Am J Cancer Res 2026;16(10):4556-4578
www.ajcr.us /ISSN:2156-6976/ajcr0174977

Original Article

Single-cell transcriptomics identifies DDX39A as a potential regulator of malignant epithelial cell states in colorectal cancer

Guocai Zhong[1,2], Hao Cheng[3], Rui Zhang[3], Dongxue Wang[1,2], Jiazhi Li[1,2], Jingyu Li[1,2], Yanfeng Wang[4], Dandan Bi[3], Jinling Zang[1,2], Lei Xu[1,2], Yunxia Su[1,2], Anni Wang[3], Ziwei Liang[3], Ende Li[3], Lijie Wang[3], Tingting Yang[3]

[1]Department of Pathology, Heilongjiang Provincial Third Hospital, Beian, Heihe, Heilongjiang, China; [2]Department of Pathology, Beian Hospital, Beidahuang Group, Beian, Heihe, Heilongjiang, China; [3]Department of Pathology, The Fourth Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China; [4]General Hospital of Beidahuang Group, Harbin, Heilongjiang, China

Received June 25, 2026; Accepted October 14, 2026; Epub October 15, 2026; Published October 30, 2026

Abstract: Colorectal cancer (CRC) exhibits substantial intratumoral heterogeneity, yet the molecular mechanisms that maintain aggressive epithelial cell states remain incompletely understood. Here, we combined single-cell RNA sequencing, weighted gene co-expression network analysis (WGCNA), metabolic profiling, and experimental validation to investigate the functional diversity of malignant epithelial populations in CRC. We identified EPC3 as a highly proliferative malignant epithelial population with pronounced metabolic activity and features associated with tumor progression. Transcriptomic network analysis revealed a proliferation-associated gene module (M23), within which DDX39A emerged as a key candidate associated with the EPC3 phenotype. Clinical analysis revealed that DDX39A was significantly upregulated in CRC tissues and was associated with advanced tumor stage, metastasis, and poor prognosis. Functional experiments demonstrated that DDX39A promotes CRC cell proliferation, migration, invasion, and tumor growth in vivo, while modulating glycolytic activity and mitochondrial respiration, suggesting its involvement in regulating metabolic adaptation in CRC cells. Mechanistically, DDX39A was found to promote glycolytic reprogramming through regulation of HIF-1α expression, thereby influencing cell cycle progression and maintaining the malignant epithelial phenotype to facilitate CRC progression. Collectively, these findings identify DDX39A as a potential regulator linking metabolic remodeling with malignant cell-state maintenance and suggest its potential value as a biomarker and therapeutic target for colorectal cancer.

Keywords: Colorectal cancer, DDX39A, epithelial heterogeneity, metabolic reprogramming

Introduction

Colorectal cancer (CRC) remains one of the most frequently diagnosed and deadly malignancies worldwide [1, 2]. Although advances in surgery, chemotherapy, targeted therapy, and immunotherapy have improved CRC management, patients with metastatic or recurrent CRC continue to have poor outcomes, with a five-year survival rate of less than 15% [1-3]. High intratumoral heterogeneity and therapy resistance are considered key determinants of unfavorable outcomes [1, 2, 4]. Elucidating the mechanisms driving CRC progression may facilitate the development of new therapeutic strategies.

CRC progression involves substantial changes in the spatial architecture of the tumor microenvironment, particularly at the invasive margin [5, 6]. At the tumor-stroma interface, normal epithelial structures are gradually replaced by malignant epithelial cells and stromal components, creating a dynamic region that supports tumor growth and invasion [7-10]. This area also promotes microenvironmental remodeling and enables tumor cells to adapt to changing conditions [4, 10, 11]. Recent single-cell studies have suggested that this spatial niche is enriched for epithelial subpopulations with distinct functional states [12-14], however, how these populations emerge and the molecular mechanisms that

https://doi.org/10.62347/SCUE5656

maintain their distinct states remain poorly understood.

At the molecular level, dysregulated RNA metabolism has emerged as a critical regulatory layer in tumorigenesis [15-17]. DEAD-box RNA helicases (DDX family proteins) participate in multiple aspects of RNA metabolism, including RNA splicing, nuclear export, and post-transcriptional regulation, thereby shaping gene expression programs involved in cell proliferation, stress responses, and therapeutic resistance [18-22]. DDX39A, a member of the DEAD-box RNA helicase family and participates in mRNA splicing and nuclear export, contributing to the regulation of transcriptional homeostasis [23-25]. Although dysregulated RNA processing has been implicated in tumor proliferation and metabolic reprogramming [26, 27], the role of DDX39A in cancer biology remains poorly characterized. Given the potential coupling between RNA processing, cell cycle regulation, and energy metabolism, DDX39A may play a critical role in linking post-transcriptional regulation to tumor functional state transitions.

In this study, we investigated epithelial heterogeneity in CRC at single-cell resolution and identified a highly proliferative and metabolically active epithelial subpopulation. Further analyses identified DDX39A as a candidate regulator of cell state, with a potential role in maintaining malignant epithelial states through HIF-1α-mediated glycolysis and cell cycle regulation.

Materials and methods

Single-cell RNA sequencing data processing, cell annotation, and malignant epithelial cell identification

Publicly available scRNA-seq data were obtained from the Gene Expression Omnibus (GEO; GSE144735). The dataset comprises 18 samples from six patients with colorectal cancer, with tumor core, tumor border, and matched normal mucosa collected from each patient. In the original dataset, sequencing reads were aligned to the GRCh38 reference genome and UMI counts were quantified using Cell Ranger (v2.1.0). Cells with > 1,000 UMIs, > 200 and < 6,000 detected genes, and < 20% mitochondrial UMI counts were retained. A total of 29,003 cells passed the initial quality-control filters. Cells with discordant annotations between the Seurat multi CCA and RCA pipelines were subsequently removed, together with potential doublets identified by abnormally high gene counts, leaving 27,414 high-quality cells for analysis.

The processed UMI count matrix was imported into R (v4.3.1) and analyzed using Seurat (v4.3.0). Expression values were normalized with Normalize Data, followed by identification of highly variable genes and data scaling. Harmony was used to reduce sample-specific batch effects. Principal component analysis was performed before UMAP embedding, and cell clusters were identified using the Louvain algorithm. Cell identities were assigned according to established marker genes and further checked using SingleR (v1.4.1), CellMarker, and the Azimuth reference atlas, followed by manual curation. Copy-number variation (CNV) profiles of epithelial cells were estimated using inferCNV (v1.10.1), with normal epithelial and stromal cells as reference populations. The inferCNV expression cutoff was set to 0.1. A CNV score was calculated from the inferred chromosomal alterations for each epithelial cell. Cells showing broad chromosomal gains or losses and a CNV score greater than the mean plus two standard deviations of the reference-cell distribution were classified as malignant epithelial cells; cells without these alterations were considered non-malignant. The relative enrichment of each cell population across clinical groups was evaluated using the ratio of observed to expected cell numbers (Ro/e). The expected cell number was calculated as $Eij = (Ni \times Nj)/N$, and the Ro/e index was defined as $Ro/e = Oij/Eij$, where Oij represents the observed cell number. $Ro/e > 1$ indicates relative enrichment, whereas $Ro/e < 1$ indicates relative depletion. Transcriptomic data from The Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx) project were obtained through UCSC Xena and GEPIA2 for independent validation and comparative analyses.

Cell-cell communication analysis

Cell-cell communication was analyzed using CellChat (v1.5.0). Normalized expression data and annotated cell-type identities were used to infer ligand-receptor interactions based on the

CellChatDB human database. Communication probabilities were calculated for each cell-cell pair, and interactions with $P < 0.05$ were considered significant. The number and strength of significant interactions were compared across cell populations. Communication networks were visualized using the netVisual_circle and netAnalysis_signalingRole_network functions.

Weighted gene co-expression network analysis (WGCNA)

Because of the sparse nature of single-cell expression data, epithelial cells were aggregated into pseudobulk profiles by patient and epithelial subtype before network analysis. WGCNA was then performed with the WGCNA package (v1.72-1) after lowly expressed genes were removed. A soft-thresholding power of $\beta = 6$ was used to obtain an approximately scale-free network ($R^2 > 0.9$). Gene modules were identified from the signed co-expression network using dynamic tree cutting, with the minimum module size set to 100 genes. Module eigengenes were correlated with epithelial subtypes to identify subtype-related modules. Genes with module membership (MM) > 0.7 and gene significance (GS) > 0.5 were considered hub genes. GO and KEGG enrichment analyses were carried out using cluster Profiler (v4.4.0), and FDR < 0.05 was considered significant.

Gene set enrichment and pathway analysis

GSEA was performed with the clusterProfiler and fgsea packages using Hallmark and KEGG gene sets from MSigDB (v7.5.1). Genes were ranked according to log2 fold change, and pathways with FDR < 0.05 were considered significantly enriched. Gene-set activity at the single-cell level was evaluated using module scores derived from the corresponding pathway genes. EPC3 marker genes were identified using the FindAllMarkers function in Seurat, with an adjusted $P < 0.05$ and an average $\log_2$ fold change > 0.25. These genes were used to define the EPC3 signature. In the GSE17538 cohort, an EPC3 score was calculated for each patient by single-sample gene set enrichment analysis (ssGSEA) using the GSVA package. Higher scores indicated greater enrichment of the EPC3 transcriptional program.

Immune infiltration analysis

Immune cell composition was estimated using CIBERSORT and TIMER2.0 based on TCGA transcriptomic data. The association between gene expression levels and immune infiltration scores was assessed using Spearman's correlation analysis. Correlation matrices and heatmaps were visualized using the ggcorrplot and Complex Heatmap packages in R.

Pan-cancer expression and validation

To assess gene expression across multiple cancer types, pan-cancer analyses were conducted using The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) datasets accessed through the GEPIA2 and UCSC Xena platforms. Normalized expression values ($\log_2$[TPM+1]) were obtained and analyzed across 33 cancer types. Differences between tumor and corresponding normal tissues were evaluated using the Wilcoxon signed-rank test. Statistical significance was defined as an adjusted *p*-value < 0.05. Data visualization was performed using the ggplot2 package (v3.4.2) in R.

Clinical cohort and ethical approval

Eighty paired CRC tissues and matched adjacent normal tissues were collected from patients undergoing surgical resection at Beian Hospital of Beidahuang Group. The study protocol was approved by the Ethics Committee of Beian Hospital of Beidahuang Group (Approval No. 2026-049), and written informed consent was obtained from all participants. None of the patients received preoperative radiotherapy or chemotherapy. Clinical and pathological information, including tumor stage, histological grade, metastasis, and recurrence status, was collected and linked to corresponding tissue samples for downstream analyses.

Cell culture

Human colorectal cancer cell lines (HCT116, HCT15, RKO, HT29) and the normal colonic epithelial cell line NCM460 were obtained from the American Type Culture Collection (ATCC). Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) or RPMI-1640 medium supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin, and main-

tained at 37°C in a humidified incubator with 5% CO_2.

Lentiviral transduction and generation of stable cell lines

DDX39A knockdown and overexpression cell lines were generated by lentiviral transduction. Three shRNAs targeting human DDX39A were cloned into lentiviral vectors, with target sequences as follows: shDDX39A-1 (5'-GGACCGGTTTGAAGTTAATGT-3'), shDDX39A-2 (5'-GCAGCAGTACTACGTCAAACT-3'), and shDDX39A-3 (5'-GGAGCTTCAGCCTAAAGAATG-3'). For DDX39A overexpression, the full-length human DDX39A coding sequence (NM_005804.4) was inserted into a lentiviral expression vector. For HIF1A overexpression, the full-length human HIF1A coding sequence (NM_001530.4) was inserted into a lentiviral expression vector. Lentiviral vectors were packaged in HEK293T cells by co-transfection with packaging plasmids. The resulting viral supernatants were collected and used to infect colorectal cancer cells. Stable cell populations were obtained by puromycin selection (2 μg/mL) for 7-10 days. DDX39A expression levels were assessed by Western blotting.

CCK-8 assay

Cell proliferation was assessed using the Cell Counting Kit-8 (CCK-8, Dojindo, Japan). Cells were seeded into 96-well plates at appropriate densities and cultured for 24, 48, 72, and 96 h under standard conditions. At each time point, CCK-8 reagent was added according to the manufacturer's instructions, and cells were incubated for 1-2 h. Absorbance was measured at 450 nm using a microplate reader. All experiments were performed in triplicate.

Colony formation assay

Cells were seeded in 6-well plates at 1,000 cells per well and cultured for 14 days, with the medium replaced every 3-4 days. Colonies were then fixed with 4% paraformaldehyde for 15-20 min and stained with crystal violet. After washing with PBS, colonies containing more than 50 cells were counted under a light microscope. Experiments were performed independently three times.

Wound healing assay

Cells were seeded into 6-well plates and cultured to 90-100% confluence. A sterile 200-μL pipette tip was used to create a linear scratch wound. Detached cells were removed by washing with phosphate-buffered saline (PBS), and images were captured at 0 h and 24-48 h under a phase-contrast microscope. Cell migration was quantified by measuring wound closure using ImageJ software, and migration rate was calculated as: wound closure (%) = [(initial wound area - remaining wound area)/initial wound area] × 100%. All experiments were performed in triplicate.

Transwell migration and invasion assay

Cell migration and invasion were assessed using Transwell chambers with 8.0-μm pore polycarbonate membranes. For migration assays, cells were seeded into uncoated upper chambers in serum-free medium; for invasion assays, the membranes were precoated with Matrigel diluted 1:8 in serum-free medium. A total of 2×10^5 cells were added to each upper chamber, with medium containing 10% FBS in the lower chamber as a chemoattractant. After 36 h, cells remaining on the upper surface were removed with a cotton swab. Migrated or invaded cells on the lower surface were fixed with 4% paraformaldehyde, stained with crystal violet, and counted in five randomly selected microscopic fields. Experiments were performed independently three times.

Western blotting (WB)

Protein was extracted using RIPA lysis buffer supplemented with protease and phosphatase inhibitors, and protein concentration was determined by BCA assay. Equal amounts of protein were separated by SDS-PAGE and transferred onto PVDF membranes. After blocking, membranes were incubated overnight at 4°C with primary antibodies against DDX39A (83083-5-RR, 1:8000), HIF-1α (20960-1-AP, 1:2000), GLUT1 (21829-1-AP, 1:1000), HK2 (22029-1-AP, 1:1000), PGK1 (17811-1-AP, 1:5000), PKM2 (15822-1-AP, 1:2000), Cyclin D1 (26939-1-AP, 1:5000), MDM2 (19058-1-AP, 1:200), p53 (10442-1-AP, 1:5000), and GAPDH (60004-1-Ig, 1:50000). All primary antibodies were purchased from Proteintech (Wuhan, China).

Membranes were then incubated with HRP-conjugated goat anti-rabbit IgG (SA00001-2, 1:5000) or goat anti-mouse IgG (SA00001-1, 1:5000), both from Proteintech. Protein bands were visualized using enhanced chemiluminescence (ECL), and band intensities were quantified with ImageJ and normalized to GAPDH.

Subcutaneous (s.c.) tumor model

All animal experiments were approved by the Animal Ethics Committee of the Fourth Affiliated Hospital of Harbin Medical University (2026-DWSYLLCZ-155) and conducted in accordance with the ARRIVE guidelines. Male BALB/c nude mice (5-6 weeks old) were purchased from Beijing Vital River Laboratory Animal Technology Co., Ltd. (Charles River, Beijing, China) and maintained under specific pathogen-free conditions with a 12-h light/dark cycle and free access to food and water. CRC cells with stable DDX39A overexpression, DDX39A knockdown, or the corresponding controls were resuspended in sterile PBS, and 5×10^6 cells in 0.1 mL were injected subcutaneously into the left shoulder of each mouse. Mice were randomly assigned to groups (n = 5 per group), and tumor measurements were performed blinded to group allocation. Once tumors reached approximately 100 mm^3, tumor length and width were measured every 3 days using calipers, and tumor volume was calculated as $V = \text{length} \times \text{width}^2 \times \pi/6$. Body weight and general condition were monitored during the experiment. Humane endpoints included a tumor volume > 1,500 mm^3, body weight loss > 15%, tumor ulceration, or severe distress. Three weeks after inoculation, mice were euthanized by CO_2 inhalation followed by cervical dislocation to ensure death, and tumors were excised, photographed, weighed, and fixed in 10% neutral-buffered formalin for histological and immunohistochemical analyses.

3D spheroid formation assay

Cells were seeded into honeycomb-patterned low-attachment plates at 500 cells per well and cultured for 14 days under standard conditions. The medium was refreshed every 2-3 days. At the end of the culture period, spheroids were photographed under an inverted microscope, and spheroid number and diameter were quantified using ImageJ. Experiments were performed independently three times.

Immunohistochemistry (IHC)

Paraffin-embedded sections were deparaffinized, rehydrated, and subjected to antigen retrieval. After blocking endogenous peroxidase activity and nonspecific binding, sections were incubated overnight at 4°C with antibodies against DDX39A (Proteintech, 83083-5-RR, 1:500), TP53 (10442-1-AP, 1:500), HK2 (22029-1-AP, 1:1000), and GLUT1 (21829-1-AP, 1:5000). All primary antibodies were purchased from Proteintech (Wuhan, China). Staining was developed with DAB after incubation with HRP-conjugated secondary antibodies, followed by hematoxylin counterstaining. IHC results were scored according to staining intensity and the percentage of positive cells. Staining intensity was graded as 0 (negative), 1 (weak), 2 (moderate), 3 (moderately strong), or 4 (strong). The percentage of positive cells was scored as 1 (0-25%), 2 (26-50%), 3 (51-75%), or 4 (76-100%). The final IHC score was obtained by multiplying the intensity score by the proportion score, giving a range of 0-16. The median score was used to divide cases into high- and low-expression groups.

Cellular metabolic analysis (OCR and ECAR assays)

Oxygen consumption rate (OCR) and extracellular acidification rate (ECAR) were measured using the Seahorse XF96 Analyzer (Agilent Technologies). Cells were seeded into Seahorse XF96 cell culture microplates at 2×10^4 cells per well and cultured overnight before analysis. For the mitochondrial stress test, oligomycin (1.5 μM), FCCP (1.0 μM), and rotenone/antimycin A (0.5 μM) were sequentially injected. For the glycolysis stress test, glucose (10 mM), oligomycin (1.0 μM), and 2-deoxyglucose (2-DG; 50 mM) were sequentially added. Glycolysis and glycolytic capacity were calculated from the ECAR profiles. After each assay, nuclei were stained with Hoechst 33342 and counted to determine the cell number in each well. OCR and ECAR values were normalized to cell number and expressed per 10^4 cells.

RNA extraction, nuclear-cytoplasmic fractionation, and RT-qPCR

Total RNA was extracted from cells using TRIzol reagent according to the manufacturer's instructions. For nuclear-cytoplasmic RNA frac-

tionation, nuclear and cytoplasmic RNA were isolated using the PARIS™ Kit (Invitrogen, Thermo Fisher Scientific, AM1921). Briefly, cells were harvested, washed with ice-cold PBS, and lysed in cell fractionation buffer. After centrifugation at 4°C, the cytoplasmic supernatant and nuclear pellet were collected separately, followed by RNA purification from each fraction. The purity of the nuclear and cytoplasmic fractions was assessed using U6 snRNA and GAPDH mRNA as nuclear and cytoplasmic markers, respectively. cDNA was synthesized by reverse transcription, and quantitative real-time PCR was performed using SYBR Green chemistry. GAPDH was used as the internal control for total HIF1A mRNA analysis. The primers for HIF1A were 5'-TCCAAGAAGCCCTAACGTGT-3' (forward) and 5'-TGATCGTCTGGCTGCTGTAA-3' (reverse). Relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method. For nuclear-cytoplasmic distribution analysis, HIF1A mRNA abundance was determined separately in the nuclear and cytoplasmic fractions, and the cytoplasmic-to-nuclear ratio was calculated to assess its subcellular distribution.

Statistical analysis

Statistical analyses were performed using R (v4.3.1) and GraphPad Prism. Data are presented as mean ± SD unless otherwise stated. Two-group comparisons were performed using an unpaired two-tailed Student's t-test or Wilcoxon rank-sum test, as appropriate. Comparisons among multiple groups were analyzed by one-way ANOVA followed by Tukey's post hoc test. Repeated measurements over time were analyzed using two-way repeated-measures ANOVA with Tukey's multiple-comparisons test. Correlations were assessed by Spearman's correlation analysis. Survival curves were generated using the Kaplan-Meier method and compared with the log-rank test. The median DDX39A expression level was used to define high- and low-expression groups. Multivariable Cox proportional hazards regression was used to evaluate prognostic associations, with hazard ratios (HRs) and 95% confidence intervals (CIs) reported. FDR correction was applied to high-throughput analyses where appropriate. All tests were two-sided, and $P < 0.05$ was considered statistically significant.

Result

Comprehensive single-cell landscape reveals dynamic cellular remodeling during tumor progression

To characterize the cellular composition of CRC, we performed single-cell RNA sequencing with stringent quality control procedures to obtain high-quality transcriptional profiles for subsequent analysis (Supplementary Figure 1A). Unsupervised clustering using UMAP identified major cellular populations, including epithelial cells, fibroblasts, immune cells (T cells, myeloid cells, plasma cells), endothelial cells, smooth muscle cells, and mast cells, confirming accurate lineage annotation based on canonical marker genes (**Figure 1A**, Supplementary Figure 1B). Comparison of normal, paratumor (border), and tumor tissues revealed substantial changes in cellular composition during CRC progression (**Figure 1B**). Tumor tissues showed increased proportions of epithelial and fibroblast populations, accompanied by decreased proportions of immune cell populations, particularly T cells and myeloid cells. Quantitative analysis of cell proportions further demonstrated differences in cellular composition among the three tissue types (**Figure 1C**) In addition to changes in cell composition, Ro/e index analysis was performed to evaluate differences in cellular functional states. Epithelial and fibroblast populations exhibited increased proliferative and structural activity in tumor tissues, whereas immune compartments showed globally reduced functional states (**Figure 1D**). Together, these findings indicate that CRC progression is accompanied by extensive alterations in cellular composition and functional states, including expansion of epithelial populations, activation of stromal components, and changes in immune cell distribution.

Malignant epithelial cell subsets exhibit distinct genomic and metabolic characteristics

To further characterize epithelial heterogeneity, malignant epithelial cells were subjected to subclustering analysis, which identified seven distinct malignant epithelial subsets (Malignant_EPC1-EPC7) withdifferent transcriptional profiles (**Figure 2A**). Among these subpopulations, Malignant_EPC3displayed the highest

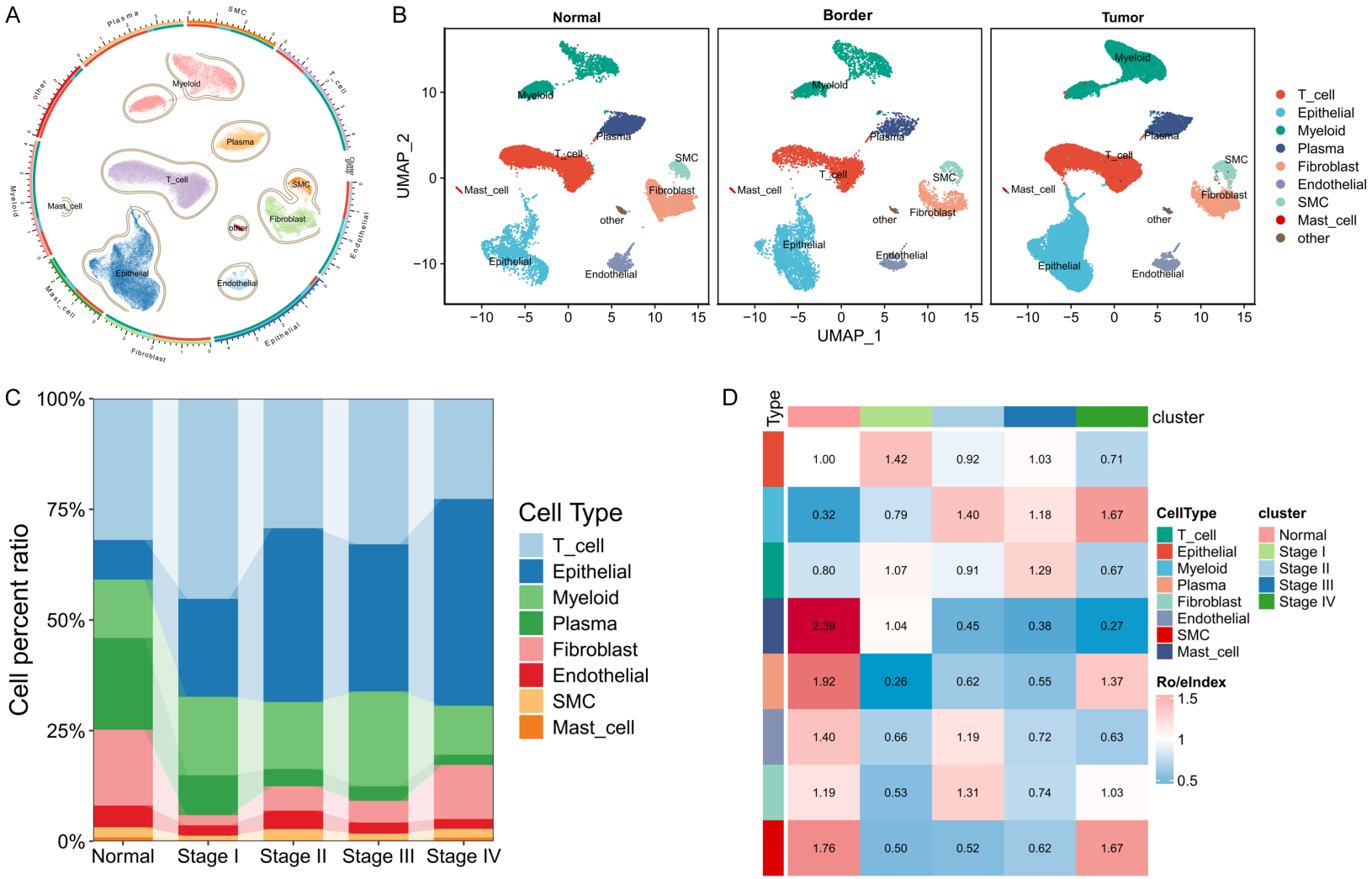


Figure 1. Single-cell landscape reveals cellular remodeling during tumor progression. A. Overview of single-cell composition across all samples, depicting major cell lineages; B. UMAP visualization of cells from normal, border, and tumor tissues, identifying major populations including epithelial, fibroblast, T, myeloid, plasma, endothelial, smooth muscle, and mast cells; C. Proportional changes of major cell types across clinical stages; D. Functional scoring heatmap illustrating dynamic shifts in cellular activity during tumor progression, characterized by increased epithelial and fibroblast activity and decreased immune activity. Statistical comparisons of cell-type proportions and functional scores were performed using the Kruskal-Wallis test followed by Dunn's multiple comparison test. Abbreviations: UMAP, uniform manifold approximation and projection; Ro/e, relative observed-to-expected index.

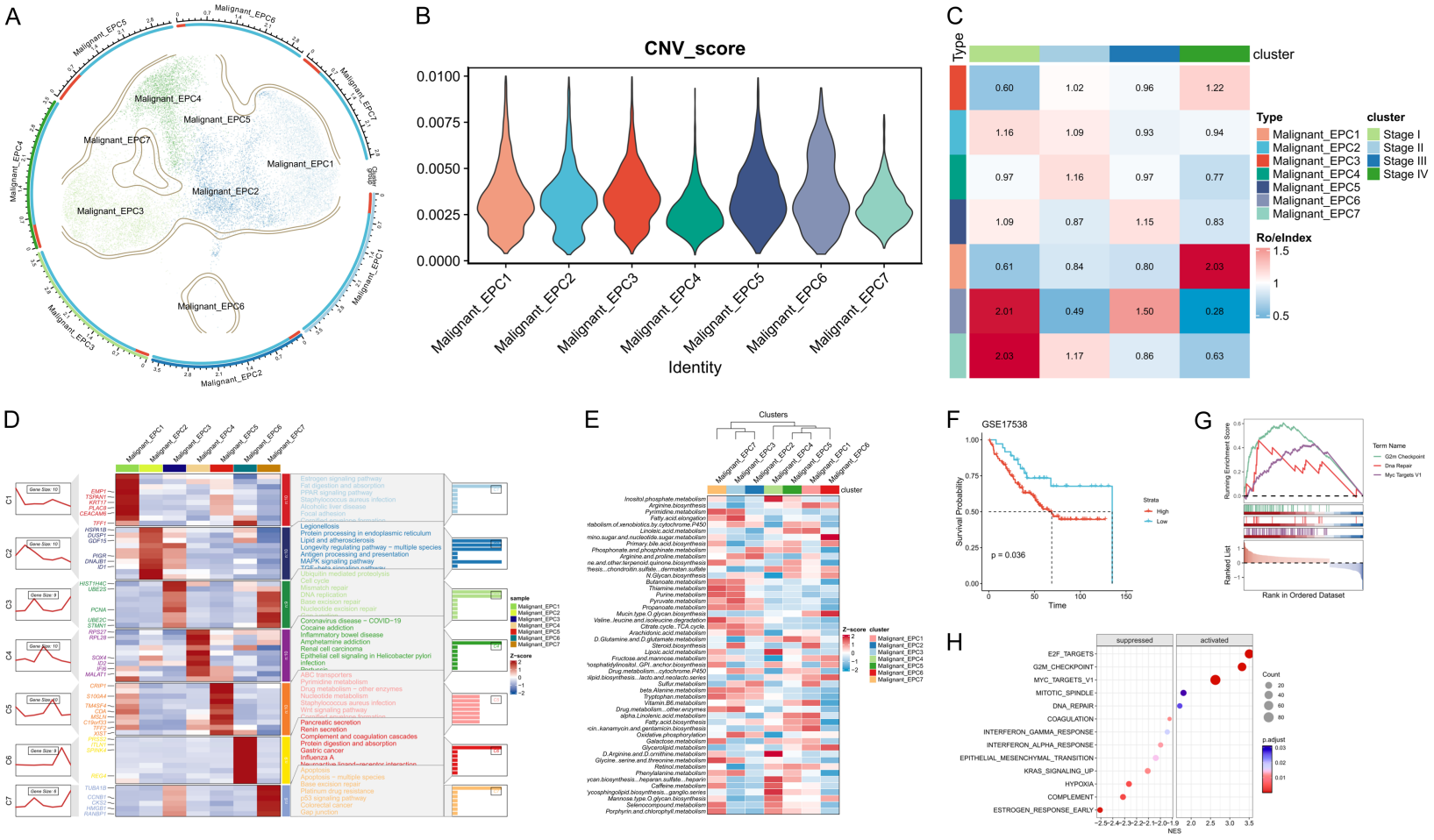


Figure 2. Malignant epithelial subsets exhibit distinct genomic and metabolic heterogeneity. A. UMAP visualization identifying seven malignant epithelial subclusters (EPC1-EPC7) with distinct transcriptional states; B. Comparison of CNV burden across malignant epithelial clusters, highlighting EPC3 as the highest; C. Ro/e index indicating elevated proliferative activity in EPC3 cells; D. Marker gene expression defining epithelial subsets and functional pathway enrichment across EPC1-EPC7, revealing distinct biological programs; E. Metabolic activity heatmap showing heterogeneous metabolic reprogramming among epithelial subclusters; F. Kaplan-Meier survival analysis in the GSE17538 cohort demonstrating worse overall survival in patients with a high EPC3 signature score. Patients were stratified into high and low expression groups based on the median EPC3 signature score, and statistical significance was determined using the log-rank test; G. GSEA revealing activation of E2F targets, G2M checkpoint, MYC signaling, and DNA repair pathways in EPC3; H. Metabolic pathway scoring indicating enr ichment of glycolysis, oxidative phosphorylation, and nucleotide biosynthesis in EPC3. Statistical comparisons among groups were performed using the Wilcoxon rank-sum test or Kruskal-Wallis test where applicable. Abbreviations: UMAP, Uniform Manifold Approximation and Projection; CNV, copy number variation; GSEA, gene set enrichment analysis; Ro/e, relative observed-to-expected index.

CNV burden and Ro/e index, suggesting increased genomic alteration and cellular activity compared with other epithelial subsets (**Figure 2B**, **2C**). CNV analysis revealed increased genomic alterations across epithelial populations from normal to tumor tissues (Supplementary Figure 2). Functionally, EPC3 exhibited a strong proliferative signature, with enrichment of cell cycle- and DNA replication-related genes, including canonical markers such as UBE2C, STMN1, and PCNA (**Figure 2D**). In contrast, other malignant epithelial subsets exhibited distinct transcriptional characteristics. EPC1-EPC2 showed enrichment of differentiation-associated genes (e.g., REG4, PIGR), while EPC6-EPC7 showed stress-response-related signatures. Metabolic profiling further revealed strong functional specialization across malignant epithelial subsets. EPC3 cells were significantly enriched in glycolysis, oxidative phosphorylation, and nucleotide biosynthesis pathways, whereas other subclusters preferentially activated fatty acid metabolism and xenobiotic processing programs (**Figure 2E**). These results indicate that EPC3 represents a metabolically highly active and biosynthetically engaged tumor cell state. Analysis of an independent CRC cohort (GSE17538) further showed that a higher EPC3 gene signature score was associated with poorer overall survival (**Figure 2F**). Gene set enrichment analysis further confirmed that EPC3 is characterized by activation of E2F targets, G2M checkpoint, MYC targets, and DNA repair pathways (**Figure 2G**, **2H**), suggesting enhanced proliferative capacity and replication stress.

Malignant epithelial cells establish immunosuppressive communication networks with the tumor microenvironment

To investigate potential interactions between malignant epithelial subpopulations and other cellular components, we performed CellChat analysis to infer ligand-receptor interactions at the single-cell level. The global communication network revealed extensive inferred interactions among malignant epithelial cells, fibroblasts, myeloid cells, endothelial cells, and immune cell populations (**Figure 3A**). Among epithelial subpopulations, Malignant_EPC3 exhibited the highest levels of both outgoing and incoming signaling interaction scores (**Figure 3B**, **3C**), suggesting enhanced predicted communication activity with other cell populations. These results represent computationally inferred interaction probabilities and do not directly indicate functional regulation. Pathway-level analysis identified several ligand-receptor interactions enriched in EPC3-associated communication patterns, including CD39-ADORA2A (ENTPD1-ADORA2A), CD96-PVR/CD155, TIGIT-PVR, and MIF-CD74/CXCR4 signaling pathways (**Figure 3D-F**). Among these pathways, the PVR-TIGIT axis showed relatively higher interaction weight, consistent with its known role in immune checkpoint regulation. Networkanalysis indicated that malignant epithelial cells and myeloid cells showed higher outgoing signaling activity, whereas T cells displayed predominantly incoming signaling patterns within the PVR-CD96/TIGIT-related network (**Figure 3G**, **3H**). Collectively, these inferred communication patterns suggest that Malignant_EPC3 exhibits distinct communication characteristics and is associated with specific ligand-receptor signaling interactions within the tumor microenvironment. The complete ligand-receptor interaction landscape across epithelial and stromal compartments is shown in Supplementary Figure 3, demonstrating heterogeneous communication patterns among different cellular populations.

WGCNA identifies a proliferation-associated M23 module linked to Malignant_EPC3

To identify key molecular programs underlying malignant epithelial heterogeneity, weighted gene co-expression network analysis (WGCNA) was performed on malignant epithelial cells, resulting in 24 co-expression modules with distinct correlation patterns across epithelial subpopulations (**Figure 4A**, **4B**). Among these modules, the M23 module showed the strongest positive correlation with the Malignant_EPC3 subpopulation, suggesting its close association with the highly proliferative epithelial state identified in previous analyses (**Figure 4C**). Consistently, genes within the M23 module showed higher expression levels in EPC3 cells (**Figure 4D**), further supporting its association with the EPC3 phenotype. Functional annotation revealed that M23 genes were predominantly associated with cell cycle progression and DNA replication, including processes such as chromosome segregation and mitotic spindle organization, as well as key cancer-

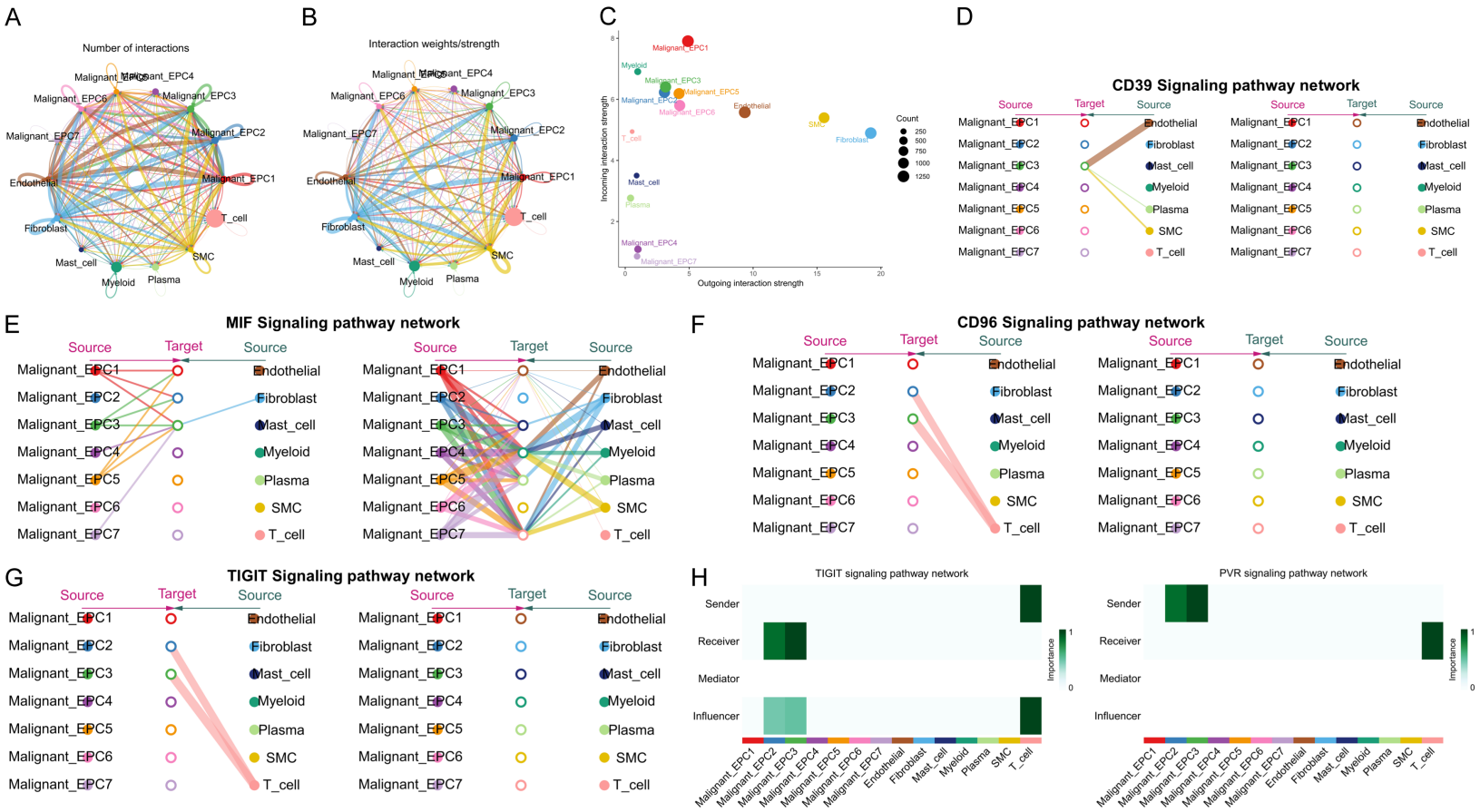


Figure 3. Malignant epithelial cells establish an immunosuppressive communication network with the tumor microenvironment. A. CellChat-inferred global cell-cell communication network showing dense interactions among malignant epithelial cells, fibroblasts, myeloid cells, and endothelial cells; B, C. Analysis of incoming and outgoing signaling strength across cell populations, indicating that Malignant_EPC3 exhibits the highest communication activity; D-F. Identification of key immunoregulatory ligand-receptor pathways enriched in EPC3-mediated interactions, including CD39 (ENTPD1-ADORA2A), MIF-CD74/CXCR4, and CD96-PVR/TIGIT signaling axes; G. Network topology of the PVR-CD96/TIGIT signaling pathway showing EPC3 and tumor-associated myeloid cells as major signal senders and mediators, while T cells predominantly act as receivers; H. Interaction strength heatmap of the PVR-TIGIT axis, highlighting it as the dominant immune checkpoint-related communication pathway. The statistical significance of the inferred cell-cell communication probabilities was evaluated using a permutation test implemented within the CellChat workflow.

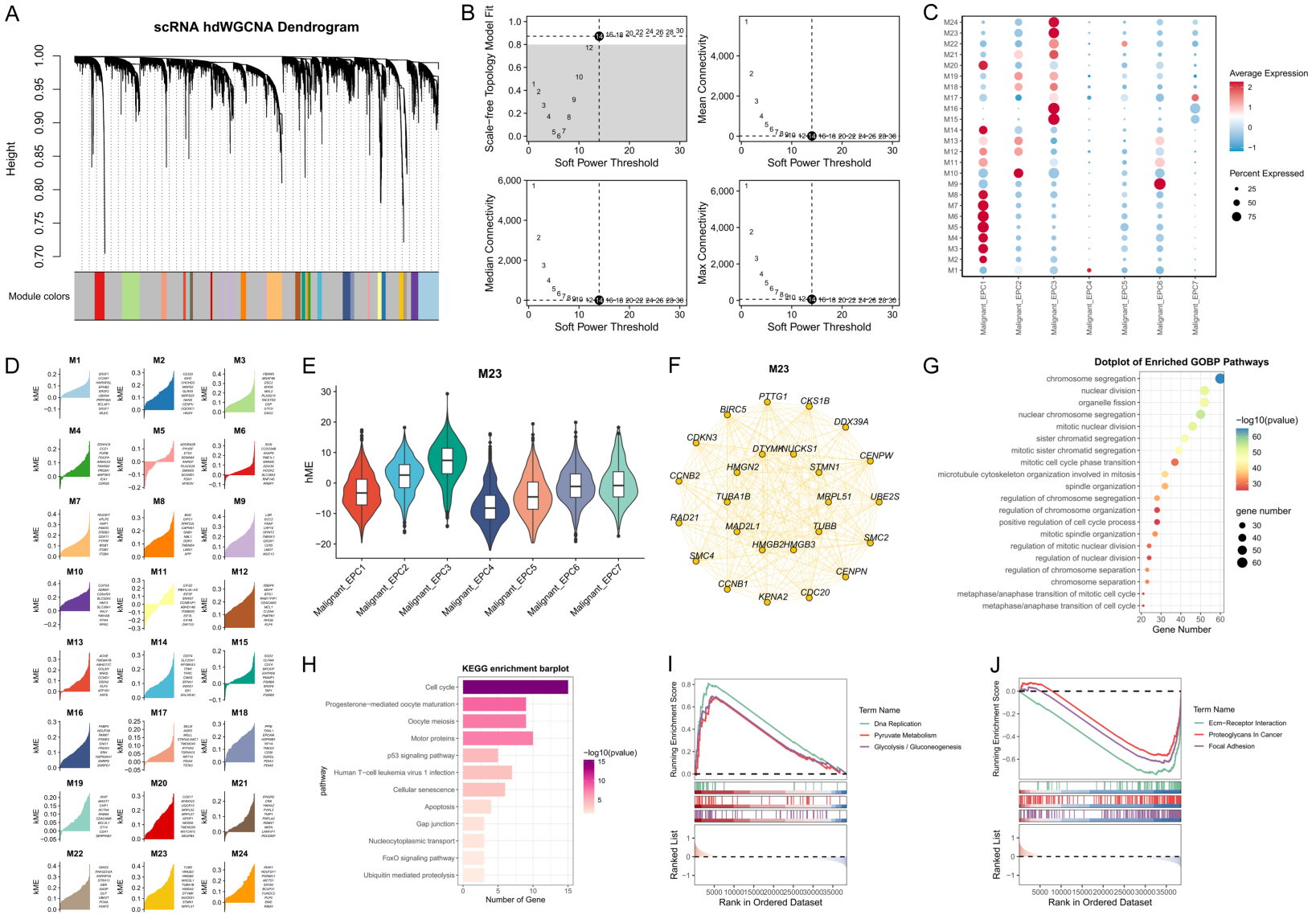


Figure 4. WGCNA identifies a proliferation-associated M23 module linked to Malignant_EPC3. A. WGCNA soft-thresholding analysis determined β = 6 as the optimal power to achieve scale-free network topology; B. Hierarchical clustering identified 24 gene co-expression modules with distinct expression patterns across malignant epithelial subclusters; C. Module-trait correlation analysis showing that the M23 module is most strongly positively associated with Malignant_EPC3; D. Average expression of the M23 module across cell subclusters, with significantly higher levels in EPC3; E. GO enrichment analysis showing that M23 genes are

mainly involved in cell cycle, chromosome segregation, and mitotic spindle organization; F. KEGG pathway analysis indicating enrichment in cell cycle, p53 signaling, and ubiquitin-mediated proteolysis pathways; G, H. Hub gene network analysis of the M23 module identifying key genes including CCNB1, CDK1, BIRC5, UBE2C, CENPW, and DDX39A; I, J. GSEA comparing DDX39A-high and DDX39A-low tumors, showing enrichment of DNA replication and glycolysis pathways in the high-expression group, and ECM-receptor interaction and focal adhesion pathways in the low-expression group. Statistical significance for module-trait correlations and functional enrichments was determined using Pearson's correlation and hypergeometric tests, respectively.

related pathways including cell cycle regulation, p53 signaling, ubiquitin-mediated proteolysis, and apoptosis (**Figure 4E**, **4F**). These findings suggest that the M23 module is closely associated with proliferative characteristics of malignant epithelial cells. To further identify central regulators within this module, hub gene analysis was performed, revealing 25 highly connected genes. Among these hub genes, DDX39A was identified as one of the highly connected genes within the M23 network, together with canonical cell cycle-related genes such as CCNB1, CDK1, and UBE2C (**Figure 4G**, **4H**). These findings suggest that DDX39A is associated with the proliferative transcriptional characteristics of EPC3 cells. GSEA revealed that DDX39A-high tumors were significantly enriched in DNA replication, glycolysis/gluconeogenesis, and pyruvate metabolism pathways, whereas DDX39A-low tumors showed enrichment in extracellular matrix (ECM)-receptor interaction and focal adhesion signaling (**Figure 4I**, **4J**). Together, these results indicate that DDX39A expression is associated with proliferative and metabolic pathway activity in malignant epithelial cells, consistent with the transcriptional characteristics of EPC3.

DDX39A is overexpressed in tumors and associated with malignant progression and poor prognosis

Pan-cancer analysis showed that DDX39A was significantly upregulated in several epithelial malignancies, including colorectal, liver, lung, breast, and gastric cancers (all $P < 0.001$; Supplementary Figure 4A). In CRC, DDX39A expression was also correlated with several immune cell populations (Supplementary Figure 4B), although these associations do not indicate direct immune regulation.

We next examined the clinical relevance of DDX39A in CRC. IHC showed stronger DDX39A staining in tumor tissues than in matched adjacent normal tissues, which was confirmed by quantitative scoring (**Figure 5A**, **5B**). Western blotting further showed increased DDX39A protein levels in tumor tissues (**Figure 5C**). DDX39A and TP53 protein levels showed an inverse pattern by IHC, with a significant negative correlation between their expression levels (**Figure 5D**). Clinicopathological analysis further indicated that DDX39A expression increased with advanced tumor stage and higher histological grade (**Figure 5E**). Moreover, patients with metastasis or tumor recurrence exhibited significantly higher DDX39A levels than those without these events (**Figure 5F**), indicating its association with aggressive tumor behavior. Finally, Kaplan-Meier survival analysis revealed that high DDX39A expression was significantly associated with poorer overall survival and shorter progression-free survival (**Figure 5G**). Multivariable Cox regression further showed that high DDX39A expression remained significantly associated with shorter PFS (HR = 32.601, 95% CI: 3.407-311.992, P = 0.002) and OS (HR = 4.145, 95% CI: 1.158-14.838, P = 0.029) after adjustment for clinicopathological variables (Supplementary Table 1). These findings support the prognostic relevance of DDX39A in CRC.

DDX39A promotes proliferation, migration, and invasion in colorectal cancer cells

To examine the role of DDX39A in CRC cells, we first compared its basal expression across different cell lines. DDX39A expression was higher in HCT15 cells and lower in HCT116 cells (**Figure 6A**). Based on basal DDX39A expression, HCT15 cells were selected for DDX39A knockdown (sh-DDX39A), whereas HCT116 cells were selected for DDX39A overexpression (OE-DDX39A) (**Figure 6B**). CCK-8 assays showed that Sh-DDX39A reduced cell viability in HCT15 cells, while OE-DDX39A increased cell viability in HCT116 cells (**Figure 6C**). Consistently, colony formation assays confirmed reduced clonogenic capacity following Sh-DDX39A and increased colony formation upon OE-DDX39A (**Figure 6D**). In addition, wound healing assays revealed that Sh-DDX39A

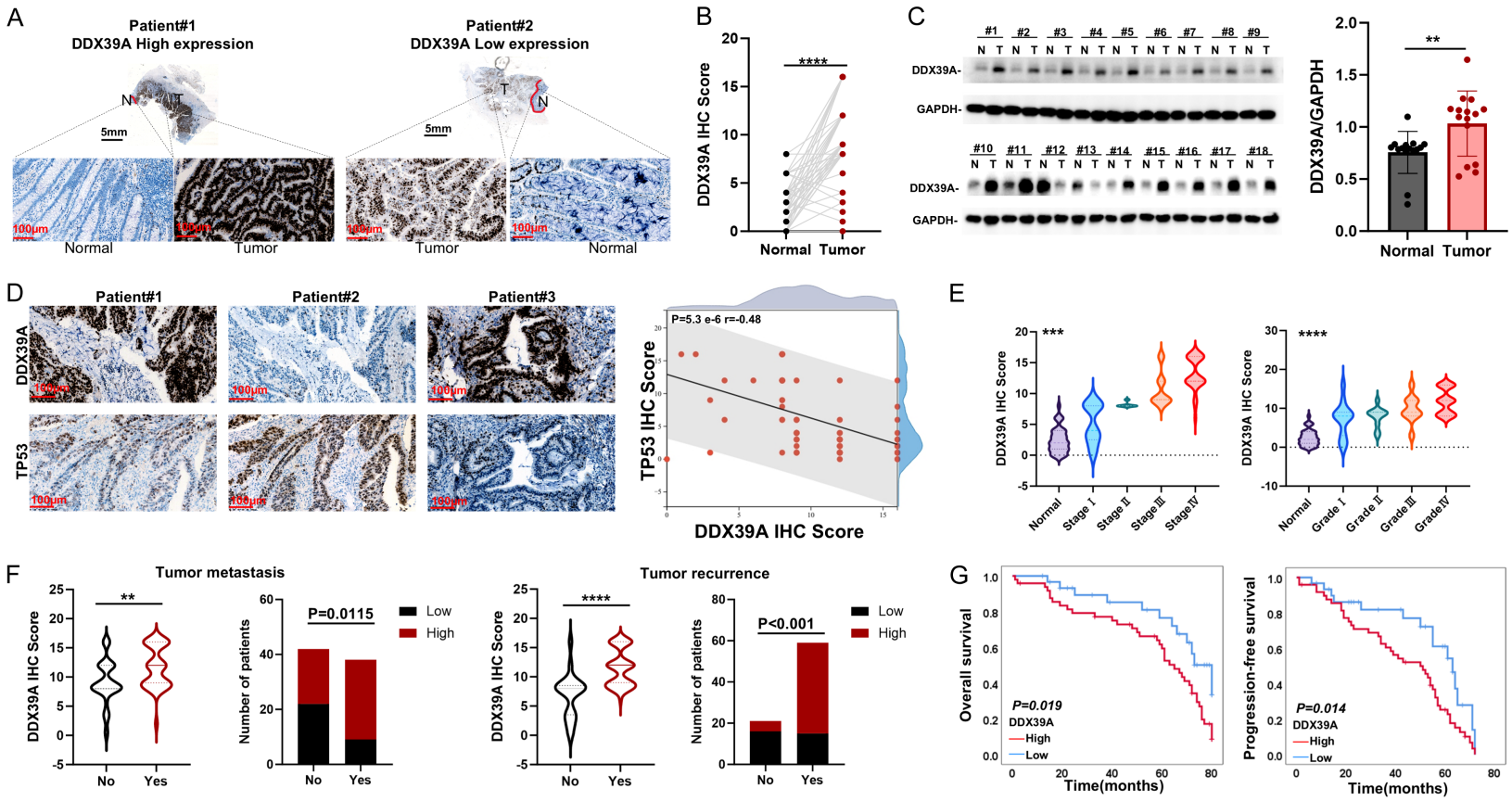


Figure 5. DDX39A is overexpressed in tumor tissue and associated with poor prognosis and malignant progression. A. Representative IHC images showing higher DDX39A expression in tumor tissues compared with matched normal tissues in two representative patients (Scale bars: 5 mm (×0.63) for gross tissue images and 100 μm (×200) for IHC images). B. Quantification of DDX39A IHC scores in paired normal and tumor tissues; Each dot represents an individual patient sample. C. Western blot analysis of DDX39A protein expression in tumor and matched normal tissues. GAPDH was used as the loading control, and the relative protein expression levels were quantified by densitometric analysis. D. Representative IHC staining of DDX39A and TP53 expression in tumor tissues from three representative patients. The correlation between DDX39A and TP53 expression levels was evaluated by Pearson's analysis (×200; Scale bar = 100 μm). E. Violin plots showing elevated DDX39A IHC scores across higher tumor stages and histological grades. F. Association between DDX39A expression and tumor metastasis and recurrence. G. Kaplan-Meier survival analysis showing the association between DDX39A expression levels and overall survival (OS) and progression-free survival (PFS). Patients were classified into high- and low-DDX39A expression groups based on the median DDX39A expression level. Statistical significance was determined using the indicated statistical tests. Data are presented as mean ± SD unless otherwise specified. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.

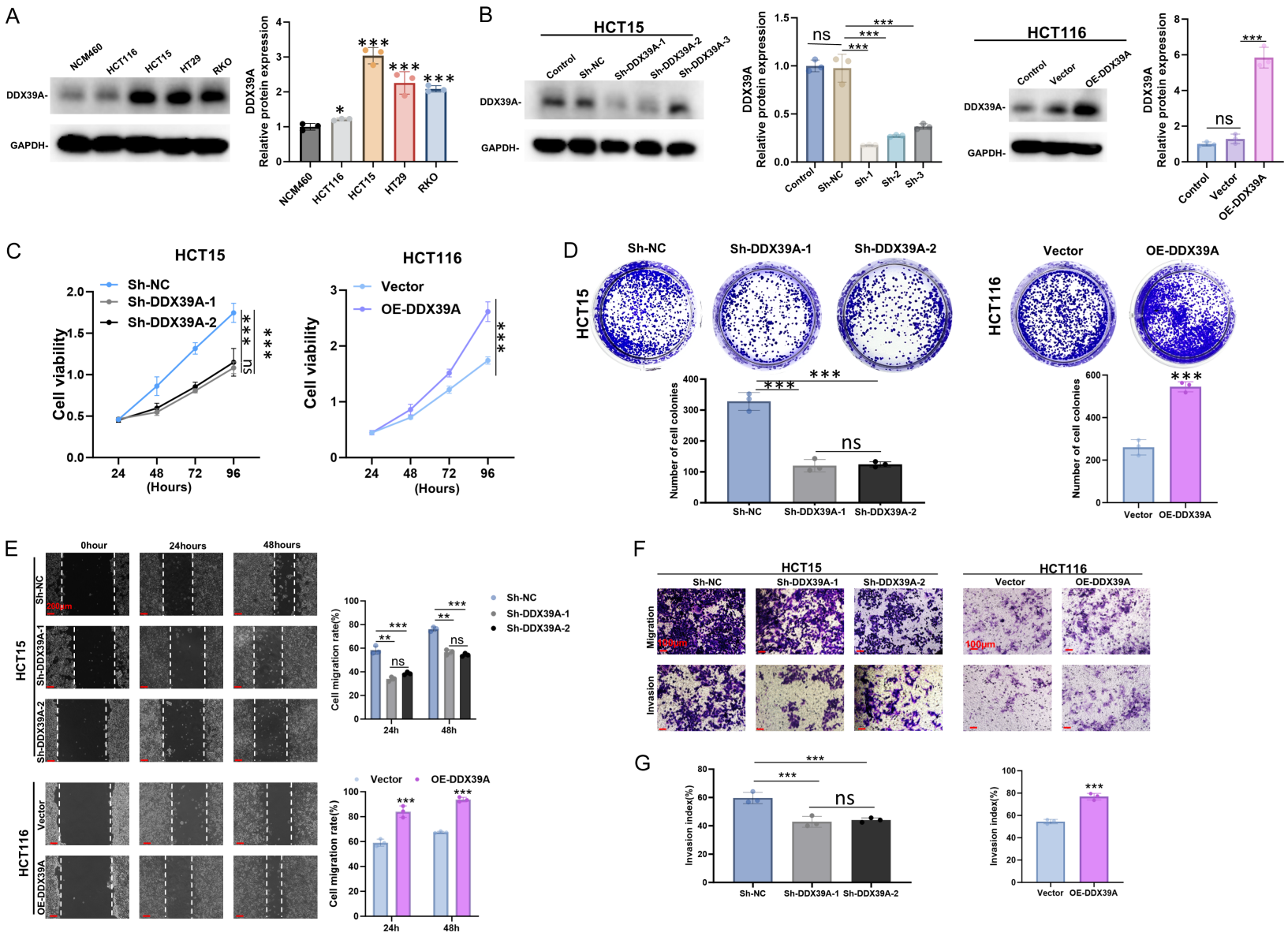


Figure 6. DDX39A promotes proliferation, migration, and invasion in colorectal cancer cells. A. Relative DDX39A protein expression in normal colonic epithelial cells and colorectal cancer cell lines, quantified by densitometry with GAPDH as loading control. B. Construction and validation of HCT15 DDX39A knockdown and

HCT116 overexpression cell models. The efficiency of DDX39A knockdown and overexpression was confirmed by Western blot analysis. C. CCK-8 assay showing that DDX39A knockdown inhibits cell proliferation, whereas overexpression promotes proliferation in HCT15 and HCT116 cells. D. Colony formation assays showing reduced proliferative capacity after DDX39A knockdown and increased proliferation upon overexpression. E. Wound healing assays demonstrating that DDX39A enhances cell migration (×100; Scale bar = 200 μm). F, G. Transwell assays showing that DDX39A promotes cell migration and invasion in HCT15 and HCT116 cells (×200; Scale bar = 100 μm). All experiments were independently performed at least three times ($n \geq 3$), and representative images are shown. Detailed experimental procedures are described in the Materials and Methods section. Data are presented as mean ± SD. Statistical significance was determined using the indicated statistical tests. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.

significantly impaired migration in HCT15 cells, whereas OE-DDX39A enhanced migration in HCT116 cells (**Figure 6E**). Transwell assays further demonstrated that DDX39A promotes both migration and invasion in CRC cells (**Figure 6F**, **6G**). Collectively, these results indicate that DDX39A acts as an oncogenic regulator of CRC progression by enhancing proliferation, migration, and invasion.

DDX39A regulates HIF1A mRNA distribution and HIF-1α-associated glycolytic activity

GeneMANIA analysis placed HIF1A within the DDX39A-associated interaction network (**Figure 7A**). Sh-DDX39A markedly reduced HIF-1α protein expression in HCT15 cells, whereas OE-DDX39A increased HIF-1α protein levels in HCT116 cells (**Figure 7B**, **7C**). In contrast, total HIF1A mRNA levels were not significantly affected by either Sh-DDX39A or OE-DDX39A (**Figure 7D**). Nuclear-cytoplasmic RNA fractionation further showed that Sh-DDX39A decreased the cytoplasmic-to-nuclear ratio of HIF1A mRNA, whereas OE-DDX39A increased this ratio (**Figure 7E**), suggesting that DDX39A may facilitate the cytoplasmic distribution of HIF1A mRNA. Consistent with the changes in HIF-1α, Sh-DDX39A reduced the expression of GLUT1, HK2, PGK1, PKM2, and CyclinD1 in HCT15 cells, whereas OE-DDX39A increased their expression in HCT116 cells (**Figure 7F**, **7G**). To further examine the role of HIF-1α, HIF1A was overexpressed in sh-DDX39A HCT-15 cells (**Figure 7H**). HIF1A overexpression increased colony formation and restored the expression of GLUT1, HK2, PGK1, PKM2, and Cyclin D1 (**Figure 7I**, **7J**). In addition, inhibition of glycolysis with 2-DG reduced Cyclin D1 expression in both HCT116 cells and HIF1A-overexpressing sh-DDX39A HCT15 cells (**Figure 7K**) and suppressed 3D spheroid formation (**Figure 7L**). These findings suggest that DDX39A may enhance HIF-1α expression by promoting the cytoplasmic distribution of HIF1A mRNA, thereby supporting glycolytic activity and CRC cell growth.

Seahorse analysis showed that DDX39A knockdown reduced OCR and basal respiration in HCT15 cells, whereas DDX39A overexpression increased basal respiration in HCT116 cells (**Figure 8A**, **8B**). Similarly, ECAR analysis showed that DDX39A knockdown decreased basal glycolysis, while DDX39A overexpression increased glycolytic activity (**Figure 8C**, **8D**). IHC analysis of CRC tissues further showed that DDX39A expression was positively correlated with HK2 and GLUT1 expression (**Figure 8E**). Representative staining also showed an inverse expression pattern between DDX39A and TP53 in tumor tissues. These findings further support an association between DDX39A expression and increased glycolytic activity in CRC.

DDX39A promotes tumor growth in vivo

To further determine the oncogenic role of DDX39A in vivo, xenograft tumor models were established using HCT15 and HCT116 cells. In the HCT15 model, Sh-DDX39A reduced tumor growth, as shown by smaller tumors, lower tumor volumes, and reduced tumor weights compared with the sh-NC group (**Figure 9A-C**). IHC staining further revealed a decreased proliferation index in Sh-DDX39A tumors (**Figure 9D**). In contrast, in the HCT116 model, OE-DDX39A produced the opposite effect, with increased tumor volume and weight compared with vector controls (**Figure 9E-G**). A higher proliferation index was also observed in OE-DDX39A tumors (**Figure 9H**). Collectively, these results demonstrate that DDX39A promotes CRC tumor growth in vivo, consistent with its oncogenic functions observed in vitro.

Discussion

Colorectal cancer (CRC) is a highly heterogeneous malignancy characterized by diverse cel-

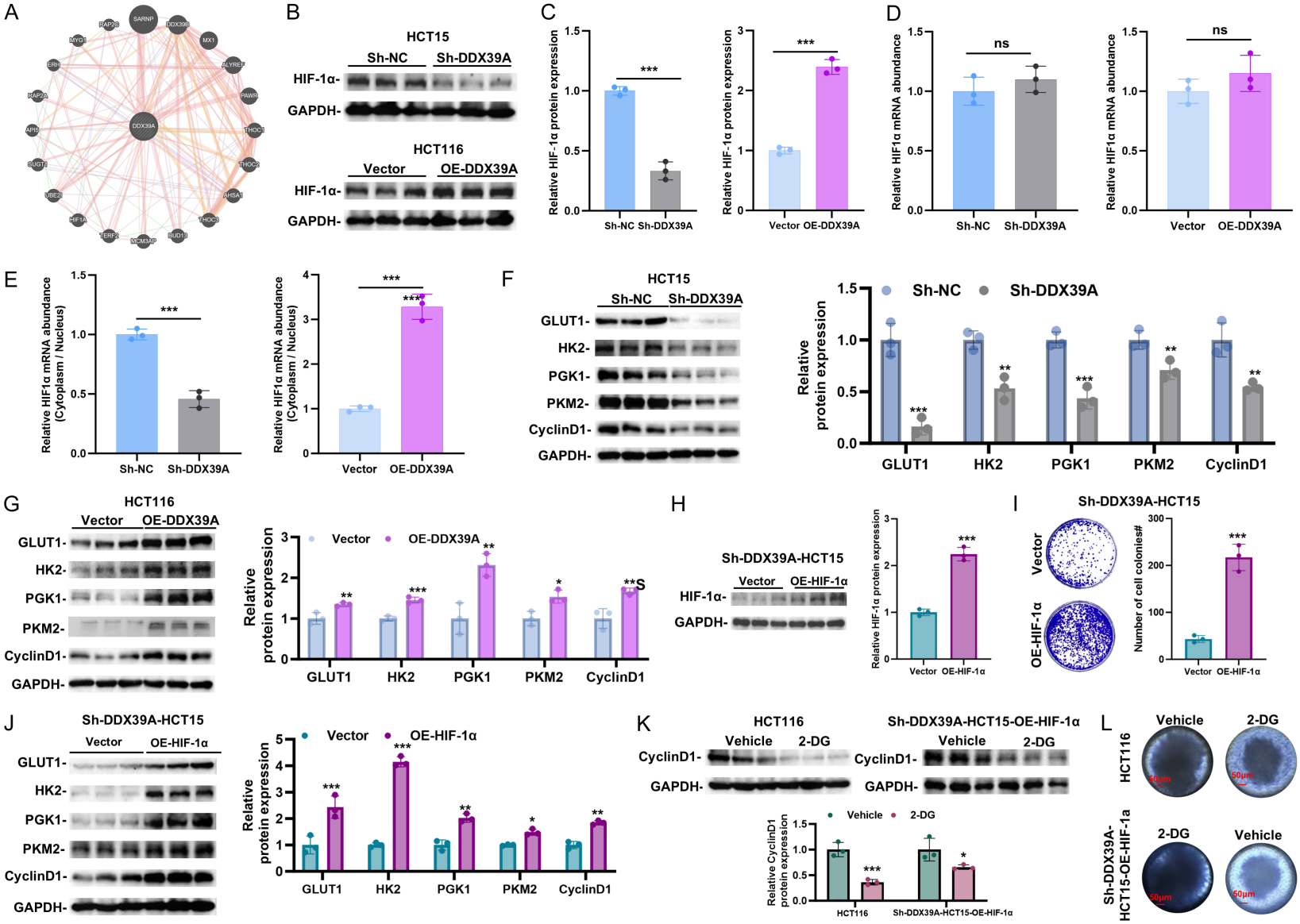
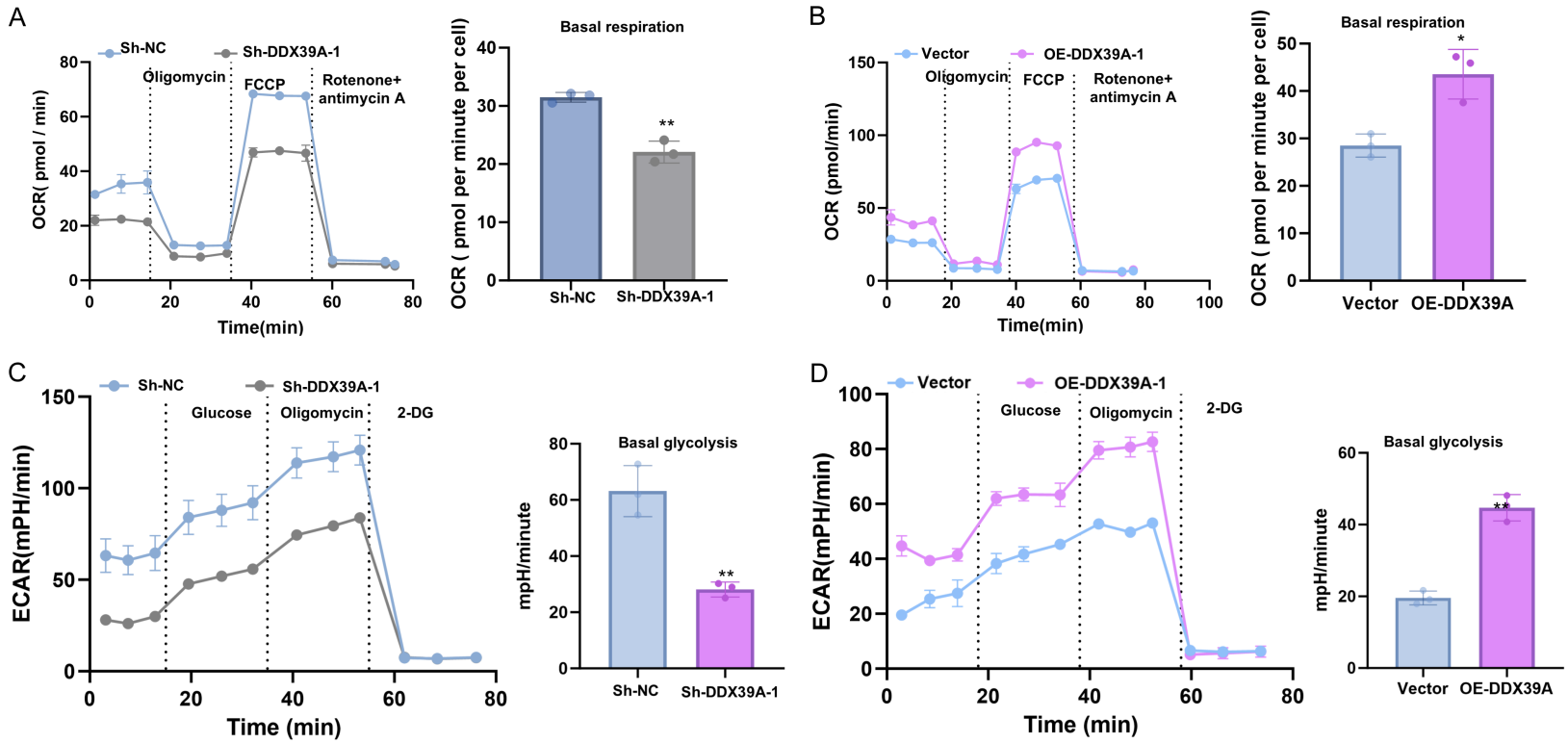

A
SARNP
DDX39B
MX1
ALYREF
PAWR
THOC1
THOC2
AHSA1
THOC5
BUD13
NCBP3
TERF2
HIF1A
UBE2I
SUGT1
PAIP2A
APEX1
ERH
MYC
PAIP2
DDX39A
B
HCT15
Sh-NC
Sh-DDX39A
HIF-1α-
GAPDH-
HCT116
Vector
OE-DDX39A
HIF-1α-
GAPDH-
C
Relative HIF-1α protein expression
Sh-NC Sh-DDX39A
Relative HIF-1α protein expression
Vector OE-DDX39A
D
Relative HIF1α mRNA abundance
ns
Sh-NC
Sh-DDX39A
Relative HIF1α mRNA abundance
ns
Vector
OE-DDX39A
E
Relative HIF1α mRNA abundance (Cytoplasm / Nucleus)
Sh-NC
Sh-DDX39A
Relative HIF1α mRNA abundance (Cytoplasm / Nucleus)
Vector
OE-DDX39A
F
HCT15
Sh-NC
Sh-DDX39A
GLUT1-
HK2-
PGK1-
PKM2-
CyclinD1-
GAPDH-
Sh-NC
Sh-DDX39A
Relative protein expression
GLUT1
HK2
PGK1
PKM2
CyclinD1
G
HCT116
Vector
OE-DDX39A
GLUT1-
HK2-
PGK1-
PKM2-
CyclinD1-
GAPDH-
Vector
OE-DDX39A
Relative protein expression
GLUT1
HK2
PGK1
PKM2
CyclinD1
H
Sh-DDX39A-HCT15
Vector
OE-HIF-1α
HIF-1α-
GAPDH-
Relative HIF-1α protein expression
Vector
OE-HIF-1α
I
Sh-DDX39A-HCT15
Vector
OE-HIF-1α
Number of cell colonies#
Vector
OE-HIF-1α
J
Sh-DDX39A-HCT15
Vector
OE-HIF-1α
GLUT1-
HK2-
PGK1-
PKM2-
CyclinD1-
GAPDH-
Vector
OE-HIF-1α
Relative protein expression
GLUT1
HK2
PGK1
PKM2
CyclinD1
K
HCT116
Vehicle
2-DG
CyclinD1-
GAPDH-
Sh-DDX39A-HCT15-OE-HIF-1α
Vehicle
2-DG
CyclinD1-
GAPDH-
Vehicle
2-DG
Relative CyclinD1 protein expression
HCT116
Sh-DDX39A-HCT15-OE-HIF-1α
L
HCT116
Vehicle
2-DG
Sh-DDX39A-HCT15-OE-HIF-1α
2-DG
Vehicle

Figure 7. DDX39A regulates HIF-1α expression and glycolysis-related phenotypes in CRC cells. A. Protein-protein interaction network centered on DDX39A. B, C. Western blotting and quantification of HIF-1α in HCT15 cells following DDX39A knockdown and in HCT116 cells following DDX39A overexpression. D. Total HIF1A mRNA levels measured by RT-qPCR after DDX39A knockdown or overexpression. E. Cytoplasmic-to-nuclear ratios of HIF1A mRNA determined by nuclear-cytoplasmic RNA fractionation and RT-qPCR. F, G. Western blotting and quantification of GLUT1, HK2, PGK1, PKM2, and Cyclin D1 following DDX39A knockdown or overexpression. H. Western blot validation of HIF-1α overexpression in sh-DDX39A HCT15 cells. I. Colony formation of sh-DDX39A HCT15 cells following HIF1A overexpression. Cells were seeded at 1,000 cells per well and cultured for 14 days before staining and colony counting. J. Western blotting and quantification of GLUT1, HK2, PGK1, PKM2, and Cyclin D1 following HIF1A overexpression in sh-DDX39A HCT15 cells. K. Cyclin D1 expression in HCT116 and HIF1A-overexpressing sh-DDX39A HCT15 cells treated with vehicle or 5 mM 2-DG for 48 h. L. Representative 3D spheroid images after vehicle or 5 mM 2-DG treatment for 48 h (×400; Scale bar = 50 μm). All experiments were independently performed at least three times (n ≥ 3), and representative images are shown. Detailed experimental procedures are described in the Materials and Methods section. Data are presented as mean ± SD. Statistical significance was determined using the indicated statistical tests. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.



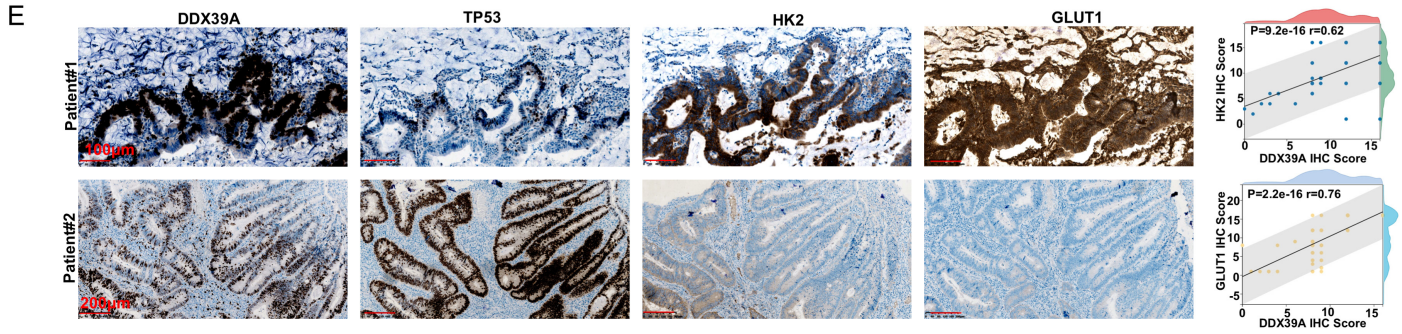


Figure 8. DDX39A is associated with metabolic reprogramming and glycolytic activation in colorectal cancer. A. Seahorse analysis showing reduced OCR in Sh-DDX39A cells; B. OCR measurement showing increased mitochondrial respiration in DDX39A overexpression cells; C. ECAR assay showing decreased glycolytic activity after DDX39A knockdown; D. ECAR assay showing enhanced glycolytic flux upon DDX39A overexpression; E. IHC staining of DDX39A, TP53, HK2, and GLUT1 in clinical samples, and correlation analysis showing positive association between DDX39A and glycolysis-related markers (HK2 and GLUT1) (patient#1: ×200; Scale bar = 100 μm); (patient#2: ×100; Scale bar = 200 μm). All cellular experiments were independently repeated at least three times (n ≥ 3), representative results are displayed. Detailed experimental procedures are provided in the Materials and Methods section. Data are presented as mean ± SD. Group differences were assessed using unpaired two-tailed Student's t-test. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. Abbreviations: OCR, oxygen consumption rate, ECAR, extracellular acidification rate.

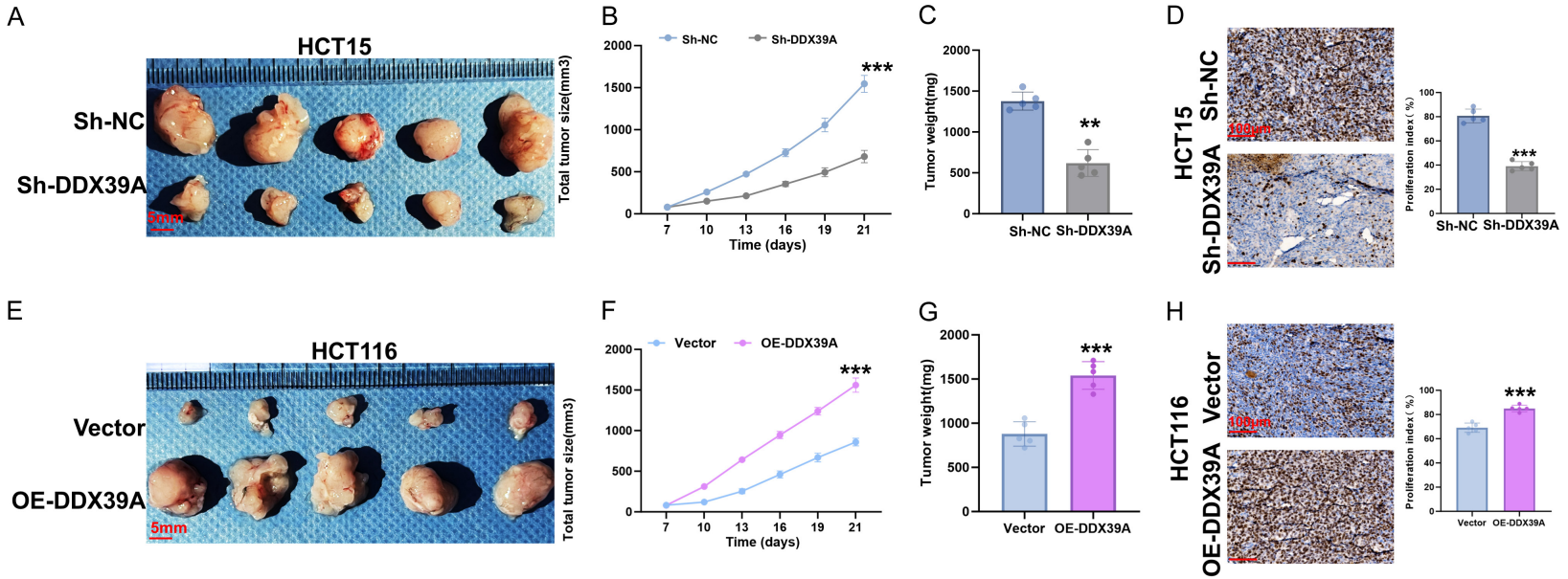


Figure 9. DDX39A promotes tumor growth in vivo. A. Representative images of xenograft tumors derived from HCT15 cells with Sh-DDX39A or Sh-NC; B. Tumor growth curves showing reduced tumor volume in the Sh-DDX39A group; C. Tumor weight comparison between Sh-NC and Sh-DDX39A groups; D. Proliferation index showing decreased tumor cell proliferation after DDX39A depletion (×200; Scale bar = 100 μm); E. Representative xenograft images from HCT116 vector and DDX39A overexpression groups; F. Tumor growth curves showing enhanced tumor growth upon DDX39A overexpression; G. Tumor weight comparison between vector and OE-DDX39A groups; H. Proliferation index showing increased tumor cell proliferation in DDX39A overexpression tumors. Each group contained 5 nude mice (n = 5) (×200; Scale bar = 100 μm). Representative images are shown. Detailed experimental procedures are described in the Materials and Methods section. Data are presented as mean ± SD. Tumor growth curves were analyzed by two-way repeated-measures ANOVA; other inter-group comparisons were performed using unpaired two-tailed Student's *t*-test. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

lular states that contribute to tumor progression and therapeutic resistance [1-3, 12, 28]. Despite advances in surgery, chemotherapy, targeted therapy, and immunotherapy, clinical outcomes for patients with advanced CRC remain unsatisfactory, highlighting the need to better understand the cellular and molecular programs associated with disease progression [2, 9, 12]. In this study, integrated single-cell transcriptomic analysis revealed a distinct malignant epithelial subpopulation, Malignant_EPC3, characterized by enhanced proliferative activity and metabolic features. WGCNA further identified a proliferation-associated module (M23) closely linked to the EPC3 phenotype, within which DDX39A was recognized as a highly connected hub gene. Clinical and functional analyses demonstrated that DDX39A expression was elevated in CRC tissues and associated with unfavorable clinical characteristics, while DDX39A alteration influenced CRC cell proliferation, migration, invasion, and tumor growth. Mechanistically, our findings suggest that DDX39A may regulate metabolic adaptation through modulation of HIF-1α expression. DDX39A alteration affected HIF-1α protein abundance without markedly changing total HIF1A mRNA levels, and nuclear-cytoplasmic fractionation suggested that DDX39A may influence HIF1A mRNA cytoplasmic distribution. Restoration of HIF1A expression partially rescued the proliferative and glycolytic defects induced by DDX39A depletion, supporting the involvement of the DDX39A/HIF-1α axis in maintaining metabolic activity of CRC cells.

Recent studies have highlighted the important roles of DEAD-box RNA helicases in cancer initiation and progression, owing to their diverse functions in RNA metabolism, including RNA splicing, nuclear export, and post-transcriptional regulation [4, 17, 19, 20, 24]. Dysregulation of these RNA regulatory processes can influence tumor cell proliferation, stress responses, and cellular adaptation. As an essential component of the RNA metabolic network, DDX39A has gradually emerged as a key molecular node linking post-transcriptional regulation to malignant phenotypes; however, its functional roles and underlying mechanisms remain context-dependent across different cancer types [24]. In this study, DDX39A was identified through WGCNA as a highly connected gene within the EPC3-associated M23 module and was found to be upregulated in CRC tissues. Its expression was associated with advanced tumor characteristics, metastasis, and unfavorable prognosis. Functional assays further demonstrated that DDX39A promotes CRC cell proliferation, migration, invasion, and in vivo tumor growth. Moreover, DDX39A expression was associated with glycolytic activity and mitochondrial respiration, suggesting a potential involvement of DDX39A in the metabolic adaptation of malignant epithelial cells. These findings suggest that DDX39A not only participates in proliferative programs in CRC but may also drive malignant phenotypes through integration of metabolic and cell cycle regulatory networks. Previous studies have also suggested a broad involvement of DDX39A in tumor-associated RNA regulatory networks [29-32]. In ovarian cancer, large-scale transcriptomic analyses revealed that DDX39A is significantly upregulated and associated with reduced overall survival [31]. Mechanistically, DDX39A promotes tumor progression by regulating alternative splicing of ITGA6, generating the oncogenic ITGA6A isoform, thereby enhancing proliferation and metastasis [31]. In glioblastoma, DDX39A is also highly expressed and predominantly localized in tumor cells, where it regulates alternative splicing of WISP1 pre-mRNA, stabilizing its expression and activating the AKT signaling pathway [32]. Moreover, tumor-derived WISP1 acts as a paracrine factor that

promotes the polarization of immunosuppressive tumor-associated macrophages, suggesting that DDX39A may indirectly shape the tumor immune microenvironment through RNA splicing-mediated regulation of secreted factors [32]. In hepatocellular carcinoma, DDX39A has been reported to sense oncogenic MYC signaling and form a positive feedback loop via PD2-mediated splicing regulation, which can be pharmacologically targeted by digoxin, further supporting its therapeutic relevance [30]. This study further revealed a potential mechanism by which DDX39A regulates metabolic phenotypes through modulation of HIF-1α expression. Notably, alteration of DDX39A expression affected HIF-1α protein abundance without substantially changing total HIF1A mRNA levels, suggesting that DDX39A may regulate HIF-1α expression through post-transcriptional mechanisms. Nuclear-cytoplasmic fractionation further indicated that DDX39A influenced the cytoplasmic distribution of HIF1A mRNA, providing a possible explanation for the discrepancy between HIF1A transcript and protein levels. Moreover, restoration of HIF1A expression partially rescued the impaired proliferation and glycolytic activity caused by DDX39A depletion, supporting the involvement of the DDX39A/HIF-1α axis in regulating CRC metabolic adaptation. Collectively, our findings extend previous knowledge of DDX39A by demonstrating that, in CRC, DDX39A is not only associated with malignant epithelial cell states but may also regulate metabolic adaptation through HIF-1α-related post-transcriptional regulation. This provides a potential molecular link between RNA regulatory machinery and tumor metabolic phenotypes.

Increasing evidence indicates that metabolic adaptation is a critical feature of cancer progression and therapeutic resistance, with glycolysis and mitochondrial oxidative phosphorylation (OXPHOS) serving as complementary metabolic pathways that support tumor cell survival and biosynthetic demands [15, 28, 33-37]. For example, novel mitochondria-targeted metal complexes, such as α-tocopherol succinate (α-TOS)-functionalized Ir(III) complex Ir2, simultaneously inhibit OXPHOS and glycolysis, thereby inducing reactive oxygen species (ROS) accumulation, mitochondrial depolarization, and DNA oxidative damage, ultimately activating apoptosis and autophagy pathways to suppress tumor proliferation [38]. In addition, nanoparticle-based NOX1/4 inhibition strategies (HANP/GKT831) also coordinately suppress glycolysis and OXPHOS, enhance radiosensitivity through metabolic and redox stress, and remodel the tumor immune microenvironment by increasing antitumor immune cell infiltration [39, 40]. In this context, RNA regulatory factors, including DEAD-box RNA helicases, have emerged as important modulators of tumor metabolic states by controlling post-transcriptional gene regulation and cellular adaptation [16, 18]. Their dysregulation can reshape tumor cell phenotypes through alterations in RNA splicing and gene expression networks [15, 16]. HIF-1α is a well-established regulator of tumor metabolic adaptation, particularly through the transcriptional activation of glycolytic programs [41, 42]. However, the upstream mechanisms controlling HIF-1α expression beyond transcriptional regulation remain incompletely understood. In this study, we found that DDX39A may regulate HIF-1α expression at the post-transcriptional level in CRC cells. Alteration of DDX39A expression affected HIF-1α protein abundance without markedly changing total HIF1A mRNA levels, while nuclear-cytoplasmic fractionation suggested that DDX39A influenced the cytoplasmic distribution of HIF1A mRNA. Moreover, restoration of HIF1A expression partially rescued the glycolytic and proliferative defects induced by DDX39A depletion, supporting a role for the DDX39A/HIF-1α axis in regulating metabolic adaptation of CRC cells.

The development and progression of CRC are closely associated with long-term self-renewal of the intestinal epithelium and persistent genotoxic stress, which together increase the susceptibility of intestinal stem cells to malignant transformation under conditions of accumulated DNA damage and impaired repair [2-4, 43, 44]. TP53, a canonical tumor suppressor, plays an essential role in maintaining genomic stability and coordinating cellular responses to DNA damage and stress conditions [44]. In this study, immunohistochemical analysis revealed an inverse correlation between DDX39A expression and TP53 protein levels in CRC tissues. However, TP53 immunoreactivity reflects p53 protein expression patterns rather than directly indicating intact p53 signaling activity, as increased TP53 staining may also result

from the accumulation of mutant p53 proteins. Since TP53 mutation status was not evaluated in the present study, this association should be interpreted cautiously and limited to the correlation between DDX39A expression and TP53 protein levels. These findings suggest that DDX39A-associated malignant epithelial characteristics are accompanied by alterations in TP53-related molecular features, although the underlying biological relationship between DDX39A and TP53 requires further investigation.

This study has several limitations. First, although our findings identify a potential association between DDX39A and HIF-1α-mediated metabolic adaptation, the detailed RNA regulatory mechanisms of DDX39A remain to be further elucidated. Future studies incorporating transcriptomic and RNA-processing analyses are required to identify its direct RNA targets and downstream regulatory networks. Second, CellChat-based communication analysis provides computational predictions rather than direct evidence of immune regulation, and the immunodeficient xenograft model limits the evaluation of DDX39A-mediated immune microenvironmental effects. Third, the relationship between DDX39A and TP53 should be interpreted cautiously, as TP53 immunostaining reflects protein expression patterns rather than functional activity, and TP53 mutation status was not assessed in this study. Finally, larger independent cohorts and multivariable clinical analyses are needed to further validate the prognostic significance of DDX39A in CRC.

Conclusion

This study provides a single-cell resolution landscape of malignant epithelial heterogeneity in colorectal cancer and identifies EPC3 as a distinct epithelial state characterized by enhanced proliferative activity and metabolic features. Integrated transcriptomic analysis and functional validation revealed DDX39A as a candidate molecular node associated with the EPC3 phenotype. Mechanistically, our findings suggest that DDX39A contributes to CRC metabolic adaptation through regulation of HIF-1α expression, potentially involving post-transcriptional modulation of HIF1A mRNA distribution. These findings expand the understanding of how RNA regulatory mechanisms are linked to malignant epithelial states and highlight DDX39A as a potential biomarker and therapeutic candidate for colorectal cancer.

Acknowledgements

ChatGPT (GPT-5.6, OpenAI) was used solely for language polishing and grammar correction. The manuscript drafting, study design, data collection, statistical analysis, interpretation of results, and conclusions were independently completed by the authors. No AI tools were involved in generating scientific content, experimental data, or figures. The authors take full responsibility for the integrity and accuracy of this work.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Guocai Zhong, Department of Pathology, Beian Hospital, Beidahuang Group, No. 1 Beidahuang Street, Beian, Heihe 164000, Heilongjiang, China. Tel: +86-15046957177; E-mail: 174443750@qq.com

References

[1] Siegel RL, Wagle NS, Star J, Kratzer TB, Smith RA and Jemal A. Colorectal cancer statistics, 2026. CA Cancer J Clin 2026; 76: e70067.

[2] Eng C, Yoshino T, Ruíz-García E, Mostafa N, Cann CG, O'Brian B, Benny A, Perez RO and Cremolini C. Colorectal cancer. Lancet 2024; 404: 294-310.

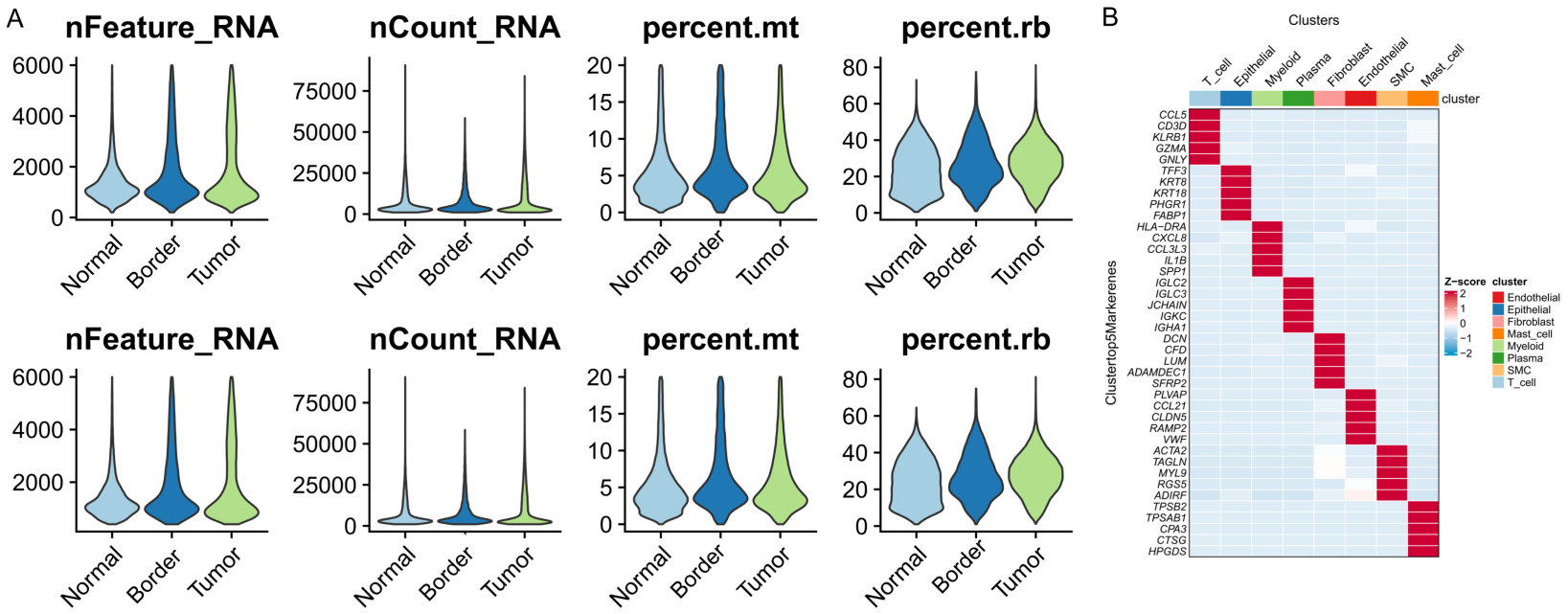
[3] Screening for colorectal cancer. CA Cancer J Clin 2026; 76: e70087.

[4] Cascianelli S, Barbera C, Ulla AA, Grassi E, Lupo B, Pasini D, Bertotti A, Trusolino L, Medico E, Isella C and Masseroli M. Multi-label transcriptional classification of colorectal cancer reflects tumor cell population heterogeneity. Genome Med 2023; 15: 37.

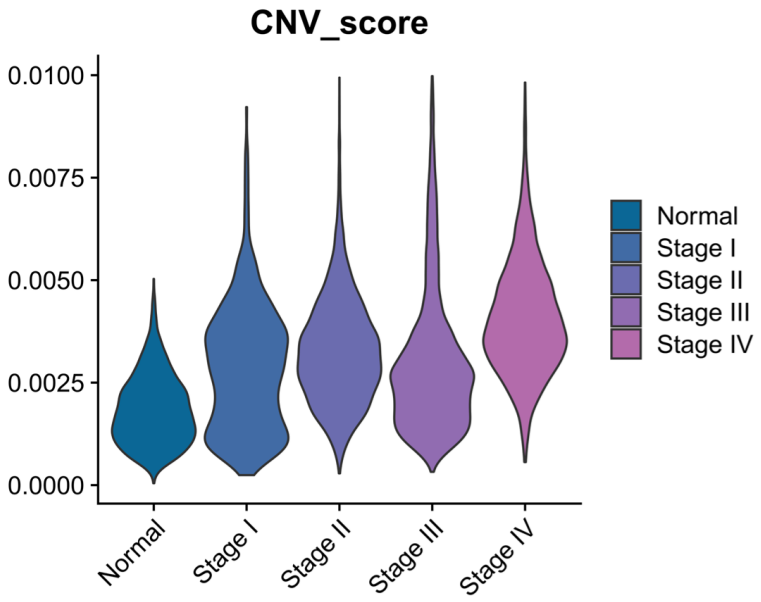
[5] Roelands J, Kuppen PJK, Ahmed EI, Mall R, Masoodi T, Singh P, Monaco G, Raynaud C, de Miranda NFCC, Ferraro L, Carneiro-Lobo TC, Syed N, Rawat A, Awad A, Decock J, Mifsud W, Miller LD, Sherif S, Mohamed MG, Rinchai D Van den Eynde M, Sayaman RW, Ziv E, Bertucci F, Petkar MA, Lorenz S, Mathew LS, Wang K, Murugesan S, Chaussabel D, Vahrmeijer AL, Wang E, Ceccarelli A, Fakhro KA, Zoppoli G, Ballestrero A, Tollenaar RAEM, Marincola FM, Galon J, Khodor SA, Ceccarelli M, Hendrickx W and Bedognetti D. An integrated tumor, immune and microbiome atlas of colon cancer. Nat Med 2023; 29: 1273-1286.

[6] Patel AS and Yanai I. A developmental constraint model of cancer cell states and tumor heterogeneity. Cell 2024; 187: 2907-2918.
[7] Huang Z, Chen J, Zhu T, Li J, Ng HY, Zhou Y, Gu X, Xu S and Jia R. Cancer-associated fibroblasts in the tumor microenvironment: heterogeneity, crosstalk mechanisms, and therapeutic implications. Mol Cancer 2025; 25: 19.
[8] Le J, Dian Y, Zhao D, Guo Z, Luo Z, Chen X, Zeng F and Deng G. Single-cell multi-omics in cancer immunotherapy: from tumor heterogeneity to personalized precision treatment. Mol Cancer 2025; 24: 221.
[9] Dentro SC, Leshchiner I, Haase K, Tarabichi M, Wintersinger J, Deshwar AG, Yu K, Rubanova Y, Macintyre G, Demeulemeester J, Vázquez-García I, Kleinheinz K, Livitz DG, Malikic S, Donmez N, Sengupta S, Anur P, Jolly C, Cmero M, Rosebrock D, Schumacher SE, Schumacher SE, Fan Y, Fittall M, Drews RM, Yao X, Watkins TBK, Lee J, Schlesner M, Zhu H, Adams DJ, McGranahan N, Swanton C, Getz G, Boutros PC, Imielinski M, Beroukhim R, Sahinalp SC, Ji Y, Peifer M, Martincorena I, Markowetz F, Mustonen V, Yuan K, Gerstung M, Spellman PT, Wang W, Morris QD, Wedge DC and Van Loo P; PCAWG Evolution and Heterogeneity Working Group and the PCAWG Consortium. Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. Cell 2021; 184: 2239-2254, e2239.
[10] Lei S, Ning W, Zhang H, Chen Q, Xu L, Tian S, Li X, Liu W, Xiang M, Liu F, Wang G, Wang L and Wang Z. Deep visual multi-omics reveals intratumor heterogeneity of gene expression and immunophenotype in colorectal cancer. Sci Bull (Beijing) 2026; 71: 2721-2725.
[11] Zeng D, Yu Y, Qiu W, Ou Q, Mao Q, Jiang L, Wu J, Wu J, Luo H, Luo P, Gu W, Huang N, Zheng S, Li S, Lai Y, Huang X, Fang Y, Zhao Q, Zhou R, Sun H, Zhang W, Zhang W, Bin J, Liao Y, Yamamoto M, Tsukamoto T, Nomura S, Shi M and Liao W. Immunotyping the tumor microenvironment reveals molecular heterogeneity for personalized immunotherapy in cancer. Adv Sci (Weinh) 2025; 12: e2417593.
[12] Prabhakar S and Tan IB. Identification of two intrinsic epithelial subtypes of colorectal cancer. Nat Genet 2022; 54: 924-925.
[13] Willson J. Hidden epithelial subtypes refine colorectal cancer classification. Nat Rev Cancer 2022; 22: 495.
[14] Hu H, Wu J and Lu Y. Unveiling metabolic pathway dysregulation in the malignant transformation of polyps to colorectal cancer: a single-cell analysis of epithelial cell trajectories. Int J Surg 2026; 112: 3071-3089.
[15] Harries LW. Dysregulated RNA processing and metabolism: a new hallmark of ageing and provocation for cellular senescence. FEBS J 2023; 290: 1221-1234.
[16] Sun Y, Tang X, Ye B and Ding K. DNA and RNA sequencing recapitulated aberrant tumor metabolism in liver cancer cell lines. J Hepatocell Carcinoma 2021; 8: 823-836.
[17] Borden KLB. Cancer cells hijack RNA processing to rewrite the message. Biochem Soc Trans 2022; 50: 1447-1456.
[18] Tapescu I and Cherry S. DDX RNA helicases: key players in cellular homeostasis and innate antiviral immunity. J Virol 2024; 98: e0004024.
[19] Bohnsack KE, Yi S, Venus S, Jankowsky E and Bohnsack MT. Cellular functions of eukaryotic RNA helicases and their links to human diseases. Nat Rev Mol Cell Biol 2023; 24: 749-769.
[20] Bourgeois CF, Mortreux F and Auboeuf D. The multiple functions of RNA helicases as drivers and regulators of gene expression. Nat Rev Mol Cell Biol 2016; 17: 426-438.
[21] Lin WL, Chen JK, Wen X, He W, Zarceno GA, Chen Y, Chen S, Paull TT and Liu HW. DDX18 prevents R-loop-induced DNA damage and genome instability via PARP-1. Cell Rep 2022; 40: 111089.
[22] Xu K, Sun S, Yan M, Cui J, Yang Y, Li W, Huang X, Dou L, Chen B, Tang W, Lan M, Li J and Shen T. DDX5 and DDX17-multifaceted proteins in the regulation of tumorigenesis and tumor progression. Front Oncol 2022; 12: 943032.
[23] Song L, Xie H, Fan H, Zhang Y, Cheng Z, Chen J, Guo Y, Zhang S, Zhou X, Li Z, Liao H, Han J, Huang J, Zhou J, Fang D and Liu T. Dynamic control of RNA-DNA hybrid formation orchestrates DNA2 activation at stalled forks by RNA-PII and DDX39A. Mol Cell 2025; 85: 506-522, e507.
[24] Banerjee S, Nagasawa CK, Widen SG and Garcia-Blanco MA. Parsing the roles of DExD-box proteins DDX39A and DDX39B in alternative RNA splicing. Nucleic Acids Res 2024; 52: 8534-8551.
[25] Xu Z, Nie C, Liao J, Ma Y, Zhou XA, Li X, Li S, Lin H, Luo Y, Cheng K, Mao Z, Zhang L, Pan Y, Chen Y, Wang W and Wang J. DDX39A resolves replication fork-associated RNA-DNA hybrids to balance fork protection and cleavage for genomic stability maintenance. Mol Cell 2025; 85: 490-505, e411.
[26] Liu L, Wang P, Guo C, Song L, Chen L, Qi H, Zheng Y, Xing X and Wang C. TRAIP enhances progression of tongue squamous cell carcinoma through EMT and Wnt/β-catenin signaling by interacting with DDX39A. BMC Cancer 2024; 24: 1481.
[27] Wang X, Jin J, Yan H, Liu J, Huang S, Bai H, Guo M, Cheng X, Deng T, Ba Y, Gu Y, Gao X and Hu D. The mRNA export pathway licenses viral

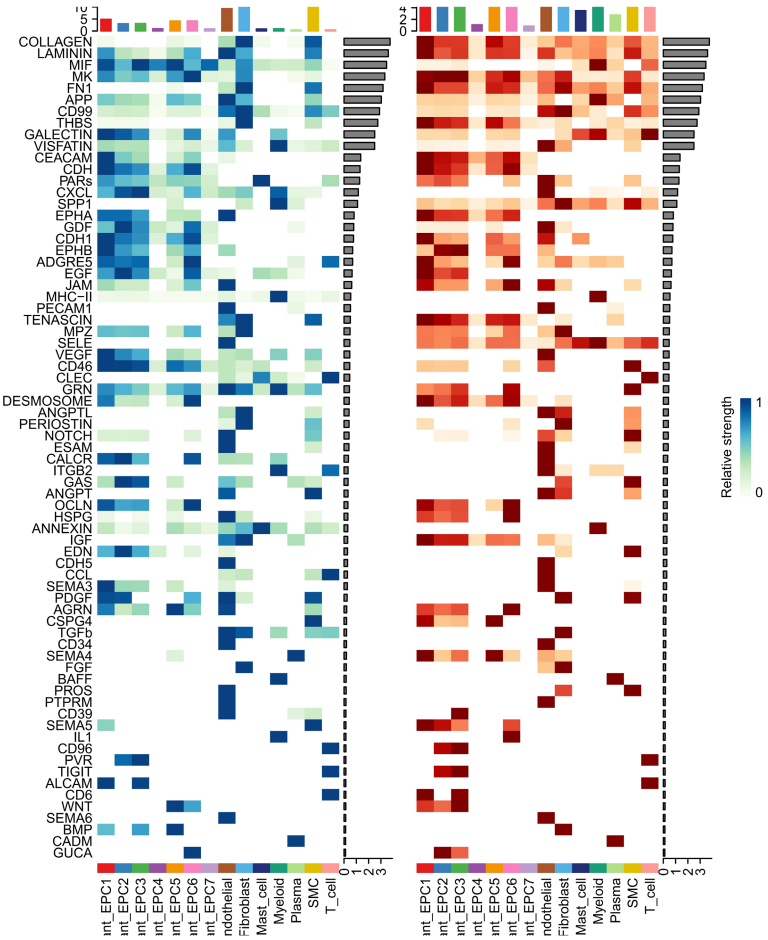
mimicry response and antitumor immunity by actively exporting nuclear retroelement transcripts. Sci Transl Med 2025; 17: eado4370.
[28] DeBerardinis RJ. Tumor microenvironment, metabolism, and immunotherapy. N Engl J Med 2020; 382: 869-871.
[29] Hui B, Hu W, Chen X, Wang Y, Li J, Sun Y, Sun X and Ren J. Role of DDX39A in modulating the tumor progression and radiosensitivity in esophageal squamous cell carcinoma. Cell Oncol (Dordr) 2026; 49: 73.
[30] Chang C, Li L, Su L, Yang F, Zha Q, Sun M, Tao L, Wang M, Song K, Jiang L, Gao H, Liang Y, Xu C, Yong C, Wang M, Huang J, Liu J, Jin W, Lv W, Dong H, Li Q, Bu F, Yan S, Qi H, Zhao S, Zhu Y, Wang Y, Shi J, Qiao Y, Xu J, Chabot B and Chen J. Intron retention of DDX39A driven by SNRPD2 is a crucial splicing axis for oncogenic MYC/spliceosome program in hepatocellular carcinoma. Adv Sci (Weinh) 2024; 11: e2403387.
[31] Li Y, Chen Z, Gao Y, Liu Y, Gao Q, Pu Y, Xiao H, Diao Y, Guo L, Yuan C, Yan S, Yang N and Kong B. SNRPB-mediated regulation of DDX39A splicing promotes ovarian cancer progression by regulating α6 integrin subunit expression. Oncogene 2025; 44: 2170-2185.
[32] Zhang Y, Xue Z, Zhang N, Zhu Y, Wu Y, Lv M, Zhang Z, Mu F, Xing W, Tang Z, Wang C, Xue Z, Zhou W, Liu X, Li X, Bjerkvig R, Huang B, Han M, Wang J and Wang D. The multifunctional RNA helicase DDX39A drives glioblastoma progression by modulating WISP1 alternative splicing that induces an immunosuppressive macrophage polarization. Oncogene 2026; 45: 1026-1041.
[33] Peng W and Zeng Z. Epigenetic activation of PTCD3 promotes CRC glutamine metabolism and metastasis via IGF2BP2-mediated SLC38A2 m6A modification. FASEB J 2025; 39: e70558.
[34] van den Berg NI, Elphick M, Mulder K, Bouricha O, Sadeghi-Alavijeh O, Fu X and Turajlic S. Immunometabolic gatekeeping: how tissue metabolism conditions tumor immunity. Cancer Cell 2026; 44: 689-695.
[35] Zhang H, Fan J, Kong D, Sun Y, Zhang Q, Xiang R, Lu S, Yang W, Feng L and Zhang H. Immunometabolism: crosstalk with tumor metabolism and implications for cancer immunotherapy. Mol Cancer 2025; 24: 249.
[36] Rauseo D, Contreras-Baeza Y, Salazar M, Galarza AJ, Holtheuer-Gallardo S, Faurand H, Cárcamo-Lemus N, Suárez R, Asenjo JL, von Faber-Castell A, Silva F, Mora-González V, Dernic J, Ravotto L, Wyss MT, Baeza-Lehnert F, Ruminot I, Alvarez-Navarro C, San Martín A, Weber B, Sandoval PY and Barros LF. Mitochondrial lactate venting limits oxidative stress. Cell Metab 2026; 38: 1130-1140, e1136.
[37] Zhang J, Han ZQ, Wang Y and He QY. Alteration of mitochondrial protein succinylation against cellular oxidative stress in cancer. Mil Med Res 2022; 9: 6.
[38] Zhang X, Liang N, Gong X, Kawashima Y, Cui F and Sun S. Tumor-targeting micelles based on folic acid and α-tocopherol succinate conjugated hyaluronic acid for paclitaxel delivery. Colloids Surf B Biointerfaces 2019; 177: 11-18.
[39] Gedla MK, Kotla RSP, Kumar R, Kamesh V, S J, Kamalakannan D and Ramasamy J. Aptamers targeting immune checkpoints for tumor immunotherapy: a comprehensive review. Front Oncol 2026; 16: 1762902.
[40] Zhang L, Liu T, Chen M, Gao S, Staley CA, Yang L and Zhu L. Dual inhibition of oxidative phosphorylation and glycolysis using a hyaluronic acid nanoparticle NOX inhibitor enhanced response to radiotherapy in colorectal cancer. Biomaterials 2025; 323: 123437.
[41] Huang C, Li Y, Li Z, Xu Y, Li N, Ge Y, Dong J, Chang A, Zhao T, Wang X, Wang H, Yang S, Xie K, Hao J and Ren H. LIMS1 promotes pancreatic cancer cell survival under oxygen-glucose deprivation conditions by enhancing HIF1A protein translation. Clin Cancer Res 2019; 25: 4091-4103.
[42] Wu Q, You L, Nepovimova E, Heger Z, Wu W, Kuca K and Adam V. Hypoxia-inducible factors: master regulators of hypoxic tumor immune escape. J Hematol Oncol 2022; 15: 77.
[43] Wei M, Huang X, Liao L, Tian Y and Zheng X. SENP1 decreases RNF168 phase separation to promote DNA damage repair and drug resistance in colon cancer. Cancer Res 2023; 83: 2908-2923.
[44] Chalabi M, Verschoor YL, Tan PB, Balduzzi S, Van Lent AU, Grootscholten C, Dokter S, Büller NV, Grotenhuis BA, Kuhlmann K, Burger JW, Huibregtse IL, Aukema TS, Hendriks ER, Oosterling SJ, Snaebjornsson P, Voest EE, Wessels LF, Beets-Tan RG, Van Leerdam ME, Schumacher TN, van den Berg JG, Beets GL and Haanen JB. Neoadjuvant immunotherapy in locally advanced mismatch repair-deficient colon cancer. N Engl J Med 2024; 390: 1949-1958.



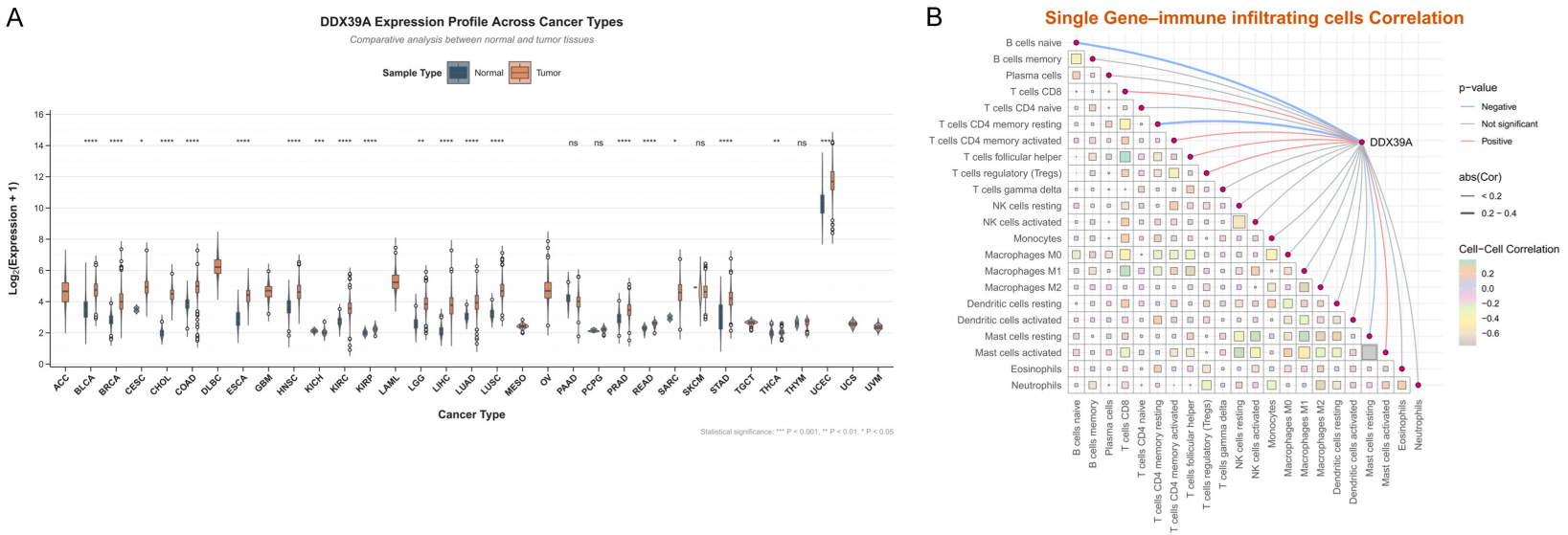
Supplementary Figure 1. A. Quality control metrics of single-cell RNA-seq data in Normal, Border, and Tumor samples, including nFeature_RNA, nCount_RNA, percent.mt, and percent.rb. Violin plots show the distribution of each metric across sample groups. B. Heatmap of cell type annotation based on canonical marker genes, showing distinct expression patterns across major cell populations. Z-scores indicate normalized expression levels.



Supplementary Figure 2. Distribution of inferred copy number variation (CNV) scores across normal and different tumor stages (Stage I-IV).



Supplementary Figure 3. Heatmap of gene expression across epithelial and stromal/immune cell clusters. Color indicates relative expression level (blue-red).



Supplementary Figure 4. Pan-cancer DDX39A expression and its correlation with immune infiltration. A. Pan-cancer DDX39A expression in normal and tumor tissues. B. Correlation between DDX39A expression and immune-cell infiltration. Line color indicates correlation direction, and line thickness indicates correlation strength. Spearman correlation analysis was applied. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Detailed procedures are described in the Materials and Methods section.

Supplementary Table 1. Multivariate Cox hazard ratio regression analysis of COAD

Variables		Progression-free Survival			Overall Survival		
		HR	95% CI	*P*	HR	95% CI	*P*
Age (Year)	≤55	2.286	0.691-7.564	0.176	0.974	0.506-1.876	0.9377
	>55						
Sex	Male	0.551	0.173-1.749	0.311	0.782	0.383-1.597	0.499
	Female						
Clinical Grade	I-II	0.046	0.007-0.307	0.001*	1.412	0.483-4.132	0.499
	III-IV						
Clinical Stage	I-II	2.316	0.469-11.438	0.303	2.427	1.031-5.716	0.042*
	III-IV						
Recurrence	No	-	-	-	1.264	0.671-2.382	0.469
	Yes						
DDX39A	Low	32.601	3.407-311.992	0.002*	4.145	1.158-14.838	0.029*
	High						

*P < 0.05 was considered statistically significant.