Dovepress
Taylor & Francis Group

ORIGINAL RESEARCH

Copper Metabolism-Related Genes as Biomarkers in Colon Adenoma and Cancer

Taikun Zhang[1], Ying Fu ⓘ[2]

[1]Department of Gastroenterology, The Second Affiliated Hospital of Guizhou University of Traditional Chinese Medicine, Guiyang, Guizhou, People's Republic of China; [2]Department of Medical Insurance, The Second Affiliated Hospital of Guizhou University of Traditional Chinese Medicine, Guiyang, Guizhou, People's Republic of China

Correspondence: Ying Fu, The Second Affiliated Hospital of Guizhou University of Traditional Chinese Medicine, No. 83 Feishan Street, Yunyan District, Guizhou, 550001, People's Republic of China, Tel +86-18385944836, Email fuying916@gzy.edu.cn

Purpose: To elucidate the role of copper (Cu) metabolism in the progression of colon adenoma (CA) to colorectal cancer (CRC) and to identify potential biomarkers and therapeutic targets through comprehensive bioinformatics analysis.

Patients and Methods: Datasets associated with colon adenoma were retrieved from the Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) between CA samples and normal controls (NC) were intersected with genes related to copper metabolism (CMRGs) and DEGs between CRC and CA. Five machine-learning algorithms were employed to identify biomarkers. The degree of immune infiltration was evaluated using single-sample Gene Set Enrichment Analysis (ssGSEA), and the expression profiles of these biomarkers across various cell types were further characterized using single-cell RNA sequencing (scRNA-seq). The expression levels of the identified genes were validated using quantitative polymerase chain reaction (qPCR) and data from the Human Protein Atlas (HPA) database.

Results: Five biomarkers were identified: ZEB1, ABCA1, SLC24A3, CAV1, and FLNA. Functional enrichment analysis revealed significant pathway alterations in the low-expression groups of CAV1 (eg, phagosome pathway) and FLNA (eg, ribosome pathway). Significant differences in the infiltration abundance of macrophages and mast cells were observed between CA and NC. scRNA-seq analysis demonstrated that these biomarkers were expressed in fibroblasts, lymphocytes, goblet cells, B cells, and macrophages. The consistency of gene expression between patient samples and public datasets was confirmed through qPCR and HPA data.

Conclusion: This study explores the role of copper metabolism in colon adenoma progression using bioinformatics. Five genes (ZEB1,ABCA1, SLC24A3, CAV1, FLNA) were identified as potential biomarkers. These genes correlate with immune infiltration and may serve as diagnostic and therapeutic targets. Further clinical validation is needed.

Keywords: copper metabolism, colon adenoma, machine learning, biomarkers, single-cell RNA sequencing

Introduction

Colorectal cancer (CRC) ranks as the third most frequently diagnosed malignant tumor globally and remains a major contributor to cancer-related deaths.[1] As of 2019, the global incidence of CRC has risen to 2.19 million cases, with Chinese patients comprising approximately 30% of the worldwide CRC population.[2] According to recent projections, the annual number of new CRC cases is expected to rise to 2.5 million by 2035, posing a growing public health challenge worldwide.[3] The treatment of CRC imposes a significant financial burden, with direct medical costs estimated at $106.4 billion, accounting for 62.4% of the total economic impact.[4] The rapid increase in cases and high economic costs highlight the urgent need for effective prevention, early detection, and treatment. The progression from normal colonic epithelium to adenoma and invasive carcinoma, driven by acquired molecular events, usually takes 5–10 years.[5,6] Colonic adenoma (CA) is a benign epithelial tumor arising in the colon's inner lining, marked by abnormal cell proliferation and the formation of small masses. Approximately 10% of CA cases may progress to CRC, influenced by the accumulation of gene mutations, epigenetic abnormalities, and environmental factors.[7,8] Colon cancer progression is closely linked to various genetic mutations, with mutations in the APC, KRAS, and TP53 genes being particularly

prevalent in CRC.[9] APC mutations lead to the over-activation of the Wnt signaling pathway, a critical driver of tumorigenesis.[10] Similarly, mutations in the KRAS gene facilitate uncontrolled tumor cell proliferation.[11] Furthermore, TP53 mutations compromise the cell's ability to respond to DNA damage, thereby accelerating cancer cell development and progression.[12] Mutations in DNA repair genes (eg, MLH1, MSH2, MSH6) are strongly linked to hereditary nonpolyposis colorectal cancer (HNPCC). They disrupt DNA repair, increasing mutation accumulation and driving CRC development.[13] Epigenetic alterations also play a pivotal role in CRC progression. For instance, mutations and reduced expression of the histone H3K36 trimethyltransferase SETD2 are linked to the advancement of CRC.[14] Additionally, aberrant DNA methylation, particularly the hypermethylation of promoters associated with microsatellite instability (MSI), leads to reduced or absent expression of DNA mismatch repair proteins, further contributing to CRC pathogenesis.[15] Meanwhile, environmental factors cannot be ignored. Unhealthy dietary habits, obesity, and long - term social stress can cause DNA damage. Once the DNA was damaged and the repair mechanism is compromised due to epigenetic changes, the risk of developing CRC significantly increases.[16] Evidence suggests that the detection and excision of adenomas substantially decrease CRC incidence and mortality.[17] Therefore, early CA detection is critical. Although colonoscopy remains effective for identifying and excising adenomas before their progression to CRC, it has notable limitations, including a missed diagnosis rate of up to 25%.[18] Stool-based screening, commonly employed for CRC detection, lacks sensitivity for adenomas, making it unsuitable for early detection of these lesions.[8,19] This underscores the urgent need to identify novel characteristic genes associated with CA.

Copper (Cu) is a vital trace element required for the growth and development of all eukaryotes.[20] It plays a critical role in redox enzymatic activities that are essential for key metabolic processes, signaling pathways, and various biological functions.[21] The link between disrupted Cu metabolism and the pathogenesis of various malignancies has attracted significant research attention. Studies indicate that inflammatory cytokines, such as interleukin (IL)-17, enhance Cu uptake through inflammatory responses, thereby fostering colon tumor development.[22] Furthermore, evidence shows that Cu levels in colon tumor tissues are significantly elevated compared to normal colon mucosa.[23] Studies have shown that copper metabolism-related genes (CMRGs) are associated with immune infiltration in colon adenocarcinoma (COAD). They can be utilized to classify COAD subtypes and distinguish patients with different molecular character-istics, serving as valuable biomarkers for immunotherapy in COAD.[24] This highlights Cu metabolism's critical role in colon tumor initiation and progression. However, research on CMRGs in the context of colon cancer development remains limited.

Bioinformatics analysis serves as a powerful tool for predicting molecular pathways and gene interactions, and it has been extensively used to identify novel candidate genes and pathways across various cancer types.[25] By leveraging computational methods to analyze extensive genomic and transcriptomic data, bioinformatics has facilitated the identi-fication of novel candidate genes and pathways associated with the development of colon carcinogenesis.[26,27] This study aims to conduct a comprehensive bioinformatics analysis to identify signature genes that can effectively distinguish between colon adenomas (CAs), colorectal cancers (CRCs), and normal colons (NCs). Based on these identified genes, we expect to develop a robust classification model that can accurately differentiate between CA, CRC, and NC, thereby enhancing both the specificity and sensitivity of colorectal disease diagnosis.Furthermore, the study will incorporate single-cell data analysis from NC, CA, and CRC patient samples, with the aim of elucidating the expression patterns of these signature genes across various cell types, including fibroblasts, lymphocytes, cuprocytes, B-cells, and macrophages. This analysis is expected to provide valuable insights into the cellular mechanisms that underpin disease progression and to shed light on the roles of these genes within the tumor microenvironment. By identifying these key genes and their interactions, the study seeks to offer novel perspectives on the molecular mechanisms driving CRC progression, thereby advancing the field of colorectal cancer research and informing future diagnostic strategies.

Materials and Methods
Data Source
The RNA sequencing (RNA-seq) datasets GSE117606, GSE20916 and GSE41657, along with the single-cell RNA-seq (scRNA-seq) dataset GSE201348, were retrieved from the Gene Expression Omnibus (GEO) database (available at

https://www.ncbi.nlm.nih.gov/geo/). The GSE117606 dataset, detected using the GPL25373 platform (Affymetrix HT HG-U133 + PM Array Plate), includes tissue samples from 65 normal controls (NC), 69 patients with CA, and 74 with CRC. The GSE20916 dataset, detected on the GPL570 platform (Affymetrix Human Genome U133 Plus 2.0 Array), consists of tissue samples from 44 NCs, 55 patients with CA, and 46 patients with CRC. The GSE41657 dataset, detected using the GPL6480 platform (Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version)), includes tissue samples from 12 NC, 51 patients with CA, and 25 with CRC. The GSE201348 dataset, processed with the GPL24676 platform (Illumina NovaSeq 6000), contains tissue samples from 8 NCs, 43 patients with CA, five patients with CRC, and 16 unaffected CRC individuals. A total of 1,874 CMRGs were sourced from the GeneCards database (accessible at https://www.genecards.org/) using the search term "Cu metabolism" under the criteria: category = protein-coding and relevance score ≥1. Furthermore, the Cancer Genome Atlas (TCGA)-CRC datasets, comprising 598 CRC patient tissue samples, were accessed from the TCGA database (available at https://gdc.cancer.gov/) for feature gene survival analysis.

Differential Expression Analysis

Differential expression analysis was the process of identified genes with significant differences in expression between sample groups. To obtained the genes with significant differences in the samples, the differentially expressed genes (DEGs) between CA and NC samples in the GSE117606 dataset were identified using the limma package, applying thresholds of $|logFC| > 0.5$ and adj. $p < 0.05$.[28] The top 50 DEGs1 were visualized using the pheatmap package (version 1.0.12).[29] Differentially expressed CMRGs (DE-CMRGs1) were subsequently identified by intersecting DEGs1 with CMRGs through the VennDiagram package (version 1.6.2).[30] Similarly, DEGs2 between CRC and CA samples in the GSE117606 dataset were also determined using the same limma parameters, and DE-CMRGs2 were obtained by overlapping DEGs2 with CMRGs. To explore common functions and relevant pathways, ingenuity pathway analysis (IPA) software was employed to conduct classical pathway enrichment analysis for DE-CMRGs1 and DE-CMRGs2 separately.

Diagnostic Model Construction and Validation

In order to comprehensively validated the diagnostic performance and prediction ability of the model and biomarkers, multiple algorithms were integrated for key marker screening. Given the strong reliability of integrating multiple algorithms for the identification of key markers,[31,32] five machine learning techniques—random forest (RF), support vector machines (SVM), logistic regression (LR), gradient boosting decision tree (GBDT), and artificial neural network (ANN)—were applied to the overlapping genes of DE-CMRGs1 and DE-CMRGs2 to develop diagnostic models for distinguishing between samples from patients with CA, CRC, and NC. The predictive performance of each model, in terms of the area under the curve (AUC) value, was assessed using the pROC package (version 1.17.0.1).[33] The model yielding the highest AUC was selected as the optimal model, and the top 10 most important variables within this model were identified and ranked. Subsequently, biomarkers were refined by overlapping the top 10 variables from the three best-performing models, and the diagnostic value of each individual feature gene was analyzed across different subgroups (CA vs NC; CRC vs CA; CRC vs NC). Additionally, the collective diagnostic accuracy of all biomarkers was evaluated using a logistic regression model through the rms package. The expression levels of biomarkers among samples from NC, CA, and CRC in the GSE117606 dataset were also examined, and their associations with CRC survival in the TCGA-CRC cohort were investigated using Kaplan-Meier (K-M) analysis. Optimal expression thresholds for different groups were calculated using the survminer (version 0.4.9) and survival (version 3.3–1) packages. To further facilitate clinical decision-making regarding CA outcomes, a nomogram was constructed using logistic regression, wherein the score for each feature gene was computed and aggregated into a total score to predict the risk of CA. The diagnostic accuracy of the nomogram was evaluated through calibration curves, decision curve analysis (DCA), and clinical impact curves. The expression levels and diagnostic relevance of the biomarkers were further validated in an external dataset (GSE20916 and GSE41657), with the diagnostic performance of individual biomarkers and the full set analyzed in a manner consistent with the procedures used for the GSE117606 and GSE41657dataset.

Single Gene Set Enrichment Analysis (GSEA)

To comprehensively investigate the pathways influenced by the biomarkers and their associated functions, GSEA was performed using the ClusterProfiler package (version 3.18.0).[34] Initially, correlation coefficients were calculated through correlation analysis between the biomarkers and other genes. GSEA was then conducted based on the ranking of these correlation coefficients. The results were visualized using GSEA spine plots generated by the Enrichplot package (version 1.10.2).[35]

Immune Infiltration Analysis

In order to further explored the pathways targeted by biomarkers and the functions, a single-sample GSEA (ssGSEA) was also performed to analyze immune infiltration, specifically assessing the tumor microenvironment (TME) cell composition in samples from NC and patients with CA and CRC. The proportions of immune cells among these three groups were compared using the rank sum test. Additionally, Spearman correlation analysis was employed to explore the relationships between biomarkers and immune cells.

scRNA-Seq Analysis

The main objective of single-cell RNA sequencing (scRNA-seq) analysis was to study the heterogeneity of gene expression at the single-cell level, to reveal cell-to-cell variability, as well as to explore cell development, differentiation, and function.[36] To explore the characteristic gene expression patterns during cell state transitions, as well as cellular changes in disease states, we performed scRNA-seq analysis. For the scRNA-seq data, processing began with the CreateSeuratObject function of the Seurat package (version 4.0.5). The data were filtered using minimum criteria of 100 cells and 100 features.[37] The percentages of erythrocyte and mitochondrial genes were calculated using the PercentageFeatureSet function. Data normalization was performed using the NormalizeData function, and the FindVariableFeatures function was utilized to identify genes with high intercellular expression variation. To ensure uniformity, the data were further processed using ScaleData, followed by principal component analysis (PCA). The JackStraw and ScoreJackStraw functions were used for linear dimensionality reduction to assess data quality. The information content of each principal component was ranked, and a PCA ElbowPlot was generated. The JackStrawPlot was employed to compare the distribution of p-values for each principal component. Unsupervised cluster analysis was subsequently performed using the Seurat package, with cell types identified through auxiliary annotation using the SingleR function on the Cellmarker website. The expression levels of biomarkers across different categories were calculated, and their relationships with cell categories were visualized using heatmaps and scatter plots.

Construction of the ceRNA Network

The miRWalk database was used to predict miRNAs targeting the biomarkers, with criteria including an energy score below −15 and a binding probability above 0.95. The Starbase database (accessible at https://starbase.sysu.edu.cn/index.php) was employed to predict the targeting relationships between lncRNAs and miRNAs, using parameters such as a clipExpNum greater than 1, a degraExpNum greater than 1, and a gene type of lncRNA. This information was used to construct the ceRNA regulatory network in Cytoscape (version 3.8.2), analyzing the relationships between lncRNAs, miRNAs, and mRNAs.

Potential Therapeutic Drug Prediction

The Drug-Gene Interaction database (DGIdb, accessible at https://dgidb.genome.wustl.edu/) was utilized to predict drugs related to the biomarkers, aiming to identify small-molecule compounds with potential therapeutic effects for patients with CA or CRC. The gene-drug interactions were visualized using Cytoscape, providing a comprehensive overview of potential therapeutic targets.

Expression Verifications of Biomarkers

To further characterize the mRNA and protein expression of the biomarkers in clinical samples, human tissue samples were obtained, including 10 pairs of frozen samples from patients with CA and NC,[38] from the Second Affiliated Hospital of Guizhou University of Traditional Chinese Medicine for qPCR experiments. This study was approved by the Ethics Committee of the Second Affiliated Hospital of Guizhou University of Traditional Chinese Medicine, and written

informed consent was obtained from all participants. Total RNA was extracted from the 10 pairs of frozen samples, each samples 50mg, and RNA concentration was measured using the NanoPhotometer N50 (Table S1). Reverse transcription was performed with the SureScript First-Strand cDNA Synthesis Kit from Saville. The reverse transcription system was shown in Table S2. The qPCR reactions were carried out in a CFX96 real-time quantitative fluorescence PCR instrument using a system comprising 3 μL of cDNA, 5 μL of 2x Universal Blue SYBR Green qPCR Master Mix, and 1 μL each of upstream and downstream primers. The amplification protocol included an initial denaturation at 95°C for 1 minute, followed by 40 cycles of denaturation at 95°C for 20 seconds, annealing at 55°C for 20 seconds, and extension at 72°C for 30 seconds. Relative gene expression levels were calculated using the $2-\Delta\Delta Ct$ method, with glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as the internal reference gene. The primer sequences for PCR were shown in Table 1. Additionally, the protein levels of the biomarkers in tissues from patients with CRC were assessed using immunohistochemistry (IHC) data available from the Human Protein Atlas (HPA) database (accessible at https://www.proteinatlas.org/).

Statistical Analysis

Statistical analysis was performed using R software. Differences in gene expression and immune cell proportions between different groups were analyzed via the Wilcoxon test, while the qPCR expression data was analyzed using a *t*-test. Statistical significance was set at adj. $p < 0.05$.

Results

Identification of DE-CMRGs1 and DE-CMRGs2

The differentially expressed genes (DEGs) between samples from patients with CA and NC, a total of 1,343 DEGs1 were identified between samples from patients with CA and NC, with 467 genes up regulated and 876 genes down regulated (Figure 1A). To visualize these findings, a heat map depicting the top 50 DEGs1 was generated (Figure S1A). By intersecting the set of 1,343 DEGs1 with the 1,874 CMRGs, 218 DE-CMRGs were identified, as illustrated in the Venn diagram (Figure 1B). Similarly, between samples from patients with CRC and CA, 982 DEGs were identified, including 641 up regulated genes and 341 down regulated genes (Figures 1C and S1B). By overlapping these 982 DEGs with the 1,874 CMRGs, 185 DE-CMRGs2 were identified, as shown in the Venn diagram (Figure 1D).

Identification of DE-CMRGs1 and DE-CMRGs2 Enrichment Pathways

IPA was performed to investigate the functions and associated pathways of DE-CMRGs1 and DE-CMRGs2. The analysis revealed that DE-CMRGs1 were involved in the activation of aryl hydrocarbon receptor signaling while inhibiting the TME pathway, oncostatin M signaling, and LXR/RXR activation. In contrast, DE-CMRGs2 were found to activate the TME pathway, oncostatin M signaling, pulmonary fibrosis idiopathic signaling, the role of chondrocytes in the

Table 1 Sequence of Primers

| Primer	Sequences
ZEB1 F	TTGGCTTGGCAACAGTATTTCG
ZEB1 R	CTCCAGCGAGCATTGCAGTAAG
ABCA1 F	TAAAGTGGCGGTAGATGGTAATGT
ABCA1 R	AAGGAAGACTGATGGCTGAAATAA
SLC24A3 F	AGGCCCAGACCTGTAAATGC
SLC24A3 R	GAGGAATGCCCACGCAAATC
CAV1 F	GCTTGACCTCATGGACCTCGTA
CAV1 R	GACCACCGCACCTTGTCTTCT
FLNA F	CTGTCGGAGCGGGACATCT
FLNA R	TGTTTAGGGTCGCGGTTGAC
GAPDH F	CCCATCACCATCTTCCAGG
GAPDH R	CATCACGCCACAGTTTCCC
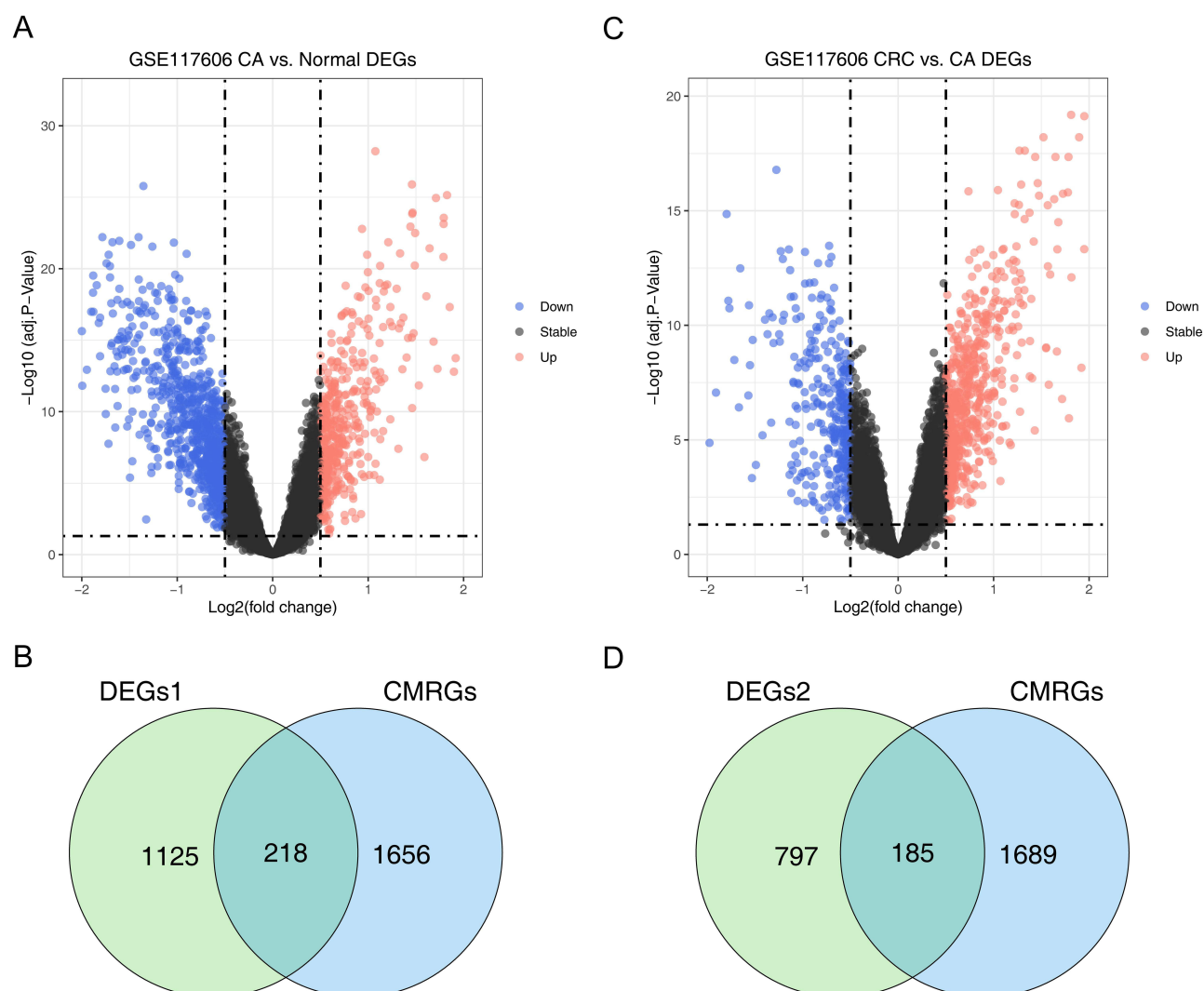
Figure 1 Differentially expressed copper metabolism-related genes (DE-CMRGs) were identified by overlapping the differentially expressed genes (DEGs1) between colon adenoma (CA) and normal control (NC) with copper metabolism-related genes (CMRGs), as well as overlapping the DEGs2 between CRC and CA with CMRGs. The analysis yielded the following results: (**A**) A volcano plot displaying 1,343 DEGs1 between samples from patients with CA and NC (|logFC| > 0.5 and p < 0.05), red for up-regulated genes, blue for down-regulated genes; (**B**) Venn diagram illustrating the 218 DE-CMRGs1; (**C**) A volcano plot showing 982 DEGs2 between samples from patients with CRC and CA (|logFC| > 0.5 and p < 0.05), red for up-regulated genes, blue for down-regulated genes; (**D**) Venn diagram depicting the 185 DE-CMRGs2.

rheumatoid arthritis signaling pathway, and hypoxia-inducible factor-alpha (HIFα) signaling. Additionally, DE-CMRGs2 significantly inhibited LXR/RXR activation and matrix metalloproteases (Figure 2A and B).

Identification and Validation of Biomarkers

We took the intersection of DE-CMRGs1 and DE-CMRGs2 to obtain the candidate biomarkers. The Venn diagram revealed that 86 candidate biomarkers were identified as common between DE-CMRGs1 and DE-CMRGs2 (Figure 3A). Among the machine learning models tested, the RF model demonstrated the highest AUC value of 0.953 for distinguishing between samples from patients with CA and NC (Figure 3B). For comparisons between samples from patients with CRC and NC, the GBDT model performed best, achieving an AUC value of 0.935 (Figure S2A). Similarly, the GBDT model also yielded the highest AUC value of 0.929 when comparing samples from patients with CRC and CA (Figure S2B). The top 10 genes from each model were selected and ranked, as shown in Table 2. Overlapping these 30 genes across the models led to the identification of five biomarkers: ZEB1, ABCA1, SLC24A3, CAV1, and FLNA (Figure S2C).

A

Analysis: DE-CMRG1 - 2022-09-06 05:47 PM
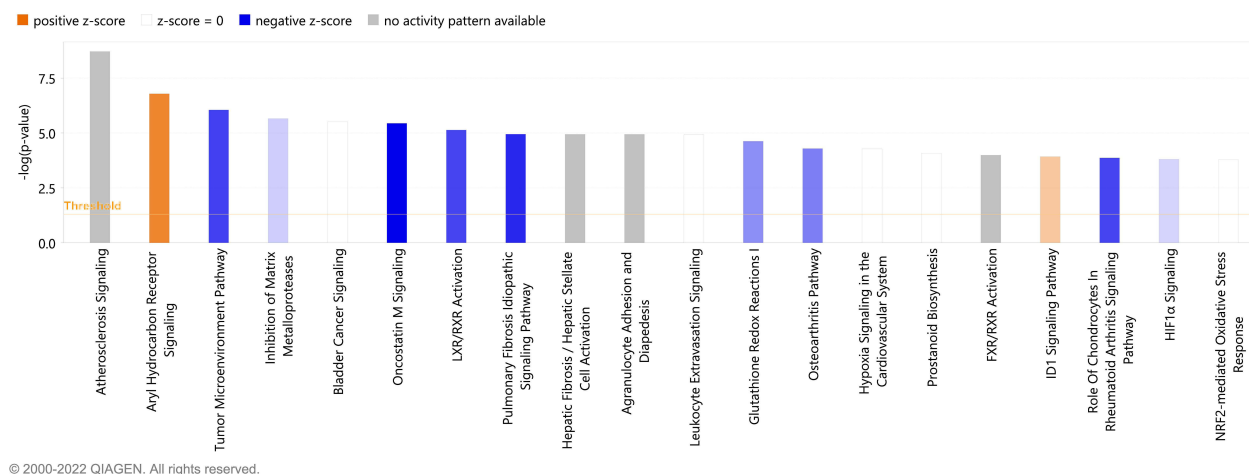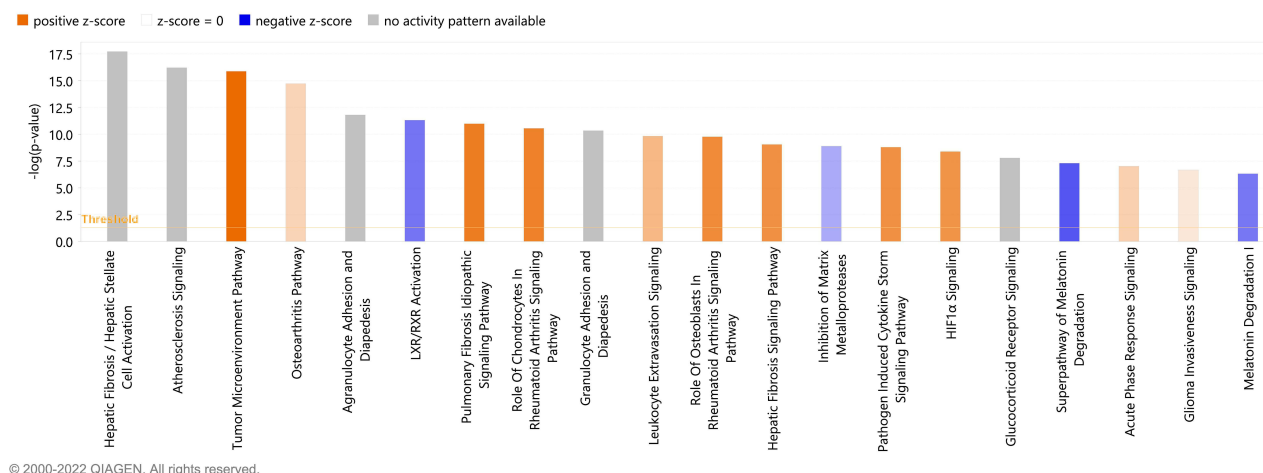
B

Analysis: DE-CMRG2 - 2022-09-06 05:48 PM

Figure 2 Function enrichment analysis for DE-CMRGs1/DE-CMRGs2. Ingenuity Pathway Analysis (IPA) diagrams of DE-CMRGs1 (**A**) and DE-CMRGs2 (**B**).

Five diagnostic models were constructed separately using the candidate hub genes and compared based on the AUC values, and the model with the largest AUC value was selected as the best model. The analysis of each individual feature gene's diagnostic value revealed that ZEB1, ABCA1, SLC24A3, CAV1, and FLNA exhibited strong discriminatory power between CA and NC, as well as CRC and CA, while only CAV1 and FLNA effectively distinguished between CRC and NC (Figure 3C and Figure S2D and E). Additionally, the diagnostic performance of the five-gene signature demonstrated excellent predictive accuracy for NC, CA, and CRC outcomes in both the GSE117606 and external GSE20916 validation sets (Figure 3D). The nomogram based on these five biomarkers, shown in Figure 3E, also exhibited high predictive accuracy for CA, as confirmed by the calibration curve (Figure S2F). The DCA indicated that the nomogram provided clinical benefits across a wide high-risk threshold range (0–1), surpassing the predictive abilities of individual genes such as ZEB1, ABCA1, SLC24A3, CAV1, and FLNA (Figure S2G). Furthermore, the clinical impact curve confirmed the nomogram's relatively accurate predictive capabilities (Figure S2H). Given the significant predictive
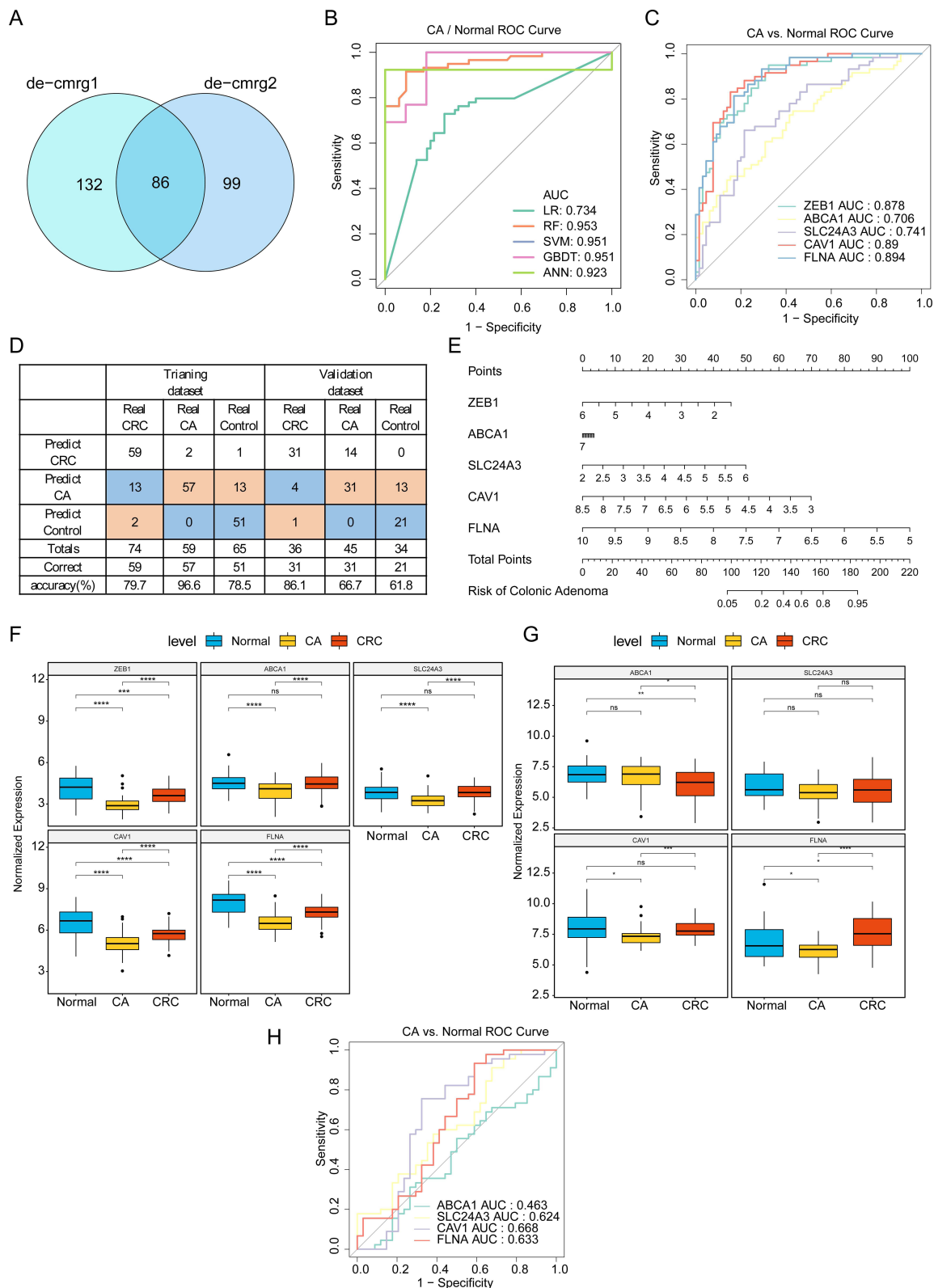
Figure 3 Diagnostic value of the biomarkers in the GSE117606 dataset. (**A**) Venn diagram showing 86 candidate biomarkers; (**B**) Diagnostic value of five machine-learning models between CA and NC in GSE117606; (**C**) Diagnostic value of five biomarkers between CA and NC in GSE117606; (**D**) Construction of the nomogram based on five biomarkers for clinical use in GSE117606; (**E**) Statistics for the predictive accuracy of biomarkers on NC, CA, CRC outcomes in both GSE117606 and GSE20916; (**F**) Expressions of five biomarkers in GSE117606; (**G**) Expressions of five biomarkers in GSE20916; (**H**) Diagnostic value of five biomarkers between CA and NC in GSE20916 (ns, not significant, * p <0.05, ** p <0.01, *** p <0.001, **** p <0.0001).

Table 2 Importance Ranks for the Top 10 Model Genes in the Three Optimal Models in CA/NC/CRC

Variable	Permutation	Dropout_loss	Variable	Permutation	Dropout_loss	Variable	Permutation	Dropout_loss
PAM	0	0.296172296	ZEB1	0	0.305809715	FLNA	0	0.179280846
ZEB1	0	0.298679048	SLC24A3	0	0.306774937	FBN1	0	0.179725533
CSRP1	0	0.299734117	SLC4A4	0	0.308207895	SLC24A3	0	0.180048873
ABCA1	0	0.299891041	FLNA	0	0.308933436	FN1	0	0.180340091
COL5A1	0	0.301992472	FGFR1	0	0.310169343	ABCA1	0	0.180340091
SLC24A3	0	0.302253564	ABCA1	0	0.311664029	NRP1	0	0.180340091
IGFBP7	0	0.302727832	SLC26A2	0	0.313494742	PAM	0	0.180547646
CAV1	0	0.305157007	CAV1	0	0.313789334	ZEB1	0	0.180896265
FBN1	0	0.305252788	NEXN	0	0.313999463	NEXN	0	0.181556524
FLNA	0	0.308301015	CSRP1	0	0.314118929	CAV1	0	0.182679888
CA vs.Normal (RF)			CRC vs.CA (GBM)			CRC vs.Normal (GBM)		

value of clinical features, age and gender were incorporated into the five-gene-based nomogram (Figure S3A). This enhanced nomogram demonstrated improved accuracy and clinical benefits compared to individual factors alone (Figure S3B and Figure S3C), with the clinical impact curve further validating these results, consistent with the previous nomogram (Figure S3D).

As illustrated in Figure 3F, all five genes were expressed at lower levels in samples from patients with CA compared to NC and those with CRC. Notably, K-M survival analysis indicated significant differences in survival rates among patients with CRC exhibiting high versus low FLNA expression, where higher FLNA expression correlated with poorer prognosis (Figure S4). In the external validation using the GSE20916 dataset, the expression of four biomarkers (ABCA1, SLC24A3, CAV1, and FLNA) was confirmed, as ZEB1 was not detected. Among these, ABCA1, CAV1, and FLNA exhibited significant differences in expression across samples from patients with CRC, CA, and NC, with FLNA expression aligning with the findings from the GSE117606 dataset (Figure 3G). In the GSE41657 dataset, significant differences were observed in the expressions of CAV1 and SLC24A3 among the samples from patients with CRC, CA, and NC. The expression of CAV1 was consistent with the results of the GSE117606 dataset (Figure S5). Moreover, ROC curve analysis demonstrated that CAV1 and FLNA could effectively distinguish CRC from CA (AUC > 0.7) (Figure 3H, Figure S6A and B). In the GSE41657 dataset, ROC curve analysis demonstrated that CAV1 could be effectively used to distinguish CRC from CA (AUC > 0.7) (Figure S6C–E).

GSEA for Biomarkers

We analyzed the correlation between the biomarkers and the remaining genes to obtain the correlation coefficient. Then, the coefficients of the genes were ranked, and then the ranked results were used to perform GSEA to obtain the biological process of biomarkers enrichment. The GSEA results revealed significant enrichment of various pathways and biological processes associated with the expression levels of CAV1 and FLNA. In the high CAV1 expression group, key pathways such as the Apelin signaling pathway, cyclic guanosine monophosphate-protein kinase G (cGMP-PKG) signaling pathway, adherens junctions, mitogen-activated protein kinases (MAPK) signaling pathway, neuroactive ligand-receptor interactions, proteoglycans in cancer, regulation of the actin cytoskeleton, and vascular smooth muscle contraction were notably enriched. Conversely, in the low CAV1 expression group, neuroactive ligand-receptor interaction were significantly enriched (Figure 4A). For FLNA, the low-expression group showed significant enrichment in ribosomal processes and ribosome biogenesis in eukaryotes. In contrast, the high FLNA expression group exhibited enrichment in pathways such as the Apelin signaling pathway, cGMP-PKG signaling pathway, adherens junctions, MAPK signaling pathway, oxytocin signaling pathway, proteoglycans in cancer, regulation of the actin cytoskeleton, and vascular smooth muscle contraction (Figure 4B). In summary, these biomarkers, particularly CAV1 and FLNA, were predominantly enriched in pathways related to the cGMP-PKG signaling pathway, Apelin signaling pathway, focal adhesion, proteoglycans in cancer, vascular smooth muscle contraction, Salmonella infection, neuroactive ligand-receptor interaction, and regulation of the actin cytoskeleton (Figures 4C–E).
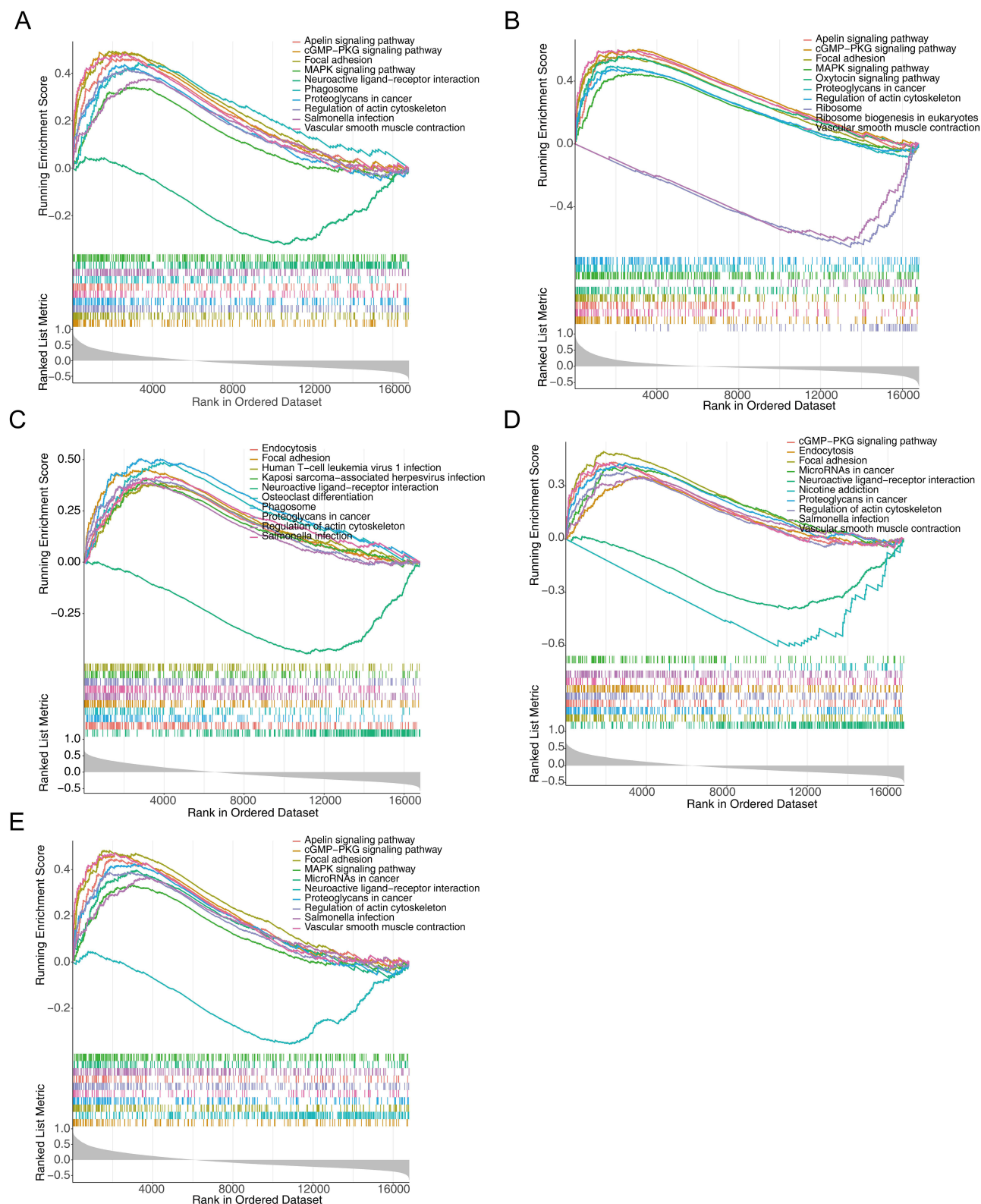
Figure 4 Single gene set enrichment analysis (GSEA) for five biomarkers. (**A**) CAV1; (**B**) FLNA; (**C**) ABCA1; (**D**) SLC24A3; (**E**) ZEB1.

Effect of Biomarkers on Immune Heterogeneity

Using ssGSEA to calculate the 24 TME cell score for each sample, the composition of 24 TME cell types in samples from patients with NC, CA, and CRC is illustrated in Figure 5A. Analysis of immune cell proportions revealed significant differences in B cells, macrophages, mast cells, and T cells between samples from patients with CA and NC, as identified by the rank-sum test (Figure 5B). Spearman correlation analysis further highlighted significant positive correlations between the biomarkers and mast cells, as well as eosinophils. Additionally, CAV1, ABCA1, and ZEB1 showed significant positive correlations with both B cells and T cells (cor > 0.2, adj.p < 0.05). In contrast, FLNA exhibited significant negative correlations with macrophages (cor < −0.1, adj. p < 0.05) (Figure 5C).

Identification of Feature Gene Expression in Different Cell Subsets

Using single-cell RNA-seq to analyzed the expression patterns of biomarkers during cell state transitions and changes in cells during disease states, after data filtering, the features of the scRNA-seq dataset are presented in Figure S7A. A total of 2,000 highly variable genes were selected for cell type identification (Figure S7B). Following normalization, the results of PCA are shown in Figure S7C, from which 30 principal components were selected for further analysis based on a p-value of < 0.05 (Figure S7D). Before clustering, a resolution of 0.06 was determined to be optimal, resulting in well-defined clusters (Figure S7E and Figure S7F). The visualization of the 13 clusters using uniform manifold approximation and projection (UMAP) and t-distributed stochastic neighbor embedding (t-SNE) is depicted in Figure 6A. Heat maps displaying the top five marker genes for each cluster were generated (Figure 6B).

Cell types were then identified for each cluster. Cluster 2, characterized by higher expression levels of marker genes such as FCGBP, MUC2, and CLCA1, was identified as the goblet cell cluster (Figure 6C). Cluster 3, with elevated expression of marker genes like FTL, TMSB10, and TPT1, was identified as the myeloid cell cluster (Figure 6D). Cluster 4, showing higher expression of marker genes such as BLNK, MS4A1, and CD37, was identified as the B cell cluster (Figure 6E). Consequently, the 13 clusters were further classified into 10 distinct cell subsets (Table S3). The distribution of cell counts across each sample is shown in Figure 6f, highlighting the presence of epithelial cells, goblet cells, myeloid cells, B cells, fibroblasts, LGR5+ stem cells, lymphocytes, dendritic cells, macrophages, and neurons as the predominant cell types in samples from patients with NC, CA, and CRC.
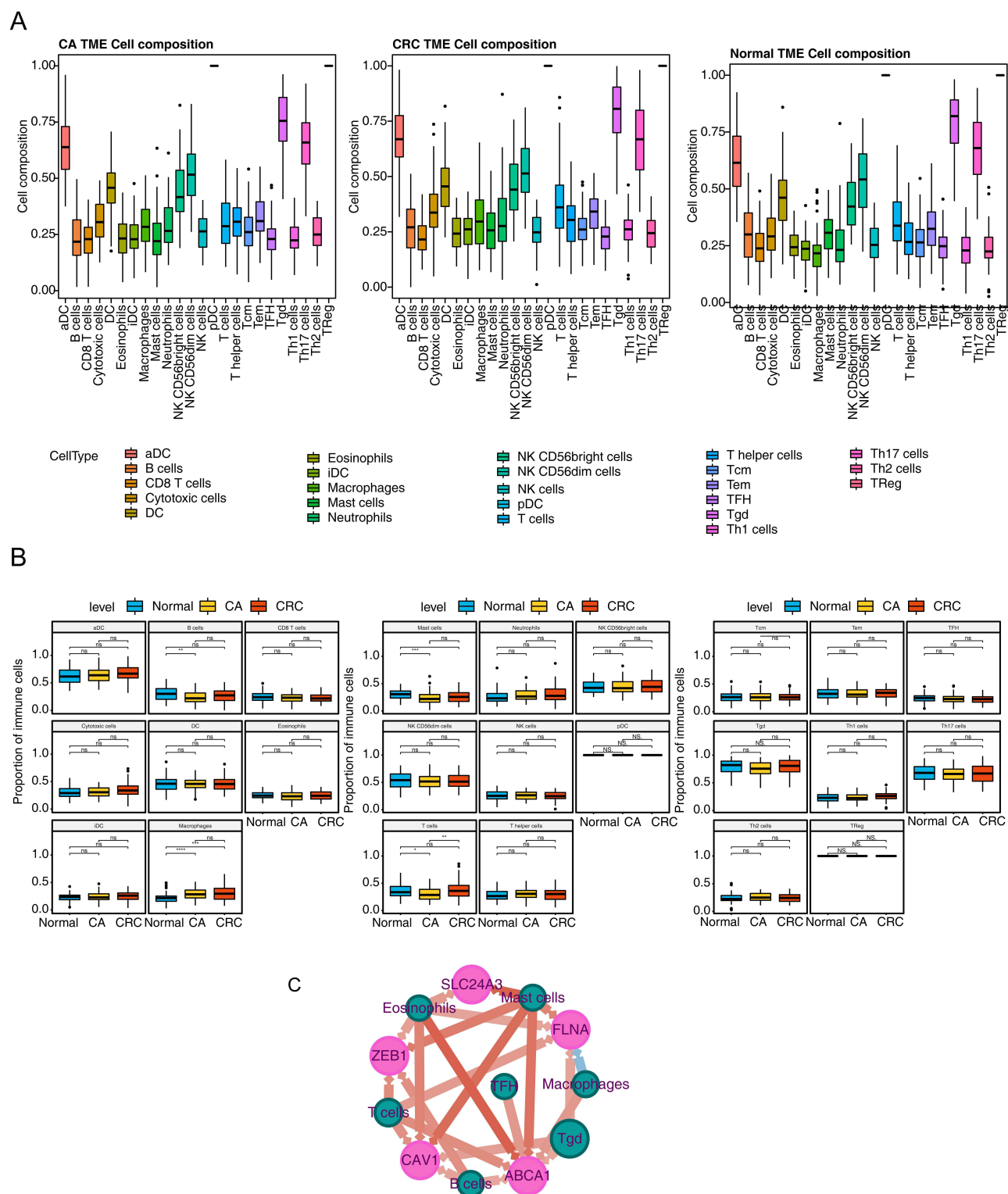
Dot plots were then employed to visualize the marker genes for each cell type, as depicted in Figure 7A. The distribution of the 10 identified cell subsets is presented in Figure 7B. Additionally, a heat map and a violin plot were used to illustrate the expression levels of the biomarkers within each cell subset, as shown in Figure 7C and D. It is noteworthy that nearly all biomarkers were predominantly expressed in fibroblasts, with FLNA displaying the highest expression levels (Figure 7C and D). Furthermore, the t-SNE cluster map provided a detailed visualization of feature gene expressions across fibroblasts, lymphocytes, goblet cells, B cells, and macrophages (Figure 7E).

ceRNA Network and Potential Therapeutic Drugs for Biomarkers

A comprehensive analysis of the five biomarkers identified 57 miRNAs and 21 lncRNAs, which facilitated the construction of a ceRNA network (Figure 8A). The findings indicated that AC138035.1 modulates CAV1 expression via hsa-let-7a-5p or hsa-let-7b-5p, AC240565.2 potentially influences ZEB1 expression through hsa-let-7b-5p, and LINC01001 similarly regulates CAV1 via hsa-let-7b-5p. Additionally, utilizing the DGIdb database, 12 potential drugs were identified targeting the four biomarkers (ABCA1, ZEB1, FLNA, and CAV1). Notably, the analysis revealed a strong binding affinity between ABCA1 and probucol, ZEB1 and salinomycin, as well as FLNA and PTI-125 (Figure 8B).

Verification of Feature Gene Expression Using qPCR and the HPA Database

The expression of the biomarkers (ZEB1, ABCA1, SLC24A3, CAV1, and FLNA) was validated through qPCR and analysis using the HPA database. qPCR results revealed that the mRNA levels of these genes in samples from patients with CA were significantly lower compared to those with NC (Figure 9), corroborating the findings from the bioinformatics analysis. Additionally, protein levels of the five biomarkers were reduced in samples from patients with CRC relative to those with NC, consistent with the expression patterns observed in the publicly available GSE117606 dataset (Figure 10A–E).

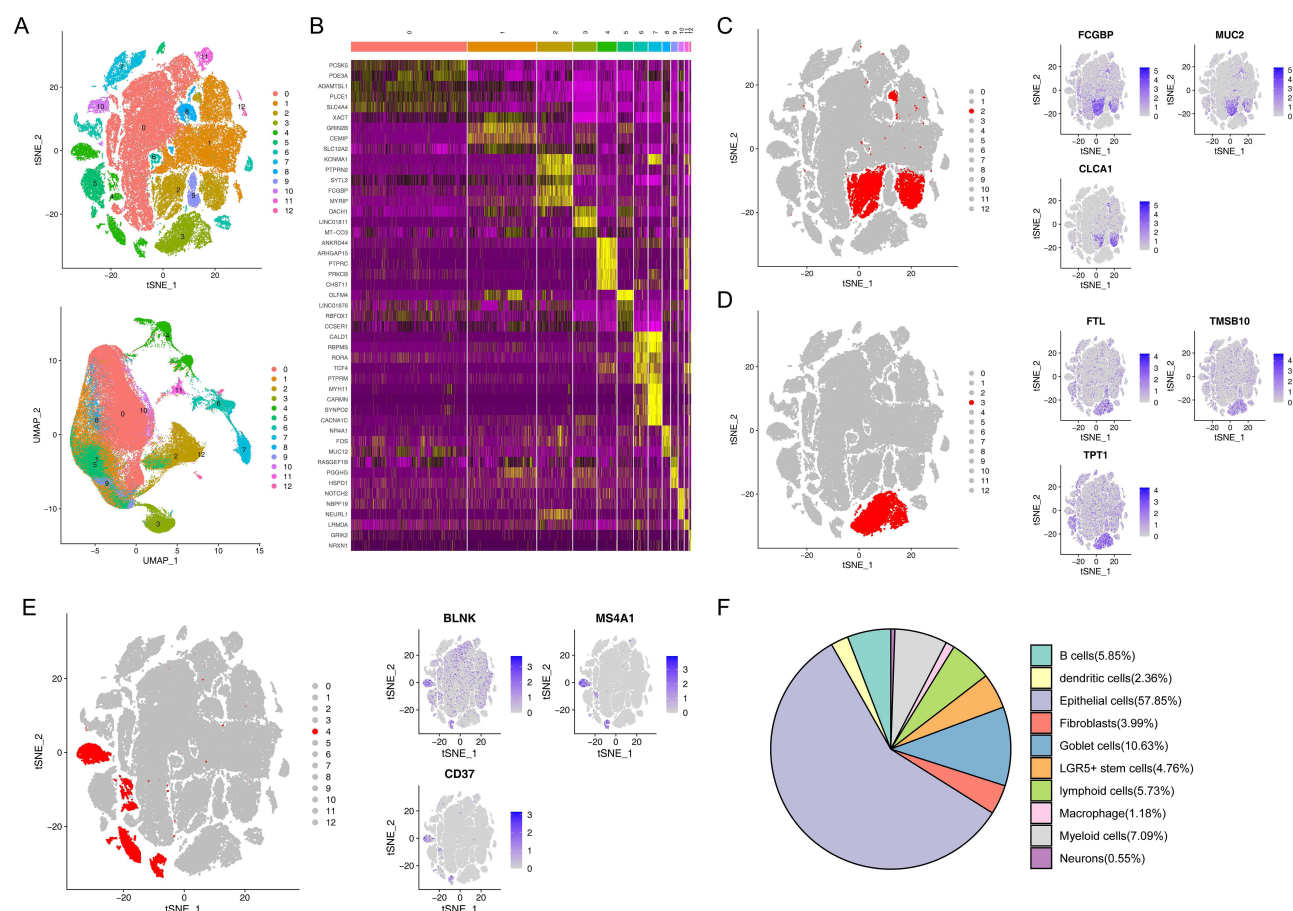**Figure 5** Immune heterogeneity analysis. (**A**) Composition of the 24 immune cell types in samples from patients with CA, CRC, and NC; (**B**) Differences in the proportions of immune cells between different groups, yellow for CA, red for CRC, blue for NC; (**C**) Correlation network of biomarkers and immune cells, connected lines represent relationships with mutualisms (ns, not significant, * p <0.05, ** p <0.01, *** p <0.001, **** p <0.0001).

Figure 6 Single-cell RNA-sequencing analysis to identify the different cell subsets in GSE201348. (**A**) t-Distributed Stochastic Neighbor Embedding (tSNE) plot and Uniform Manifold Approximation and Projection (UMAP) plot for 13 principal components (PCs) identified using principal component analysis (PCA). Each PC captures some biological differences, with the earlier PCs containing more information about these differences; (**B**) Expression heatmap for the top 5 marker genes among 13 clusters; (**C**) t-SNE plots labeling cluster 2 as goblet cells based on the expression patterns of FCGBP, MUC2, and CLCA1; (**D**) t-SNE plots labeling cluster 3 as myeloid cells based on the expression patterns of FTL, TMSB10, and TPT1; (**E**) t-SNE plots labeling cluster 4 as B cells according to the expression patterns of BLNK, MS4A1, and CD37; (**F**) A pie chart showing the proportion of 10 cell subsets in samples from patients with NC, CA, and CRC.

Discussion

Early detection of CA is essential for preventing the progression to CRC. Thus, identifying CMRGs linked to both CA and CRC is of substantial clinical importance, as it enables timely interventions. The role of copper metabolism and homeostasis in the development of CA and CRC has garnered increasing interest. Building on the potential relevance of CMRGs in colon cancer,[22,23] a comprehensive bioinformatics analysis was undertaken to identify key genes associated with copper metabolism that contribute to the progression of CA and CRC.

Through this analysis, five biomarkers—ZEB1, ABCA1, SLC24A3, CAV1, and FLNA—were identified. Evaluating the diagnostic value of each gene individually and collectively revealed that ZEB1, ABCA1, SLC24A3, CAV1, and FLNA possess strong discriminatory power between CA and NC.

Zinc finger E-box binding homeobox 1 (ZEB1) plays a pivotal role in the epithelial-mesenchymal transition (EMT) of tumor cells, a process essential for metastasis in colon cancer cell lines.[37,39–41] Silencing ZEB1 has been shown to inhibit the proliferation and metastatic potential of colon cancer cell lines, such as LoVo.[42] Studies have shown that during the adenoma (CA) stage, ZEB1 expression is generally downregulated, which may be related to maintaining normal epithelial architecture and function, thereby restricting abnormal cell migration and proliferation.[43] However, as CA progresses to colorectal cancer (CRC), ZEB1 expression increases significantly.[44] Elevated ZEB1 expression promotes epithelial–mesenchymal transition (EMT), facilitating tumor cell detachment from the primary site and contributing to cancer metastasis.[45] ZEB1-mediated regulation of colorectal cancer by cancer-associated fibroblasts (CAFs) has revealed that heterogeneous expression and defects
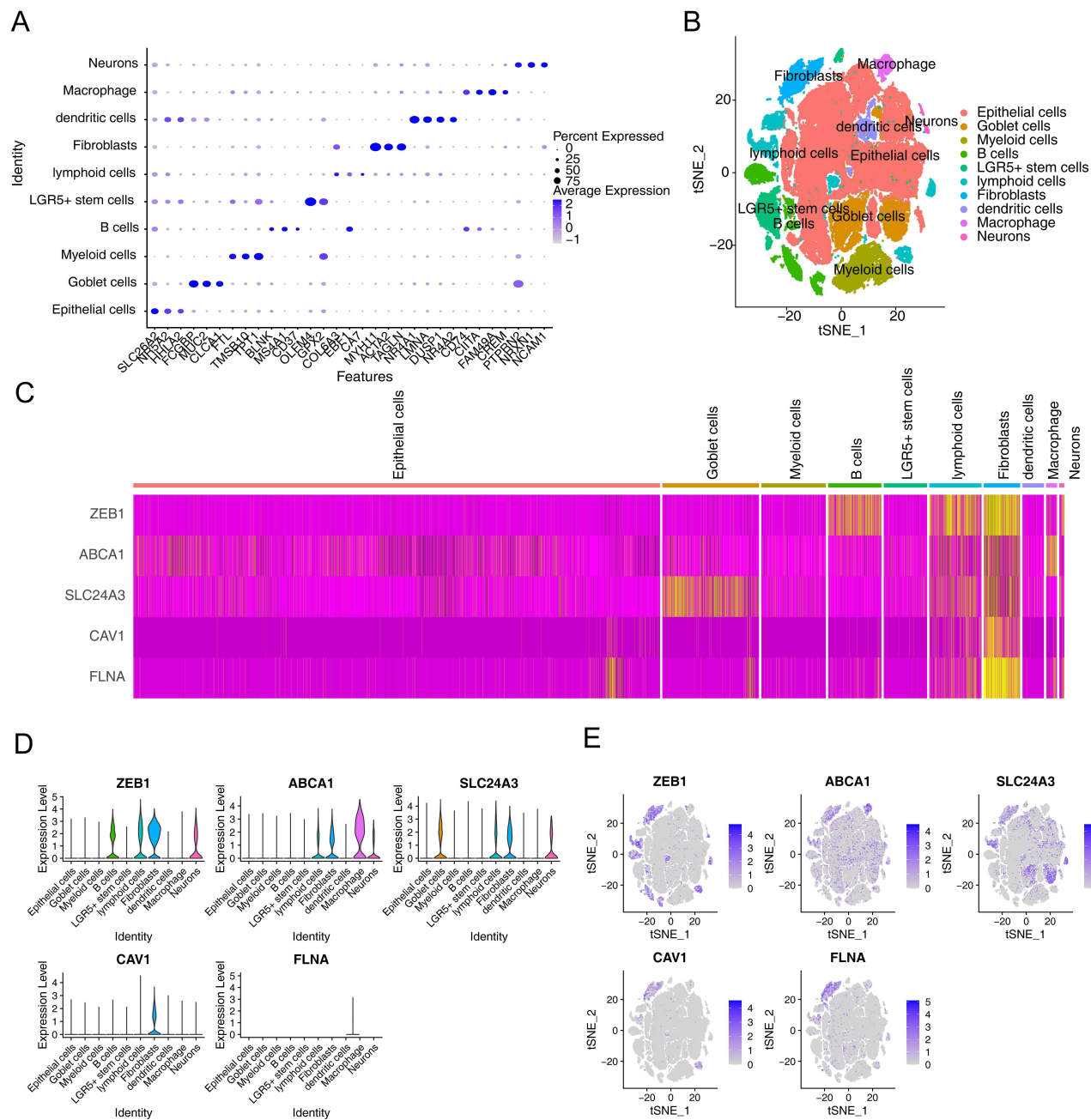
Figure 7 Single-cell RNA-sequencing analysis for biomarkers. (**A**) Dot plots of marker gene expressions for 10 cell subsets, the larger the diameter, the higher the expression; (**B**) t-SNE plots for 10 cell subsets; (**C**) Expression heatmap for five biomarkers among 10 cell subsets; (**D**) Violin plot for feature gene expression in 10 cell subsets; (**E**) t-SNE cluster map of five biomarkers in different cell subsets, the bluer the color, the higher the expression level.

in ZEB1 within these fibroblasts not only compromise the stability of the extracellular matrix (ECM) and the barrier function of CAFs but also collectively contribute to carcinogenesis, suppression of anti-tumor immunity, and promotion of metastasis. This occurs through the upregulation of cytokines mediated by Nuclear Factor kappa-light-chain-enhancer of activated B cells (NF-κB) and the activation of immune checkpoints, underscoring the multifaceted role of ZEB1 in the tumor microenvironment and its significance as a potential therapeutic target.[46] Copper played a critical role in cellular redox reactions, and its metabolic imbalance could lead to the generation of oxidative stress.[47] ZEB1 has been shown to repress E-cadherin (CDH1), thereby promoting epithelial–mesenchymal transition (EMT), and may also influence the expression of antioxidant enzymes such as superoxide dismutase (SOD) and glutathione peroxidase (GPX), altering the cellular redox state.[48] In colorectal cancer
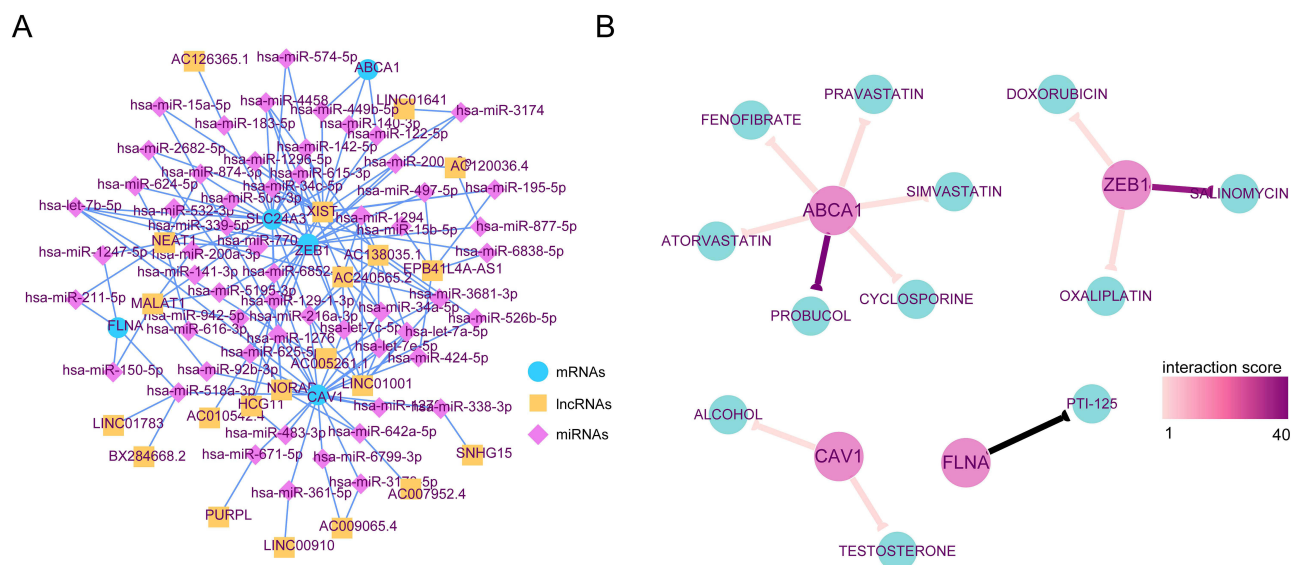
Figure 8 Exploration of the biomarkers. (**A**) LncRNA-miRNA-mRNA network relevant to biomarkers; (**B**) Feature gene-associated targeted drug network using the Drug-Gene Interaction database (DGIdb) database. The pink circle represents the biomarkers, the blue circle represents the targeted drugs, the lines represent the interaction scores, and the color of the lines deepen as the interaction score increases.
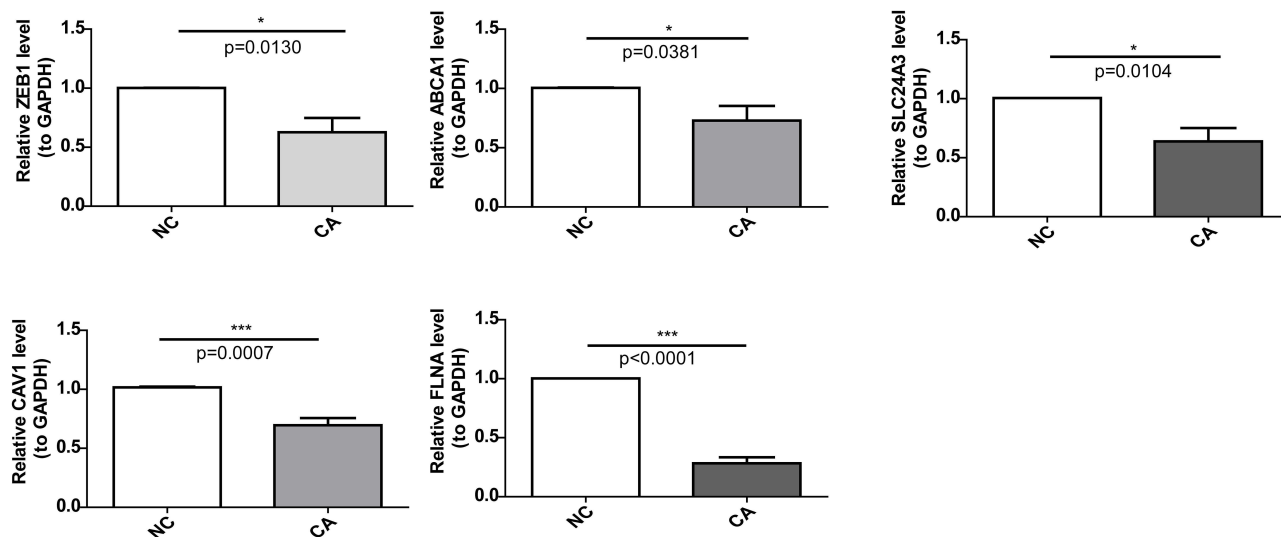

Figure 9 Validation of feature gene expression through the quantitative polymerase chain reaction (qPCR) experiments. * p <0.05, *** p <0.001.

(CRC), oxidative stress may contribute to disease pathogenesis by affecting processes such as intestinal epithelial cell proliferation, apoptosis, and inflammation.[49] Therefore, dysregulated copper metabolism may promote the progression of colorectal adenoma (CA) and CRC through oxidative stress, with ZEB1 potentially serving as a mechanistic link in this process. Our study observed that ZEB1 expression was downregulated in CA samples but upregulated in CRC, reflecting its dual role in cancer progression. However, this phenomenon warrants further investigation. ATP-binding cassette transporter A1 (ABCA1) is a transmembrane protein that mediates the reverse transport of cholesterol from cells to the circulatory system.[50,51] Wu et al proposed that inducing ABCA1 expression could be a novel cancer treatment strategy by reducing intracellular cholesterol levels through reverse cholesterol transport.[52] In the tumor microenvironment, copper metabolism dysregulation can lead to intracellular changes that may indirectly affect ABCA1-mediated cholesterol transport, thereby influencing tumor cell growth and metastasis.[53,54] Although the direct relationship between ABCA1 and epithelial–mesenchymal transition (EMT) remains unclear, aberrant cholesterol metabolism can alter membrane properties and impact cell–cell
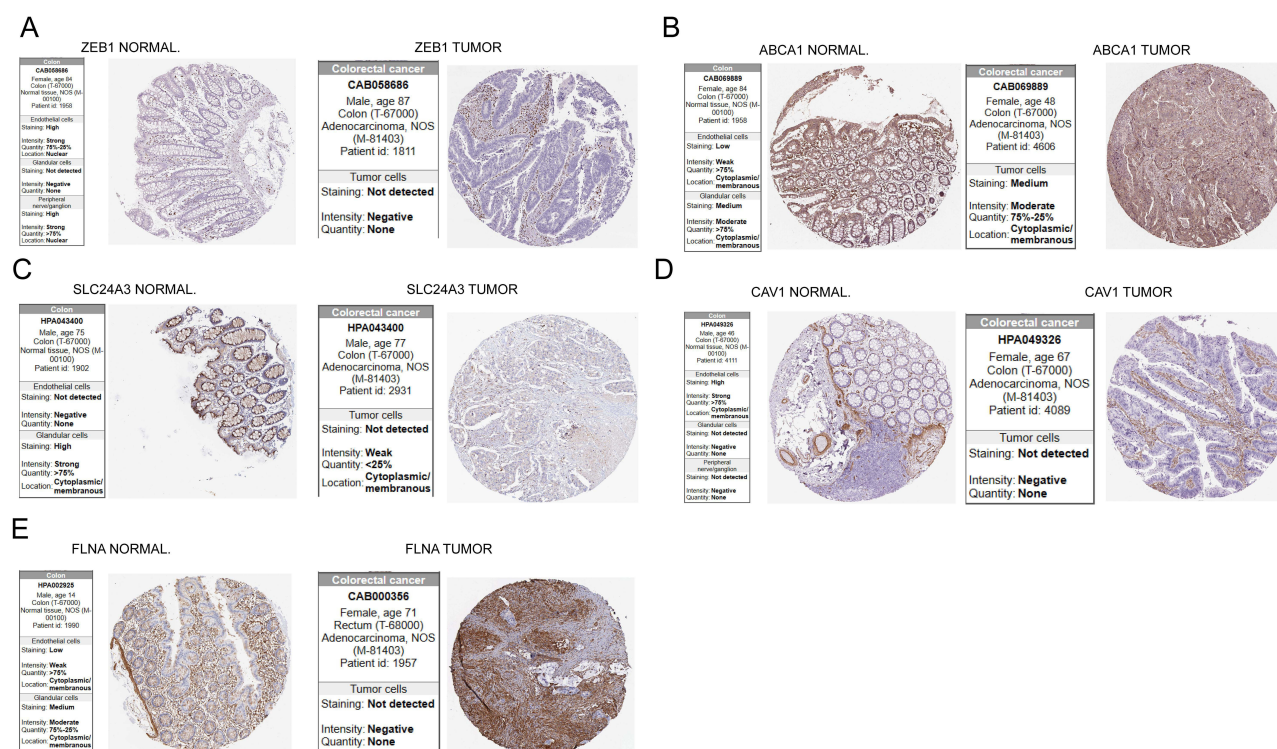
Figure 10 Exploration of the protein expressions of the biomarkers in CRC tissue samples and normal colon samples. (**A**) ZEB1; (**B**) ABCA1; (**C**) SLC24A3; (**D**) CAV1; (**E**) FLNA.

adhesion, potentially interacting with EMT processes. Copper-mediated regulation of ABCA1 may serve as a mechanistic link in this context.[55,56] Abnormal ABCA1 expression has been documented in colon cancer tissues and cells, where dysregulated intracellular cholesterol, driven by ABCA1 overexpression in advanced CRC, may contribute to tumor growth and metastasis. Conversely, silencing ABCA1 has been found to promote colon cancer cell proliferation and inhibit apoptosis.[50,57] This study similarly observed downregulation of ABCA1 in patients with CA, which reversed during the progression to CRC, paralleling the pattern seen with ZEB1. This phenomenon supports a potential role for ABCA1 in tumour progression. Caveolin-1 (CAV1), a membrane protein implicated in various cellular processes including endocytosis, extracellular matrix organization, and cholesterol distribution, has been identified as an oncogenic factor in several cancers, including liver, colon, breast, kidney, and lung cancers.[58,59] However, CAV1 expression varies depending on tumor type and stage. While it is downregulated in certain sarcomas and adenocarcinomas, it is significantly upregulated in advanced tumors like squamous cell carcinoma, a finding consistent with our results.[59] It is important to highlight that there are various copper - transport proteins, such as ATP7A and ATP7B, within cells to maintain copper homeostasis. Given the close relationship between the cell membrane environment housing CAV1 and these transport proteins, it is highly possible that CAV1 interacts with copper - transport proteins.[60] Research indicates that copper - homeostasis imbalance, either excess or deficiency, can lead to pathological changes like atherosclerosis.[61] These links imply a potentially complex connection between CAV1, copper - transport proteins, and copper homeostasis in tumor - related physiological and pathological processes. In colorectal adenoma (CA), CAV1 expression is usually low. Low CAV1 expression may limit the activation of cell - proliferation and migration - promoting signaling pathways, keeping the cell relatively stable.[62] Our study shows that low CAV1 expression is significantly related to neural - activity ligand - receptor interaction and the ribosome pathway. Neural - activity ligands like neurotransmitters may have their receptors' expression and function on the cell surface influenced by copper - ion levels.[63] When CA progresses to colorectal cancer (CRC), CAV1 expression increases. High CAV1 expression may boost endocytosis, affecting the number and distribution of cell - surface receptors and thus cell response to growth - factor - like signaling molecules.[64] Notably, Friedrich et al reported that the absence of CAV1, especially when combined with certain tumor suppressors and adenomatous polyposis in mouse models, may increase the risk of CRC.[59,65] In the study by Li et al, HSPA8 was found to be

transcriptionally upregulated in BRAF V600E mutant CRC, promoting the degradation of CAV1 and the release of β-catenin into the nucleus, thereby activating the Wnt/β-catenin pathway and driving metastasis and progression in BRAF V600E CRC.[66] However, the expression of CAV1 appears to be context-dependent. In our study, we observed that CAV1 expression was downregulated in colon adenomas (CA) but upregulated in CRC. Filamin A (FLNA), the most abundant and broadly distributed member of the actin-binding protein family, has diverse functions.[67,68] As essential cofactors for numerous enzymes, copper ions participate in multiple intracellular metabolic pathways and may indirectly modulate the expression and function of Filamin A (FLNA).[54,69] Our study reveals a significant association between ribosomal biogenesis pathways and FLNA underexpression. Copper-dependent enzymes, such as copper-zinc superoxide dismutase (SOD1), may contribute to rRNA stability or processing, thereby influencing ribosomal assembly.[70] In colorectal cancer (CRC), dysregulated copper metabolism could potentially drive tumor progression through ribosomal pathway activation.[71] Furthermore, within the tumor microenvironment, altered copper ion concentrations may indirectly affect FLNA's role in epithelial-mesenchymal transition (EMT) by modulating cytoskeleton-associated signaling pathways.[72] In the context of CRC, the clinical expression and biological role of FLNA remain controversial. Tian et al found that FLNA is downregulated in patients with CA and is closely associated with CRC development and progression, suggesting it as a prognostic marker.[73] Conversely, other research indicates that FLNA overexpression and phosphorylation at ser2152 are linked to snail-induced EMT and cell adhesion, proposing FLNA as a potential target for combined therapies in CRC progression.[74] This could explain the observed upregulation of FLNA in CA and its downregulation in CRC samples in this study. Cheng demonstrated that FLNA directly regulates metastasis and epithelial-mesenchymal transition (EMT) in chemotherapy-resistant CRC cells, suggesting its potential as a reliable CRC marker. This is consistent with our findings, which identified FLNA as a primary marker for CRC diagnosis.[75] Additionally, SLC24A3, a sodium-potassium-calcium exchanger, is associated with poor prognosis when expressed at high levels, indicating its utility as a prognostic marker, particularly in cervical cancer.[76,77] This study is the first to suggest a potential role for SLC24A3 in the initiation and progression of colon cancer. However, further research is required to elucidate the roles and underlying molecular mechanisms of these five characteristic genes in CA and CRC.

To further elucidate the relationship between characteristic genes and key immune cells in CA, immune cell infiltration was analyzed. Previous studies, such as those by Mo et al, have shown that the downregulation of the ITGA4 gene is positively correlated with increased proportions of mast cells and eosinophils in CA.[78] The protective and antitumor roles of mast cells in early-stage intestinal tumor models have also been demonstrated.[79] Consistent with these results, the biomarkers identified in this study exhibited significant positive correlations with the infiltration levels of these immune cells. Additionally, the association between a higher proportion of M0 macrophages, low FLNA expression, and poor prognosis in CRC, as documented in previous literature,[80] was confirmed by the current study in CA samples. These results underscore the close relationship between characteristic genes and immune cell infiltration in both CA and CRC. Subsequent scRNA-seq analysis revealed that these biomarkers were expressed across various cell types, including fibroblasts, lymphocytes, goblet cells, B cells, and macrophages. Notably, fibroblasts exhibited the highest expression of nearly all characteristic genes, with FLNA showing particularly elevated levels. The critical role of fibroblasts in tumor initiation and progression, where fibroblast-related characteristics serve as prognostic markers, has been well-documented.[81] Hiroki Kobayashi et al further highlighted the importance of cancer-associated fibroblasts in CRC progression and their link to poor patient prognosis.[82] These findings reinforce the diagnostic value of the selected characteristic genes in CRC.

To delve deeper into the pathways and functions associated with the identified characteristic genes, GSEA was conducted. The analysis revealed significant enrichment of these genes in several pathways, including the cGMP-PKG signaling pathway, Apelin signaling pathway, focal adhesion, and proteoglycans in cancer. Therapeutic activation of the cGMP-PKG pathway has been proposed as a promising strategy for colon cancer prevention and treatment, with PKG playing a central role in intestinal homeostasis and tumor suppression by inhibiting the β-catenin, T-cell factor (TCF), and SOX9 signaling pathways, thereby blocking cell proliferation and tumor angiogenesis.[83] Additionally, increasing evidence supports the anti-tumor signaling capabilities of PKG1 in colon cancer cell lines.[84] The Apelin pathway, which activates G-protein coupled receptors and Apelin receptors, also exerts various biological functions.[85,86] Notably, Picault et al found that the Apelin gene was upregulated in approximately half of the CA cases, suggesting that the Apelin signaling pathway may contribute to colon cancer progression by activating anti-apoptotic pathways.[87] The focal

adhesion signaling network, composed of integrins, growth factor receptors, and other survival-promoting molecules, represents another potential target for cancer treatment.[88] Moreover, studies have shown that syndecan-2, a cell surface proteoglycan, is upregulated in CRC and promotes cell migration, while the expression of the antineoplastic molecules syndecan-1 and syndecan-4 is reduced.[89] These results suggest that the dysregulation of these biological processes is closely linked to CRC onset and development, highlighting the potential of the five biomarkers as indicators of pathways associated with CRC progression in patients with CA. The incorporation of these genes into diagnostic strategies and treatment plans holds significantly to enhance our ability to manage CRC. Their expression levels in specific cell types, notably fibroblasts and immune cells, suggest that they may also significantly influence the tumour microenvironment and impact response to treatment.

Moreover, an upstream regulatory network involving lncRNA-miRNA-mRNA interactions was constructed for the characteristic genes linked to disease progression. Vahid et al proposed that MALAT1, regulated by the lncRNA-p53 network, could serve as an early detection biomarker for CRC.[90] In this study, MALAT1 was found to competitively bind to miR-141-3p, thereby influencing the expression of SLC24A3. Additionally, Mezgebe et al demonstrated that miRNA let-7c, which interacts with CAV1 and ZEB1 in our study, was significantly upregulated in adenomas compared to healthy samples.[91] This finding may explain the downregulation of CAV1 and ZEB1 in CA samples. The constructed network offers a foundation for further exploration of the underlying mechanisms governing these genes in CA and CRC.

Among the drugs predicted to target the characteristic genes involved in disease progression, probucol, which targets ABCA1, has been shown to effectively inhibit ABCA1-mediated cholesterol efflux,[92] a process implicated in the development of various tumors. CAV1 demonstrates a strong binding affinity for testosterone, and studies have shown that both endogenous and exogenous testosterone can exacerbate azoxymethane (AOM) or dextran sodium sulfate (DSS)-induced colitis and its carcinogenic effects.[93] Another potential drug, salinomycin, targets ZEB1 and has shown promise in eradicating cancer stem cells (CSC) and treating drug-resistant cancer cells, suggesting its potential as an effective chemotherapy agent.[94] Wang et al found that salinomycin significantly reduced tumor growth and CSC-related Wnt target gene expression, including LGR5, in colorectal tumor xenografts and APCmin/+ transgenic mice.[95] Therefore, the predicted drugs highlighted above hold considerable promise as potential therapeutic candidates for the treatment and prevention of CA and CRC.

While our study provides valuable insights, it is important to acknowledge its limitations. The current sample may not fully capture the heterogeneity of all colon adenoma (CA) patients. While this study provides valuable insights into the roles of copper metabolism and the immune microenvironment in CA, several limitations should be acknowledged. The current sample cohort may not fully capture the heterogeneity of CA patients. The qPCR validation was performed with a limited sample size and lacks long-term clinical follow-up. Subsequent studies should expand the sample size to include individuals of diverse ethnicities, ages, genders, and clinical stages, coupled with longitudinal monitoring of gene expression patterns and disease progression to enhance the generalizability and reliability of the findings. Furthermore, the stability of signature gene expression may vary across different disease stages, necessitating longitudinal studies to validate their clinical utility. We plan to systematically analyze existing datasets, incorporating clinical staging, family history, and other variables to construct a nomogram for evaluating associations and interactions with the five key genes, thereby improving predictive accuracy. Addressing these limitations is essential for translating these research findings into clinical applications.

To enhance the scientific and translational value of the study, animal models and cell culture experiments should be conducted in the future to validate the biological functions of these signature genes by knocking out or overexpressing them, thereby confirming their roles in the pathogenesis of CA. Secondly, the dynamic trends of cancer characteristics were effectively captured by monitoring the biomarkers through longitudinal studies using the Smoothing Cubic B-Spline Model,[96] which combines multi-omics data and reinforcement learning state behaviour reward state Behaviour (SARSA) hybrid cancer prediction model, which lays a solid foundation for subsequent bioinformatics analyses.[97–99] Combined with multi-centre and large-sample clinical studies, we will assess its clinical application value in different stages of the disease and enhance the generality of the findings.

Through the efforts in the above directions, we expect to overcome the limitations of current studies, further reveal the molecular mechanisms of CA, promote clinical translation, and open up new avenues for precision diagnosis and treatment of CA.

Conclusions

In conclusion, this study successfully identified and validated Cu metabolism-related biomarkers associated with CA and CRC through comprehensive bioinformatics approaches. The research also examined the correlation between the biomarkers and immune infiltration, alongside predicting potential targeted drugs for CA treatment. Additionally, a ceRNA regulatory network was constructed to elucidate lncRNA-miRNA-mRNA interactions, offering valuable insights for future mechanistic studies. Overall, this study lays a solid foundation for understanding the molecular mechanisms underlying these diseases.

Abbreviations

TCGA, The Cancer Genome Atlas; LUAD, Lung Adenocarcinoma; GEO, Gene Expression Omnibus; FRGs, Ferroptosis-Related Genes; ERGs, Epithelial-Mesenchymal Transition-Related Genes; DEGs, Differentially Expressed Genes; PCA, Principal Component Analysis; CA, Colon Adenoma; CRC, Colorectal Cancer; Cu, Copper; CMRGs, Copper Metabolism-Related Genes; DEGs1, Differentially Expressed Genes 1; NC, Normal Controls; DEGs2, Differentially Expressed Genes 2; ssGSEA, Single-Sample Gene Set Enrichment Analysis; TME, Tumor Microenvironment; GSEA, Gene Set Enrichment Analysis; IPA, Ingenuity Pathway Analysis; AUC, Area Under the Curve; RF, Random Forest; SVM, Support Vector Machines; LR, Logistic Regression; GBDT, Gradient Boosting Decision Tree; ANN, Artificial Neural Network; qPCR, Quantitative Polymerase Chain Reaction; HPA, Human Protein Atlas; MSI, Microsatellite Instability; HNPCC, Hereditary Nonpolyposis Colorectal Cancer; SETD2, Histone H3K36 Trimethyltransferase; MLH1, MutL Homolog 1; MSH2, MutS Homolog 2; MSH6, MutS Homolog 6; scRNA-seq, Single-Cell RNA Sequencing; UMAP, Uniform Manifold Approximation and Projection; t-SNE, t-Distributed Stochastic Neighbor Embedding; DE-CMRGs1, Differentially Expressed Copper Metabolism-Related Genes 1; DE-CMRGs2, Differentially Expressed Copper Metabolism-Related Genes 2; cGMP-PKG, Cyclic Guanosine Monophosphate-Protein Kinase G; MAPK, Mitogen-Activated Protein Kinases; EMT, Epithelial-Mesenchymal Transition; NF-κB, Nuclear Factor kappa-light-chain-enhancer of activated B cells; ECM, Extracellular Matrix; CAFs, Cancer-Associated Fibroblasts; DGIdb, Drug-Gene Interaction database; LASSO, Least Absolute Shrinkage and Selection Operator; IoT, Internet of Things; AFRLS, Adaptive Federated Reinforcement Learning-enabled System; SARSA, State Action Reward State Action.

Data Sharing Statement

The datasets analysed during the current study are available in the GEO database (https://www.ncbi.nlm.nih.gov/geo/), reference number [GSE117606, GSE20916 and GSE201348].

Ethics Approval and Informed Consent

This study was in compliance with the Declaration of Helsinki, was reviewed and approved by [the Ethics Committee of The Second Affiliated Hospital of Guizhou University of Traditional Chinese Medicine] with the approval number [KYW202209], dated [6/9/2022]. Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patients to publish this paper.

Consent for Publication

The authors agree to publish.

Acknowledgments

https://doi.org/10.2147/IJGM.S521512 **3039**

Author Biography

Professor Taikun Zhang, a chief physician with nearly three decades of experience, specializes in gastroenterology. Professor Zhang is a member of the Ultrasound Endoscopy and ERCP groups. He has authored over 20 academic papers, including seven in key national journals, and has led research projects funded by the Health Department of Guangdong Province, the Science and Technology Bureau of Dongguan City, and the Health and Family Planning Commission of Guizhou Province. Professor Zhang's practice focuses on the treatment of digestive system diseases, including chronic gastritis and gastrointestinal tumors, and he is skilled in endoscopic procedures such as EMR, ESD, EVL, and PEG.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This research received no external funding.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Matsuda T, Fujimoto A, Igarashi Y. Colorectal cancer: epidemiology, risk factors, and public health strategies. *Digestion*. 2025;106(2):91–99. doi:10.1159/000543921
2. Sharma R, Abbasi-Kangevari M, Abd-Rabu R; GBD. Colorectal cancer collaborators. global, regional, and national burden of colorectal cancer and its risk factors, 1990-2019: a systematic analysis for the global burden of disease study 2019. *Lancet Gastroenterol Hepatol*. 2022;7:627–647. doi:10.1016/S2468-1253(22)00044-9
3. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut*. 2017;66(4):683–691. doi:10.1136/gutjnl-2015-310912
4. Wang H, Li YJ, Lei L, et al. Estimating the economic burden of colorectal cancer in China, 2019-2030: a population-level prevalence-based analysis. *Cancer Med*. 2024;13:e6787.
5. Alaiyan B, Ilyayev N, Stojadinovic A, et al. Differential expression of colon cancer associated transcript1 (CCAT1) along the colonic adenoma-carcinoma sequence. *BMC Cancer*. 2013;13:196. doi:10.1186/1471-2407-13-196
6. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, La D Jr, Kinzler KW. Cancer genome landscapes. *Science*. 2013;339:1546–1558. doi:10.1126/science.1235122
7. Fearon ER. Molecular genetics of colorectal cancer. *Annu Rev Pathol*. 2011;6:479–507. doi:10.1146/annurev-pathol-011110-130235
8. Lacalamitaa A, Piccinno E, Scalavino V, Bellotti R, Giannelli G, Serino G. A gene-based machine learning classifier associated to the colorectal adenoma-carcinoma sequence. *Biomedicines*. 2021;9:1937. doi:10.3390/biomedicines9121937
9. Peng H, Liu G, Bao Y, et al. Prognostic factors of colorectal cancer: a comparative study on patients with or without liver metastasis. *Front Oncol*. 2021;11:626190. doi:10.3389/fonc.2021.626190
10. Fu T, Coulter S, Yoshihara E, et al. FXR regulates intestinal cancer stem cell proliferation. *Cell*. 2019;176:1098–1112.e18. doi:10.1016/j.cell.2019.01.036
11. Zhu G, Pei L, Xia H, Tang Q, Bi F. Role of oncogenic KRAS in the prognosis, diagnosis and treatment of colorectal cancer. *Mol Cancer*. 2021;20:143. doi:10.1186/s12943-021-01441-4
12. Dariya B, Aliya S, Merchant N, Alam A, Nagaraju GP. Colorectal cancer biology, diagnosis, and therapeutic approaches. *Crit Rev Oncol*. 2020;25:71–94. doi:10.1615/CritRevOncog.2020035067
13. Rashid MU, Naeemi H, Muhammad N, et al. Prevalence and spectrum of MLH1, MSH2, and MSH6 pathogenic germline variants in Pakistani colorectal cancer patients. *Hered Cancer Clin Pract*. 2019;17:29. doi:10.1186/s13053-019-0128-2
14. Ma C, Liu M, Feng W, et al. Loss of SETD2 aggravates colorectal cancer progression caused by SMAD4 deletion through the RAS/ERK signalling pathway. *Clin Transl Med*. 2023;13:e1475.
15. Taieb J, Svrcek M, Cohen R, Basile D, Tougeron D, Phelip JM. Deficient mismatch repair/microsatellite unstable colorectal cancer: diagnosis, prognosis and treatment. *Eur J Cancer*. 2022;175:136–157. doi:10.1016/j.ejca.2022.07.020
16. Oliveira ML, Biggers A, Oddo VM, et al. A perspective review on diet quality, excess adiposity, and chronic psychosocial stress and implications for early-onset colorectal cancer. *J Nutr*. 2024;154:1069–1079. doi:10.1016/j.tjnut.2024.03.002
17. Atkin W, Wooldrage K, Brenner A, et al. Adenoma surveillance and colorectal cancer incidence: a retrospective, multicentre, cohort study. *Lancet Oncol*. 2017;18:823–834. doi:10.1016/S1470-2045(17)30187-0
18. Thayalasekaran S, Alkandari A, Varytimiadis L, et al. To cap/cuff or ring: do distal attachment devices improve the adenoma detection? *Expert Rev Gastroenterol Hepatol*. 2019;13:119–127. doi:10.1080/17474124.2019.1551131

19. Imperiale TF, Gruber RN, Stump TE, Emmett TW, Monahan PO. Performance characteristics of fecal immunochemical tests for colorectal cancer and advanced adenomatous polyps: a systematic review and meta-analysis. *Ann Intern Med*. 2019;170:319–329. doi:10.7326/M18-2390

20. Nevitt T, Ohrvik H, Thiele DJ. Charting the travels of copper in eukaryotes from yeast to mammals. *Biochim Biophys Acta*. 2012;1823:1580–1593. doi:10.1016/j.bbamcr.2012.02.011

21. Maung MT, Carlson A, Olea-Flores M, et al. The molecular and cellular basis of copper dysregulation and its relationship with human pathologies. *FASEB J*. 2021;35:e21810.

22. Liao Y, Zhao J, Bulek K, et al. Inflammation mobilizes copper metabolism to promote colon tumorigenesis via an IL-17-STEAP4-XIAP axis. *Nat Commun*. 2020;11:900. doi:10.1038/s41467-020-14698-y

23. Kim ES, Lim CS, Chun HJ, et al. Detection of Cu(I) and Zn(II) ions in colon tissues by multi-photon microscopy: novel marker of antioxidant status of colon neoplasm. *J Clin Pathol*. 2012;65:882–887. doi:10.1136/jclinpath-2012-200666

24. Lin J, Luo B, Yu X, Yang Z, Wang M, Cai W. Copper metabolism patterns and tumor microenvironment characterization in colon adenocarcinoma. *Front Oncol*. 2022;12:959273. doi:10.3389/fonc.2022.959273

25. Al Mehedi Hasan M, Maniruzzaman M, Shin J. Identification of key candidate genes for IgA nephropathy using machine learning and statistics based bioinformatics models. *Sci Rep*. 2022;12:13963. doi:10.1038/s41598-022-18273-x

26. Liu Y, Liang Y, Su Y, et al. Exploring the potential mechanisms of Yi-Yi-Fu-Zi-Bai-Jiang-San therapy on the immune-inflamed phenotype of colorectal cancer via combined network pharmacology and bioinformatics analyses. *Comput Biol Med*. 2023;166:107432. doi:10.1016/j.compbiomed.2023.107432

27. Chen L, Lu D, Sun K, et al. Identification of biomarkers associated with diagnosis and prognosis of colorectal cancer patients based on integrated bioinformatics analysis. *Gene*. 2019;692:119–125. doi:10.1016/j.gene.2019.01.001

28. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47. doi:10.1093/nar/gkv007

29. Li GM, Zhang CL, Rui RP, Sun B, Guo W. Bioinformatics analysis of common differential genes of coronary artery disease and ischemic cardiomyopathy. *Eur Rev Med Pharmacol Sci*. 2018;22:3553–3569. doi:10.26355/eurrev_201806_15182

30. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinf*. 2011;12:35. doi:10.1186/1471-2105-12-35

31. Zhao S, Wang L, Ding W, et al. Crosstalk of disulfidptosis-related subtypes, establishment of a prognostic signature and immune infiltration characteristics in bladder cancer based on a machine learning survival framework. *Front Endocrinol*. 2023;14:1180404. doi:10.3389/fendo.2023.1180404

32. Zhao S, Zhang L, Ji W, et al. Machine learning-based characterization of cuprotosis-related biomarkers and immune infiltration in Parkinson's disease. *Front Genet*. 2022;13:1010361. doi:10.3389/fgene.2022.1010361

33. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf*. 2011;12:77. doi:10.1186/1471-2105-12-77

34. Yu G, Wang LG, Han Y, He, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16:284–287. doi:10.1089/omi.2011.0118

35. Xu Q, Xu H, Deng R, et al. Multi-omics analysis reveals prognostic value of tumor mutation burden in hepatocellular carcinoma. *Cancer Cell Int*. 2021;21:342. doi:10.1186/s12935-021-02049-w

36. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. 2016;352:189–196. doi:10.1126/science.aad0501

37. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184:3573–3587.e3529. doi:10.1016/j.cell.2021.04.048

38. Pastushenko I, Blanpain C. EMT transition states during tumor progression and metastasis. *Trends Cell Biol*. 2019;29:212–226. doi:10.1016/j.tcb.2018.12.001

39. Liu X, Wang H, Wang X, et al. Identification and verification of inflammatory biomarkers for primary Sjögren's syndrome. *Clin Rheumatol*. 2024;43(4):1335–1352. doi:10.1007/s10067-024-06901-y

40. Feigelman G, Simanovich E, Brockmeyer P, Rahat MA. Knocking-Down CD147/EMMPRIN Expression in CT26 colon carcinoma forces the cells into cellular and angiogenic dormancy that can be reversed by interactions with macrophages. *Biomedicines*. 2023;11:768. doi:10.3390/biomedicines11030768

41. Wang X, Lai Q, He J, et al. LncRNA SNHG6 promotes proliferation, invasion and migration in colorectal cancer cells by activating TGF-β/Smad signaling pathway via targeting UPF1 and inducing EMT via regulation of ZEB1. *Int J Med Sci*. 2019;16(1):51–59. doi:10.7150/ijms.27359

42. Luo D, Ge W. MeCP2 promotes colorectal cancer metastasis by modulating ZEB1 Transcription. *Cancers*. 2020;12:349. doi:10.3390/cancers12030758

43. Li J, Xia L, Zhou Z, et al. MiR-186-5p upregulation inhibits proliferation, metastasis and epithelial-to-mesenchymal transition of colorectal cancer cell by targeting ZEB1. *Arch Biochem Biophys*. 2018;640:53–60. doi:10.1016/j.abb.2018.01.002

44. Kobayashi Y, Yokoi A, Hashimura M, et al. Nucleobindin-2 mediates transforming growth factor-β1-driven phenotypes in zinc finger e-box binding homeobox 1-high uterine carcinosarcoma. *Am J Pathol*. 2023;193(8):1116–1128. doi:10.1016/j.ajpath.2023.04.011

45. Zhang GJ, Zhou T, Tian HP, Liu ZL, Xia SS. High expression of ZEB1 correlates with liver metastasis and poor prognosis in colorectal cancer. *Oncol Lett*. 2013;5(2):564–568. doi:10.3892/ol.2012.1026

46. Menche C, Schuhwerk H, Armstark I, et al. ZEB1-mediated fibroblast polarization controls inflammation and sensitivity to immunotherapy in colorectal cancer. *EMBO Rep*. 2024;25(8):3406–3431. doi:10.1038/s44319-024-00186-7

47. Gupta R, Verma N, Tewari RK. Micronutrient deficiency-induced oxidative stress in plants. *Plant Cell Rep*. 2024;43(9):213. doi:10.1007/s00299-024-03297-6

48. Drápela S, Bouchal J, Jolly MK, Culig Z, Souček K. ZEB1: a critical regulator of cell plasticity, DNA damage response, and therapy resistance. *Front Mol Biosci*. 2020;7:36. doi:10.3389/fmolb.2020.00036

49. Rossin D, Calfapietra S, Sottero B, Poli G, Biasi F. HNE and cholesterol oxidation products in colorectal inflammation and carcinogenesis. *Free Radic Biol Med*. 2017;111:186–195. doi:10.1016/j.freeradbiomed.2017.01.017

50. Aguirre-Portolés C, Feliu J, Reglero G, Ramírez de Molina A. ABCA1 overexpression worsens colorectal cancer prognosis by facilitating tumour growth and caveolin-1-dependent invasiveness, and these effects can be ameliorated using the BET inhibitor apabetalone. *Mol Oncol*. 2018;12:1735–1752. doi:10.1002/1878-0261.12367

51. Oram JF, Lawn RM. ABCA1. The gatekeeper for eliminating excess tissue cholesterol. *J Lipid Res*. 2001;42:1173–1179. doi:10.1016/S0022-2275(20)31566-2

52. Wu K, Zou L, Lei X, Yang X. Roles of ABCA1 in cancer. *Oncol Lett*. 2022;24:349. doi:10.3892/ol.2022.13469

53. Linke JA, Munn LL, Jain RK. Compressive stresses in cancer: characterization and implications for tumour progression and treatment. *Nat Rev Cancer*. 2024;24(11):768–791. doi:10.1038/s41568-024-00745-z

54. Kamiya T. Copper in the tumor microenvironment and tumor metastasis. *J Clin Biochem Nutr*. 2022;71(1):22–28. doi:10.3164/jcbn.22-9

55. Yang Y, Li S, Wang Y, Zhao Y, Li Q. Protein tyrosine kinase inhibitor resistance in malignant tumors: molecular mechanisms and future perspective. *Signal Transduct Target Ther*. 2022;7(1):329. doi:10.1038/s41392-022-01168-8

56. Luo L, Zhang W, You S, et al. The role of epithelial cells in fibrosis: mechanisms and treatment. *Pharmacol Res*. 2024;202:107144. doi:10.1016/j.phrs.2024.107144

57. Bi DP, Yin CH, Zhang XY, Yang NN, Xu JY. MiR-183 functions as an oncogene by targeting ABCA1 in colon cancer. *Oncol Rep*. 2016;35:2873–2879. doi:10.3892/or.2016.4631

58. Nwosu ZC, Ebert MP, Dooley S, Meyer C. Caveolin-1 in the regulation of cell metabolism: a cancer perspective. *Mol Cancer*. 2016;15:71. doi:10.1186/s12943-016-0558-7

59. Burgermeister E, Liscovitch M, Röcken C, Schmid RM, Ebert MP. Caveats of caveolin-1 in cancer progression. *Cancer Lett*. 2008;268:187–201. doi:10.1016/j.canlet.2008.03.055

60. Ruturaj MM, Maji S, Maji S, Rodriguez-Boulan E, Schreiner R, Gupta A. Regulation of the apico-basolateral trafficking polarity of the homologous copper-ATPases ATP7A and ATP7B. *J Cell Sci*. 2024;137(5):jcs261258. doi:10.1242/jcs.261258

61. Wang F, Gu HM, Zhang DW. Caveolin-1 and ATP binding cassette transporter A1 and G1-mediated cholesterol efflux. *Cardiovasc Hematol Disord Drug Targets*. 2014;14(2):142–148. doi:10.2174/1871529X14666140505122802

62. Castillo Bennett J, Silva P, Martinez S, Torres VA, Quest AFG. Hypoxia-induced caveolin-1 expression promotes migration and invasion of tumor cells. *Curr Mol Med*. 2018;18(4):199–206.

63. D'Ambrosi N, Rossi L. Copper at synapse: release, binding and modulation of neurotransmission. *Neurochem Int*. 2015;90:36–45. doi:10.1016/j.neuint.2015.07.006

64. Li J, Mi J, Yuan Q. Fatty acid metabolism-related enzymes in colorectal cancer metastasis: from biological function to molecular mechanism. *Cell Death Discov*. 2024;10(1):350. doi:10.1038/s41420-024-02126-9

65. Friedrich T, Richter B, Gaiser T, et al. Deficiency of caveolin-1 in Apc(min/+) mice promotes colorectal tumorigenesis. *Carcinogenesis*. 2013;34:2109–2118. doi:10.1093/carcin/bgt142

66. Li B, Ming H, Qin S, et al. HSPA8 activates wnt/β-catenin signaling to facilitate BRAF V600E colorectal cancer progression by CMA-Mediated CAV1 Degradation. *Adv Sci*. 2024;11:e2306535.

67. Hargrave TB, Aitken RJ, Elton RA. Prognostic significance of the zona-free hamster egg test. *Br J Urol*. 1988;62:603–608. doi:10.1111/j.1464-410X.1988.tb04436.x

68. Hartwig JH, Stossel TP. Macrophages. *J Biol Chem*. 1975;250:5696–5705. doi:10.1016/S0021-9258(19)41235-0

69. Huang M, Zhang Y, Liu X. The mechanism of cuproptosis in Parkinson's disease. *Ageing Res Rev*. 2024;95:102214. doi:10.1016/j.arr.2024.102214

70. Bellia F, Lanza V, Naletova I, et al. Complexes with carnosine conjugates of hyaluronic acids at different dipeptide loading percentages behave as multiple sod mimics and stimulate nrf2 translocation and antioxidant response in in vitro inflammatory model. *Antioxidants*. 12(8):1632

71. Qin K, Yu S, Liu Y, et al. USP36 stabilizes nucleolar Snail1 to promote ribosome biogenesis and cancer cell survival upon ribotoxic stress. *Nat Commun*. 2023;14(1):6473. doi:10.1038/s41467-023-42257-8

72. Yan T, Shi J. Angiogenesis and EMT regulators in the tumor microenvironment in lung cancer and immunotherapy. *Front Immunol*. 2024;15:1509195. doi:10.3389/fimmu.2024.1509195

73. Tian ZQ, Shi JW, Wang XR, Li Z, Wang GY. New cancer suppressor gene for colorectal adenocarcinoma: filamin A. *World J Gastroenterol*. 2015;21:2199–2205. doi:10.3748/wjg.v21.i7.2199

74. Wieczorek K, Wiktorska M, Sacewicz-Hofman I, et al. Filamin A upregulation correlates with Snail-induced epithelial to mesenchymal transition (EMT) and cell adhesion but its inhibition increases the migration of colon adenocarcinoma HT29 cells. *Exp Cell Res*. 2017;359:163–170. doi:10.1016/j.yexcr.2017.07.035

75. Cheng M, Jiang Y, Yang H, Zhao D, Lx L, Liu X. FLNA promotes chemoresistance of colorectal cancer through inducing epithelial-mesenchymal transition and smad2 signaling pathway. *Am J Cancer Res*. 2020;10:403–423.

76. Yu SH, Cai JH, Chen DL, et al. LASSO and bioinformatics analysis in the identification of key genes for prognostic genes of gynecologic cancer. *J Pers Med*. 2021;11:1177. doi:10.3390/jpm11111177

77. Zhou S, Jin Q, Yao H, et al. Pain-related gene solute carrier family 24 member 3 is a prognostic biomarker and correlated with immune infiltrates in cervical squamous cell carcinoma and endocervical adenocarcinoma: a study via integrated bioinformatics analyses and experimental verification. *Comput Math Methods Med*. 2023;2023:4164232. doi:10.1155/2023/4164232

78. Mo J, Zhang J, Huang H. The early predictive effect of low expression of the ITGA4 in colorectal cancer. *J Gastrointest Oncol*. 2022;13:265–278. doi:10.21037/jgo-22-92

79. Sinnamon MJ, Carter KJ, Sims LP, Lafleur B, Fingleton B, Matrisian LM. A protective role of mast cells in intestinal tumorigenesis. *Carcinogenesis*. 2008;29:880–886. doi:10.1093/carcin/bgn040

80. Wang D, Liufu J, Yang Q, Dai S, Wang J, Xie B. Identification and validation of a novel signature as a diagnostic and prognostic biomarker in colorectal cancer. *Biol Direct*. 2022;17:29. doi:10.1186/s13062-022-00342-w

81. Xu H, Pan Y. A prognostic fibroblast-related risk signature in colorectal cancer. *Aging*. 2021;13:24251–24270. doi:10.18632/aging.203677

82. Kobayashi H, Geniec KA, Lannagan TRM, et al. The origin and contribution of cancer-associated fibroblasts in colorectal carcinogenesis. *Gastroenterology*. 2022;162:890–906. doi:10.1053/j.gastro.2021.11.037

83. Browning DD, Kwon IK, Wang R. cGMP-dependent protein kinases as potential targets for colon cancer prevention and treatment. *Future Med Chem*. 2010;2:65–80. doi:10.4155/fmc.09.142

84. Kwon IK, Wang R, Thangaraju M, et al. PKG inhibits TCF signaling in colon cancer cells by blocking beta-catenin expression and activating FOXO4. *Oncogene*. 2010;29:3423–3434. doi:10.1038/onc.2010.91

85. de Oliveira AA, Vergara A, Wang X, Vederas JC, Oudit GY. Apelin pathway in cardiovascular, kidney, and metabolic diseases: therapeutic role of apelin analogs and apelin receptor agonists. *Peptides*. 2021;147:170697. doi:10.1016/j.peptides.2021.170697

86. Yao F, Lv YC, Zhang M, et al. Apelin-13 impedes foam cell formation by activating Class III PI3K/Beclin-1-mediated autophagic pathway. *Biochem Biophys Res Commun*. 2015;466:637–643. doi:10.1016/j.bbrc.2015.09.045

87. Picault FX, Chaves-Almagro C, Projetti F, Prats H, Masri B, Audigier Y. Tumour co-expression of apelin and its receptor is the basis of an autocrine loop involved in the growth of colon adenocarcinomas. *Eur J Cancer*. 2014;50:663–674. doi:10.1016/j.ejca.2013.11.017

88. Eke I, Cordes N. Focal adhesion signaling and therapy resistance in cancer. *Semin Cancer Biol*. 2015;31:65–75. doi:10.1016/j.semcancer.2014.07.009

89. Vicente CM, DA Silva DA, Sartorio PV, et al. Heparan sulfate proteoglycans in human colorectal cancer. *Anal Cell Pathol*. 2018;8389595.

90. Chaleshi V, Irani S, Alebouyeh M, Mirfakhraie R, Aghdaei HA. Association of lncRNA-p53 regulatory network (lincRNA-p21, lincRNA-ROR and MALAT1) and p53 with the clinicopathological features of colorectal primary lesions and tumors. *Oncol Lett*. 2020;19:3937–3949. doi:10.3892/ol.2020.11518

91. Gebrekristos M, Melson J, Jiang A, Buckingham L. DNA methylation and miRNA expression in colon adenomas compared with matched normal colon mucosa and carcinomas. *Int J Exp Pathol*. 2022;103:74–82. doi:10.1111/iep.12432

92. Favari E, Zanotti I, Zimetti F, Ronda N, Bernini F, Rothblat GH. Probucol inhibits ABCA1-mediated cellular lipid efflux. *Arterioscler Thromb Vasc Biol*. 2004;24:2345–2350. doi:10.1161/01.ATV.0000148706.15947.8a

93. Song CH, Kim N, Nam RH, et al. Testosterone strongly enhances azoxymethane/dextran sulfate sodium-induced colorectal cancer development in C57BL/6 mice. *Am J Cancer Res*. 2021;11:3145–3162.

94. Dewangan J, Srivastava S, Rath SK. Salinomycin: a new paradigm in cancer therapy. *Tumour Biol*. 2017;39:1010428317695035. doi:10.1177/1010428317695035

95. Wang Z, Zhou L, Xiong Y, et al. Salinomycin exerts anti-colorectal cancer activity by targeting the β-catenin/T-cell factor complex. *Br J Pharmacol*. 2019;176:3390–3406. doi:10.1111/bph.14770

96. Seyal N, Abdullah SN. Cluster analysis on longitudinal data of patients with kidney dialysis using a smoothing cubic b-spline model. *Int J Math Stat Comput Sci*. 2024;2:85–95. doi:10.59543/ijmscs.v2i.8337

97. Mohammed MA, Lakhan A, Abdulkareem KH, et al. Federated-reinforcement learning-assisted iot consumers system for kidney disease images. *IEEE Trans Consum Electron*. 2024;1–11.

98. Ali AM, Mohammed MA. A comprehensive review of artificial intelligence approaches in omics data processing: evaluating progress and challenges. *Int J Math Stat Comput Sci*. 2023;2:114–167. doi:10.59543/ijmscs.v2i.8703

99. Mohammed MA, Lakhan A, Abdulkareem KH, Garcia-Zapirain B. A hybrid cancer prediction based on multi-omics data and reinforcement learning state action reward state action (SARSA). *Comput Biol Med*. 2023;154:106617. doi:10.1016/j.compbiomed.2023.106617