aim to identify potential biomarkers and therapeutic targets for gastric cancer, thereby providing new insights for early detection and prognostic evaluation.

Methods

Data Resources

(1) The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/) is one of the largest public resources for cancer genomics. Its data types range from gene expression, miRNA, lncRNA, copy number variations, DNA methylation and SNPs. In this study, we obtained processed STAD transcriptome data from TCGA, comprising 136 MSI samples and 275 MSS samples.

(2) Gene Expression Omnibus (GEO) is a public repository for high-throughput gene expression data, accessible at https://www.ncbi.nlm.nih.gov/geo/info/datasets.html. It is developed and maintained by the National Center for Biotechnology Information (NCBI) in the United States. In this study, we retrieved the single-cell data dataset GSE163558 from the GEO database, which contains comprehensive expression profiles derived from gastric cancer tissues of three patients.

(3) Source: eQTL data were obtained from the eQTLGen Consortium (https://www.eqtlgen.org/). This collaborative project primarily investigates the genetic regulation of gene expression in blood and aims to identify the genetic contributors to complex diseases. The consortium is currently in its second phase, focusing on large-scale, genome-wide expression analysis in blood samples. For GWAS data, the participants included in this study were predominantly of European ancestry. Summary statistics were retrieved from the MR-Base database, specifically the dataset indexed as ebi-a-GCST90018849. The current GWAS catalog provides published studies, major association results, and complete summary statistics, and its datasets are harmonized with the latest GenomeAssembly and dbSNP Build. In this dataset, the case group comprised 1029 individuals while the control group included 475,087 individuals.

Differential Expression Analysis

We used the Limma package in R to analyze the transcriptomic data and identify the genes exhibiting significant differences in expression between the groups. This tool enables the investigation of disease-related molecular mechanisms and facilitates the detection of genes with distinct expression patterns between healthy and diseased samples. Differentially expressed genes were selected based on a p-value < 0.05 and an absolute log2 fold change >0.585. The results were subsequently visualized using volcano maps and heat maps.

Functional Analysis

To comprehensively investigate the functional relevance of these genes, we performed functional annotation using the Metascape platform (https://www.Metascape.org). For selected genes, Gene Ontology (GO) pathway anal-

yses were also conducted. Enrichment significance was defined by a minimum overlap of three genes and a p-value $\leq$ 0.01.

Mendelian Randomization Analysis

The MR-Base database (https://www.mrbase.org/) provides comprehensive summary statistics from numerous GWAS. By extracting outcome identifiers from the MR-Base database, we obtained the corresponding eQTL associations, all derived from GWAS summary statistics (https://gwas.mrcieu.ac.uk/). We conducted MR analyses to evaluate whether genetically predicted expression levels of differentially expressed genes (DEGs) exert causal effects on gastric cancer risk. Cis-eQTLs from the eQTLGen Consortium were used as instrumental variables, selected based on genome-wide significance ($p < 1 \times 10^{-8}$) and linkage disequilibrium thresholds ($r^2 < 0.001$, 10,000 kb window). Of the 300 identified DEGs, 161 had genome-wide significant and LD-independent cis-eQTLs available in the eQTLGen dataset, enabling their inclusion as valid instruments for MR analysis. GWAS summary statistics for gastric cancer (ebi-a-GCST90018849; 1029 cases and 475,087 controls) were used as the outcome dataset. MR analyses followed the core instrumental-variable assumptions of relevance, independence, and exclusion restriction. Causal effects were estimated using inverse-variance weighted (primary approach), MR-Egger, weighted median, weighted mode, and Wald ratio (for single-SNP instruments) methods, with effect sizes reported as odds ratios. Horizontal pleiotropy and heterogeneity were assessed using the MR-Egger intercept, Cochran's Q statistic, and leave-one-out analyses. Colocalization analyses were performed using the coloc framework to determine whether gene expression and gastric cancer shared a causal variant, with PPH4 >0.9 considered strong evidence of colocalization. All analyses were conducted in R (version 4.2; R Foundation for Statistical Computing, Vienna, Austria).

Although cis-eQTLs derived from whole blood are widely used in MR studies owing to their large sample size and strong statistical power, it should be noted that eQTL effects exhibit varying degrees of tissue specificity. Given that our focus is gastric cancer, we therefore carefully assessed whether the instrumental variables employed in this study satisfied the three core assumptions of Mendelian randomization.

First, to satisfy the relevance assumption, only genome-wide significant cis-eQTLs ($p < 1 \times 10^{-8}$) were included, and linkage disequilibrium clumping ($r^2 < 0.001$, 10 Mb window) was applied to guarantee that the selected variants accounted for sufficient variation in gene expression.

Secondly, with respect to the exclusion restriction assumption, multiple sensitivity analyses—including MR-Egger regression, Cochran's Q test, weighted median and weighted mode estimators, and leave-one-out analysis—revealed no evidence of horizontal pleiotropy or single-variant–driven effects, indicating that the causal estimates were unlikely to be substantially biased by alternative pathways.

For the independence assumption, linkage disequilibrium (LD) clumping helped minimize confounding arising from correlated variants; however, because a comprehensive assessment of all possible environmental or clinical confounders was not feasible, residual confounding cannot be entirely excluded. This limitation is acknowledged in the Discussion section.

Importantly, large-scale cross-tissue studies, such as the Genotype-Tissue Expression (GTEx) project, have demonstrated that a substantial proportion of cis-eQTL effects are shared across multiple tissues, including blood and gastrointestinal tissues [11,12]. Therefore, despite the tissue mismatch, blood-derived cis-eQTLs can reasonably approximate the genetically regulated component of gene expression in gastric tissue, particularly when stomach-specific eQTL datasets remain limited in sample size.

Co-Location Analysis

We performed colocalization analysis using the coloc framework by integrating eQTL summary statistics with gastric cancer GWAS data. This analysis focuses on a 100-kilobase window centered on the index SNP to estimate posterior probability. In the coloc results, H3 represents the probability that two traits (gene expression and gastric cancer) are associated but driven by distinct causal variants, whereas H4 indicates the possibility that both traits share the same causal variant. A threshold of SNP.PP.H4 >0.9 was used as evidence supporting colocalization.

Gene Set Enrichment Analysis

We performed GSEA by stratifying patients into high- and low-expression groups based on the expression levels of the target gene. This approach allowed us to evaluate differences in signaling pathways between the two groups. Gene sets from the Molecular Signatures Database (MsigDB) were used as the reference background of annotated pathways. We identified significantly enriched pathways (adjusted p value < 0.05) across subgroups and ranked them according to their normalized enrichment scores. GSEA is widely used in studies that link disease classification with underlying biological mechanisms.

Gene Set Variation Analysis

GSVA is an unsupervised, non-parametric method that enables the identification of activated gene sets within transcriptomic data [13]. This approach transforms gene-level variation into pathway-level by calculating enrichment scores for each gene set, thereby providing insights into the underlying biological activities within samples. In this study, we obtained gene sets from the MSigDB, ver-

sion 7.0, and computed their enrichment scores using the GSVA algorithm to explore differences in biological functions across samples.

Analysis of Regulatory Network of Key Genes

In this study, R package "RcisTarget" was used to predict transcription factors. All computations performed by RcisTarget are based on motif enrichment analysis. The normalized enrichment score (NES) for each motif was determined using the full set of motifs in the database. In addition to the annotations provided in the original dataset, supplementary annotation files were incorporated by leveraging motif similarity and gene sequence information. To evaluate motif enrichment within a specific gene set, the first step involves calculating the area under the curve (AUC) of each motif-gene set pair. This process calculation is derived from a recovery curve generated by comparing the gene set against the ranked motifs list. The NES of each motif is then obtained from the distribution of AUC values across all motifs within the gene set. For gene motif mapping, the analysis was performed exclusively using the RC ISTarget. HG19. Motifdbs. Cisbponly.500bp database.

Construction of miRNA Network

MicroRNAs (miRNAs) are small noncoding RNA molecules that regulate gene expression by promoting mRNA degradation or inhibiting translation. Therefore, we sought to investigate whether specific miRNAs influence the transcription regulation or degradation of disease-related genes. Key miRNAs associated with these genes were identified using the mircode database, and their interactions were visualized as a gene-miRNA regulatory network using Cytoscape software.

Production of Nomogram

To develop a prognostic model for overall survival (OS), we used the TCGA-STAD cohort as the training dataset. Clinical variables—including age, gender, tumor grade, T stage, N stage, M stage, and the expression levels of *PDK4* and *RGMA*—were considered candidate predictors. Multivariable Cox proportional hazards regression was performed to identify independent prognostic factors, with variables retained in the final nomogram based on their clinical relevance and statistical significance. A prognostic nomogram was constructed using the rms package in R. Model performance was assessed using the concordance index (C-index), which yielded a value of 0.71, indicating moderate discriminatory ability. Internal validation and calibration analyses were subsequently conducted. The baseline survival function from the multivariate Cox model was applied to estimate 3- and 5-year predicted OS probabilities for each patient. Calibration curves were generated by stratifying patients into deciles of predicted risk (approximately 100 patients per group) and comparing the mean predicted survival probabilities with Kaplan–Meier survival estimates. To minimize overfitting and adjust for optimism, 1000 bootstrap resamples (B = 1000) were performed.

Single Cell Analysis

We initially processed the single-cell gene expression data using Seurat toolkit to first filter out lowly expressed genes. The data were then normalized to ensure comparability across cells. Dimensionality reduction was performed using the t-SNE algorithm to visualize the spatial relationships among distinct cell clusters. Finally, cell-type annotation was conducted with reference to the annotation files provided by the Celldex toolkit, enabling the identification of cell populations most relevant to disease development.

Clinical Sample Collection

We collected tissue samples consisting of 20 gastric cancer specimens and 20 adjacent normal tissues from patients undergoing gastrointestinal surgery at the Second Affiliated Hospital of Chongqing Medical University. This study was approved by the hospital's ethics committee, and written informed consent was obtained from all participants. All procedures were conducted in accordance with the principles outlined in the Declaration of Helsinki.

Validation of Protein Expression by Immunohistochemistry

Immunohistochemistry (IHC) was performed to assess protein expression. Formalin-fixed, paraffin-embedded tissue sections were first deparaffinized in xylene and rehydrated through a graded ethanol series. Antigen retrieval was conducted by heating the sections in citrate buffer (pH 6.0) using a pressure cooker. Endogenous peroxidase activity was quenched by incubating the sections with 3% hydrogen peroxide for 15 minutes. The sections were then incubated overnight at 4 °C with primary antibodies against PDK4 (1:800 dilution, Proteintech, Wuhan, China; Cat No.12949-1-AP) and RGMA (1:500 dilution, Proteintech, Wuhan, China; Cat No.12387-1-AP). After washing with PBS, HRP-polymer anti-rabbit secondary antibody (AiFang Bio, Changsha, China; Cat. No. AFIHC003) was applied at room temperature. Signal detection was performed using 3,3′-Diaminobenzozidine (DAB), followed by hematoxylin counterstaining. Finally, the slides were dehydrated, cleared, and coverslipped. Negative controls, in which primary antibodies were omitted, were included in each experiment to ensure staining specificity.

Statistical Analysis

All statistical analyses were performed using R (version 4.2), Graph Pad Prism (version 10.4.1) or SPSS (version 27.0.1). A p-value < 0.05 was considered statistically significant. For immunohistochemistry, PDK4 and RGMA staining in 20 pairs of primary gastric cancer and adjacent

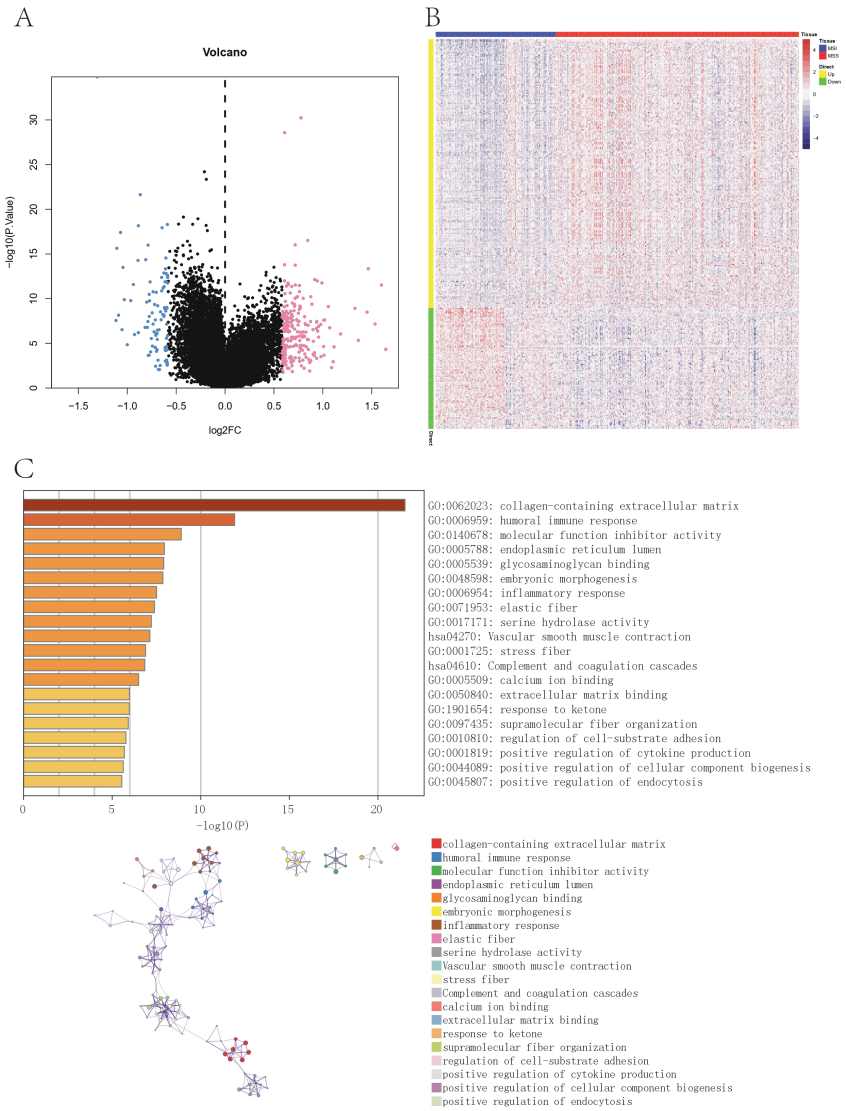


Fig. 1. Gene differential expression analysis and Gene Ontology (GO) pathway analysis. (A) Volcano plot showing differentially expressed genes (DEGs) identified by Limma. Upregulated genes ($log_2FC > 0.585$, adjusted $p < 0.05$) are shown in pink; downregulated genes ($log_2FC < –0.585$, adjusted $p < 0.05$) are shown in blue; nonsignificant genes are shown in black. (B) Heatmap. Each column represents an individual sample, and each row represents a gene. Colors indicate normalized expression levels (red: high expression; blue: low expression), demonstrating clear separation between gastric cancer and adjacent normal tissues. (C) Circle plot and network visualizing the biological processes enriched by GO analyses.

non-tumorous tissues was quantified as the percentage of positively stained area. Paired comparisons between tumor and adjacent tissues were conducted using the Wilcoxon signed-rank test.

Results

Differential Expression Analysis and Functional Analysis

We obtained the TCGA-STAD dataset from the TCGA public database, which included a total of 411 samples, comprising 136 MSI and 275 MSS cases. To identify the genes differentially expressed between the two groups, we applied the Limma package. Differentially expressed genes (DEGs) were defined using a p-value < 0.05 and an absolute log2 fold change ($|logFC|$) >0.585. A total of 300 DEGs were identified, of which 207 were upregulated and 93 were downregulated (Fig. 1A,B). Functional enrichment analyses were then performed on these DEGs. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment results indicated that the DEGs were primarily associated with biological processes such as ex-

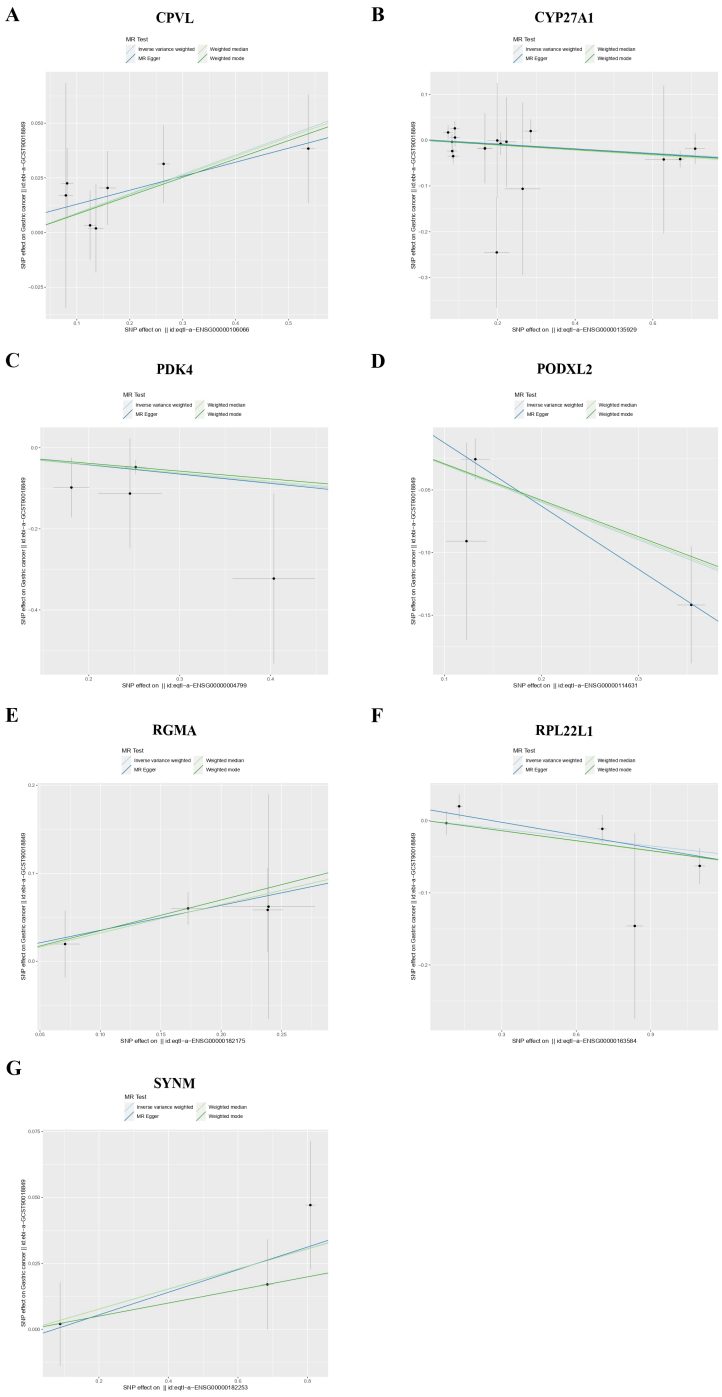


Fig. 2. Mendelian randomization analysis of gene expression and gastric cancer risk. (A–G) Causal relationships corresponding to positive expression quantitative trait loci (eQTL) results for 7 gene pairs (*CPVL*, *CYP27A1*, *PDK4*, *PODXL2*, *RGMA*, *RPL22L1*, and *SYNM*).

tracellular matrix organization, humoral immune response, and molecular function inhibitor activity (Fig. 1C).

MR Analysis and Colocalization Analysis

We performed MR analyses on 161 DEGs using cis-eQTL instruments when available. The inverse-variance weighted (IVW) method served as the primary analytical approach, identifying seven genes significantly associated with gastric cancer risk. *CPVL* (*OR* = 1.09, 95% *CI*: 1.02–1.66, $p = 0.0074$), *RGMA* (*OR* = 1.38, 95% *CI*: 1.16–1.65, $p = 0.0004$) and *SYNM* (*OR* = 1.04, 95% *CI*: 1.07–1.0787, $p = 0.0457$) were associated with increased risk, whereas

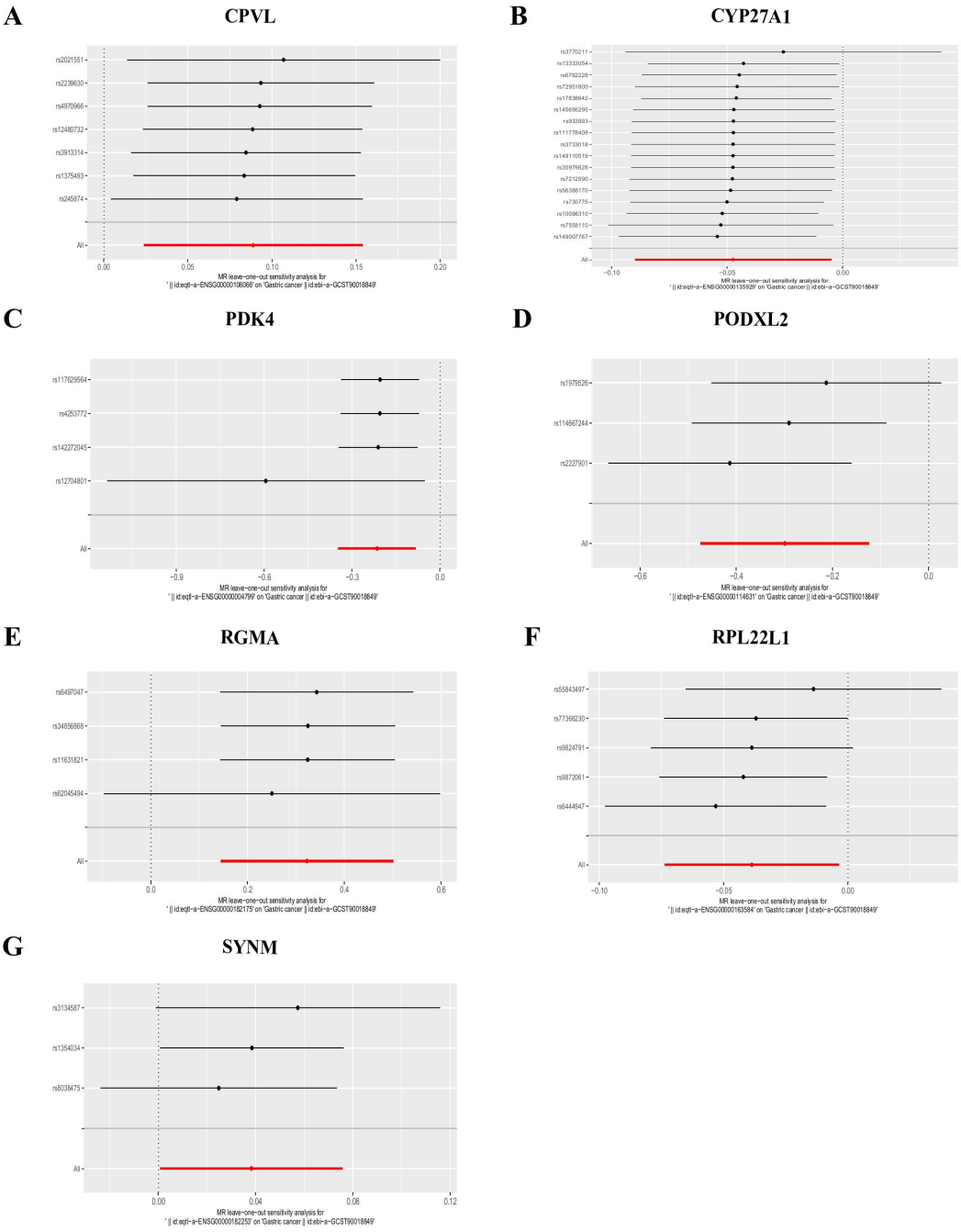


Fig. 3. Sensitivity analysis of 7 pairs of genes for causal associations with positive eQTLs. The results for *CPVL* (A), *CYP27A1* (B), *PDK4* (C), *PODXL2* (D), *RGMA* (E), *RPL22L1* (F), and *SYNM* (G) are shown.

CYP27A1 (*OR* = 0.95, 95% *CI*: 0.91–1.00, *p* = 0.0287), *PDK4* (*OR* = 0.81, 95% *CI*: 0.71–0.92, *p* = 0.0014) and *PODXL2* (*OR* = 0.74, *p* = 0.0008) and *RPL22L1* (*OR* = 0.96, 95% *CI*: 0.93–1.00, *p* = 0.0303) exhibited protective effects (Fig. 2A–G). Sensitivity analyses, including MR-Egger intercept testing, Cochran's Q test and leave-one-out analysis, revealed no evidence of directional pleiotropy or substantial heterogeneity, supporting the robustness of the results (Fig. 3A–G). Furthermore, colocalization analysis provided strong evidence for shared causal variants for *PDK4* (PPH4 = 0.94) and *RGMA* (PPH4 = 0.92), identifying them as highly reliable candidate genes for subsequent functional investigation (Fig. 4A,B).

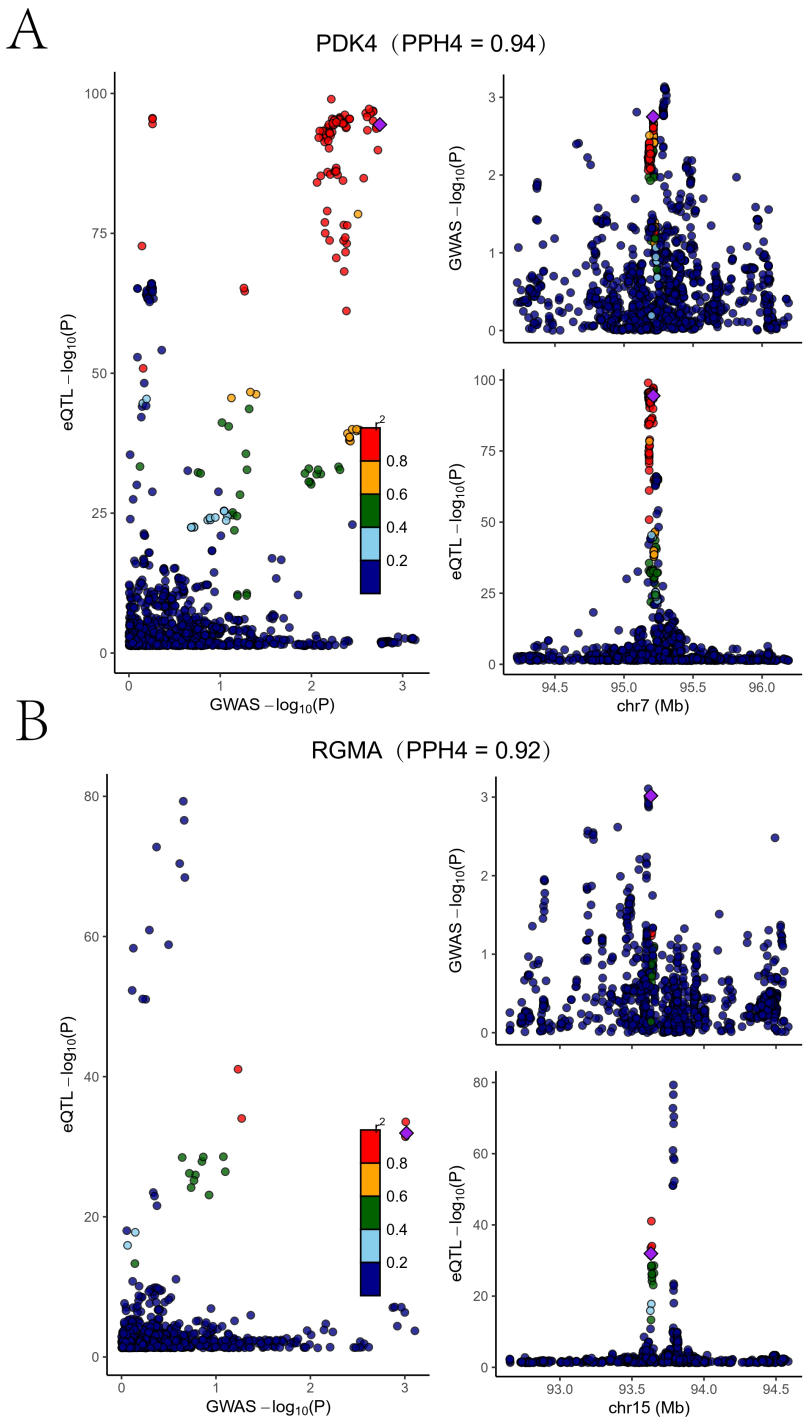


Fig. 4. Colocalization analysis of cis-eQTL signals and gastric cancer (GC) GWAS associations for *PDK4* (A) and *RGMA* (B). Scatterplots show the strength of association between GWAS and eQTL signals. The lead SNP in each region is highlighted with a diamond symbol, representing the variant with the strongest association signal. The posterior probability supporting a shared causal variant (PPH4) was 0.94 for *PDK4* and 0.92 for *RGMA*, indicating strong evidence of colocalization. GWAS, genome-wide association studies; PDK4, pyruvate dehydrogenase kinase 4; RGMA, repulsive guidance molecule A.

GSEA and GSVA

We next investigated the signaling pathways associated with the key genes to explore how they may contribute to disease progression. GSEA results indicated that *PDK4*-related pathways included the glucagon signaling pathway, TGF-beta signaling pathway, C-type selectin receptor signaling pathway (Fig. 5A,B). *RGMA* was

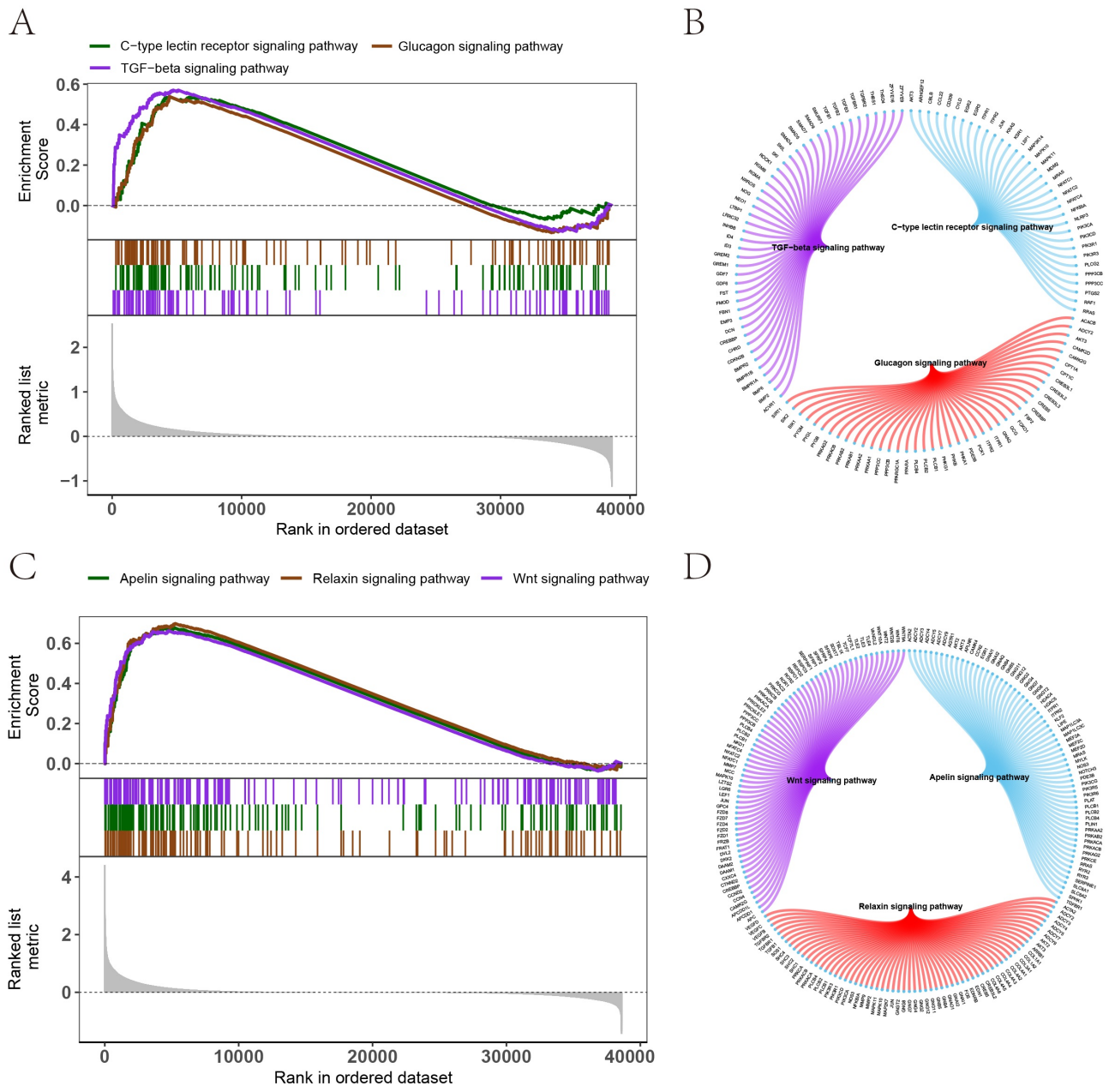


Fig. 5. GSEA results. (A,B) GSEA results showing the enriched pathways of *PDK4*. (C,D) Enrichment pathway of the *RGMA* shown by the GSEA results. GSEA, Gene Set Enrichment Analysis.

enriched in pathways such as the apelin signaling pathway, relaxin signaling pathway and Wnt signaling pathway (Fig. 5C,D). GSVA analysis further revealed that increased *PDK4* expression was primarily associated with pathways, including IL2_STAT5_SIGNALING and HEDGEHOG_SIGNALING (Fig. 6A), whereas elevated *RGMA* expression was associated with HEDGEHOG_SIGNALING and NOTCH_SIGNALING (Fig. 6B).

Key Gene Regulatory Network Analysis and miRNA Network Construction

We examined the regulatory mechanisms of the two key genes and found that they were controlled by several shared transcription factors. To identify these controllers, we applied a cumulative recovery curve-based approach. Through subsequent motif-transcription factor annotation and the analysis of the key gene sets, we identified the motif cisbp_M2092 as the most strongly enriched, with the highest normalized enrichment score (NES: 6.76). This study shows all the motifs found and their related transcription factors (Fig. 7A,B). In addition, using the miRcode

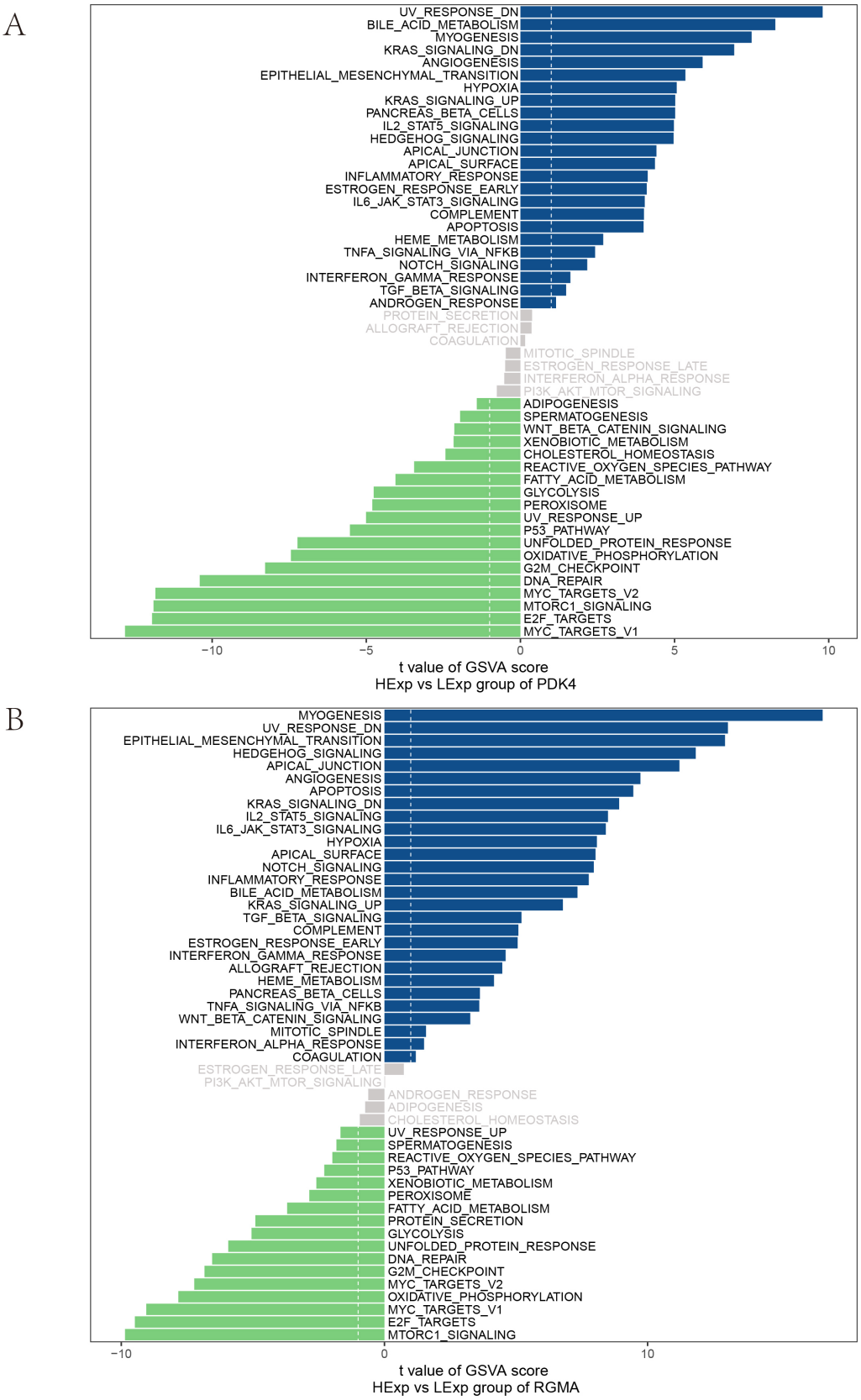
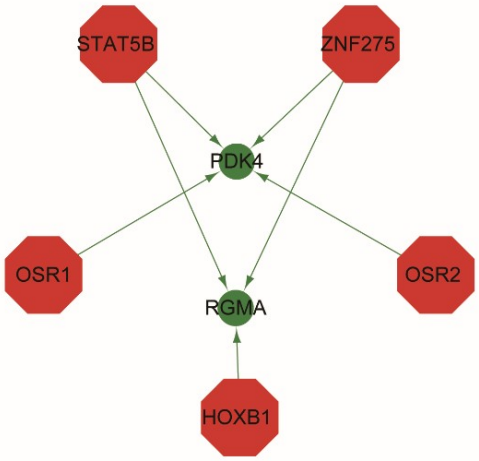


Fig. 6. GSVA results. (A) Signaling pathways enriched with high *PDK4* expression. (B) Signaling pathways enriched with high *RGMA* expression. GSVA, Gene Set Variation Analysis.

A



B

	logo	geneSet	motif	NES	AUC	TF_highConf	nEnrGenes	enrichedGenes
	All	All	All	All	All	All	All	All
1		key_gene	cisbp_M2092	6.76	0.754		2	PDK4;RGMA
2		key_gene	cisbp_M1726	6.61	0.739		2	PDK4;RGMA
3		key_gene	cisbp_M3990	5.69	0.641	STAT5B (directAnnotation).	2	PDK4;RGMA
4		key_gene	cisbp_M4858	5.49	0.62	ZNF275 (inferredBy_Orthology).	2	PDK4;RGMA
5		key_gene	cisbp_M4719	5.27	0.596		2	PDK4;RGMA

C

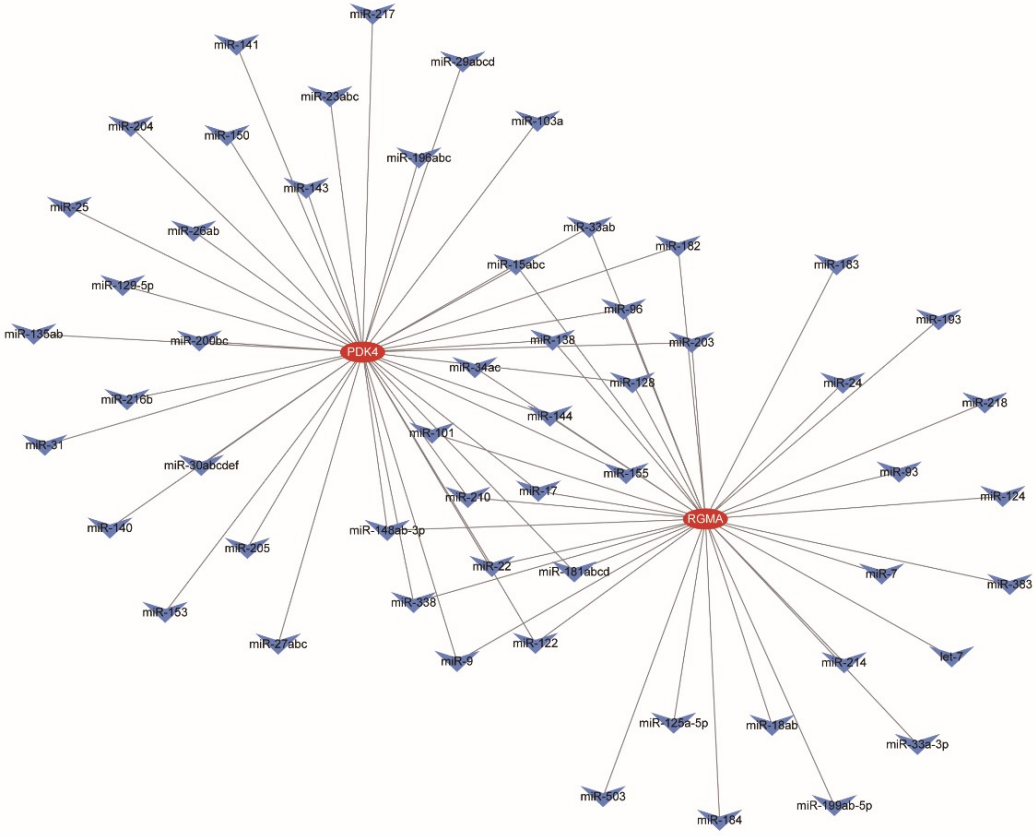


Fig. 7. Construction of a gene regulatory network. (A,B) Enrichment analysis revealing transcription factors and main functional categories related to key genes. (C) A diagram showing the relationship between transcription factors and the genes they control.

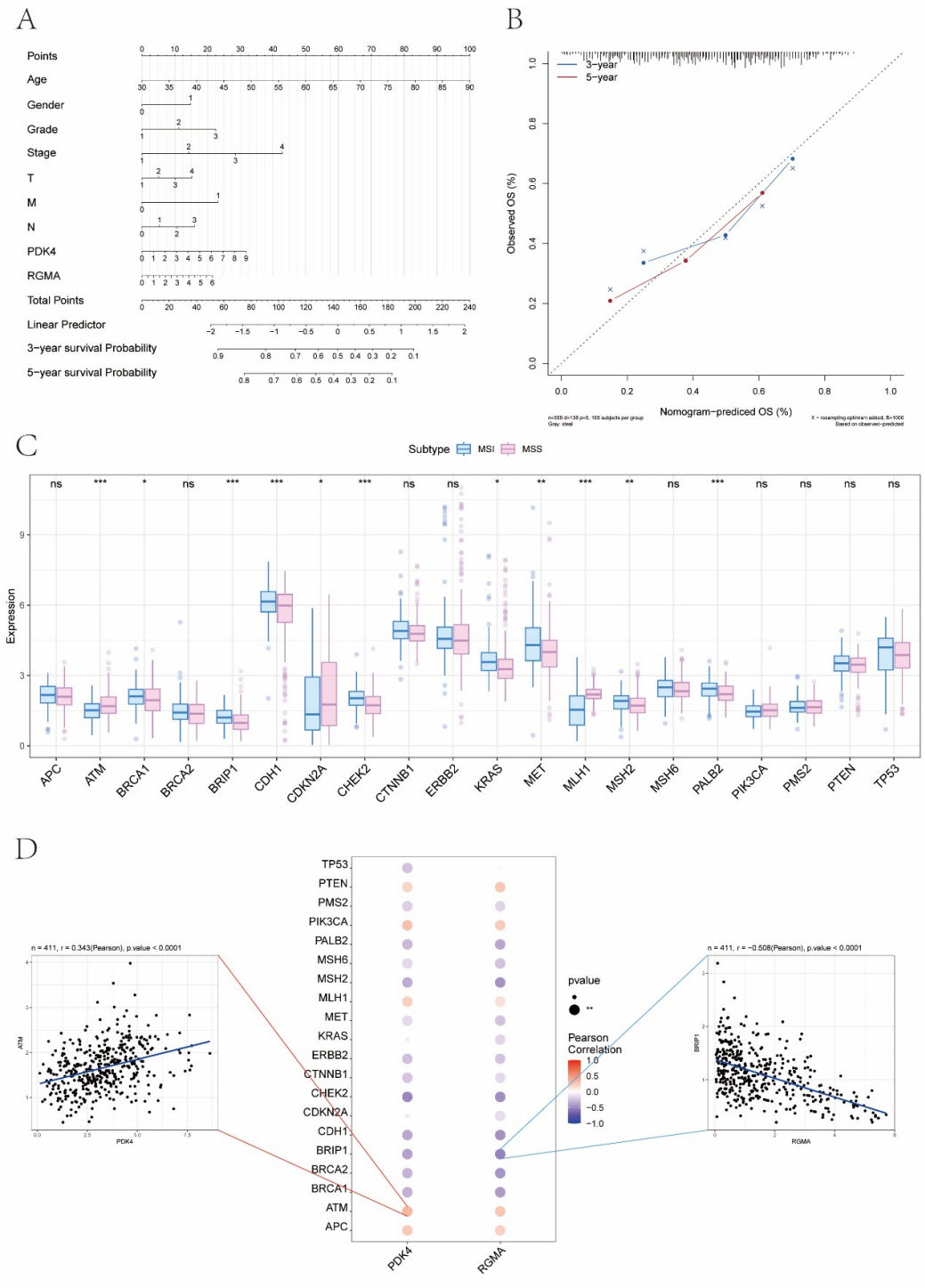


Fig. 8. Construction and validation of a prognostic nomogram and molecular characterization associated with *PDK4* and *RGMA* in gastric cancer. (A) Nomogram combining *PDK4*, *RGMA*, and clinical variables (age, gender, pathological stage) for predicting 3- and 5-year overall survival in gastric cancer patients from the TCGA cohort. (B) Calibration curves of the Cox-based nomogram for predicting 3- and 5-year overall survival. The gray dashed line represents the ideal reference line; solid lines and dots represent the apparent calibration, and crosses (X) indicate the bootstrap-corrected estimates based on 1000 resamples (B = 1000). (C) Boxplots showing the expression of tumor-related genes retrieved from the GeneCards database in MSI (blue) and MSS (pink) gastric cancer samples from the TCGA-STAD cohort. Asterisks denote statistical significance levels between MSI and MSS (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$), whereas 'ns' indicates no significant difference. (D) Analysis of the correlation between *PDK4* and *RGMA* and tumor-related genes. TCGA, The Cancer Genome Atlas.

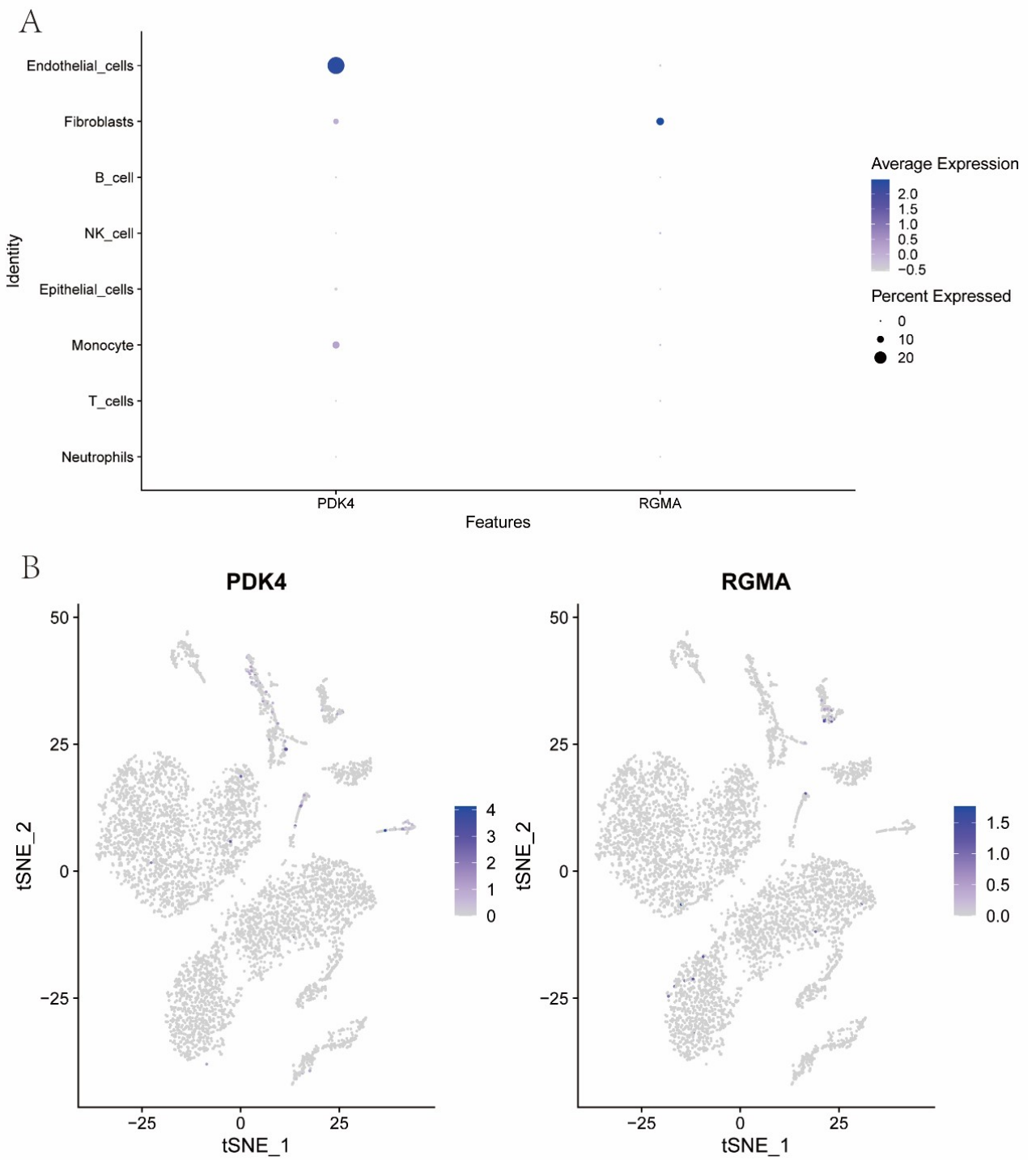


Fig. 9. Single-cell expression patterns of *PDK4* and *RGMA* across major cell types in gastric cancer. (A) The activities of those key genes in neutrophils, T cells, monocytes, epithelial cells, NK cells, B cells, fibroblasts and endothelial cells. (B) t-SNE plot showing the expression patterns of *PDK4* and *RGMA* in eight cell types.

database, we predicted miRNAs targeting the two key genes and identified 56 miRNAs forming 72 mRNA-miRNA interaction pairs. These interactions were visualized using Cytoscape (Fig. 7C).

Modeling of Risk Prediction

We constructed a prognostic nomogram based on multivariable regression analysis to illustrate the relationship between key gene expression levels and clinical outcomes in gastric cancer. The nomogram demonstrated that both clinical variables and gene expression contributed differen-

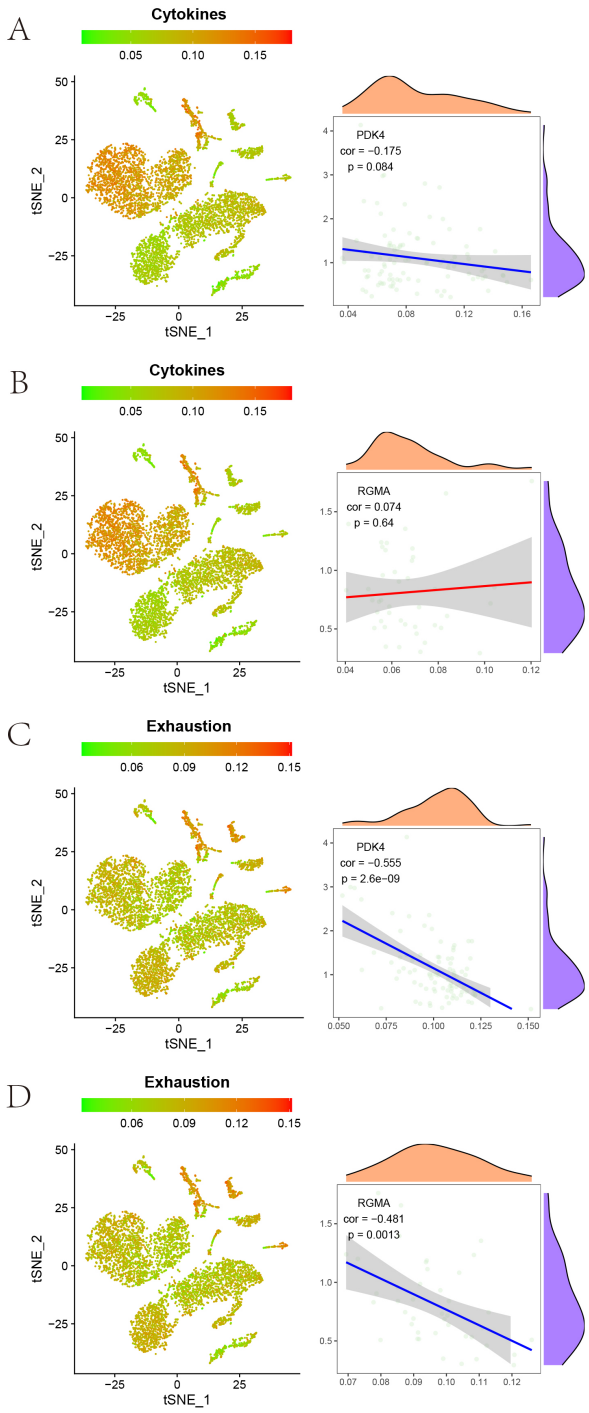


Fig. 10. Associations between *PDK4* and *RGMA* expression and cytokine- and exhaustion-related pathway activities at the single-cell level. (A–D) For each panel, AUCell AUC scores of cytokine- or exhaustion-related gene sets are mapped onto the t-SNE embedding (left), and scatterplots (right), showing the correlation between *PDK4* or *RGMA* expression and the corresponding pathway AUC score. 'cor' denotes the correlation coefficient, and '*p*' denotes the associated *p*-value from the correlation test.

tially to the overall scoring system (Fig. 8A). To assess the predictive performance of the model, calibration analyses were performed for 3- and 5-year OS (Fig. 8B). The calibration curves demonstrated excellent agreement between the nomogram-predicted survival probabilities and the Kaplan–Meier–observed outcomes. The bootstrap-corrected curves (B = 1000) further indicated minimal optimism, supporting the strong stability and reliability of the model's predictions. Overall, the model exhibited strong predictive accuracy, as the estimated survival closely approximated the observed clinical data.

Correlation of Key Genes With Tumor-related Genes

We first identified tumor-related genes using the GeneCards database (https://www.genecards.org/). By comparing the expression levels between MSI and MSS patients in the TCGA-STAD cohort, we observed significant subtype-associated differences across multiple genes (*ATM*, *BRCA1*, *BRIP1*, *CDH1*, *CDKN2A*, *CHEK2*, *KRAS*, *MET*, *MLH1*, *MSH2*, and *PALB2*; all $p < 0.05$), whereas several other tested genes did not display significant differences (Fig. 8C). Given the broad gene panel, we focused on representative genes with clearer biological interpretability and the most direct links to our key genes in the subsequent correlation analysis. Furthermore, correlation analyses between the key genes and tumor regulatory genes revealed closely associated expression patterns. Specifically, *PDK4* exhibited a positive correlation with *ATM* (r = 0.343, $p < 0.0001$), whereas *RGMA* showed a significant negative correlation with *BRIP1* (r = –0.508, $p < 0.0001$) (Fig. 8D).

Single-cell Analysis

We obtained the single-cell dataset GSE163558 from the NCBI GEO database. Using t-SNE and the SingleR package in R (version 4.2; R Foundation for Statistical Computing, Vienna, Austria), cells were classified into neutrophils, T cells, monocytes, epithelial cells, NK cells, B cells, fibroblasts and endothelial cells, and the expression patterns of key genes were examined across these cell types (Fig. 9A,B). In addition, we visualized the expression distributions of tumor regulatory genes (*CDH1*, *MET* and *TP53*) together with the key genes across the eight cell populations (**Supplementary Fig. 1A–F**). Correlation analysis based on AUCell_calcAUC showed that *PDK4* displayed a weak and nonsignificant correlation with the cytokine pathway (r = –0.175, $p = 0.084$), while *RGMA* showed almost no correlation (r = 0.074, $p = 0.64$) (Fig. 10A,B). In contrast, both genes were strongly and negatively correlated with the T-cell exhaustion pathway (*PDK4*: r = –0.555, $p = 2.6 \times 10^{-9}$, *RGMA*: r = –0.481, $p = 0.0013$) (Fig. 10C,D). Developmental trajectories of all cell subtypes were subsequently inferred (Fig. 11A–C). We performed ssGSEA to evaluate pathway activity at the single-cell level. Enrichment scores were first calculated for each individual cell, and then averaged within annotated cell types for downstream comparison. Differential pathway activity between the groups was determined using logFC and FDR based

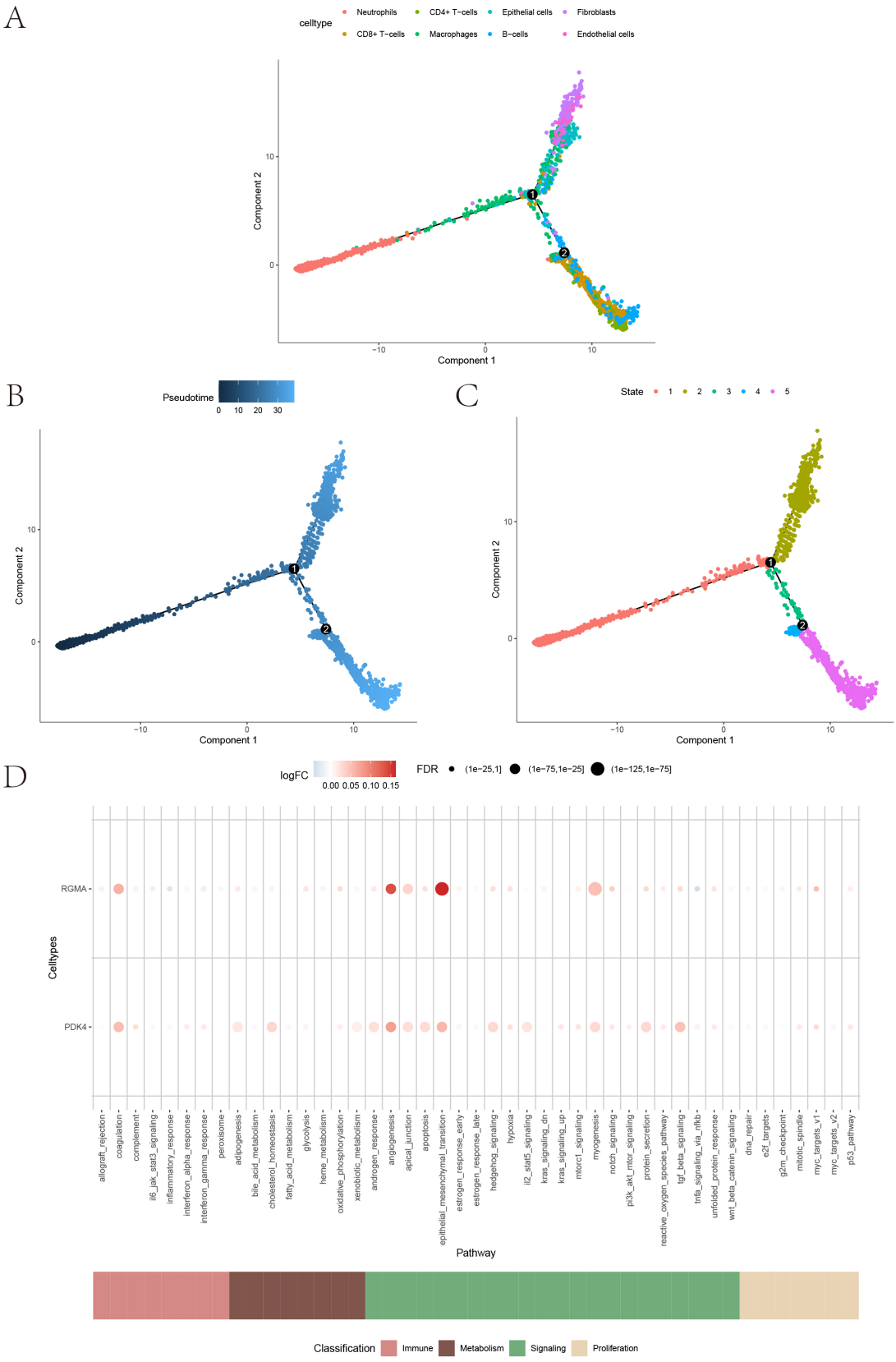


Fig. 11. Single-cell developmental trajectories and pathway activity profiling associated with *PDK4* and *RGMA*. (A–C) Developmental trajectories of all cell subtypes. (D) The ssGSEA method was used to measure the activities of immune and metabolic pathways to explore and explain the relationship between key genes and these pathways.

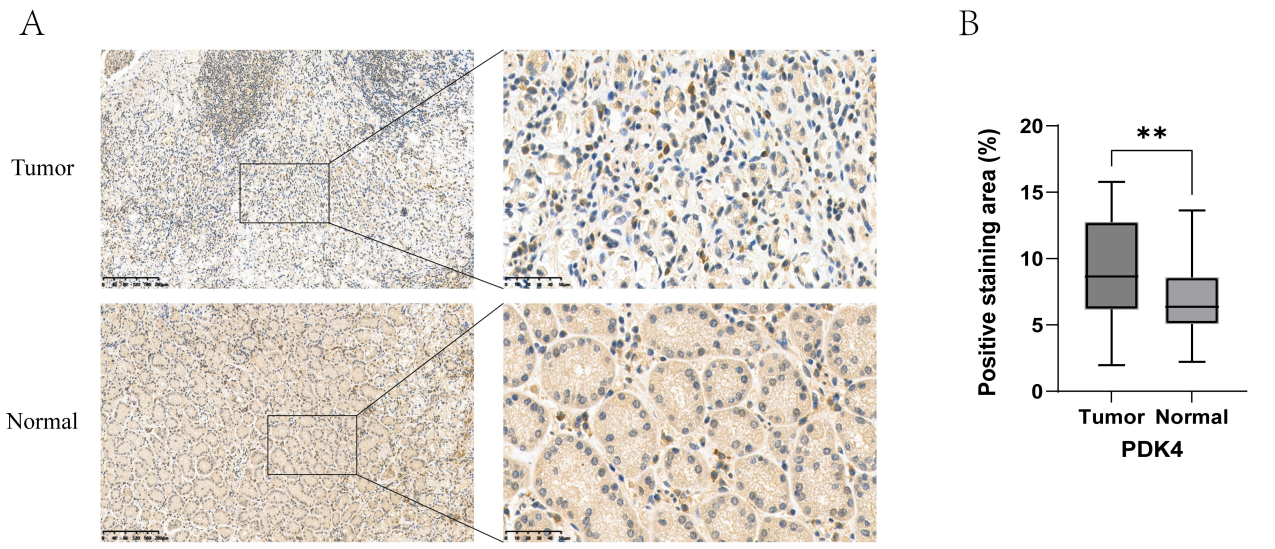


Fig. 12. Immunohistochemical validation of PDK4 protein expression in gastric cancer tissues. (A) Representative immunohistochemical staining of PDK4 in primary gastric cancer tissues and matched adjacent normal tissues (n = 20 pairs). Left panels show low-magnification images (scale bar = 200 µm); boxed regions are shown at higher magnification on the right (scale bar = 50 µm). (B) Semi-quantitative comparison of PDK4-positive staining area (%) between tumor and adjacent tissues. Data are shown as boxplots. Statistical analysis was performed using the Wilcoxon signed-rank test. ** $p < 0.01$.

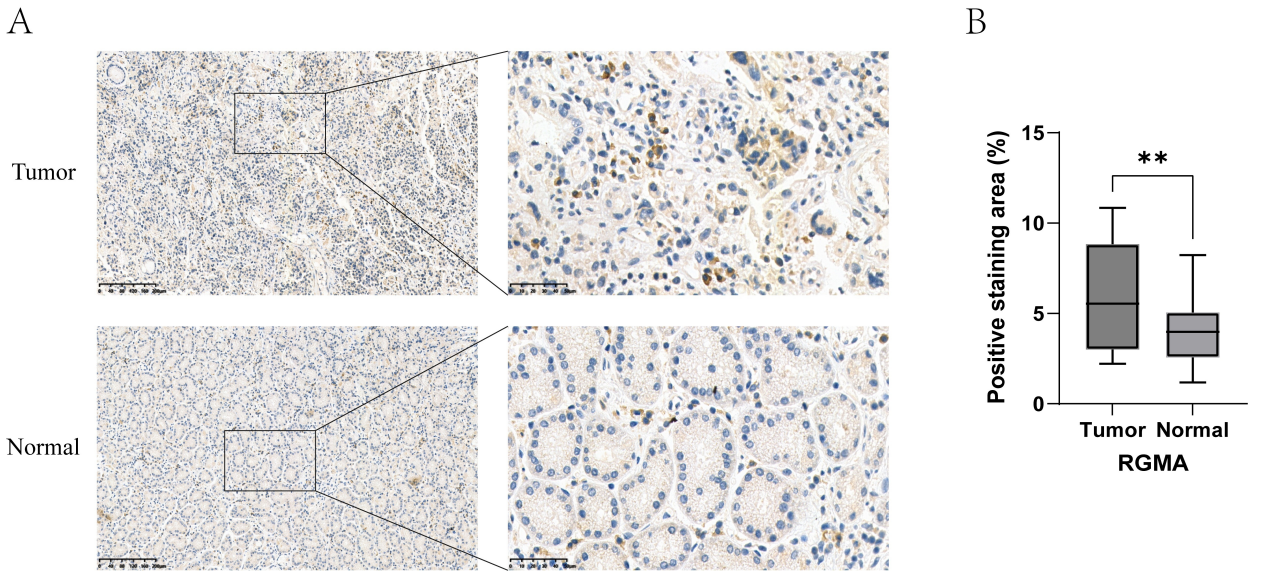


Fig. 13. Immunohistochemical validation of RGMA protein expression in gastric cancer tissues. (A) Representative immunohistochemical staining of RGMA in primary gastric cancer tissues and matched adjacent normal tissues (n = 20 pairs). Left panels show low-magnification images (scale bar = 200 µm); boxed regions are shown at higher magnification on the right (scale bar = 50 µm). (B) Semi-quantitative comparison of RGMA-positive staining area (%) between tumor and adjacent tissues. Data are shown as boxplots. Statistical analysis was performed using the Wilcoxon signed-rank test. ** $p < 0.01$.

on the distribution of enrichment scores across cells. The results were visualized in a dot-plot format where the dot color represents the magnitude of logFC and the dot size corresponds to the statistical significance (FDR) (Fig. 11D). Notably, *PDK4* and *RGMA* were most strongly associated with epithelial-mesenchymal transition (EMT) and angiogenesis pathways.

Immunohistochemistry (IHC)

To validate the transcriptomic findings, we performed immunohistochemical staining on FFPE primary gastric cancer tissues and matched adjacent normal tissues obtained from the Second Affiliated Hospital of Chongqing Medical University. Consistent with the bioinformatics predictions, the expression levels of PDK4 and RGMA

were higher in tumor tissues than in adjacent non-tumorous tissues (Figs. 12,13).

Discussion

Gastric cancer represents a major global health burden, ranking fifth in both incidence and cancer-related mortality worldwide [1]. Standard treatment strategies include gastrectomy with regional lymphadenectomy, systemic chemotherapy, and immunotherapy [14]. However, since most patients are diagnosed at advanced stages, the high metastatic potential frequently precludes curative resection, and the effectiveness of chemotherapy is often limited by drug resistance, resulting in poor clinical outcomes [15]. Therefore, identifying reliable molecular biomarkers for early detection and risk stratification is essential, as they can improve diagnostic accuracy, predict recurrence risk, and guide personalized therapeutic decision-making.

By analyzing the TCGA-STAD dataset from the TCGA public database, our differential expression analysis identified 207 upregulated genes and 93 downregulated genes. Functional enrichment analysis based on GO and KEGG revealed that these genes were involved in several cancer-associated biological processes, including extracellular matrix organization in collagen-rich tissues, humoral immune responses, and molecular function inhibitor activity [16–18].

To identify genes potentially involved in the pathogenesis of gastric cancer, we incorporated large-scale genomic datasets into our transcriptome analysis framework. Mendel's randomization method was performed using cis-eQTLs as instrumental variables, all of which satisfied the core hypotheses of relevance, independence and exclusion restriction. Predicted gene expression levels were compared with the GWAS summary statistics for gastric cancer using the IVW method as the primary analytical approach. Additional sensitivity analyses, including MR-Egger, weighted median, weighted mode, and leave-one-out methods, were conducted to evaluate the robustness of the findings. Consistent results across these methods, along with the absence of directional pleiotropy and minimal heterogeneity, strongly supported the causal relationships suggested by the MR estimates. Moreover, colocalization analysis (PPH4 >0.9 for both *PDK4* and *RGMA*) confirmed that gene expression and disease risk likely share the same causal variants, effectively mitigating concerns regarding confounding due to linkage disequilibrium.

Overall, these findings support *PDK4* and *RGMA* as credible contributors to gastric cancer development and provide a biological basis for investigating them as potential diagnostic biomarkers and therapeutic targets.

PDK4, a pyruvate dehydrogenase kinase, is a well-established metabolic regulator that promotes a shift toward glycolysis metabolism by inhibiting the pyruvate dehydrogenase complex [19]. Previous studies have demonstrated that impaired insulin signaling induces *PDK4* upregulation, thereby facilitating tumorigenic processes [20,21]. Consistent with our findings, *PDK4* dysregulation has been implicated in cancer metabolic reprogramming, altered mitochondrial dynamics, and enhanced aggressiveness in several malignancies—although its role may vary depending on cancer type and context [22,23].

RGMA is a glycophosphoinositol-anchored cell-surface protein originally identified for its essential role in neural development, particularly in guiding axonal growth and extension [24]. Beyond its neural functions, Siebold *et al.* [25] reported that *RGMA* modulates immune responses within the tumor microenvironment, a mechanism that may enable tumor cells to evade immune surveillance and thereby promote sustained growth and metastasis progression. Moreover, accumulating evidence indicates that *RGMA* suppresses tumor progression by inhibiting multiple cellular processes, including proliferation, adhesion, migration, and invasion [26,27].

Our systematic analyses demonstrated that elevated expression of both *PDK4* and *RGMA* in gastric cancer was strongly associated with immune evasion and unfavorable prognosis. At the single-cell level, both genes were expressed in immune cell populations within the tumor microenvironment (TME), suggesting their potential roles in modulating antitumor immune surveillance. *PDK4* may contribute to immune escape by altering immune cell function, whereas *RGMA* appears to promote tumor progression through angiogenesis and suppression of antitumor immunity. Collectively, these results identify *PDK4* and *RGMA* as promising biomarkers and therapeutic targets in gastric cancer. Furthermore, we developed prognostic nomogram integrating *PDK4* and *RGMA* expression with key clinical variables to estimate 3- and 5-year OS. The model exhibited excellent calibration, with predicted survival probabilities closely matching observed outcomes, indicating high reliability. Thus, this prognostic tool represents an efficient and accurate method for assessing long-term outcomes in patients diagnosed with gastric cancer.

MicroRNAs (miRNAs) are single-stranded noncoding RNA molecules that plays essential regulatory roles in eukaryotic gene expression [28–30]. Increasing evidence indicates that microRNAs are critically involved in cancer development by modulating mRNA stability and translation, thereby influencing cellular metabolism, immune evasion, and metastatic progression. For example, miR-5683 suppresses glycolysis and cell proliferation by targeting *PDK4* in gastric cancer [31]. Also, the expression of *PDK4* is associated with hsa-miR-181a-5p, a relationship significantly correlated with clinical prognosis in gastric cancer [32]. Furthermore, circHIPK2 upregulates *RGMA* expression by sponging miR-373-3p, thereby activating the BMP/Smad signaling pathway and ultimately suppressing tumor cell proliferation [33]. Conversely, miR-4472 promotes tumorigenesis by downregulating *RGMA* and induc-

ing the epidermal-mesenchymal transition [34]. Through the construction of miRNA regulatory networks, we further identified several miRNAs—including miR-155 and miR-34a—that regulate *PDK4* and *RGMA* expression at multiple levels, thereby influencing tumor initiation and progression. By modulating these key genes, such miRNAs likely contribute to shaping the immune environment of gastric cancer, facilitating immune escape and promoting disease progression.

Protein-level validation using IHC in our cohort of 20 paired gastric cancer and adjacent noncancerous tissues demonstrated significantly elevated expression of both PDK4 and RGMA in tumor samples, consistent with the transcriptomic findings. Their broad staining distribution within tumor regions further supports their potential utility as pathological markers for clinical risk stratification. Furthermore, our prognostic model incorporating PDK4, RGMA, and key clinical variables accurately predicted 3- and 5-year overall survival, underscoring its translational relevance and potential value in guiding individualized management.

Notably, in this study, the MR analyses indicated that genetically predicted higher *PDK4* expression was associated with a reduced risk of gastric cancer ($OR < 1$). This finding appears inconsistent with previous reports showing that *PDK4* is overexpressed in multiple tumors and promotes malignant progression, as well as with our own IHC observations.

It is important to emphasize that MR does not estimate the immediate or disease-stage-specific expression changes observed in tumor tissues. Instead, MR captures the causal effect of lifelong, genetically regulated differences in gene expression on disease risk. In other words, MR reflects the potential role of *PDK4* before or during the early stages of carcinogenesis, rather than its expression pattern after tumor formation [35]. By design, MR evaluates the influence of germline variants on long-term exposure and is therefore not confounded by environmental factors, inflammation, tumor microenvironment changes, or metabolic reprogramming [9].

In contrast, *PDK4* expression in tumor tissues is strongly influenced by acquired factors, including hypoxia, metabolic stress, inflammatory signaling, and epigenetic alterations [36,37]. Studies on tumor metabolic reprogramming have shown that many metabolic enzymes may exert opposite roles in early versus late stages of cancer development, and their increased expression may represent an adaptive response of tumor cells to hostile microenvironmental conditions [38,39].

Taken together, the apparent discrepancy between MR and IHC findings is biologically plausible. MR reflects the long-term causal effect of germline-regulated *PDK4* expression on cancer susceptibility, whereas IHC reflects post-tumor metabolic adaptation. These differences in biological context and temporal scale suggest that the two results are not contradictory but instead indicate that *PDK4* may exert dynamic and context-dependent roles in the initiation and progression of gastric cancer.

Several limitations should also be acknowledged. First, discrepancies between our study design and publicly available datasets may introduce bias. Additionally, the immunohistochemical validation was performed using a relatively small sample size from a single institution, and the semi-quantitative scoring approach may be influenced by observer-dependent criteria. Finally, mechanistic insights remain inferential, as functional perturbation experiments for *PDK4* and *RGMA* were not performed.

The cis-eQTL datasets used for the MR and colocalization analyses were derived from peripheral blood rather than gastric tissue. Although large cross-tissue analyses (e.g., GTEx) indicate that many cis-eQTL effects are shared across tissues, including blood and gastrointestinal organs, it remains possible that certain regulatory variants exert tissue-specific effects that do not fully translate to the stomach. Consequently, the causal estimates in this study should be interpreted as reflecting the genetically driven component of overall gene expression, rather than the precise causal effect of gastric-tissue–specific expression. In addition, MR can only indirectly evaluate the independence and exclusion restriction assumptions through sensitivity analyses, and residual confounding or undetected horizontal pleiotropy cannot be completely excluded. Therefore, while our MR findings support *PDK4* and *RGMA* as putative causal genes, they should be interpreted with methodological caution. Future studies integrating stomach-specific eQTL resources or larger multi-omics datasets will be required for validating these causal relationships.

This study involves ancestry differences between the datasets analyzed. The GWAS summary statistics used for MR were derived from European populations (1029 cases), whereas the TCGA-STAD cohort represents a mixed Asian and White population. In our design, TCGA-STAD was used only for DEG identification and prognostic model construction and did not contribute to MR instrument or outcome estimation. Thus, the MR estimates mainly reflect causal effects in European populations, where population stratification bias is relatively well controlled. Nevertheless, genetic architecture and expression patterns vary across ancestries, and therefore both the MR-based causal inferences and the TCGA-derived prognostic model may have limited generalizability across populations. Independent validation in non-European cohorts will be essential to confirm the robustness of these findings.

In addition, our Mendelian randomization analyses relied on conventional sensitivity approaches, including MR-Egger regression, Cochran's Q test, and leave-one-out analysis, to assess horizontal pleiotropy and heterogeneity. Although these methods did not indicate substantial directional pleiotropy, MR-Egger is known to have limited sta-

tistical power, especially when the number of instruments per gene is modest. We did not apply more recent robust methods such as MR-PRESSO or MR-cML in the present work. Therefore, residual horizontal pleiotropy cannot be completely ruled out, and the causal estimates should be interpreted with caution. Future studies with larger instrument sets and dedicated robust MR frameworks are needed to further validate these findings.

Additionally, although the GSE163558 dataset contains 42,968 single cells, the number of biological replicates remains limited, as the samples were obtained from only three gastric cancer patients. In single-cell studies, biological variation is primarily determined by the number of patients rather than by the number of sequenced cells. Given the well-recognized heterogeneity of gastric cancer, both at the inter-patient and intra-tumoral levels, a small patient cohort cannot fully capture the diversity of molecular and cellular states across the disease. Therefore, the single-cell analysis in this study is interpreted as exploratory and hypothesis-generating rather than definitive. The results were used to provide cellular-level context for *PDK4* and *RGMA* expression, but were not used to draw population-level inferences. Future studies incorporating larger, multicenter single-cell cohorts will be necessary to validate the cellular patterns observed in this study.

Future studies should externally validate the prognostic model in multicenter, prospective cohorts and conduct functional experiments in both cellular and animal models. Moreover, integrating spatial transcriptomics and proteomics with targeted metabolic and signaling assays—particularly those aimed at elucidating *PDK4*-driven metabolic reprogramming and *RGMA*-mediated microenvironmental signaling—will help clarify the mechanistic basis underlying these associations and assess the applicability of *PDK4* and *RGMA* as therapeutic targets and predictive biomarkers.

Conclusions

Overall, by integrating bioinformatics analyses with Mendelian randomization, we identified *PDK4* and *RGMA* as potential biomarkers for early detection and risk assessment of gastric cancer. These genes influence tumor metabolism, cell proliferation, metastasis and immune response, underscoring their important roles in gastric cancer. In summary, our study provides a novel molecular perspective and establishes a solid foundation for future mechanistic investigations and potential clinical applications.

Availability of Data and Materials

The datasets supporting the conclusions of this article are available in the following repositories: Gene Expression Omnibus (GEO) repository, http://www.ncbi.nlm.nih.gov/geo/; The Cancer Genome Atlas (TCGA) repository, https://portal.gdc.cancer.gov/; eQTLGen repository, https://www.eqtlgen.org; and MR Base repository, ebi-a-GCST90018849, https://www.kimi.com/chat/ebi-a-GCST90018849. Other datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Author Contributions

MXW conceived and designed the study, performed data curation, formal bioinformatics analyses, and visualization. MWT drafted the manuscript and was responsible for specimen data collection, data analysis, and visualization. ZZ and XL contributed to the interpretation of the data, methodological discussion, and critical evaluation of the results. JBZ contributed to the overall study design, supervised the project, provided critical intellectual input, and served as the corresponding author. All authors contributed to the critical revision of the manuscript for important intellectual content. All authors have read and approved the final version of the manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Board of the Second Affiliated Hospital of Chongqing Medical University (Approval No.281; Date of approval: 31 December 2024). Written informed consent was obtained from all participants prior to inclusion in the study.

Acknowledgment

Not applicable.

Funding

This work was supported by the funding from Nan'an District Science and Health Joint Medical Research Program: Preliminary exploration of *PDK4* combined with *RGMA* as a biomarker for early diagnosis and prognosis of gastric cancer (ID=2024-02); and Chongqing Municipal Science and Health Joint Medical Research Program: Screening novel VEGFR2 small-molecule inhibitors and research on their antiangiogenic effects and anticancer mechanisms in gastric cancer (ID=cstc2017jcyjAX0224).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.24976/Discov.Med.202638205.46.

References

[1] Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, *et al*. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2024; 74: 229–263. https://doi.org/10.3322/caac.21834.

[2] Qi J, Li M, Wang L, Hu Y, Liu W, Long Z, *et al*. National and subnational trends in cancer burden in China, 2005-20: an analysis of national mortality surveillance data. The Lancet. Public Health. 2023; 8: e943–e955. https://doi.org/10.1016/S2468-2667(23)00211-6.

[3] Ajani JA, D'Amico TA, Bentrem DJ, Chao J, Cooke D, Corvera C, *et al*. Gastric Cancer, Version 2.2022, NCCN Clinical Practice Guidelines in Oncology. Journal of the National Comprehensive Cancer Network: JNCCN. 2022; 20: 167–192. https://doi.org/10.6004/jnccn.2022.0008.

[4] Janjigian YY, Kawazoe A, Bai Y, Xu J, Lonardi S, Metges JP, *et al*. Pembrolizumab plus trastuzumab and chemotherapy for HER2-positive gastric or gastro-oesophageal junction adenocarcinoma: interim analyses from the phase 3 KEYNOTE-811 randomised placebo-controlled trial. Lancet. 2023; 402: 2197–2208. https://doi.org/10.1016/S0140-6736(23)02033-0.

[5] Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. Lancet. 2020; 396: 635–648. https://doi.org/10.1016/S0140-6736(20)31288-5.

[6] Tomar S, Siddiqui S, Pathak R, Srivastava V. Unlocking the Potential of Immunomodulators as Synergistic Immune-Based Therapies in Cancer. Discovery Medicine. 2025; 37: 411–432. https://doi.org/10.24976/Discov.Med.202537194.35.

[7] Baccili Cury Megid T, Farooq AR, Wang X, Elimova E. Gastric Cancer: Molecular Mechanisms, Novel Targets, and Immunotherapies: From Bench to Clinical Therapeutics. Cancers. 2023; 15: 5075. https://doi.org/10.3390/cancers15205075.

[8] Liu J, Yuan Q, Guo H, Guan H, Hong Z, Shang D. Deciphering drug resistance in gastric cancer: Potential mechanisms and future perspectives. Biomedicine & Pharmacotherapy. 2024; 173: 116310. https://doi.org/10.1016/j.biopha.2024.116310.

[9] Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, *et al*. Mendelian randomization. Nature Reviews. Methods Primers. 2022; 2: 6. https://doi.org/10.1038/s43586-021-00092-5.

[10] Chen LG, Tubbs JD, Liu Z, Thach TQ, Sham PC. Mendelian randomization: causal inference leveraging genetic data. Psychological Medicine. 2024; 54: 1461–1474. https://doi.org/10.1017/S0033291724000321.

[11] GTEx Consortium. Genetic effects on gene expression across human tissues. Nature. 2017; 550: 204–213. https://doi.org/10.1038/nature24277.

[12] GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020; 369: 1318–1330. https://doi.org/10.1126/science.aaz1776.

[13] Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013; 14: 7. https://doi.org/10.1186/1471-2105-14-7.

[14] Guan WL, He Y, Xu RH. Gastric cancer treatment: recent progress and future perspectives. Journal of Hematology & Oncology. 2023; 16: 57. https://doi.org/10.1186/s13045-023-01451-3.

[15] Kim SE, Park MI, Lee MH. Chemotherapy for Metastatic Gastric Cancer. The Korean Journal of Gastroenterology. 2025; 85: 1–10. https://doi.org/10.4166/kjg.2024.139.

[16] Dzobo K, Dandara C. The Extracellular Matrix: Its Composition, Function, Remodeling, and Role in Tumorigenesis. Biomimetics. 2023; 8: 146. https://doi.org/10.3390/biomimetics8020146.

[17] Tan R, Nie M, Long W. The role of B cells in cancer development. Frontiers in Oncology. 2022; 12: 958756. https://doi.org/10.3389/fonc.2022.958756.

[18] Chong X, Peng R, Sun Y, Zhang L, Zhang Z. Identification of Key Genes in Gastric Cancer by Bioinformatics Analysis. BioMed Research International. 2020; 2020: 7658230. https://doi.org/10.1155/2020/7658230.

[19] Dou X, Fu Q, Long Q, Liu S, Zou Y, Fu D, *et al*. PDK4-dependent hypercatabolism and lactate production of senescent cells promotes cancer malignancy. Nature Metabolism. 2023; 5: 1887–1910. https://doi.org/10.1038/s42255-023-00912-w.

[20] Jeong JY, Jeoung NH, Park KG, Lee IK. Transcriptional regulation of pyruvate dehydrogenase kinase. Diabetes & Metabolism Journal. 2012; 36: 328–335. https://doi.org/10.4093/dmj.2012.36.5.328.

[21] Teaney NA, Cyr NE. FoxO1 as a tissue-specific therapeutic target for type 2 diabetes. Frontiers in Endocrinology. 2023; 14: 1286838. https://doi.org/10.3389/fendo.2023.1286838.

[22] Thoudam T, Chanda D, Sinam IS, Kim BG, Kim MJ, Oh CJ, *et al*. Noncanonical PDK4 action alters mitochondrial dynamics to affect the cellular respiratory status. Proceedings of the National Academy of Sciences of the United States of America. 2022; 119: e2120157119. https://doi.org/10.1073/pnas.2120157119.

[23] Zhang Z, Han S, Ouyang S, Zeng Z, Liu Z, Sun J, *et al*. PDK4 Constitutes a Novel Prognostic Biomarker and Therapeutic Target in Gastric Cancer. Diagnostics. 2022; 12: 1101. https://doi.org/10.3390/diagnostics12051101.

[24] Yamamoto M, Itokazu T, Uno H, Maki T, Shibuya N, Yamashita T. Anti-RGMa neutralizing antibody ameliorates vascular cognitive impairment in mice. Neurotherapeutics. 2025; 22: e00500. https://doi.org/10.1016/j.neurot.2024.e00500.

[25] Siebold C, Yamashita T, Monnier PP, Mueller BK, Pasterkamp RJ. RGMs: Structural Insights, Molecular Regulation, and Downstream Signaling. Trends in Cell Biology. 2017; 27: 365–378. https://doi.org/10.1016/j.tcb.2016.11.009.

[26] Li Y, Liu HT, Chen X, Wang YW, Tian YR, Ma RR, *et al*. Aberrant promoter hypermethylation inhibits RGMA expression and contributes to tumor progression in breast cancer. Oncogene. 2022; 41: 361–371. https://doi.org/10.1038/s41388-021-02083-y.

[27] Zhao ZW, Lian WJ, Chen GQ, Zhou HY, Wang GM, Cao X, *et al*. Decreased expression of repulsive guidance molecule member A by DNA methylation in colorectal cancer is related to tumor progression. Oncology Reports. 2012; 27: 1653–1659. https://doi.org/10.3892/or.2012.1693.

[28] Wang Y, Yin Z. Prediction of miRNA-disease association based on multisource inductive matrix completion. Scientific Reports. 2024; 14: 27503. https://doi.org/10.1038/s41598-024-78212-w.

[29] Letelier P, Saldías R, Loren P, Riquelme I, Guzmán N. MicroRNAs as Potential Biomarkers of Environmental Exposure to Polycyclic Aromatic Hydrocarbons and Their Link with Inflammation and Lung Cancer. International Journal of Molecular Sciences. 2023; 24: 16984. https://doi.org/10.3390/ijms242316984.

[30] Ruan K, Fang X, Ouyang G. MicroRNAs: novel regulators in the hallmarks of human cancer. Cancer Letters. 2009; 285: 116–126. https://doi.org/10.1016/j.canlet.2009.04.031.

[31] Miao Y, Li Q, Sun G, Wang L, Zhang D, Xu H, *et al*. MiR-5683

suppresses glycolysis and proliferation through targeting pyruvate dehydrogenase kinase 4 in gastric cancer. Cancer Medicine. 2020; 9: 7231–7243. https://doi.org/10.1002/cam4.3344.

[32] Yang Y, Zhang J. Ascites-derived hsa-miR-181a-5p serves as a prognostic marker for gastric cancer-associated malignant ascites. BMC Genomics. 2024; 25: 628. https://doi.org/10.1186/s12864-024-10359-2.

[33] Lun J, Zhang Y, Yu M, Zhai W, Zhu L, Liu H, *et al*. Circular RNA circHIPK2 inhibits colon cancer cells through miR-373-3p/RGMA axis. Cancer Letters. 2024; 593: 216957. https://doi.org/10.1016/j.canlet.2024.216957.

[34] Li Y, Wang YW, Chen X, Ma RR, Guo XY, Liu HT, *et al*. MicroRNA-4472 Promotes Tumor Proliferation and Aggressiveness in Breast Cancer by Targeting RGMA and Inducing EMT. Clinical Breast Cancer. 2020; 20: e113–e126. https://doi.org/10.1016/j.clbc.2019.08.010.

[35] Daghlas I, Gill D. Mendelian randomization as a tool to inform drug development using human genetics. Cambridge Prisms. Precision Medicine. 2023; 1: e16. https://doi.org/10.1017/pcm.2023.5.

[36] Wu D, Liang J. Activating transcription factor 4: a regulator of stress response in human cancers. Frontiers in Cell and Developmental Biology. 2024; 12: 1370012. https://doi.org/10.3389/fcell.2024.1370012.

[37] Koizume S, Miyagi Y. Adaptation mechanisms in cancer: Lipid metabolism under hypoxia and nutrient deprivation as a target for novel therapeutic strategies (Review). Molecular Medicine Reports. 2025; 31: 83. https://doi.org/10.3892/mmr.2025.13448.

[38] Mao Y, Xia Z, Xia W, Jiang P. Metabolic reprogramming, sensing, and cancer therapy. Cell Reports. 2024; 43: 115064. https://doi.org/10.1016/j.celrep.2024.115064.

[39] Chuang YM, Tzeng SF, Ho PC, Tsai CH. Immunosurveillance encounters cancer metabolism. EMBO Reports. 2024; 25: 471–488. https://doi.org/10.1038/s44319-023-00038-w.