Clin. Lab. 2026;72:XXX-XXX

ORIGINAL ARTICLE

Diagnostic and Prognostic Significance of Apelin Early Ligand A (APELA) in Colorectal Cancer: Insights from Multi-Omics Analysis

Zhiqiang Zhu

Capital Institute of Pediatrics, Beijing, China

SUMMARY

Background: **Colorectal cancer (CRC) is a major global health concern and is often diagnosed at advanced stages with poor clinical outcomes. Apelin early ligand A (APELA) has been identified as a potential contributor to tumor progression, but its role in colon adenocarcinoma (COAD) remains unclear.**
Methods: **The present study integrated transcriptomic data obtained from the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) to evaluate the APELA expression in COAD. Then, the expression profile, diagnostic/prognostic value, co-expression networks, and functional pathways of APELA were evaluated by *in silico* analysis approaches, such as R and STRING. Afterwards, the associated biological processes were explored by Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Gene Set Enrichment Analysis (GSEA).**
Results: **Compared to normal controls, the APELA expression was significantly upregulated in CRC tissues. However, there was no significant difference between the normal mucosa and adenomas. Receiver operating characteristic (ROC) analysis indicated the potential of the APELA expression as a diagnostic marker. The high APELA expression correlated to the poor prognosis in COAD patients. The functional network analysis identified an association of APELA to PI3K signaling regulation and revealed that its co-expressed genes were downregulated in NK CD56 bright cells. GO and KEGG enrichment highlighted the roles of APELA expression in angiogenesis and cell migration. GSEA revealed that the APELA expression was positively associated with respiratory electron transport and protein synthesis pathways but negatively associated with extracellular matrix (ECM) remodeling.**
Conclusions: **APELA influences CRC progression through multiple pathways (ECM remodeling, metabolic adaptation, and immune modulation), suggesting its potential as a promising diagnostic and prognostic biomarker and a potential therapeutic target in CRC.**
(Clin. Lab. 2026;72:xx-xx. DOI: 10.7754/Clin.Lab.2025.250677)

Correspondence:
Zhiqiang Zhu
Capital Institute of Pediatrics
Room 327, 3rd Floor, Research Building
2 Yabao Road, Chaoyang District
Beijing, 100020
China
Phone: + 86 01085695544
Email: zzhq8888@126.com

Manuscript accepted July 9, 2025

KEYWORDS

APELA, CRC, prognosis, complex molecular networks, DeepSeek

LIST OF ABBREVIATIONS

APELA - Apelin Early Ligand A
Apelin - Apelin signaling pathway
ADP-P2Y12 - ADP-P2Y12 signaling
AUC - Area under the curve
AHD - Adult heart development
aDCs - Activated dendritic cells
BMI - Body mass index

CRC - Colorectal cancer
CEA - Carcinoembryonic antigen
COAD - Colon adenocarcinoma
CMIA - Cell migration in angiogenesis
CCT - Collagen chain trimerization
DEGs - Differentially expressed genes
DCs - Dendritic cells
ECM - Extracellular matrix
ECM_PG - ECM proteoglycans
EIF - Eukaryotic translation initiation
ETE - Eukaryotic translation elongation
EFF - Elastic fiber formation
FDR - False discovery rate
GEO - Gene Expression Omnibus
Gαβγ - G-protein complex
GO - Gene Ontology
GSEA - Gene Set Enrichment Analysis
NK - Natural killer
NPV - Negative predictive value
Naba_Col - Naba collagens
NIMECMI - Non-integrin membrane ECM interactions
PPV - Positive predictive value
pDC - Plasmacytoid dendritic cells
P-R-PI3K - Positive regulation of PI3K signaling
P/V-FGFR2 - PDGF/VEGF-FGFR2 signaling
ROC - Receiver operating characteristic
RET - Respiratory electron transport
Treg - Regulatory T cells
TCM - Trophoblast cell migration
TCGA - The Cancer Genome Atlas
iDCs - Immature dendritic cells
KEGG - Kyoto Encyclopedia of Genes and Genomes

INTRODUCTION

Colorectal cancer (CRC) is a major global health concern, and presently ranks third in incidence and second in cancer-related deaths worldwide [1]. Its epidemiology significantly varies across regions, with incidence rates stabilizing in high-income countries, while continuing to rise in low- and middle-income regions, and in younger individuals under the age of 50 [2]. CRC risk might be enhanced by multiple factors, such as genetic predisposition [3] and lifestyle influences [4], such as metabolic disorders, environmental exposure, and high-fat diet, which may increase the risk by 2- to 10-fold [5].

Early-stage CRC is often asymptomatic or presents with vague symptoms, such as changes in bowel habits and abdominal discomfort [6], leading to delayed diagnoses, and a high proportion of advanced-stage cases. It has been reported that the 5-year survival rate for stage IV CRC is merely 14% [7]. Therefore, it is essential to improve CRC early detection strategies and identify more effective therapeutic targets [8]. Although present treatment approaches that incorporate neoadjuvant chemoradiotherapy, targeted therapies, and immunotherapy have improved outcomes [9], therapeutic resistance and tumor heterogeneity continue to limit long-term success. Apelin early ligand A (APELA), which is also known as Elabela or Toddler, is a peptide ligand of the APJ receptor, and was originally recognized for its essential role in cardiovascular development during embryogenesis [10]. Although APELA exhibits dual roles in cancer biology and can act as a tumor suppressor in prostate cancer [11], it can promote tumorigenesis in gliomas [12]. Furthermore, its role in CRC has not yet been elucidated.

The present study analyzed multi-omics datasets by integrating advanced bioinformatics tools, including DeepSeek. The data used for the analysis were retrieved from the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO), which contain information on the expression patterns, clinical relevance, and functional implications of APELA in COAD. The present study aimed to define the diagnostic and prognostic potential of APELA and explore its mechanistic contributions to CRC progression, with focus on pathways involved in angiogenesis, metabolic reprogramming, and immune modulation.

MATERIALS AND METHODS

Preparation of colon cancer data

The RNA sequencing data derived from the TCGA-COAD project (https://portal.gdc.cancer.gov) were processed for multi-omics analysis using STAR-aligned pipelines and TPM-normalized expression matrices. The dataset included 521 tumor samples, 41 adjacent normal tissues, and clinical metadata for 461 cases (with 24 paired tumor-normal samples).

In order to validate the findings, three GEO datasets (https://www. ncbi.nlm.nih.gov/geo/) were used for the analysis under identical protocols: GSE39582, including 584 tumor samples, 19 adjacent normal tissues, and clinical metadata for 565 cases; GSE198692, including 10 normal mucosa and 10 adenoma samples; GSE37 364, including 38 normal controls, 29 adenomas, and 27 tumors (a summary of these datasets are presented in Table 1).

Differential expression analysis

Volcano plots

The patients were divided into APELA high- and low-expression groups based on their median expression thresholds. Then, the APELA-associated differentially expressed genes (DEGs) were identified and visualized by volcano plots using ggplot2 (v3.4.4), with significance thresholds set at $|\log^2 (\text{fold change})| > 2$ and adjusted $p < 0.05$.

Paired sample analysis

The difference in APELA expression between tumor-normal paired samples (TCGA-COAD) was evaluated using Wilcoxon signed-rank test (R packages stats v4.2.1 and car v3.1-0). Then, the expression values

were presented in log^2-transformed (log^2 [value + 1]) in the absence of preprocessing filters.

Disease-state comparisons
The disparities in APELA expression profiles of the TCGA-COAD cohorts were evaluated by Wilcoxon rank-sum test (R packages stats v4.2.1 and car v3.1-0), with log^2-transformed values and no preprocessing.

Estimation of the adenoma-carcinoma progression
The estimation was performed by comparing the data of the following GSE datasets: Student's *t*-test was performed to compare normal mucosa (n = 10) and adenomas (n = 10) in GSE198692, and Kruskal-Wallis test was performed to compare the expression differences across the normal (n = 38), adenoma (n = 29), and tumor (n = 27) subgroups in GSE37364.

Assessment of diagnostic utility
Receiver operating characteristic (ROC) curve analysis was performed to evaluate the diagnostic performance of APELA in COAD using the R package pROC (v1.17.0.1). The key metrics included the area under the curve (AUC), optimal cutoff value, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and Youden's index. Diagnostic accuracy was interpreted as follows: AUC 0.50 - 0.70, low; AUC 0.70 - 0.90, moderate; AUC ≥ 0.90, high.

Survival analysis and prognostic value
Cox proportional hazards models were applied to assess the prognostic relevance of APELA in both the TCGA-COAD and GSE39582 datasets. The survival evaluation was performed using the survival (v3.3.1) R package, with preprocessing steps that included the following: filtering, exclusion of normal samples and those that lack clinical annotations; transformation: expression data were normalized using log^2 (value + 1); prognostic endpoints were based on the curated survival data obtained from a previously published study [13].

Clinicopathological association analysis
Nonparametric statistical tests were performed to compare the APELA expression across clinical and pathological variables (T, N and M stage, gender, race, age, body mass index [BMI], and carcinoembryonic antigen [CEA] level) in the TCGA-COAD cohort. Then, the data was log^2-transformed, and samples with missing values were excluded from the analysis.

Single-gene correlation analysis
The genome-wide Pearson correlation coefficients were calculated to identify genes that co-expressed with APELA in TCGA-COAD. The p-values were adjusted using the Benjamini-Hochberg adjusted method. The preprocessing procedures were consistent with those described above.

Protein-protein interaction (PPI) network construction
APELA-centered PPI networks were generated using the STRING database (https://string-db.org/). Initial functional clustering was performed using K-means clustering (k = 3), followed by a refined analysis with k = 5 to prioritize hub genes among the top 50 interactors.

Functional enrichment analysis [14,15]
Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were conducted using the clusterProfiler (v4.4.4) R package, with Benjamini-Hochberg adjusted $p < 0.05$ as the significance threshold. Gene annotation mapping was performed using org.Hs.eg.db (v3.16.0) for Ensembl-to-Entrez ID conversion.
Gene Set Enrichment Analysis (GSEA) was conducted to identify pathways differentially enriched between the high- and low-APELA expression groups using the MSigDB c2.cp gene set with a false discovery rate (FDR) threshold of $q < 0.25$.

Immune infiltration analysis
Single-sample GSEA (ssGSEA) was conducted using the GSVA (v1.46.0) R package to estimate the relative abundance of 24 immune cell types, based on validated immune gene signatures [16,17]. The preprocessing steps included log^2 transformation and the exclusion of duplicate and normal samples.
The key immune populations included the following: activated dendritic cells (aDCs), B cells, CD8 T cells, cytotoxic cells, dendritic cells (DCs), eosinophils, immature dendritic cells (iDCs), macrophages, mast cells, neutrophils, natural killer (NK) cells (CD56 bright and CD56 dim), plasmacytoid dendritic cells (pDC), total T cells, helper T cells, central memory (Tcm), effector memory (Tem), follicular helper (TFH), gamma delta (Tgd), Th1, Th2, Th17, and regulatory T cells (TReg). The correlations between the APELA expression and immune infiltration levels were visualized using heatmaps.

Comprehensive statistical analysis
Nonparametric comparisons were conducted using the Wilcoxon rank-sum test for independent samples and Wilcoxon signed-rank test for paired comparisons. The correlations were assessed using Spearman's rank correlation coefficient (ρ).
For the survival modeling, a two-stage Cox regression approach was applied: the univariate analysis identified the candidate variables, followed by multivariate modeling, in order to adjust for potential confounders and confirm the independent prognostic factors. Then, Kaplan-Meier survival curves were constructed, and intergroup differences were assessed using log-rank test.
All statistical analyses were performed using R (v4.2.1), with significance thresholds defined as follows:

Table 1. Summary of the colorectal cancer transcriptomic datasets derived from TCGA and GEO.

Dataset	Normal	Adenoma	Tumor	Clinical metadata
TCGA-COAD	41	-	521	461
GSE39582	19	-	585	566
GSE198692	10	10	-	-
GSE37364	38	29	27	-

TCGA Cancer Genome Atlas, GEO Gene Expression Omnibus, COAD colon adenocarcinoma.

Table 2. ROC curve analysis summary.

Variable	Cutoff	Sen.	Spe.	Acc.	PPV	NPV	YI
COAD	0.160	0.758	0.976	0.775	0.997	0.256	0.734
GEO39582	2.845	0.696	0.684	0.696	0.985	0.070	0.380

ROC receiver operating characteristic, cutoff optimal cutoff value, Sen. Sensitivity, Spe. Specificity, Acc. Accuracy, PPV positive predictive value, NPV negative predictive value, YI Youden's index, COAD colon adenocarcinoma.

Table 3. Statistical description.

Dataset	Group	n	Events	Censored	Censoring rate (%)	Median OS	95% CI
TCGA-COAD (median)	low	239	65	174	0.728	2,475 days	1,518-?
	high	238	38	200	0.840	?	1,910-?
TCGA- COAD (optimal)	low	189	58	131	69.30%	2,134 days	1,348-?
	high	288	45	243	84.40%	?	2,003-?
GSE39582	low	144	57	87	60.40%	97 months	76-?
	high	418	134	284	67.90%	145 months	106-?

OS overall survival, CI confidence interval, TCGA-COAD The Cancer Genome Atlas-Colon Adenocarcinoma, ? not calculable.

$p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), and $p < 0.0001$ (****).

RESULTS

Analysis of APELA differential expression

Through the cross-validated analysis of independent CRC cohorts (TCGA-COAD and GEO-GSE39582), the results, which included the volcano plot (Figure 1A), Wilcoxon signed-rank test (Figure 1B), Wilcoxon rank-sum test (Figure 1C), and Mann-Whitney *U* test (Figure 1D), revealed the consistent overexpression of APELA in malignant specimens in both datasets, when compared to the histologically normal counterparts ($p < 0.01$).

Comparison of APELA expression between colonic benign and malignant tumors

As presented in Figure 2, the comparative analysis revealed a statistically robust differential APELA expression between the normal (n = 38) and tumor (n = 27) cohorts ($p = 4.81 \times 10^{-7}$). Furthermore, the comparison results revealed the higher expression of APELA in colon cancer (n = 38), when compared to colonic benign tumors (adenoma, n = 29) ($p = 6.4 \times 10^{-5}$, Kruskal-Wallis test).

ROC curve analysis

The diagnostic utility of the APELA expression in distinguishing CRC from normal tissue was assessed using ROC curve analysis (Figure 3). The results revealed the potential of APELA as a diagnostic biomarker for CRC, indicating its translational relevance for clinical bio-

Table 4. Baseline characteristics of COAD patients.

Characteristics	Low expression of APELA	High expression of APELA	p-value
n	239	239	
Pathologic T stage, n (%)			
T1	6 (1.30%)	5 (1.00%)	0.814
T2	38 (8.00%)	45 (9.40%)	
T3	162 (34.00%)	161 (33.80%)	
T4	32 (6.70%)	28 (5.90%)	
Pathologic N stage, n (%)			
N0	141 (29.50%)	143 (29.90%)	0.163
N1	48 (10.00%)	60 (12.60%)	
N2	50 (10.50%)	36 (7.50%)	
Pathologic M stage, n (%)			
M0	175 (42.20%)	174 (41.90%)	0.983
M1	33 (8.00%)	33 (8.00%)	
Pathologic stage, n (%)			
Stage I	38 (8.10%)	43 (9.20%)	0.900
Stage II	97 (20.80%)	90 (19.30%)	
Stage III	68 (14.60%)	65 (13.90%)	
Stage IV	33 (7.10%)	33 (7.10%)	
Gender, n (%)			
Female	119 (24.90%)	107 (22.40%)	0.272
Male	120 (25.10%)	132 (27.60%)	
Age, n (%)			
≤ 65	88 (18.40%)	106 (22.20%)	0.094
> 65	151 (31.60%)	133 (27.80%)	
Race, n (%)			
Asian	4 (1.30%)	7 (2.30%)	0.419
White	113 (36.90%)	119 (38.90%)	
Black or African American	35 (11.40%)	28 (9.20%)	
BMI, n (%)			
≤ 25	39 (15.20%)	48 (18.80%)	0.516
> 25	83 (32.40%)	86 (33.60%)	
CEA level, n (%)			
≤ 5	99 (32.70%)	97 (32.00%)	0.529
> 5	50 (16.50%)	57 (18.80%)	

COAD colon adenocarcinoma, APELA Apelin early ligand A, BMI body mass index, CEA carcinoembryonic antigen.

marker panels. The detailed diagnostic parameters, including the sensitivity, specificity, accuracy, and Youden index, are presented in Table 2.

Kaplan-Meier curve

The survival trajectory analysis depicted in Figure 4 demonstrates the significantly prolonged survival duration in patients stratified within the high-APELA expression cohort (log-rank $p < 0.05$). The comprehensive statistical metrics are presented in Table 3.

Baseline characteristics of COAD patients

A comprehensive cohort of 454 cases was stratified according to the following clinicopathological parameters: pathologic T stage, pathologic N stage, pathologic M stage, pathologic stage, gender, race, age, BMI, and

Table 5. The k-means clustering.

Cluster	ID gene	Count description
Cluster 1	9	ADP signalling through P2Y purinoceptor 12 Apelin signaling pathway Heterotrimeric G-protein complex
Cluster 2	1	ARRB2
Cluster 3	1	SLN

Systems-level interactome mapping was initially performed using APELA and its top 50 co-expressed partners, followed by K-means clustering (k = 5), in order to elucidate the modular organization. The resultant network contained 49 nodes and 28 edges with significant interconnectivity based on the protein-protein interaction (PPI) enrichment analysis ($p = 1.69 \times 10^{-11}$, hypergeometric test). In Figure 6, these interactions are thermodynamically visualized, while the cluster-specific metrics are codified (Table 5).

Table 6. The k-means clustering.

Cluster	ID gene	Count description
Cluster 1	17	Positive regulation of phosphatidylinositol 3-kinase signaling
		Mixed, incl. FGFR2 ligand binding and activation, and PDGF/VEGF domain
Cluster 2	3	Corneal dystrophy
Cluster 3	2	LSAMP, NEGR1
Cluster 4	2	Fatty acid elongation
Cluster 5	2	EDIL3, GPC6

Table 7. GO/KEGG pathway analysis.

ID	Description	p-value	p.adjust
GO:0001659	Temperature homeostasis	5.88e-08	0.0001
GO:0043410	Positive regulation of MAPK cascade	2.15e-07	0.0002
GO:0070374	Positive regulation of ERK1 and ERK2 cascade	3.13e-07	0.0002
GO:0048568	Embryonic organ development	8.64e-07	0.0004
GO:0072012	Glomerulus vasculature development	1.26e-06	0.0004
GO:0062023	Collagen-containing extracellular matrix	3.64e-05	0.0045
GO:0031225	Anchored component of membrane	0.0016	0.0936
GO:0005788	Endoplasmic reticulum lumen	0.0022	0.0936
GO:0005161	Platelet-derived growth factor receptor binding	8.97e-06	0.0014
GO:0008143	Poly(A) binding	4.28e-05	0.0027
GO:0008266	Poly(U) RNA binding	5.54e-05	0.0027
GO:0008187	Poly-pyrimidine tract binding	7.84e-05	0.0027
GO:0070717	Poly-purine tract binding	8.73e-05	0.0027
hsa04540	Gap junction	6.37e-07	5.41e-05
hsa04020	Calcium signaling pathway	1.98e-05	0.0008
hsa05218	Melanoma	0.0001	0.0036
hsa01521	EGFR tyrosine kinase inhibitor resistance	0.0002	0.0039
hsa05215	Prostate cancer	0.0004	0.0069

The Gene Ontology (GO) analysis comprised of three domains: molecular function (MF), cellular component (CC), and biological process (BP). The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed the significant enrichment of APELA-associated co-expressed mRNAs in 12 signaling pathways. Among these, the Apelin signaling pathway and neuroactive ligand-receptor interaction were the most pronounced enrichment pathways. These pathways include: Molecular function (MF): neuropeptide receptor binding; hormone activity; G protein-coupled receptor binding; signaling receptor activator activity; receptor ligand activity; biological process (BP): trophoblast cell migration; regulation of trophoblast cell migration; the cell migration involved in angiogenesis; G protein-coupled receptor signaling pathway; adult heart development.

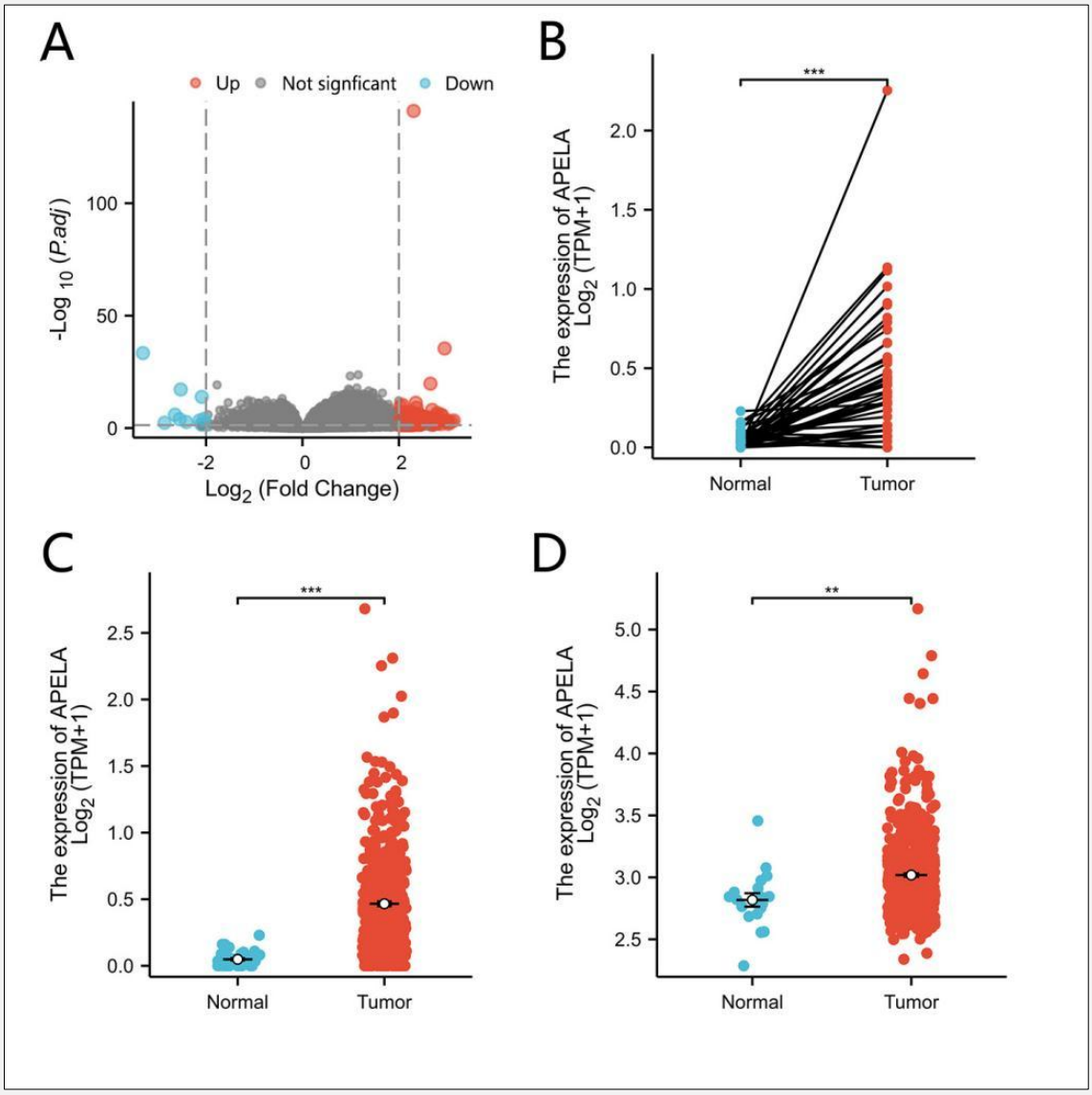

Figure 1. APELA is significantly overexpressed in colorectal cancer across multiple cohorts.

A) The volcano plot analysis of APELA expression-stratified colorectal cancer cohorts (|log2FC| > 2, FDR < 0.05) delineates 198 differentially expressed genes (DEGs). B) The paired analysis of 41 tumor-adjacent normal tissue specimens reveals the significant APELA overexpression in its malignant counterparts (p < 0.001, Wilcoxon signed-rank test). C) The comparative analysis of the TCGA-COAD dataset reveals the significantly higher APELA expression levels in tumor specimens (n = 480), when compared to adjacent normal tissues (n = 41) (Wilcoxon rank-sum test, p < 0.001). D) The evaluation of the GEO-GSE39582 cohort reveals the significant APELA overexpression in tumor specimens (n = 565), when compared to non-malignant controls (n = 19) (Mann-Whitney U-test, p < 0.01).

CEA level. The comparative analysis of these parameters between the low- and high-APELA expression groups revealed no statistically significant disparities in the clinically defined subgroups. The detailed demographic and clinicopathological characteristics are presented in Table 4.

Single-gene correlation analysis

Based on the multi-threshold correlation criteria ($|r| > 0.30$, 0.50, or 0.70) combined with concomitant significance ($p < 0.05$), 1,981 genes that had co-expression relationships with APELA were identified. Among them, there are 1,964 with $0.5 > |r| > 0.3$ and 17 with $0.7 > |r| > 0.5$.

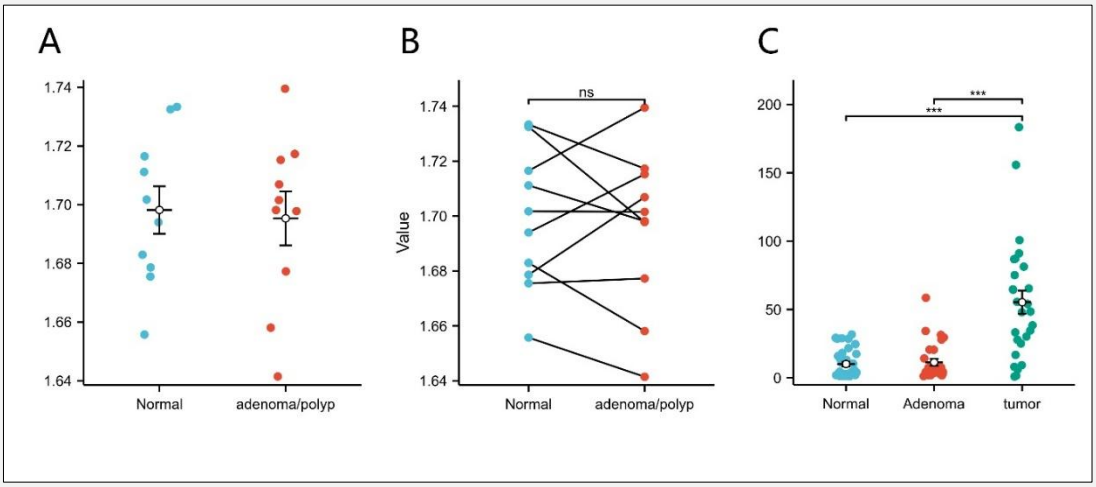

Figure 2. The APELA overexpression is specific to malignant transformation and is not observed in adenomatous colorectal tissues.

A) An elevated APELA expression was observed in tumor specimens, from both the TCGA-COAD ($p < 0.001$) and GEO-GSE39582 ($p = 0.008$) cohorts. In contrast, the comparison between normal (n = 10) and adenomatous (n = 10) tissues in the GEO-GSE198692 dataset revealed no statistically significant expression disparity ($p = 0.8168$, Student's *t*-test). B) The comparative analysis of 10 matched normal-adenoma tissue pairs in the GEO-GSE198692 cohort revealed the biological equivalence in APELA expression levels ($p = 0.6815$, Wilcoxon signed-rank test). C) The evaluation of normal (n = 38) and adenomatous (n = 29) specimens in the GEO-GSE37364 cohort uncovers the statistically indistinguishable APELA expression profiles (exact $p = 1.0000$, Fisher's permutation test). There were significant statistical differences between the normal group [16] and tumor group [10] ($p = 4.81 \times 10 - 7$), as well as between the adenoma group [39] and tumor group [16] ($p = 6.4 \times 10 - 5$).

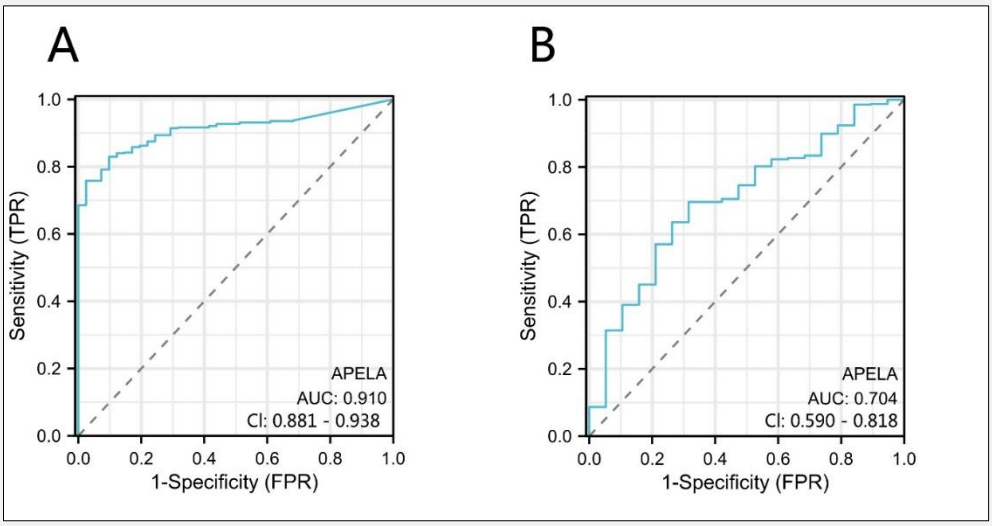

Figure 3. Diagnostic utility of the APELA expression in distinguishing colorectal cancer from normal tissues.

A) APELA expression level was used as a prognostic indicator within the TCGA-COAD cohort (41 normal vs. 480 tumor specimens), and the receiver operating characteristic (ROC) analysis indicated the superior discriminative capacity of the APELA expression (AUC = 0.910, 95% CI: 0.881 - 0.938) for clinical outcome prediction. B) The similar ROC analysis of the GEO-GSE39582 cohort (normal [n = 19] vs. tumor [n = 565]) revealed moderate discriminative capacity (AUC = 0.704, 95% CI: 0.590 - 0.818) for APELA in clinical outcome prognostication.

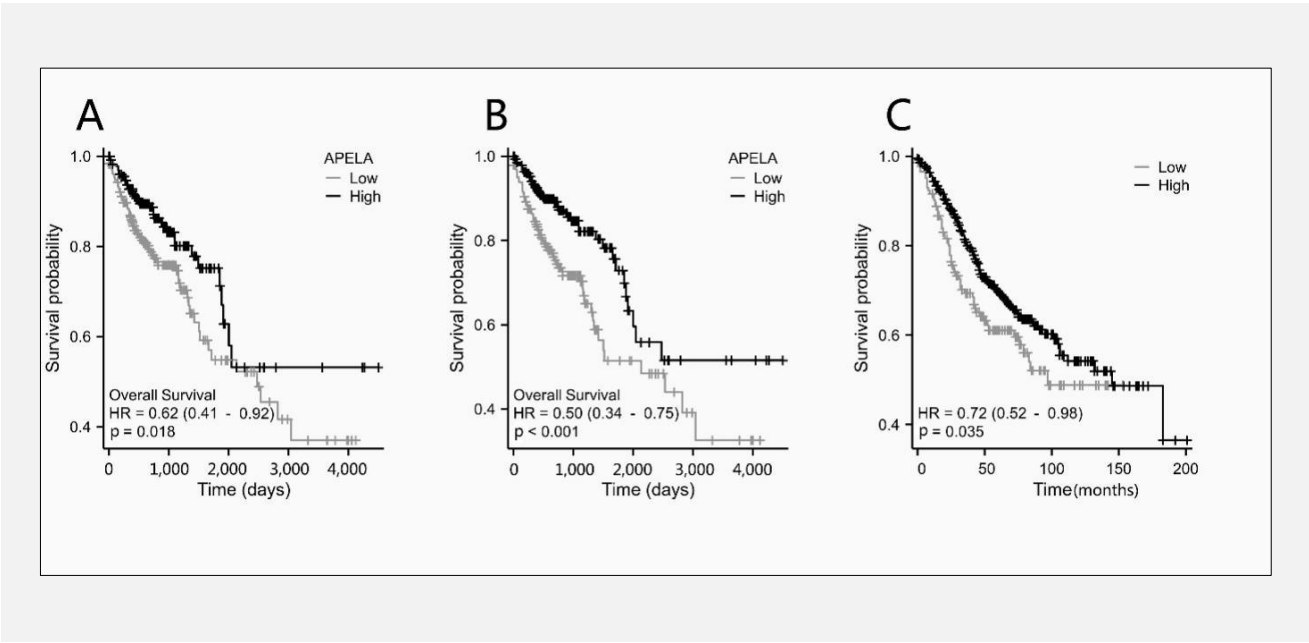

Figure 4. High APELA expression is associated with favorable overall survival in the colorectal cancer cohorts.

A) Kaplan-Meier curve analysis of the TCGA-COAD cohort stratified by median APELA expression (low = 239, high = 238) revealed prolonged survival outcomes in the high-expression cohort (log-rank $p < 0.05$; HR = 0.62, 0.41 - 0.92). B) The optimal prognostic stratification of the TCGA-COAD cohort through statistically-driven dichotomization (low = 189, high = 288) revealed a robust survival discrimination (log-rank $p < 0.001$), HR = 0.50 (0.34 - 0.75). C) Both the discovery (n = 281) and validation (n = 281) cohorts (total n = 562), using the GEO-GSE39582 dataset for analysis, resulted in a Kaplan-Meier survival curve, and this was implemented through optimal prognostic stratification (low = 144, high = 418; log-rank $p < 0.05$; HR = 0.72, 0.52 - 0.98).

PPI network construction

A systematic PPI network was used to identify the APELA-associated interactors, which included the following: APLN, APLNR, ARRB2, GNAI1, GNAI2, GNAI3, GNB2, GNB4, GNB1, and SLN. The PPI network comprised of 11 vertices and 36 edges, demonstrating significant interconnectivity (PPI enrichment $p = 1.32 \times 10^{-7}$, Fisher's exact test). Figure 5 delineates the complex interaction topology through force-directed graph visualization. The unsupervised clustering analysis delineated three molecular subgroups. The cluster-specific characteristics are comprehensively tabulated in Table 5. Figure 6 presents the expanded APELA interactome and functional clustering of associated proteins, with k-means clustering (Table 6).

Gene set enrichment analysis for APELA-associated genes

Enrichment analysis for APELA-associated genes was performed, with focus on molecules that exhibited absolute correlation coefficients ($|Cor| > 0.30$ with APELA. The statistically significant thresholds for the GO/KEGG analysis were set at adjusted p-value < 0.05 (Benjamini-Hochberg correction). The comprehensive pathway enrichment metrics are presented in Table 7.

Gene Set Enrichment Analysis

GSEA was performed to identify the enriched pathways between the high- and low-APELA expression groups. Six related pathways with the highest negative correlation with APELA were identified. These were divided into two categories: protein translation regulation (eukaryotic translation elongation, and eukaryotic translation initiation, ribosome), and mitochondrial energy metabolism (respiratory electron transport, mitochondrial OXPHOS electron transport chain, ATP synthesis via chemiosmotic coupling, and thermogenesis through uncoupling proteins). In addition, six pathways with a strong positive correlation with APELA were identified. These were divided into three categories: collagen synthesis and assembly (collagen chain trimerization and NABA collagen complexes), non-collagen matrix assembly and signal integration (extracellular matrix [ECM] proteoglycan networks, elastic fiber assembly, and core matrisome components), and matrix-cell interface signaling interactions (non-integrin membrane-ECM crosstalk). Figure 6 presents the ridge and mountain plot, which depicts the top negatively- and positively-correlated pathways associated with APELA expression.

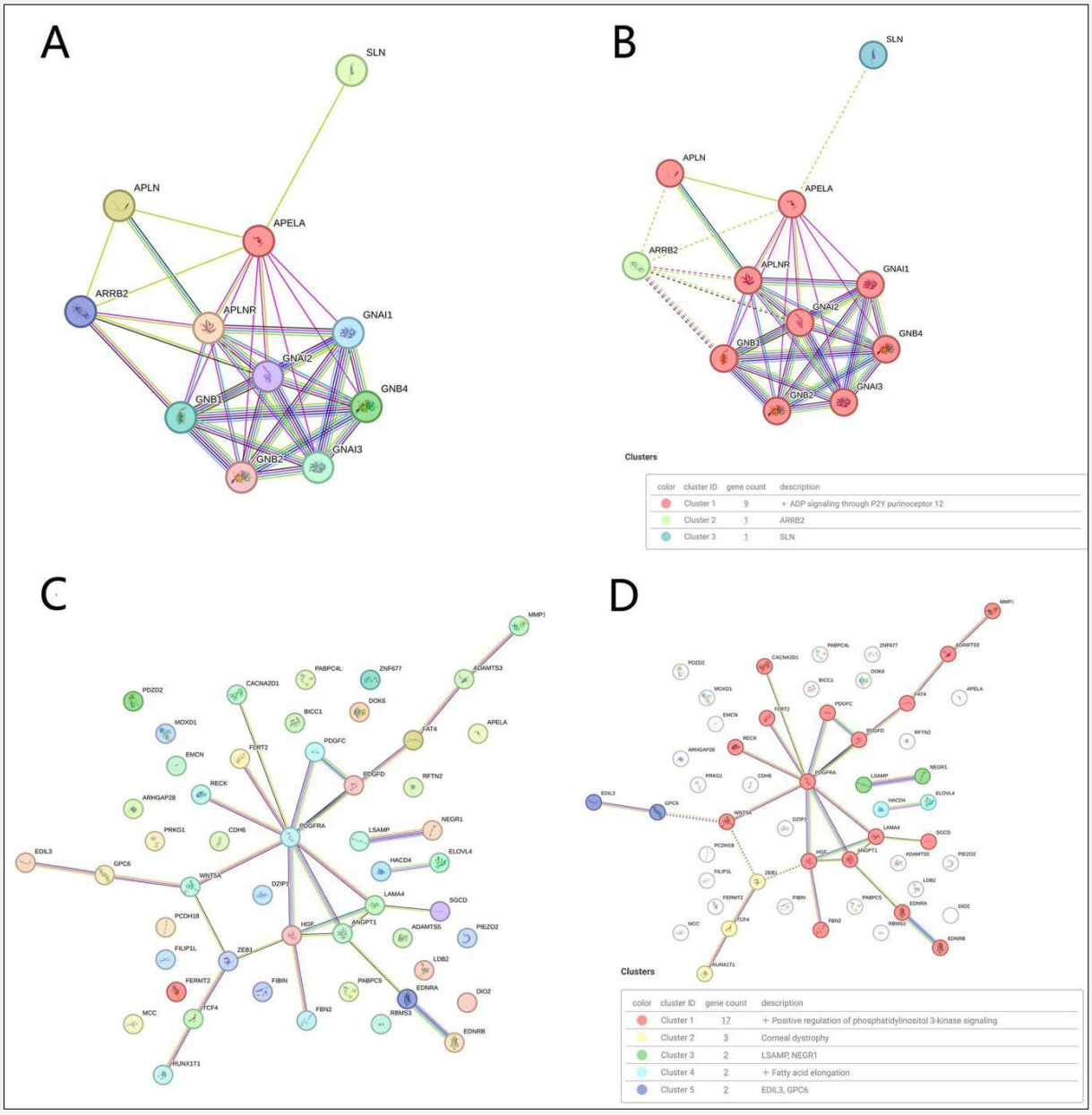

Figure 5. Protein-protein interaction network and modular clustering of APELA-associated interactome.

A) The protein-protein interaction network that involved APELA and 10 associated partners is shown. B) Cluster topology of the APELA-centric interactome following k-means partitioning (k = 3) is shown. C) The protein-protein interaction (PPI) topology that involved 50 functionally associated partners is shown (PPI enrichment p-value = 3.34 × 10 - 11). D) The unsupervised K-means partitioning (k = 5 clusters) is shown.

APELA expression and immune infiltration

Figure 7A shows that there was a significant divergence of APELA expression (FDR-corrected) across the immune cell subtypes, which was most prominent in eosinophils ($p = 1.89 \times 10^{-6}$), iDCs ($p = 2.16 \times 10^{-5}$), Tgd ($p = 6.61 \times 10^{-5}$), mast cells ($p = 7.00 \times 10^{-5}$), DCs ($p = 0.0007$), macrophages ($p = 0.0008$), neutrophils ($p = 0.0012$), Tem ($p = 0.0018$), pDCs ($p = 0.0034$), T helper cells ($p = 0.0042$), Tcm ($p = 0.0063$), and NK CD56 bright cells ($p = 0.0069$).

Despite the observable heterogeneity in APELA expression across the selected immune subsets, all 24 analyzed

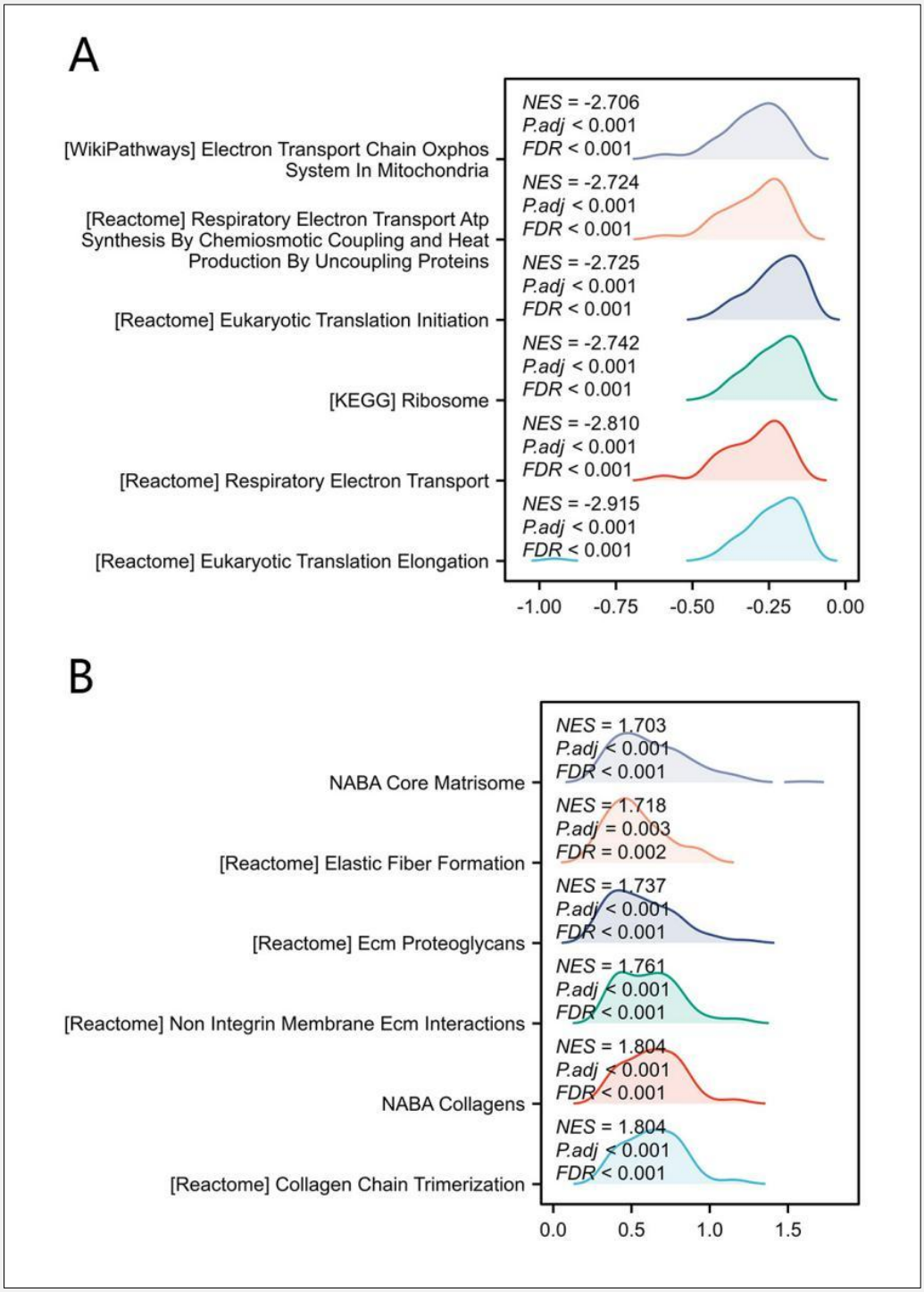

Figure 6. Ridge and mountain plot visualizations for the top negatively- and positively-correlated pathways associated with APELA expression.

A) Ridge plot visualization of the six most inversely correlated pathways: reactome eukaryotic translation elongation; reactome respiratory electron transport; KEGG ribosome; reactome eukaryotic translation initiation; reactome respiratory electron transport ATP synthesis; WikiPathways electron transport chain OXPHOS system in mitochondria. B) Mountain plot representation for the six most positively associated pathways: reactome collagen chain trimerization; NABA collagens; reactome non-integrin membrane-ECM interactions; reactome ECM proteoglycans; reactome elastic fiber formation; NABA core matrisome.

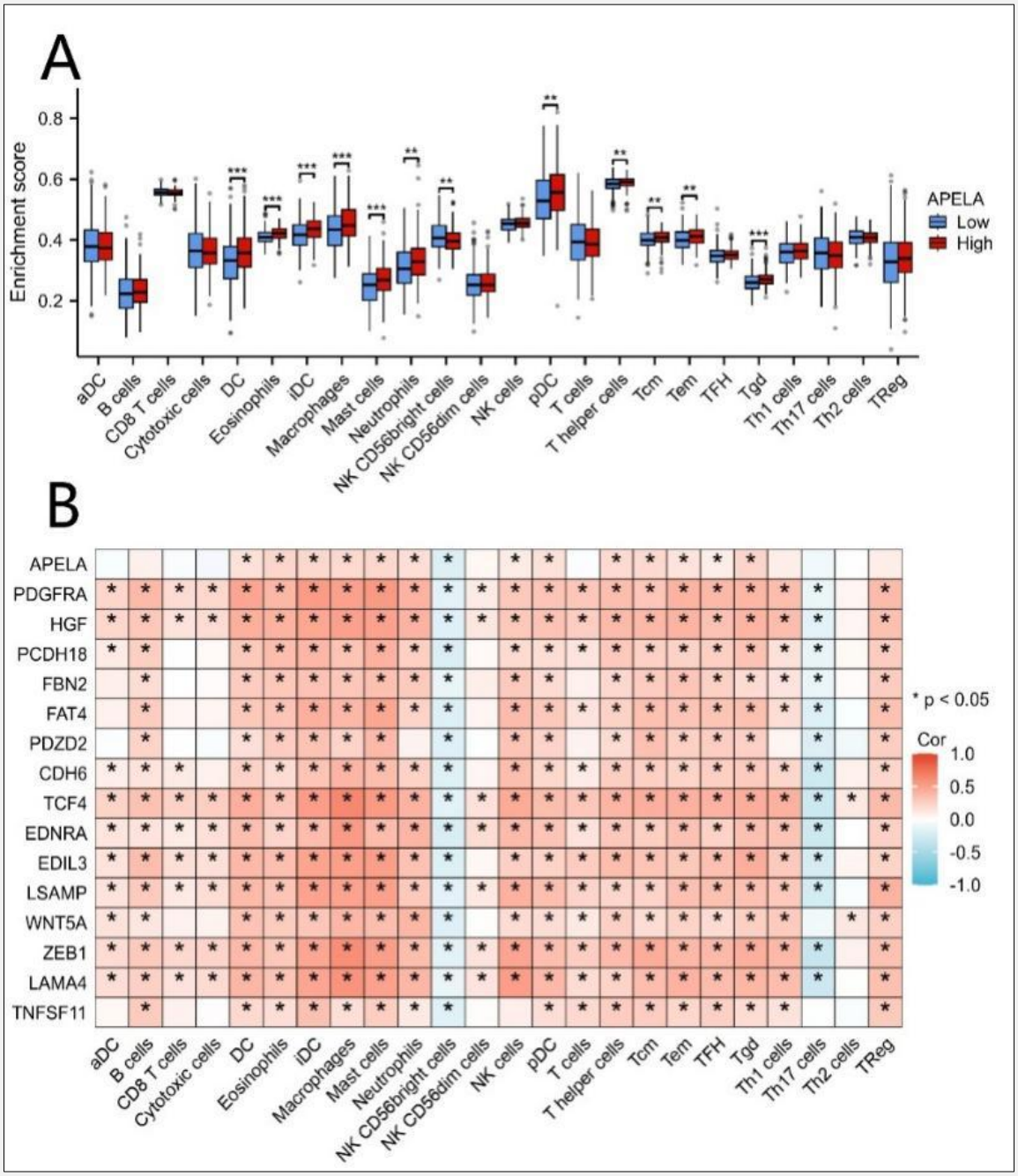

Figure 7. Immune infiltration analysis.

A) Differential analysis of tumor-infiltrating immune cell populations in colorectal cancer tissues. The quantitative comparison of tumor-infiltrating immune cell composition was performed using the Mann-Whitney U test (nonparametric Wilcoxon rank-sum methodology) in tissue specimens. B) Correlation of APELA and the top co-expressed genes with diverse immune cell populations in the tumor microenvironment. The analysis revealed that APELA and its top 15 co-expressed genes have strong association patterns with heterogeneous immune cell populations, including, but not limited, to myeloid-derived suppressor cells (MDSCs) and tumor-associated macrophages (TAMs).

populations exhibited weak intercellular covariation with APELA and co-expressed genes ($|r| < 0.3$). Notably, NK CD56 bright cells exhibited a coordinated downregulation of APELA mRNA expression and its top 15 co-expressed genes. In Th17 cells, several genes, such as PDGFRA, HGF, PCDH18, FBN2, FAT4, PDZD2, CDH6, TCF4, EDNRA, EDIL3, LSAMP, ZEB1, and LAMA4, were significantly underexpressed ($p < 0.05$, Figure 7B).

DISCUSSION

The present study was the first to systematically characterize the dual roles of APELA in CRC. APELA is significantly overexpressed in malignant tissues, when compared to its normal or adenomatous counterparts, suggesting its utility as a diagnostic biomarker.

The validation across two independent cohorts demonstrated consistently high positive predictive values (PPV > 0.98) for APELA in CRC detection, supporting its potential diagnostic value for clinical confirmation. However, the following limitations restricted its standalone diagnostic application, undermining its utility in ruling out CRC: suboptimal sensitivity (≤ 75.80%), which may lead to missed diagnosis; considerable variability in diagnostic performance across cohorts (Youden index range: 0.380 - 0.734); critically low negative predictive values (minimum NPV: 7.00%).

Prognostic refinement through algorithm-based stratification in the TCGA cohort enhanced the predictive performance of APELA, with the improvement in hazard ratio (HR) from 0.62 to 0.50 and increase in statistical significance ($p < 0.05$ to $p < 0.001$). The external validation in the GSE39582 dataset further confirmed its prognostic relevance ($HR = 0.72$, $p < 0.05$), establishing cross-platform robustness. In the clinic, patients with high APELA expression had a 49% increase in median overall survival (145 vs. 97 months), representing a notable 4-year survival benefit. Furthermore, the optimized model improved cohort efficiency ($n = 477$ to $n = 562$), and the enhanced HR underscored the strong association of APELA with reduced mortality risk.

Paradoxically, higher APELA expression is associated with improved overall survival, despite having no correlation with conventional tumor staging parameters [18-20]. This paradox reflects the context-dependent nature of APELA's function, which has been reported as tumor-suppressive in prostate cancer [11], yet oncogenic in gliomas [12].

At the mechanistic level, APELA promotes tumorigenic signaling via the APLNR-G protein-PI3K/AKT pathway and involves the following aspects:

Angiogenesis modulation: The PPI analysis confirms the direct binding of APELA to the APLNR and G-protein subunits (GNAI1-3 and GNB1-4) [21-24], leading to the activation of the PI3K/AKT cascade and upregulation of vascular endothelial growth factor (VEGF) expression [25-27]. This mechanism parallels its pro-metastatic activity observed in ovarian cancer models [28].

Metabolic reprogramming: The GSEA revealed that APELA induces a pseudohypoxic metabolic state, which is characterized by the significant suppression of mitochondrial oxidative phosphorylation ($p < 0.001$) [29] and enhancement of glycolytic activity. This shift supports the proliferative demands of the aggressive CMS4 CRC subtype.

Immune microenvironment remodeling: The loss of APELA impairs NK cell cytotoxicity [30,31] and facilitates immune escape by disrupting Th17 cell function. This includes defects in Th17 cell migration and altered metabolic profiles, which collectively contribute to immunosuppressive tumor microenvironments [32-34].

It is possible to develop noninvasive diagnostic models by combining APELA expression and ECM stiffness biomarkers, such as PRO-C3 [35]. Therapeutic strategies that target the APELA-APLNR-PI3K signaling axis, particularly the combination of ML221 and Alpelisib [36,37], show potential in overcoming chemoresistance. Notably, the subgroup with the marker "APELA high/FGFR-VEGF active" demonstrated increased sensitivity to regorafenib [38], underscoring the value of multi-omics-based stratification for precision intervention.

The present study had several limitations: a small sample size was used to compare the APELA expression across normal, adenomatous, and tumor tissues; there was an issue of data availability due to the few datasets with APELA-specific annotations in public repositories, such as the GEO; absence of *in vitro* or *in vivo* functional assays to substantiate the proposed mechanistic insights. These limitations reflect the broader underappreciation of the role of APELA in tumor biology, restricting the scope of the present diagnostic and prognostic analysis. Therefore, it is necessary to use large cohorts for experimental validation in future research.

CONCLUSION

APELA functions as a central multi-omics regulator in CRC, linking Apelin signaling, ECM remodeling, metabolic reprogramming, and immune evasion. Its dual role (serving as a potential therapeutic target and a prognostic biomarker) highlights its clinical relevance and makes it a strong candidate for further preclinical and clinical studies.

Acknowledgment:
The authors sincerely thank Ms. Lixia Wang for her valuable contributions in language editing and manuscript preparation.

Availability of Data and Materials:
The authors confirm that the data supporting the findings of the study are available within the article itself.

Declaration of Interest:
The authors state that there are no potential conflicts of interest.

References:

1. Han Z, Kong Q, Li Y. Current status and future perspectives on early detection and diagnosis of colorectal cancer in china. Cancer Screen Prev 2024;3(4):214-22. https://www.xiahepublishing.com/2835-3315/CSP-2024-00023
2. REACCT Collaborative; Zaborowski AM, Abdile A, Adamina M, et al. Characteristics of Early-Onset vs Late-Onset Colorectal Cancer: A Review. JAMA Surg 2021;156(9):865-74. (PMID: 34190968)
3. Siegel RL, Miller KD, Goding Sauer A, et al. Colorectal cancer statistics, 2020. CA Cancer J Clin 2020;70(3):145-64. (PMID: 32133645)
4. de Vries-Ten Have J, Winkels RM, Bloemhof SAG, et al. Determinants of healthy lifestyle behaviours in colorectal cancer survivors: a systematic review. Support Care Cancer 2025;33(4): 292. (PMID: 40097728)
5. Carroll KL, Frugé AD, Heslin MJ, Lipke EA, Greene MW. Diet as a Risk Factor for Early-Onset Colorectal Adenoma and Carcinoma: A Systematic Review. Front Nutr 2022;9:896330. (PMID: 35757246)
6. Bhaskaran NA, Kumar L. Treating colon cancers with a non-conventional yet strategic approach: An overview of various nanoparticulate systems. J Control Release 2021;336:16-39. (PMID: 34118336)
7. Rawla P, Sunkara T, Barsouk A. Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. Prz Gastroenterol 2019;(2):89-103. (PMID: 31616522)
8. Zhang X, Zhang H, Shen B, Sun XF. Chromogranin-A Expression as a Novel Biomarker for Early Diagnosis of Colon Cancer Patients. Int J Mol Sci 2019;20(12):2919. (PMID: 31207989)
9. Wu X, Yang H, Chen X, et al. Nano-herb medicine and PDT induced synergistic immunotherapy for colon cancer treatment. Biomaterials 2021;269:120654. (PMID: 33434712)
10. Chng SC, Ho L, Tian J, Reversade B. ELABELA: a hormone essential for heart development signals via the apelin receptor. Dev Cell 2013;27(6):672-80. (PMID: 24316148)
11. Kudryavtseva AV, Lukyanova EN, Kharitonov SL, et al. Bioinformatic identification of differentially expressed genes associated with prognosis of locally advanced lymph node-positive prostate cancer. J Bioinform Comput Biol 2019;17(1):1950003. (PMID: 30866732)
12. Ganguly D, Cai C, Sims MM, et al. APELA Expression in Glioma, and Its Association with Patient Survival and Tumor Grade. Pharmaceuticals (Basel) 2019;12(1):45. (PMID: 30917521)
13. Liu J, Lichtenberg T, Hoadley KA, et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. Cell 2018;173(2):400-416.e11. (PMID: 29625055)
14. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16(5):284-7. (PMID: 22455463)
15. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102(43):15545-50. (PMID: 16199517)
16. Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 2013;39(4):782-95. (PMID: 24138885)
17. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013;14:7. (PMID: 23323831)
18. Wang Z, Huang J. Apela Promotes Cardiomyocyte Differentiation from Transgenic Human Embryonic Stem Cell Lines. Appl Biochem Biotechnol 2019;189(2):396-410. (PMID: 31025171)
19. Liu L, Yi X, Lu C, et al. Study Progression of Apelin/APJ Signaling and Apela in Different Types of Cancer. Front Oncol 2021; 11:658253 (PMID: 33912466)
20. Zong Y, Wang Y, Hu Y, Wang Z. Clinical Significance of Apela in Acute Cardiorenal Insuffiency of Chronic Heart Failure. Kidney Blood Press Res 2024;49(1):100-13. (PMID: 38237563)
21. Jiang Y, Yan M, Wang C, et al. The Effects of Apelin and Elabela Ligands on Apelin Receptor Distinct Signaling Profiles. Front Pharmacol 2021;12:630548. (PMID: 33746758)
22. Anto S, Sun C, O'Rourke ST. Activation of APJ Receptors by CMF-019, But Not Apelin, Causes Endothelium-Dependent Relaxation of Spontaneously Hypertensive Rat Coronary Arteries. J Cardiovasc Pharmacol 2025;85(4):287-96. (PMID: 39836102)
23. Xu W, Yan J, Travis ZD, et al. Apelin/APJ system: a novel promising target for anti-oxidative stress in stroke. Front Pharmacol 2024;15:1352927. (PMID: 39881878)
24. Niebrügge N, Trovato O, Praschberger R, Lieb A. Disease-Associated Dopamine Receptor D2 Variants Exhibit Functional Consequences Depending on Different Heterotrimeric G-Protein Subunit Combinations. Biomedicines 2024;13(1):46. (PMID: 39857630)
25. Dvorak HF. Vascular permeability factor/vascular endothelial growth factor: a critical cytokine in tumor angiogenesis and a potential target for diagnosis and therapy. J Clin Oncol 2002;20(21): 4368-80. (PMID: 12409337)
26. Kerbel RS. Tumor angiogenesis. N Engl J Med 2008;358(19): 2039-49. (PMID: 18463380)
27. Eguchi R, Kawabe JI, Wakabayashi I. VEGF-Independent Angiogenic Factors: Beyond VEGF/VEGFR2 Signaling. J Vasc Res 2022;59(2):78-89. (PMID: 35152220)
28. Ho L, Tan SY, Wee S, et al. ELABELA Is an Endogenous Growth Factor that Sustains hESC Self-Renewal via the PI3K/AKT Pathway. Cell Stem Cell 2015;17(4):435-47. (PMID: 26387754)
29. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144(5):646-74. (PMID: 21376230)
30. Mamessier E, Pradel LC, Thibult ML, et al. Peripheral blood NK cells from breast cancer patients are tumor-induced composite subsets. J Immunol 2013;190(5):2424-36. (PMID: 23359508)

31. Poznanski SM, Nham T, Chew MV, et al. Expanded CD56(superbright)CD16(+) NK Cells from Ovarian Cancer Patients Are Cytotoxic against Autologous Tumor in a Patient-Derived Xenograft Murine Model. Cancer Immunol Res 2018;6(10): 1174-85. (PMID: 30018043)

32. Kuen DS, Kim BS, Chung Y. IL-17-Producing Cells in Tumor Immunity: Friends or Foes? Immune Netw 2020;20(1):e6. (PMID: 32158594)

33. Grasso CS, Giannakis M, Wells DK, et al. Genetic Mechanisms of Immune Evasion in Colorectal Cancer. Cancer Discov 2018; 8(6):730-49. (PMID: 29510987)

34. Tauriello DVF, Palomo-Ponce S, Stork D, et al. TGFβ drives immune evasion in genetically reconstituted colon cancer metastasis. Nature 2018;554(7693):538-43. (PMID: 29443964)

35. Chen IM, Willumsen N, Dehlendorff C, et al. Clinical value of serum hyaluronan and propeptide of type III collagen in patients with pancreatic cancer. Int J Cancer 2020;146(10):2913-22. (PMID: 31642523)

36. Fruman DA, Chiu H, Hopkins BD, Bagrodia S, Cantley LC, Abraham RT. The PI3K Pathway in Human Disease. Cell 2017; 170(4):605-35. (PMID: 28802037)

37. Kai F, Drain AP, Weaver VM. The Extracellular Matrix Modulates the Metastatic Journey. Dev Cell 2019;49(3):332-46. (PMID: 31063753)

38. Fukumura D, Kloepper J, Amoozgar Z, Duda DG, Jain RK. Enhancing cancer immunotherapy using antiangiogenics: opportunities and challenges. Nat Rev Clin Oncol 2018;15(5):325-40. (PMID: 29508855)

39. Hamada J, Kimura J, Ishida J, et al. Evaluation of novel cyclic analogues of apelin. Int J Mol Med 2008; 22(4):547-52. (PMID: 18813863)