



## OPEN ACCESS

## EDITED BY

Qi Zhang,  
Yale University, United States

## REVIEWED BY

Bufu Tang,  
Fudan University, China  
Yihang Qi,  
Harvard Medical School, United States  
Ximing Ran,  
Emory University, United States  
Weihao Zhao,  
Yale University, United States

## \*CORRESPONDENCE

Yihang Yuan

✉ surgeryuan@outlook.com  
Peng Sun

✉ sp2082@shtrhospital.com

Xinyi Zhang

✉ 2249317206@qq.com

<sup>†</sup>These authors have contributed equally to this work and share first authorship

RECEIVED 11 October 2024

ACCEPTED 21 November 2024

PUBLISHED 09 December 2024

## CITATION

Lin Q, Wang Z, Wang J, Xu M, Zhang X, Sun P and Yuan Y (2024) Innovative strategies to optimise colorectal cancer immunotherapy through molecular mechanism insights. *Front. Immunol.* 15:1509658. doi: 10.3389/fimmu.2024.1509658

## COPYRIGHT

© 2024 Lin, Wang, Wang, Xu, Zhang, Sun and Yuan. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Innovative strategies to optimise colorectal cancer immunotherapy through molecular mechanism insights

Quanjun Lin<sup>1†</sup>, Zhiqiang Wang<sup>1†</sup>, Jue Wang<sup>1</sup>, Ming Xu<sup>1</sup>, Xinyi Zhang<sup>2\*</sup>, Peng Sun<sup>1\*</sup> and Yihang Yuan<sup>1\*</sup>

<sup>1</sup>Department of General Surgery, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, <sup>2</sup>Clinical Medical College, Southwest Medical University, Luzhou, China

**Background:** Colorectal cancer (CRC) is a leading cause of cancer-related deaths globally. The heterogeneity of the tumor microenvironment significantly influences patient prognosis, while the diversity of tumor cells shapes its unique characteristics. A comprehensive analysis of the molecular profile of tumor cells is crucial for identifying novel molecular targets for drug sensitivity analysis and for uncovering the pathophysiological mechanisms underlying CRC.

**Methods:** We utilized single-cell RNA sequencing technology to analyze 13 tissue samples from 4 CRC patients, identifying key cell types within the tumor microenvironment. Intercellular communication was assessed using CellChat, and a risk score model was developed based on eight prognostic genes to enhance patient stratification for immunotherapeutic approaches. Additionally, *in vitro* experiments were performed on *DLX2*, a gene strongly associated with poor prognosis, to validate its potential role as a therapeutic target in CRC progression.

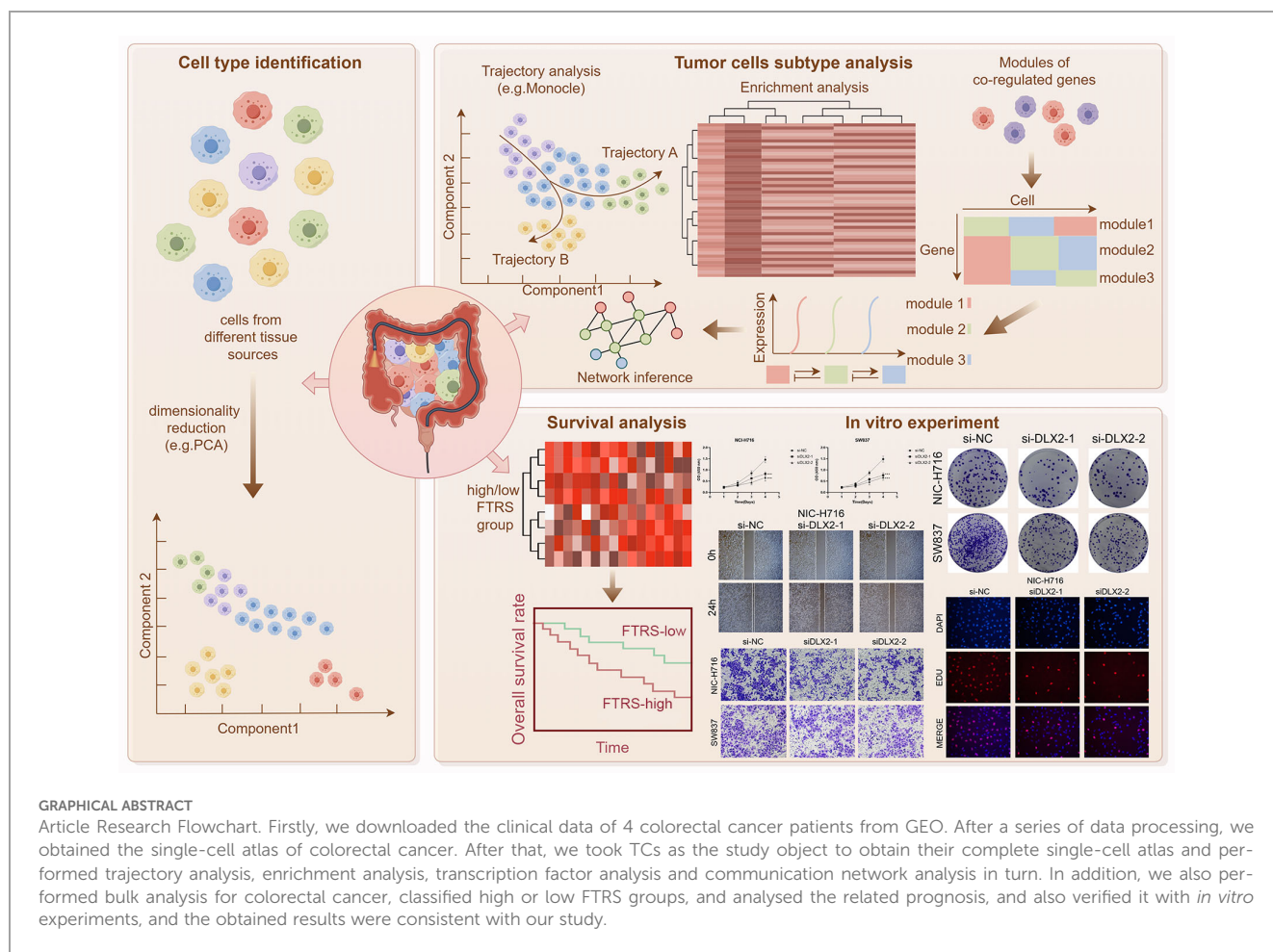
**Results:** Eight major cell types were identified across the tissue samples. Within the tumor cell population, seven distinct subtypes were recognized, with the C0 *FXVD5+* tumor cells subtype being significantly linked to cancer progression and poor prognosis. CellChat analysis indicated extensive communication among tumor cells, fibroblasts, and immune cells, underscoring the complexity of the tumor microenvironment. The risk score model demonstrated high accuracy in predicting 1-, 3-, and 5-year survival rates in CRC patients. Enrichment analysis revealed that the C0 *FXVD5+* tumor cell subtype exhibited increased energy metabolism, protein synthesis, and oxidative phosphorylation, contributing to its aggressive behavior. *In vitro* experiments confirmed *DLX2* as a critical gene associated with poor prognosis, suggesting its viability as a target for improving drug sensitivity.

**Conclusion:** In summary, this study advances our understanding of CRC progression by identifying critical tumor subtypes, molecular pathways, and

prognostic markers that can inform innovative strategies for predicting and enhancing drug sensitivity. These findings hold promise for optimizing immunotherapeutic approaches and developing new targeted therapies, ultimately aiming to improve patient outcomes in CRC.

KEYWORDS

colorectal cancer (CRC), single-cell RNA sequencing (scRNA-seq), tumor cell subtypes, FXYD5+ Tumor Risk Score (FTRS), energy metabolism in cancer



## 1 Introduction

CRC is one of the leading causes of cancer-related deaths globally and poses a major global health burden; it is the third most common malignant tumor and the second leading cause of cancer-related deaths worldwide (1, 2). It arises from the epithelial cells of the colon or rectum, often following a multi-step progression from normal tissue to adenoma (a benign tumor) and ultimately to carcinoma (a malignant tumor). This is a multi-step process that requires the accumulation of

genetic/epigenetic aberrations (3–5). From a molecular point of view, CRC is caused by mutations in targeted oncogenes, tumor suppressor genes, and genes associated with DNA repair mechanisms. CRC originates from 3 main pathways: the adenoma-cancer sequence, the serrated pathway, and the inflammatory pathway (6). Depending on the origin of the mutation, colorectal cancer can be categorized as sporadic, hereditary, and familial (7, 8). Disease progression is influenced by a variety of genetic (9), environmental (10), and lifestyle factors, of which dietary factors are the most critical, shaping

the gut microbiota to modulate regulation of colonic epithelial cell homeostasis and carcinogenesis, and preventing tumors in up to 50% of cases by improving lifestyle (7, 11–13). People with CRC usually present with breathlessness and fatigue. These are also common symptoms in patients with chronic heart failure (14–17). Currently, common treatments for CRC include a multimodal combination of surgery, radiotherapy and chemotherapy (18). However, recurrence and metastasis rates remain high. The combination of radiotherapy and immunotherapy may provide a new solution to this problem and has rapidly become the mainstay of treatment for many types of solid cancers, including CRC, but its future remains uncertain (19–22).

CRC is characterized by the heterogeneity of the tumor microenvironment (TME), with different types of cells playing different roles at different stages of tumorigenesis (23, 24). A comprehensive molecular characterization of human colon and rectal cancers has been previously revealed by deep whole genome sequencing (25). And as analytical tools continue to evolve, there are now a large number of studies that map individual cells in the TME at high resolution by scRNA-seq (26–28) and allow for clinical stratification of CRC (29). These analyses have revealed key cellular players, including epithelial cells, fibroblasts, immune cells such as T and B lymphocytes and tumor-associated macrophages. The latest emerging analytical technique, spatial transcriptomics analysis, has also been applied in CRC, helping to further advance targeted immunotherapy (30, 31).

In recent years, research has also focused on understanding the molecular drivers of CRC. Tumor cells (TCs) undergo genetic mutations that affect key signaling pathways, such as Wnt, Notch, and p53, which regulate cell growth, differentiation, and apoptosis (32–34). Furthermore, studies have shown that the epithelial-to-mesenchymal transition (EMT) process and the stemness properties of certain tumor subtypes play critical roles in CRC metastasis and recurrence (35, 36). Despite advancements in early detection and treatment (37), the prognosis for metastatic CRC remains poor (38, 39). Therefore, continued research into the molecular and cellular mechanisms driving CRC is essential for developing more effective therapies and improving patient outcomes.

In this study, we aimed to identify and validate novel molecular targets for drug sensitivity analysis by revealing the immune profile of CRC through scRNA-seq data analysis. We explored the functional roles and differentiation trajectories of tumor cell subtypes, established a prognostic model based on genetic risk scores, and identified relevant prognostic genes. Furthermore, we validated *DLX2*, a gene strongly correlated with poor prognosis, through *in vitro* experiments. This comprehensive approach enhances our understanding of CRC progression and metastasis, providing actionable strategies to optimize immunotherapeutic approaches and improve patient outcomes.

## 2 Materials and methods

### 2.1 Data source

The CRC Single-Cell RNA Sequencing (scRNA-seq) data were obtained from the GEO website (<https://www.ncbi.nlm.nih.gov/>

*geo/*), GSE number was GSE161277. The dataset includes patient-matched blood, normal, para-cancer, polyp (adenoma), and cancer tissue. As we used publicly available data from databases, we did not need to assign an ethical approval number for this study.

### 2.2 scRNA-seq

Gene expression data was imported into the R software (version 4.2.0) and analyzed using the Seurat R package (version 4.3.0) (40). Low-quality cells were rejected: (1)  $300 < nFeature < 7,500$ ; (2)  $500 < nCount < 100,000$ ; (3) mitochondrial gene expression not exceeding 25% of the total number of genes in the cell; (4) erythrocyte gene expression not exceeding 5% of the total number of genes in the cell.

The harmony R package (v.0.1.1) (41, 42) was used to eliminate the batch effect between samples (43–45). The dimensionality reduction parameter (dim value) was set to 30 and the resolution parameter was set to 1.2. All gene expression data were normalized and scaled using the *NormalizeData* and *ScaleData* functions in the Seurat R package (v.3.1.4) (46). The “*FindVariableFeatures*” function (47) was used to select the top 2,000 most variable genes for PCA (48–51). Cells were divided into clusters based on the top 30 principal components (PCs) using *FindClusters* at a resolution of 1.0. The top 30 essential PCs were selected for uniform manifold approximation and projection (UMAP) downscaling and gene expression visualization (52, 53).

### 2.3 Cell type recognition

We conducted differential expression analysis of all genes in cell clusters using Seurat’s *FindAllMarkers* function to identify the marker genes in each cluster (54–56). An adjustment of  $P\text{-value} < 0.05$ , expression percentage  $> 0.25$ , and  $|\log_2[\text{fold change (FC)}]| > 0.25$  were taken as threshold standards for identifying marker genes. Afterwards, different cell clusters were identified and annotated by the *singleR* package based on the composition patterns of the marker genes and were then manually confirmed and corrected with the *CellMarker* database.

### 2.4 Cancer preferences analysis

To assess the predilection of TCs subtypes for cancer, odds ratios (OR) were calculated using the following computational methods (57).

### 2.5 Assessment of cell stemness

*AUCell* (58) is a method for identifying cells with active genes in scRNA-seq data. It takes a gene set as input and outputs the “activity” of that gene set in each cell. It was used in this paper to rate the stemness of subtypes of TCs. Cell stemness was assessed through the utilization of the *CytoTRACE* R package (version 0.3.3) (59), allowing the temporal order of cell differentiation to be inferred (60, 61).

## 2.6 Trajectory analysis of TCs subtypes

Single-cell pseudotime trajectories were constructed using the Monocle 2 algorithm (53, 62), an R package designed for single-cell trajectories by Qiu et al (63). This algorithm employs a machine learning technique to narrow down the high-dimensional expression spectrum to a low-dimensional space, visualized as a UMAP plot. Single cells were cast into this space and sorted into a trajectory with branching points. A dynamic expression heatmap was constructed using the `plot_pseudotime_heatmap` function. Furthermore, differential expression analysis between branches was performed using the `plot_genes_branched_heatmap` function.

The slingshot R package (version 2.6.0) was used to infer cell lineages and pseudotimes (64). It defined lineage structure using clustering-based minimum spanning trees (MSTs) and applications synchronized master curves to fit branching curves for these lineages. The `getCurves` function was used to obtain smooth trajectory curves.

## 2.7 Enrichment analysis of cellular subtypes

Using the Gene Ontology (GO) (52, 65–67), Genome Enrichment Analysis (GSEA) (<http://software.broadinstitute.org/gsea/msigdb>) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (68) tools in the ClusterProfiler R package (version 4.6.0) (63), differentially expressed genes (DEGs) were enriched and analyzed. Significance of GO terms was determined based on an adjusted  $P < 0.05$  (69, 70).

## 2.8 Gene regulatory network analysis

To investigate the top 5 transcription factors (TFs) with the most prominent expression changes in each TCs subtypes, single-cell regulatory network inference and clustering analyses were performed using the pySCENIC R package (version 0.10.0) in Python (version 3.7). Initially, GRNBoost was leveraged to identify potential target genes for each TFs. Then, DNA-motif analysis helped to determine potential direct binding targets. Lastly, the activity of regulon in the cells was scored by AUCell, and the top 5 TFs with the highest score were identified. The human gene-motif rankings were derived from <https://resources.aertslab.org/cistarget/>. Regulon modules were then identified according to the Connection Specificity Index (CSI) to confirm specific associating partners (71). Hierarchical clustering with Euclidean distance was then conducted to identify different regulon modules. We then constructed a regulator association network using 0.65 as a cutoff to investigate the relationship between different regulators.

## 2.9 Cell communication analysis

The CellChat R package (version 1.6.1) was utilised to infer sophisticated intercellular interactions and construct regulatory networks based on ligand-receptor levels (72). It utilizes

the “`netVisualDiffInteraction`” function to indicate differences in the strength of intercellular communication and the “`IdentifyCommunicationPatterns`” function to indicate an estimate of the number of communication patterns. We used a significance threshold setting with a P-value cutoff of 0.05 to predict cell-cell interactions between different cell types.

## 2.10 Construction and validation of the prognostic model

To elucidate the impact of CRC-associated TCs on predicting patient survival, we employed marker genes of critical tumor cell subtypes as predictive factors for the construction of a prognostic model. Employing univariate Cox analysis and LASSO regression analysis (73–75), we identified the most significant prognostic genes. Multivariate Cox regression analysis was then conducted to calculate the risk coefficients for each prognostic gene (76–78), enabling the establishment of a risk score model:

Risk score =  $\sum_i^n X_i \times Y_i$ , where  $X$  represents the coefficient and  $Y$  represents the gene expression level.

Patients were categorized into low-risk and high-risk groups based on the optimal cutoff value calculated using the “`surv_cutpoint`” function. To observe the prognostic outcomes in different patient cohorts, we further utilized the R package “Survival” (version 3.3.1) for survival analysis of the constructed risk score model and visualized survival curves using the “`ggsurvplot`” function. The accuracy and calibration of the predictive model were evaluated and calibrated using the “`timeROC`” package (version 0.4.0) to plot ROC curves. The AUC values for the time-dependent ROC curve were computed at different time intervals to evaluate the model’s predictive accuracy over time (79).

CIBERSORT R package is a machine learning approach, which estimates the abundance of cell clusters in bulk RNA-seq data (80). We used this tool to digitally purify the transcriptomes of individual cell clusters from bulk data without isolating single cells. We extracted the transcripts per million (TPM) normalization datasets of selected cell types including  $CD8^+$  T cells, fibroblasts, myeloid cells and epithelial cells to create the signature matrix in 1000 permutations and without batch correction. We then divided the CRC patients available in the TCGA database into training and testing cohorts at a 1:1 ratio using survival-based randomisation and estimated the proportion of each cell cluster in the training and testing cohorts, respectively, using CIBERSORT. Notably, the bulk RNA-seq data from TCGA was initially normalized to TPM values.

## 2.11 Cell culture

NIC-H716 and SW837 cell lines were obtained from American Type Culture Collection (ATCC). NIC-H716 cell line was cultured in RPMI 1640 medium containing 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin under standard conditions (37°C, 5% CO<sub>2</sub>, 95% humidity). SW837 cell line was cultured in RPMI 1640 medium containing 10% fetal bovine serum (FBS), 1%

penicillin-streptomycin under standard conditions (37°C, 5% CO<sub>2</sub>, 95% humidity) (81).

## 2.12 Cell transfection

*DLX2* knockdown was achieved using small interfering RNA (siRNA) constructs from GenePharma (Suzhou, China). Transfection was performed following the manufacturer's instructions of Lipofectamine 3000 RNAiMAX (Invitrogen, USA). Cells were seeded at 50% confluency in 6-well plates and transfected with negative control (si-NC) and knockdown constructs (Si-*DLX2-1* and Si-*DLX2-2*). Lipofectamine 3000 RNAiMAX (Invitrogen, USA) was used for each transfection.

## 2.13 Cell viability assay

Cell viability of transfected NIC-H716 and SW837 cells was assessed using CCK-8 assay. Cells were seeded in 96-well plates at a density of  $5 \times 10^3$  cells per well and cultured for 24 hours. After adding 10  $\mu$ L of CCK-8 reagent (A311-01, Vazyme) per well, plates were incubated at 37°C in the dark for 2 hours. Absorbance at 450nm was measured using a microplate reader (A33978, Thermo) on days 1, 2, 3, and 4 post-transfection. Mean OD values were calculated and plotted (82).

## 2.14 Wound-healing assay

Stably transfected cells were seeded in 6-well plates and grown to confluence. A sterile 200  $\mu$ L pipette tip was used to scratch each well, which was then washed with PBS to remove cell debris and incubated in serum-free medium. Images of the scratches were captured at 0 hours and 48 hours, and the width of the scratch was measured using Image-J software (21).

## 2.15 Transwell assay

Cells were starved in serum-free medium for 24 hours prior to the experiment. Following treatment with Matrigel (BD Biosciences, USA), cell suspension was added to the upper chamber of Costar plates, with serum-containing medium in the lower chamber. Cells were incubated for 48 hours in a cell culture incubator. Following cultivation, cells were fixed with 4% paraformaldehyde and stained with crystal violet to assess their invasive capacity.

## 2.16 5-Ethynyl-2'-deoxyuridine proliferation assay

Transfected NIC-H716 and SW837 cells were seeded at  $5 \times 10^3$  cells per well in 6-well plates and cultured overnight. Subsequently,

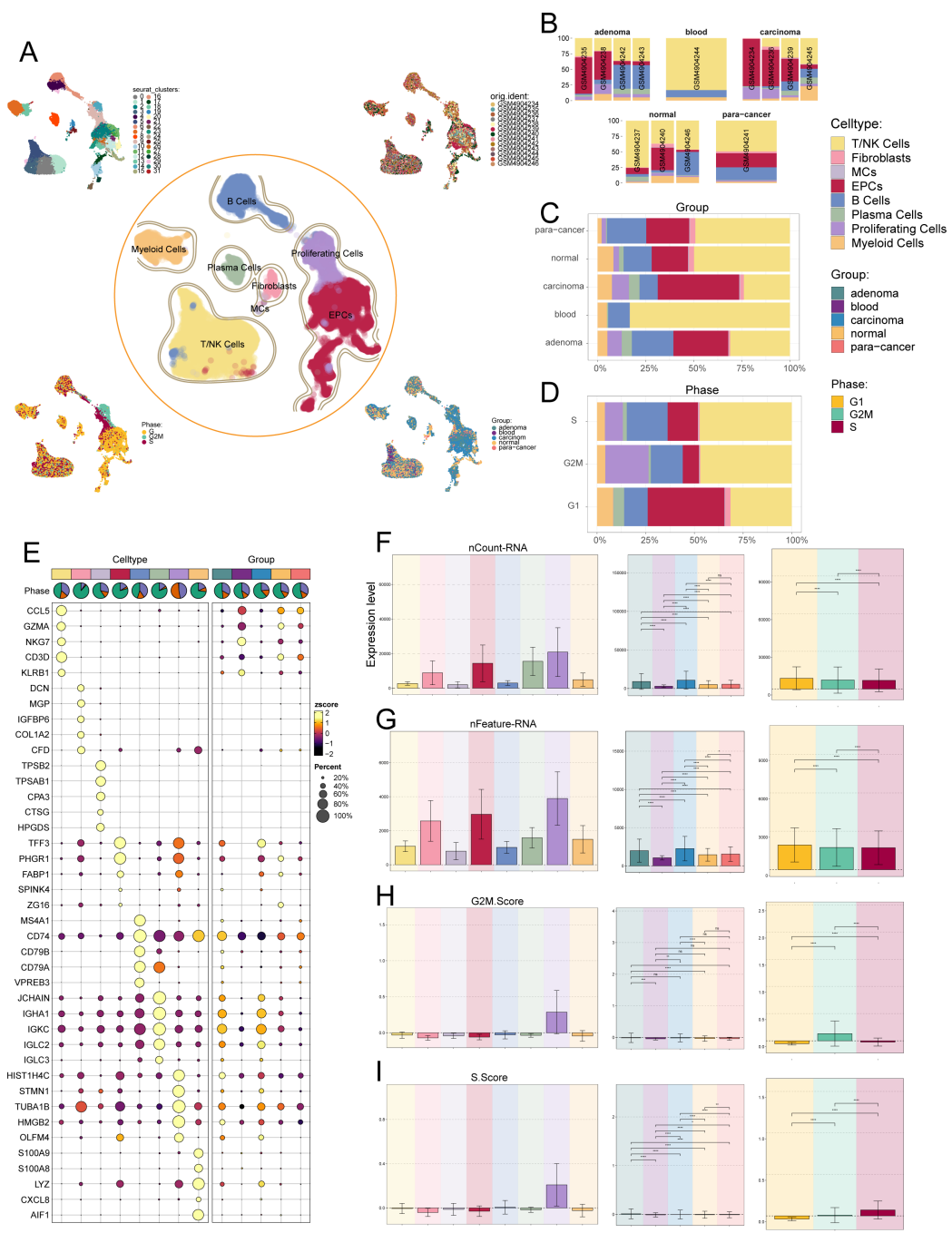
a 2x EdU working solution was prepared by mixing 10 mM EdU with serum-free medium. After a 2-hour incubation at 37°C, cells were washed with PBS, fixed with 4% paraformaldehyde for 30 minutes, permeabilized with 2 mg/mL glycine and 0.5% Triton X-100 for 15 minutes, and then stained with a mixture of 1X Apollo and 1X Hoechst 33342 for 30 minutes at room temperature. Cell proliferation was quantified using fluorescence microscopy.

## 3 Results

### 3.1 Single-cell mapping of CRC was obtained by scRNA-seq, yielding 8 cell types

To identify the major cell types during CRC progression, we collected 13 tissue samples containing adenoma, blood, carcinom, normal and para-cancer from 4 CRC patients on GEO for scRNA-seq. In addition, we checked the quality and completeness of the raw data. We then excluded genes in the samples that did not reach the minimum expression threshold. After initial quality control and removal of batch effects, we maintained a total of 42,133 cells (Figure 1A). We classified these 42,133 cells into 32 cell clusters by dimensionality reduction clustering. Based on cellular genetic profiles and typical markers, we finally identified the 32 cell clusters into 8 cell types, including T and NK cells (*CD3D*), fibroblasts (*DCN*), mast cells (MCs, *TPSB2*), epithelial cells (EPCs, *PHGR1*), B cells (*CD74*), plasma cells (*JCHAIN*), proliferating cells (*TUBA1B*) and myeloid cells (*LYZ*). From the UMAP plots and the bar plots we could learn that from normal to para-cancer to adenoma and carcinoma, the proportion of EPCs and proliferating cells roughly increases gradually, in contrast to a gradual decrease in the proportion of T and NK cells, and the proportion of B cells from normal to para-cancer to adenoma was gradually increasing but decreased in carcinoma; for cell cycle, most EPCs were in the G1 phase, which was the pre-DNA synthesis period with active material metabolism, aiming to prepare the material and energy for DNA replication in the next stage, the S phase (83), whereas T and NK cells were evenly distributed in all three periods, and proliferating cells were only present in S and G2M phases, synthesizing DNA as well as making sufficient preparations for mitosis, in line with their proliferative properties (Figures 1B–D). The bubble plot showed us the first five markers used to distinguish cell types, and the results were consistent with immunology (Figure 1E). The bar graphs showed the expression levels of nCount-RNA, nFeature-RNA, G2M.Score and S.Score for different cell types, tissue sources and cell cycle, respectively. From the results, it could be seen that proliferating cells had the highest expression levels of all four items, while EPCs, fibroblasts and plasma cells ranked behind them. In the tissue source, from normal to para-cancer to adenoma to carcinoma, the expression of both nCount-RNA and nFeature-RNA gradually increased, and there were statistically significant differences between all groups (Figures 1F–I).

Among the above cell types, EPCs attracted our attention. EPCs not only have relatively abundant RNA counts, but also tend to aggregate gradually during the developmental stages of cancer,



**FIGURE 1**

Single-cell analysis was carried out on CRC, including 32 cell populations, 8 cell types, and 5 tissue sources. **(A)** UMAP plot showed the 32 clusters of cells in CRC patients (top left); UMAP plot showed the 8 major cell types obtained by dimensionality reduction clustering of cells in CRC (middle); UMAP plot showed the distribution of sample sources (top right), cell phases (bottom left), and 5 tissue sources (bottom right) of all cells. Each point corresponded to a single cell colored according to cell cluster or cell type. **(B-D)** The bar graphs showed the proportion of different cell types in sample sources **(B)**, tissue sources **(C)** and cell phases **(D)** respectively. **(E)** Bubble plot showed differential expression of top 5 marker genes in CRC across 8 cell types. Bubble colors were based on normalized data (zscore), and sizes indicated the percentage of genes expressed in each cell type. **(F-I)** Bar graphs revealed the expression levels of nCount-RNA, nFeature-RNA, G2M.Score, and S.Score in different cell types, tissue sources and cell phases. ns stands for  $P > 0.05$  and the results are not statistically significant, \* stands for  $P \leq 0.05$ , \*\* stands for  $P \leq 0.01$ , \*\*\* stands for  $P \leq 0.001$  and \*\*\*\* represents  $P \leq 0.0001$ , a statistically significant difference.

whereas the tumor cells of CRC originate from the intestinal epithelial cells of the colon and rectum (84). The tumor cells are the most fundamental cause of tumors, so we next turned our attention to the study of CRC's tumor cells.

## 3.2 Single-cell sequencing of TCs in the CRC

Next, we identified malignant tumor cells from epithelial cells using inferCNV (Supplementary Figure S1). Due to the absence of blood vessels in the epithelial tissues, we removed the blood tissues and compiled scRNA-seq data from the remaining four tissues in the CRC for further subclustering analysis against tumor cells. The analysis yielded seven clusters, resulting in seven different tumor cell subtypes: C0 *FXYD5*+ TCs, C1 *APCDD1*+ TCs, C2 *MUC2*+ TCs, C3 *HEPACAM2*+ TCs, C4 *OTOP2*+ TCs, C5 *SLC26A3*+ TCs, and C6 *AVIL*+ TCs, and showed the distribution of the various tissue origins and cellular period distribution in the subtypes (Figure 2A). Ro/e preference plots combined with box line plots showed that C6 *AVIL*+ TCs preferred the G2M phase, while C1 *APCDD1*+ TCs preferred the S phase, and were favored by adenoma tissues, whereas in contrast, C2 *MUC2*+ TCs, C4 *OTOP2*+ TCs, and C5 *SLC26A3*+ TCs were found overwhelmingly in the normal tissue and para-cancer tissues, whereas C0 *FXYD5*+ TCs preferred carcinoma tissues (Figures 2B, C). The above tissue distribution features suggested that C1 *APCDD1*+ TCs may be a watershed for colorectal progression from adenoma to malignancy, whereas C0 *FXYD5*+ TCs was a subtype that had become cancerous. The top five marker genes obtained by enrichment in different TCs subtypes were next visualized by bubble plots. We could see that *S100A11* was significantly enriched in C0 *FXYD5*+ TCs, whereas overexpression of *S100A11* in cytoplasmic and nuclear subcellular compartments was associated with tumor metastasis and poor prognosis in patients with CRC and meta-analyses demonstrated that the expression level of *S100A11* in CRC tissues significantly elevated in CRC tissues, a marker gene for clinical use in identifying CRC (85, 86). Meanwhile, *SELENBP1* was also noteworthy, as it was highly expressed in C1 *APCDD1*+ TCs, and studies had shown that it was down-regulated in CRC but not in normal tissue or adenoma tissue (87), a typical oncogene that inhibits colorectal cancer progression by suppressing the EMT (88), whereas it was low-expressed in C0 *FXYD5*+ TCs, reflecting its intermediary role in the progression from C1 *APCDD1*+ TCs to C0 *FXYD5*+ TCs (Figure 2D). From the UMAP plots and bar graphs, we could tell that the C0 *FXYD5*+ TCs and C1 *APCDD1*+ TCs subtypes had an abundant number of RNA, as well as a higher CNVscore for C0 *FXYD5*+ TCs and C2 *MUC2*+ TCs (Figures 2E–G).

Finally, the violin plots showed the marker genes of each TCs subtype in turn (Figures 2H–N), which had shown corresponding expression advantages in their respective subtypes. *FXYD5* in C0 *FXYD5*+ TCs deserved our attention, as it had been shown to activate TGF- $\beta$ /SMAD signaling and drive EMT to promote ovarian cancer progression, and to form a positive feedback loop to drive EMT during the progression of ovarian cancer and to promote tumor growth and metastasis (89); meanwhile, recent

study also showed that *FXYD5* was upregulated in various tumor types and positively correlates with tumor progression (90).

## 3.3 Analysis of stemness genes in TCs subtypes

To explore the expression of stemness genes in TCs subtypes and to comprehend their proliferative differentiation potential, we used bubble plot to illustrate the different expression of stemness genes in TCs subtypes (Figure 3A). The differential expression analysis was intended to provide a holistic view of stemness-related gene expression, including both high and low-expression markers, to fully capture the diversity of stemness-associated features across tumor cell subtypes. Results showed the corresponding expression of stemness genes *KDM5B*, *EPAS1*, *CTNBN1*, *MYC*, *KLF4* and *CD44* in TCs subtypes and different tissue sources, with *MYC* highly expressed in C0 *FXYD5*+ TCs and *PROM1* in C1 *APCDD1*+ TCs. Following this, UMAP plot showed the AUC scores of cells stemness in different tissue sources, which were significantly higher in adenoma and carcinoma than in normal and para-cancer (Figure 3B). Next, bar graphs showed the AUC scores of stemness genes in different TCs subtypes and in the cell cycle, respectively (Figures 3C, D). We next enumerated the stemness genes enriched in different TCs subtypes and different tissue sources using bubble plots, respectively, and the results remained that *MYC* was significantly enriched in C0 *FXYD5*+ TCs, and for the tissues, *PROM1*, *NOTCH1*, and *CD44* were enriched in adenoma, whereas *KDM5B* and *MYC* were enriched in carcinoma. Stemness genes were also enriched in normal and para-cancer, but at lower levels, which suggested that adenoma and carcinoma tissues had a higher proliferative potential (Figure 3E). Finally, we visualized and compared the stemness genes with higher expression in C0 *FXYD5*+ TCs and C1 *APCDD1*+ TCs using box line plots, and the results further validated the previous results, showing a stronger proliferative capacity in C0 *FXYD5*+ TCs and C1 *APCDD1*+ TCs as well as adenoma and carcinoma (Figures 3F–I).

## 3.4 Construction of single-cell trajectories of TCs using Monocle and Slingshot for pseudotime analysis

Based on the above results, in order to further clarify the differentiation status of TCs subtypes, we first used CytoTRACE to clustered all tumor cells, and obtained UMAP plots and box plots of cell stemness prediction ordering (Figures 4A, B). From these three plots, we could learn that the C0 *FXYD5*+ TCs and C1 *APCDD1*+ TCs subtypes were in a high stemness and low differentiation state, which was consistent with the above results (Figure 3). Next, the bar graphs showed the genes that correlated with the CytoTRACE results, where greater than 0 was a positive correlation and vice versa was a negative correlation (Figure 4C). We then performed Monocle analysis of tumor cells to obtain a pseudotime atlas of tumor cell subtypes, colored by the order of pseudotime, different subtypes, and different tissue sources

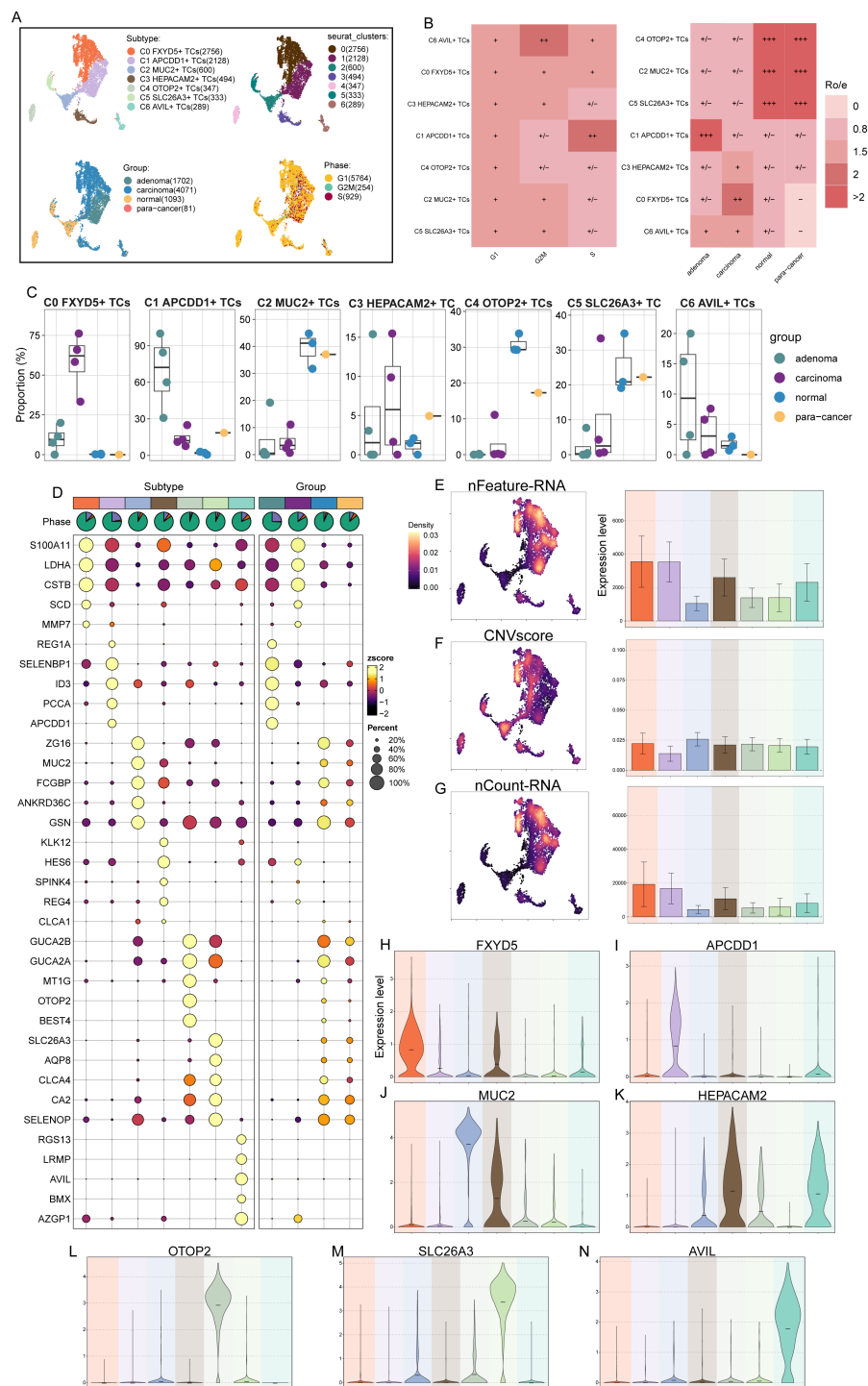
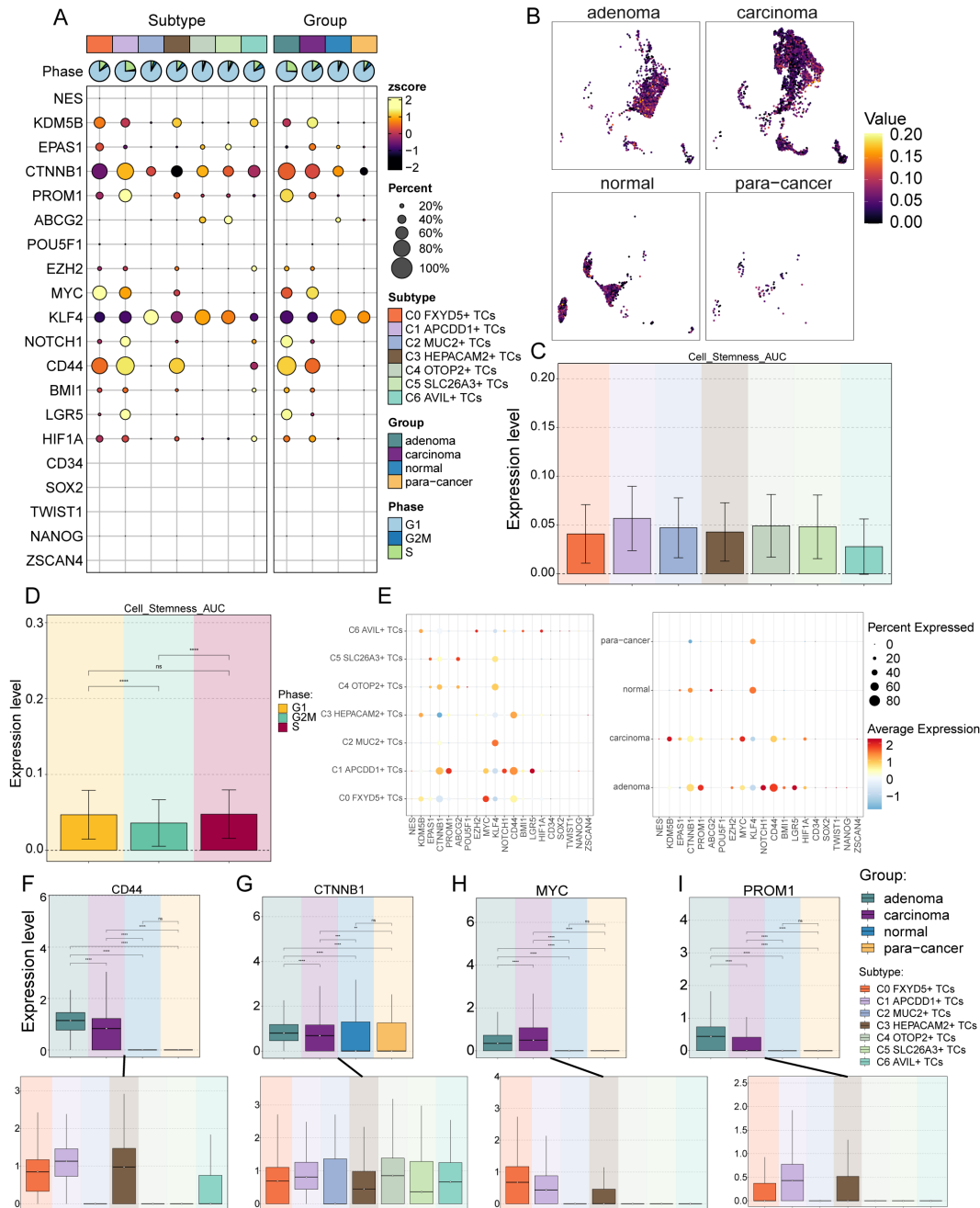


FIGURE 2

7 TCs subtypes were identified by different markers. (A) UMAP plot demonstrated the 7 cell subtypes of TCs in CRC patients and the number of cells in each cell subtype (top left); UMAP plot demonstrated the distribution of 7 clusters (top right), 4 tissue sources (bottom left) and cell phases (bottom right) in the 7 TCs subtypes respectively. Each point corresponded to a single cell colored according to cell different groups. (B) Cell phases (left) and tissue sources (right) preference of each TCs subtype estimated by Ro/e score. (C) The box line plots showed the proportion of different tissue sources in each TCs subtype. (D) Bubble plot showed differential expression of top 5 marker genes in 7 TCs subtypes. Bubble colors were based on normalized data (zscore) and sizes indicated the percentage of genes expressed in each subtype. (E-G) UMAP and bar plots revealed the expression levels of nFeature-RNA, CNVscore and nCount-RNA in different TCs subtypes. (H-N) Violin plots showed the expression levels of named gene in each of the TCs subtypes, in order of subtype nomenclature.





**FIGURE 3**  
 Analysis of cell stemness in TCs subtypes. **(A)** Bubble plot showed expression of stemness genes in 7 TCs subtypes. Bubble colors were based on normalized data (zscore) and sizes indicated the percentage of genes expressed in each subtype. **(B)** The UMAP plots showed the cell stemness AUC values for different tissue sources in TCs, respectively. **(C, D)** Bar graphs revealed the AUC values of TCs cell stemness in different subtypes **(C)** and cell phases **(D)** ns stands for  $P > 0.05$  and the results are not statistically significant and \*\*\*\* represents  $P \leq 0.0001$ , a statistically significant difference. **(E)** Bubble plots demonstrated the expression of stemness genes in different cell subtypes (left) and tissue sources (right), where the colors indicated high or low average expression levels and sizes indicated the percentage of genes expressed in each subtype and tissue source. **(F–I)** Box line plots visualized the 4 stemness genes expressed in TCs tissue sources and subtypes. ns stands for  $P > 0.05$  and the results are not statistically significant, \*\* stands for  $P \leq 0.01$ , \*\*\* stands for  $P \leq 0.001$  and \*\*\*\* represents  $P \leq 0.0001$ .

(Figures 4D–G). First, the pseudotime order speculated that the developmental trajectory started at the bottom left corner and branches out to two more trajectories when it reached state 1. Also in conjunction with the subtypes of TCs, C2 MUC2+ TCs, C4 OTOP2+ TCs, and C5 SLC26A3+ TCs subtypes were concentrated in early developmental trajectories, whereas C1 APCDD1+ TCs, C3

HEPACAM2+ TCs, and C6 AVIL+ TCs were concentrated in late developmental trajectories, and unlike C0 FXYD5+ TCs subtype, it was found in mid- to late-developmental trajectories; for the tissue origin, normal and para-cancer tissues were present at the initial developmental stage, which later i.e., developed into the adenoma and carcinoma.

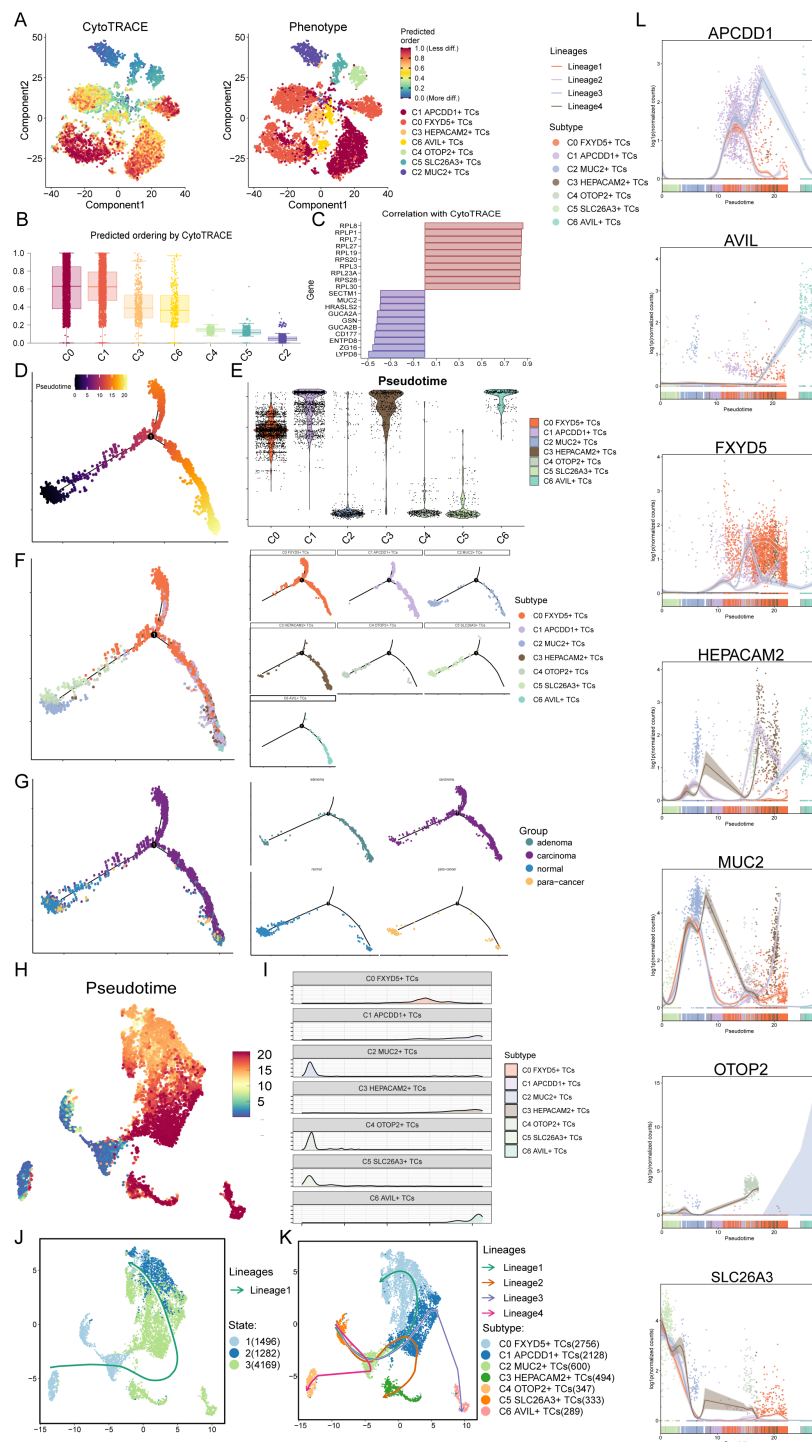


FIGURE 4

CytoTRACE and SlingShot analyses demonstrated the developmental trajectories of different subtypes of TCs. (A) The left panel demonstrated the distribution of TCs CytoTRACE scores. The color represented high or low cell stemness. The right panel indicated the distribution of TCs subtypes. The color represented different TCs subtypes. (B) Box line plot ranked the stemness of TCs subtypes according to CytoTRACE. (C) The bar graph showed the correlation of different genes with CytoTRACE, where higher than 0 indicates a positive correlation and less than 0 indicates a negative correlation. (D) Pseudotemporal trajectory plot demonstrated monocle-predicted differentiation trajectories of TCs. The distribution of pseudotime order by Monocle were shown. (E) Violin plots demonstrated the pseudotime distribution and density of different subtypes of TCs. (F, G) The pseudotime distribution of subtypes (F) and tissue sources (G) of TCs by Monocle were shown respectively. (H) Demonstration of the distribution of slingshot-predicted TCs differentiation trajectories among all TCs by UMAP plot. Plotting each spectrum according to the pseudotime value to infer the result, the color from blue to red indicates the pseudotime from naive to mature. (I) Ridgeline plots demonstrated the pseudotime distribution and density of different subtypes of TCs. (J) The distribution of differentiation trajectory of 3 states fitted by the pseudotime order in all tumor cells. (K) The distribution of four differentiation trajectories of 7 TCs subtypes fitted by the pseudotime order in all tumor cells. (L) Dynamic trend plots demonstrated the trajectories of named genes of 7 cell subtypes of tumor cells changing on four lineages obtained after slingshot visualization.

Furthermore, we then performed Slingshot analysis on tumor cells, which resulted in a pseudotime speculative map colored by pseudotime order and based on the UMAP plot obtained earlier (Figure 4H). Combined with the ridge plot, the analysis showed that, again, C2 *MUC2+* TCs, C4 *OTOP2+* TCs, and C5 *SLC26A3+* TCs subtypes were at the anterior end of the developmental trajectory, while the rest of the subtypes were located in the middle to posterior (Figure 4I). By observing how Slingshot's spectral differentiation trajectories were plotted, we could learn that tumor cells go from state 1 to state 3 to state 2; at the same time, this result yielded four spectral trajectories based on cell subtypes, as follows: Lineage1: C5 to C2 to C1 to C0; Lineage2: C5 to C2 to C1 to C3; Lineage3: C5 to C2 to C1 to C6; Lineage4: C5 to C2 to C4 (Figures 4J, K). The differences between the four trajectories were mainly in the late stage, with lineage1 ending at C0, lineage2 at C3, lineage3 at C6, and lineage4 at C4, where C0 not only had a preference for, but also a high percentage of tumor tissue. Hence, we predicted that lineage1 represented the lineage of differentiation of TCs associated with CRC tumorigenesis. Furthermore, we noted that lineage1 passed through C1, which was predominantly organized as adenoma, a precancerous, benign tumor, and previous studies had reported that participants with positive diagnostic colonoscopy results and advanced adenomas had a significantly increased risk of colorectal cancer compared to participants without adenomas (91). Therefore, studying how C1 progresses to C0 and the key factors that drive the epithelium to change from benign to malignant has important implications for how to stop CRC from occurring and worsening. Finally, the dynamic trends plots displayed the expression changes and distribution of marker genes for TCs subtypes across the four differentiation trajectories in pseudotime (Figure 4L).

### 3.5 All TCs subtypes were analyzed for enrichment, especially C0 *FXYD5+* TCs

Then, we performed GO-BP DEGs in TCs subtypes to reveal their enrichment in biological processes. Heatmap displayed the results of the top five enriched items in the seven TCs subtypes (Figure 5A). The C0 subtype was mainly associated with pathways such as cytoplasmic translation, aerobic respiration, oxidative phosphorylation, cellular respiration and energy derivation by oxidation of organic compounds; The C1 subtype was enriched in pathways such as cytoplasmic translation, ribonucleoprotein complex biogenesis, ribosome biogenesis, ribosome assembly and ribonucleoprotein complex; The C2 subtype revealed their close association with antigen processing and presentation of endogenous antigen, endogenous peptide antigen or peptide antigen; The C3 subtype showed enrichment in pathways such as protein targeting to endoplasmic reticulum (ER), establishment of protein localization to endoplasmic reticulum, cotranslational protein targeting to membrane and response to endoplasmic reticulum stress; The C4 subtype was enriched in pathways related to stress response to metal and copper ion, detoxification of copper ion and inorganic compound and regulation of cell morphogenesis; The C5 subtype

mainly exhibited enrichment in pathways such as alcohol metabolic process, viral life cycle, negative regulation of response to external stimulus, cellular hormone metabolic process and primary alcohol metabolic process; The C6 subtype revealed pathways related to actin filament organization, regulation of actin filament-based process, supramolecular fiber organization and actin cytoskeleton organization and mononuclear cell differentiation. The enrichment items derived from the C0 subtype caught our attention, and the last four items were all related to cellular energy metabolism, which corresponds to the increased energy metabolism that occurred after cancerous transformation, e.g., in order to maintain proliferative and metastatic capacity, tumor cells increased their energy-producing pathways (92), and which we conjecture mediate cytoplasmic and mitochondrial translational homeostasis in order to modulate the respiratory capacity of the mitochondria to provide the tumor cells to provide energy to adapt to different environmental stresses and growth demands (93). This implied that C0 subtype was strongly correlated with tumor tissue development and progression. Next, we utilized volcano plots to demonstrate the spectrum of DEGs among TCs subtypes (Figure 5B). The results showed that the main DEGs upregulated in C0 subtype were *TMSB10*, *S100A10*, *CLIC1*, *GSTP1* and *RPS21*. Among them, *TMSB10*, *S100A10*, *CLIC1* and *GSTP1* have been shown in previous studies to have an oncogenic mechanism in the colorectum as well as to contribute to the poor prognosis of colorectal cancer, and *TMSB10* could be used as a minimally invasive serum tumor marker for the detection of CRC, while *CLIC1* and *GSTP1* may be potential prognostic biomarkers for patients with CRC (94–97).

In addition, the bubble plot showed the results of GSEA (Figure 5C). The results showed that the C0 subtype was significantly expressed in the pathways of translation, peptide biosynthetic process, amide biosynthetic process and ribonucleoprotein complex biogenesis. All these pathways indicated that energy metabolism as well as ribose and protein synthesis were actively taking place in the C0 subtype, providing a material basis for the rapid proliferation of tumor cells. The word cloud map showed the pathways enriched in DEGs in the TCs subtypes (Figures 5D–J). Consistent with the above results, it indicated that the C0 subtype was undergoing enriched metabolic and oxidative reactions and substance synthesis. Finally, we performed a GSEA of DEGs in the C0 subtype according to GO-BP (Figure 5K). We observed that pathways associated with cytoplasmic translation, amide biosynthetic process, peptide biosynthetic process and ribonucleoprotein complex subunit organization, which were consistent with the previous enrichment items, were upregulated in the C0 subtype. While pathways associated with defense response to gram-positive bacterium, positive regulation of cyclase activity and lyase activity and regulation of cyclase activity were down-regulated in the C0 subtype. Where for positive regulation of lyase activity decreases, it may lead to increased levels of degradative enzymes, which in turn are able to degrade proteoglycan basement membrane components, promoting increased tumor cell detachment from the primary tumor and local invasion, resulting in a poor prognosis (98). Thus, the items enriched by various methods in the C0 subtype

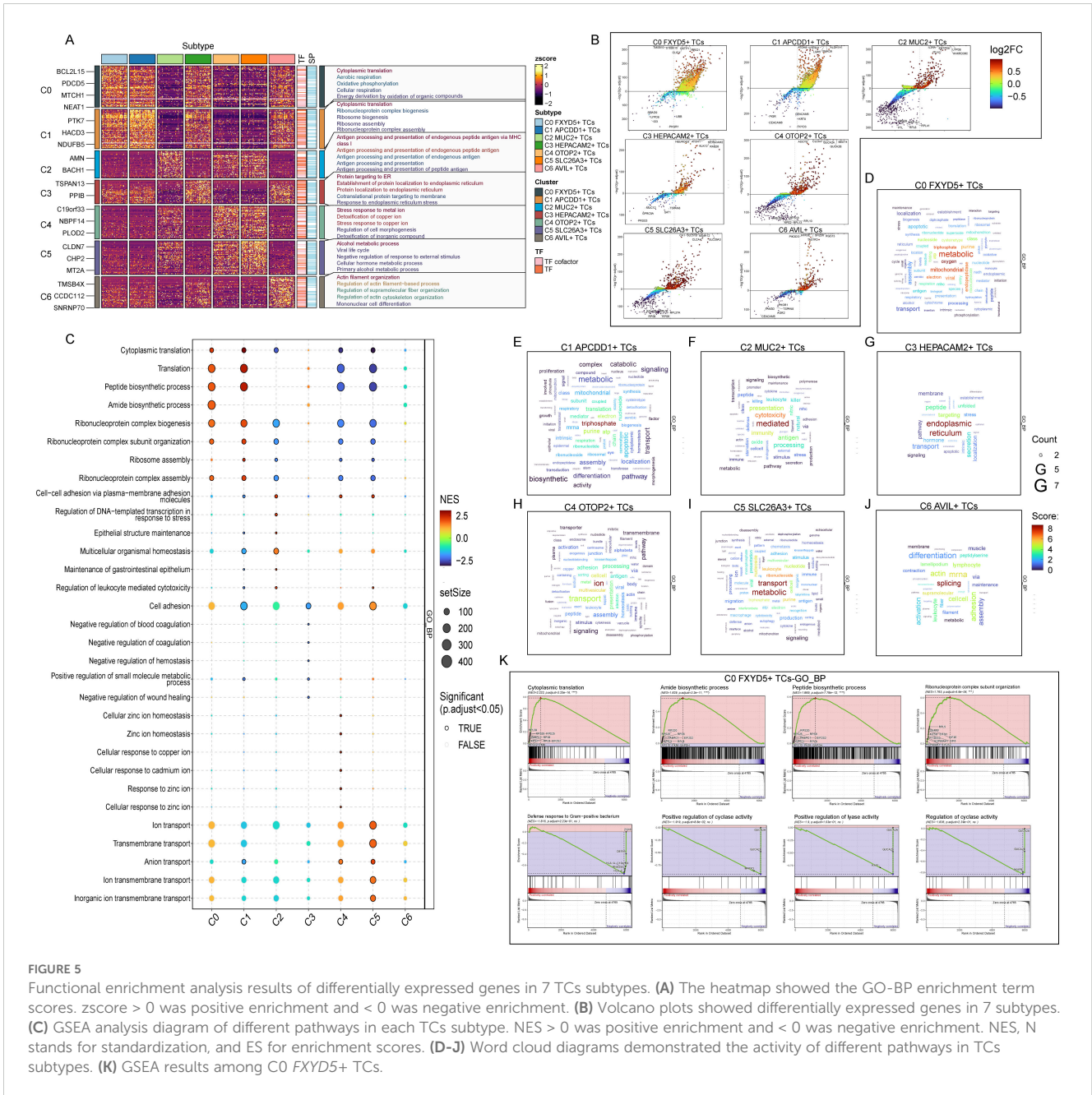


FIGURE 5

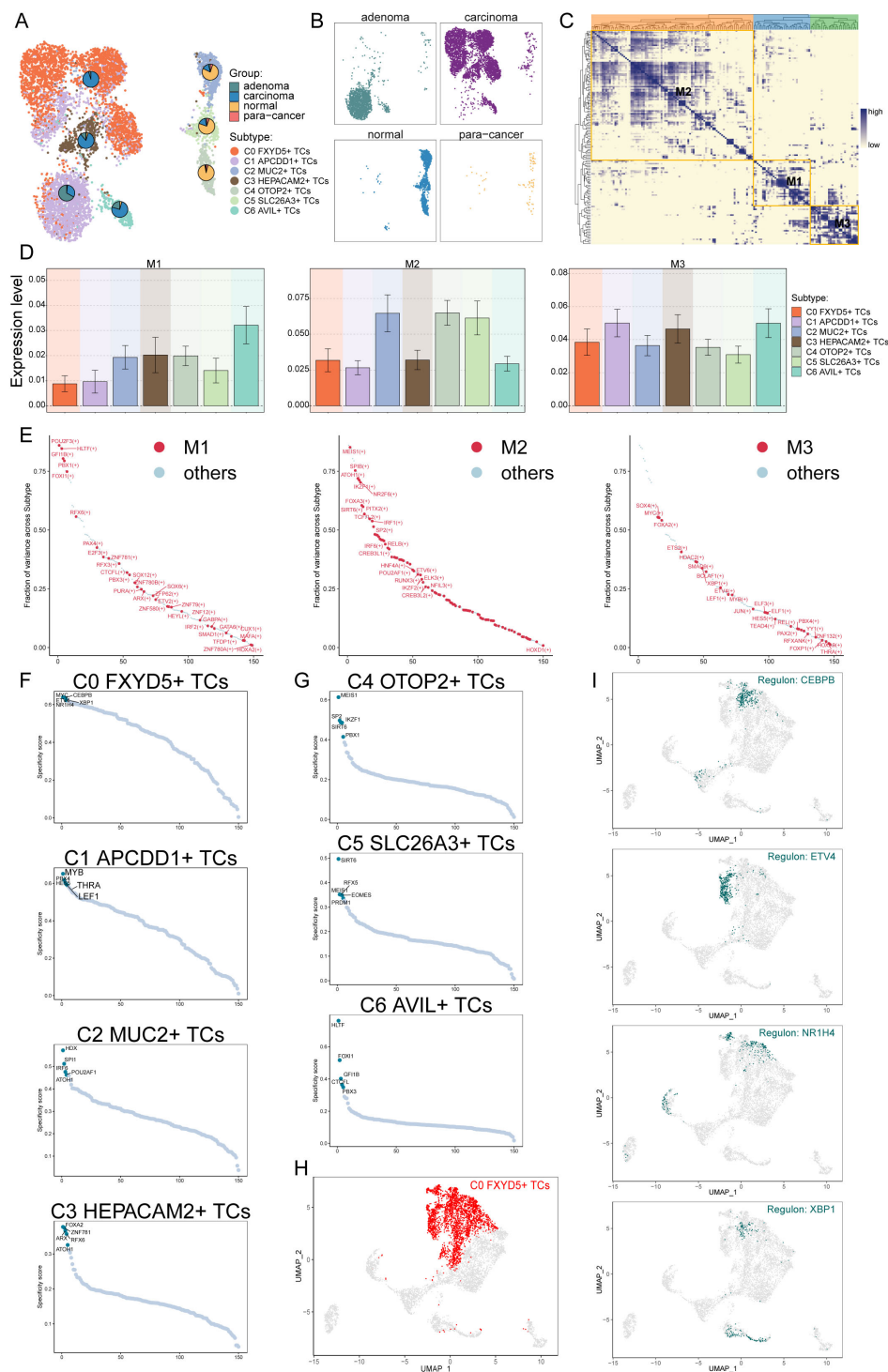
Functional enrichment analysis results of differentially expressed genes in 7 TCs subtypes. (A) The heatmap showed the GO-BP enrichment term scores. zscore > 0 was positive enrichment and < 0 was negative enrichment. (B) Volcano plots showed differentially expressed genes in 7 subtypes. (C) GSEA analysis diagram of different pathways in each TCs subtype. NES > 0 was positive enrichment and < 0 was negative enrichment. NES, N stands for standardization, and ES for enrichment scores. (D-J) Word cloud diagrams demonstrated the activity of different pathways in TCs subtypes. (K) GSEA results among C0 FXYD5+ TCs.

were suggestive that the C0 subtype might be transforming into tumors with a more aggressive and poorer prognosis, and that we could incorporate this feature into risk group of the tumor population.

### 3.6 TFs regulate the oncogenesis of C0 FXYD5+ TCs subtype

TFs act directly on the genome to regulate gene transcription by binding to specific nucleotide sequences upstream of genes, thereby affecting the biological functions of cells. Here, we analyzed the gene regulatory network of the C0 subtype using scenic analysis. First, we performed a dimensionality reduction clustering of TCs based on

the activities of regulatory factors and listed the results from different tissue sources (Figures 6A, B). Among them, the C0 subtype still mainly consisted of carcinoma tissues. Next, we observed the correlation between the TCs subtypes. Considering that different TFs can co-regulate the expression of certain genes, based on the connectivity specificity index (CSI), we categorized the TCs subtypes into three modules of regulatory factors: M1, M2 and M3 (Figure 6C). These three regulatory submodules showed TFs that may be cooperative with each other in gene regulation. As shown in the bar graphs, the M3 regulatory submodule may take a major regulatory role in the biological functions of the C0, C1, and C3 subtypes (Figure 6D). Meanwhile, the M3 regulatory submodule was also the dominant regulatory module for the C0 subtype (Figure 6E). By further analyzing the key regulators of different



**FIGURE 6**

Identification of tumor cells Gene Regulatory Network. **(A, B)** UMAP plot visualization of all TCs based on regulator activity. Colored according to cell subtype. Pie charts showed the proportion of different tissue sources in TCs subtypes. **(C)** Based on the CSI matrix, three regulatory modules of TCs subtypes were identified. **(D)** Bar graphs illustrated the expression levels of different regulator modules in subtypes of TCs. **(E)** In subtypes, the ranking of regulators in different regulation modules based on fraction of variance. **(F, G)** Rank for regulons in different TCs subtypes based on regulon specificity score (RSS). **(H)** C0 FXYD5+ TCs were highlighted in the UMAP plot (red). **(I)** Expression of key regulators *CEBPB*, *ETV4*, *NR1H4* and *XBP1* of C0 FXYD5+ TCs in all TCs.

TCs subtypes, it was found that three of the five major regulators of the C0 subtype, *ETV4*, *MYC* and *XBPI*, were present in the M3 regulatory submodule (Figures 6F, G). The expression of TFs in the C0 subtype was labeled in the UMAP plot (Figures 6H, I). It is worth noting that previous studies have indicated that *CEBPB* regulates the bile acid receptor *FXR* to accelerate colon cancer progression by modulating aerobic glycolysis (99), corresponding to the analysis of its enrichment pathway; whereas, the expression of *ETV4* was significantly correlated with the depth of infiltration, lymphovascular and venous infiltration, lymph node and distant metastasis, and the stage of lymph node metastasis of the pathologic tumors in terms of progression and recurrence, which suggests an unfavorable prognosis (100); while the expression of *XBPI* could also promote the proliferation of colon cancer cells (101). These were all factors of poor prognosis of CRC.

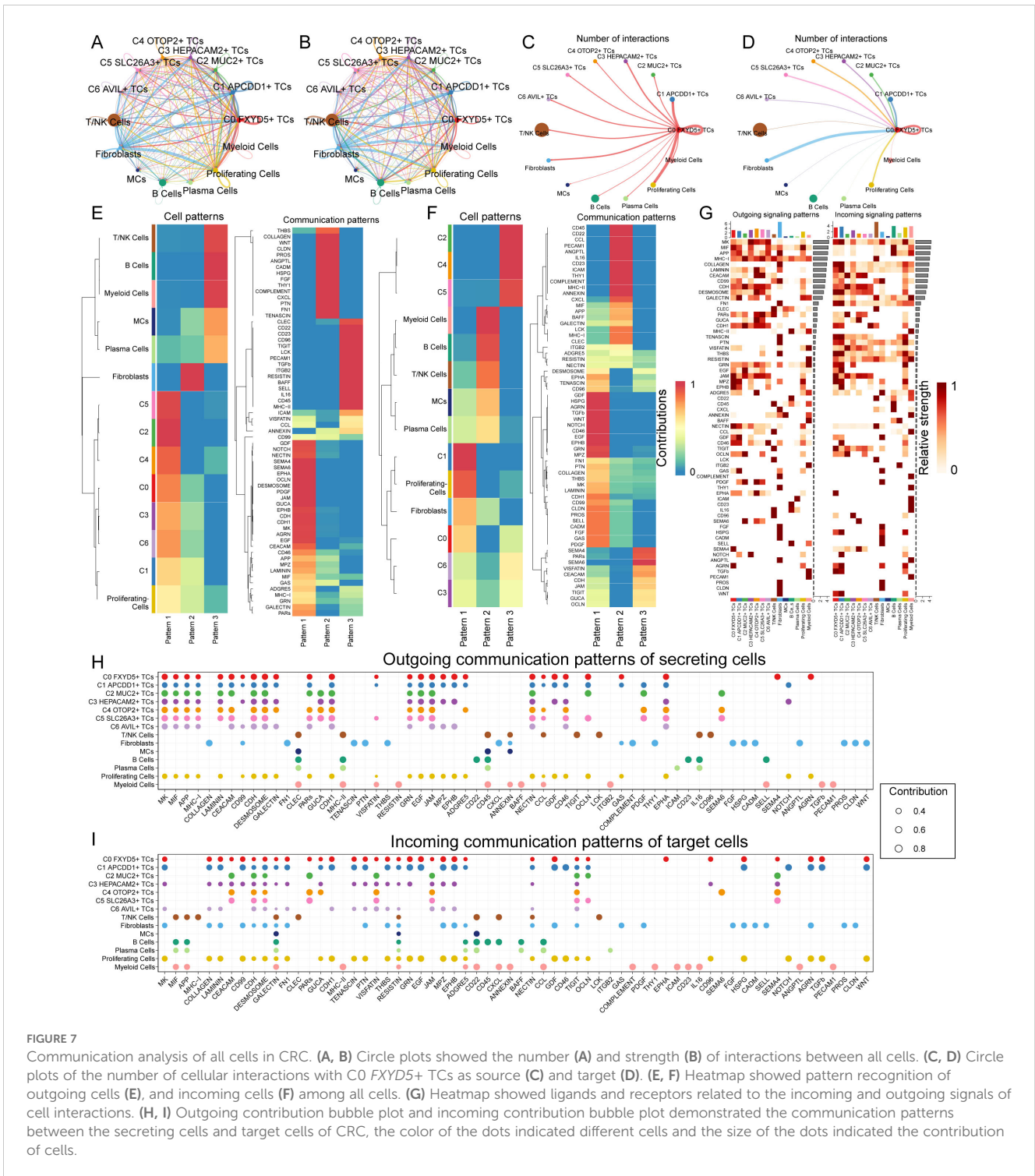
### 3.7 CellChat analyzes the communication of all cells

In order to specifically elucidate complex cellular responses, we intended to explore intercellular relationships and ligand-receptor communication networks for a better understanding of intercellular interactions. Through analyses via CellChat, we preliminarily established intercellular communication networks involving a variety of cells, such as proliferating cells, fibroblasts, T and NK cells, and different subtypes of TCs (Figures 7A, B). We also established a communication network with C0 *FXVD5+* TCs as the source cells and target cells (Figures 7C, D). After building the intercellular communication network using CellChat analysis, we accounted for the number of interactions (expressed in terms of the thickness of the connecting lines between the two types of cells) and the strength of the interactions (expressed in terms of the weight of the connecting lines, with thicker lines indicating stronger interactions). This method helped to quantify the complexity and strength of the communication pathways between different cell types in the network. From the intricate communication network diagram, we could see that C0 *FXVD5+* TCs sent positive communication signals to almost all cells, while fibroblasts and proliferating cells are the cell types that give it the most responses. Next, we used the gene expression pattern analysis method provided by CellChat to examine how cells and signaling pathways interact with each other (Figures 7E, F). First, we assessed the relationship between the inferred potential communication patterns and the population of cells secreting signaling molecules to decipher the outgoing communication patterns. Through our analysis, we identified three distinct signaling patterns: pattern 1 (tumor cells and proliferating cells), pattern 2 (fibroblasts), and pattern 3 (myeloid cells, B cells, plasma cells, MCs, and T and NK cells). Whereas for the incoming communication mode, the distinct signaling patterns: pattern 1 (C1 *APCDD1+* TCs, proliferating cells, fibroblasts, C0 *FXVD5+* TCs, C6 *AVIL+* TCs and C3 *HEPACAM2+* TCs), pattern 2 (myeloid cells, B cells, T cells and NK cells, MCs and plasma cells), and pattern 3 (C2 *MUC2+* TCs, C4 *OTOP2+* TCs and C5 *SLC26A3+* TCs). Next, the heatmaps showed the signaling molecules of various types of cells under outgoing signaling patterns and under incoming signaling

patterns, respectively (Figure 7G). Finally, to identify the major afferent and efferent signals associated with the seven TCs subtypes, we quantified the ligand-receptor network using CellChat (Figures 7H, I). With this approach, we could predict the major outgoing signals from secretory cells (signal senders) that release various cytokines or ligands. In addition, we evaluated on which cell types act as target cells (signal receivers) and how ligand-receptor-mediated communication between different cell types contributes to the progression of CRC. This analysis helps to illustrate how receptors on these cells are targeted by ligands released by the same or other cell types. In summary, we visualized signaling molecules between different cell types using multiple communication analyses that provide support for describing cellular communication.

### 3.8 Construction of risk score profiles and immune infiltration analysis

Since we focused on the C0 *FXVD5+* TCs subtype, we identified 9 genes that could be used as prognostic features based on the top 100 marker genes of the *FXVD5+* TCs subtype by univariate Cox regression analysis (Figure 8A), of which *ETS2* and *ATO1H1* were protective factors (HR < 1), and the other genes were risk factors. To address the problem of multicollinearity among these genes, we further screened them using LASSO regression analysis and multivariate regression analysis and finally identified 8 genes associated with prognosis (Figures 8B–D). The one that caught our attention was *DLX2*, which was the most dangerous prognostic factor among them with the largest HR. To further investigate the impact of *FXVD5* highly expressed TCs on CRC patients, we categorized patients in the TCGA cohort into high and low FTRS (*FXVD5+* TCs risk score) groups based on the identified 8 prognostic genes. Subsequently, we created a Nomogram survival prediction model for OS in CRC patients using autonomous prognostic factors to predict the prognosis of CRC patients (Figure 8E). The model was validated and performed well in predicting the OS C index (Figure 8F). Subsequent prognostic gene versus risk correlation plots, as well as ridge and box line plots, illustrated the clear correlation of prognostic genes with risk and variability across risk groups (Figures 8G–H). It could be seen that the difference between the two groups for our gene of interest, *DLX2*, was statistically significant and was expressed higher in the high FTRS group than in the low FTRS group. Next, to assess the accuracy of risk scores in predicting 1-, 3- and 5-year survival in patients with CRC, we performed ROC curve analyses on the three cohorts. The results showed high predictive accuracy with AUC (1-year) = 0.641, AUC (3-years) = 0.658, and AUC (5-years) = 0.691 (Figure 8I). Moreover, calibration curves were plotted to show the consistency between predicted and observed values for 1-, 3- and 5-year OS in both the training and validation cohorts (Figure 8J), which showed good agreement. The PCA plot also provided further evidence of the discrete and differential character of the high and low FTRS groups (Figure 8K). With definitive evidence, we next analyzed the calculation of coefficient values for these prognostic genes (Figure 8L). Using Kaplan-Meier survival curves, we first confirmed the poor prognosis corresponding to high expression of *DLX2*, and then we further concluded that the survival outcome was worse in the high FTRS group (Figures 8M, N). *DLX2*



was demonstrated to be associated with EMT in osteosarcoma in previous studies and its increased expression was associated with advanced gastric adenocarcinoma, but no article defining the relationship between *DLX2* and poor prognosis in colorectal cancer could be found yet (102, 103).

The CIBERSORT analysis revealed the immune infiltration in the tumor samples. With the heatmap, we could observe the estimated proportion of immune cells in the high and low FTRS groups (Figure 8O). From the box line plot, we could easily see that there

was a statistically significant difference in immune cell expression between the different groups (Figure 8P). Also, we were able to find that the relatively high expression of immune cells in the high FTRS group was positively correlated with the risk score (Figures 8Q, R). For example, T cells regulatory (Treg) relatively high expression in the high FTRS group and positive correlation with risk score represented a poor prognosis. Tregs have been documented in malignant tumors, where they promote tumor progression by suppressing effective anti-tumor immunity and are associated with

a poor prognosis in various types of human cancers (104). In addition, we used ESTIMATE to calculate the stromal score, immune score, and ESTIMATE score for the high FTRS group and low FTRS group, and found that all the scores of the high FTRS group were significantly higher than those of the low FTRS group, which indicated that the level of immune cell infiltration in the tumor samples of the high FTRS group was higher (Figure 8S). In addition, tumor purity was higher in the low FTRS group than in the high

FTRS group (Figure 8T). Since low tumor purity was associated with extensive infiltration of stromal and immune cells, it was independently associated with shorter survival time and faster recurrence, and significantly correlated with mesenchymal, invasive, and metastatic phenotypes suggesting a poor prognosis that corresponds to the prognostic curve. Finally, TIDE is a tool used to comprehensively assess tumor immune escape mechanisms and predict response to immune checkpoint inhibitor (ICI) therapy,

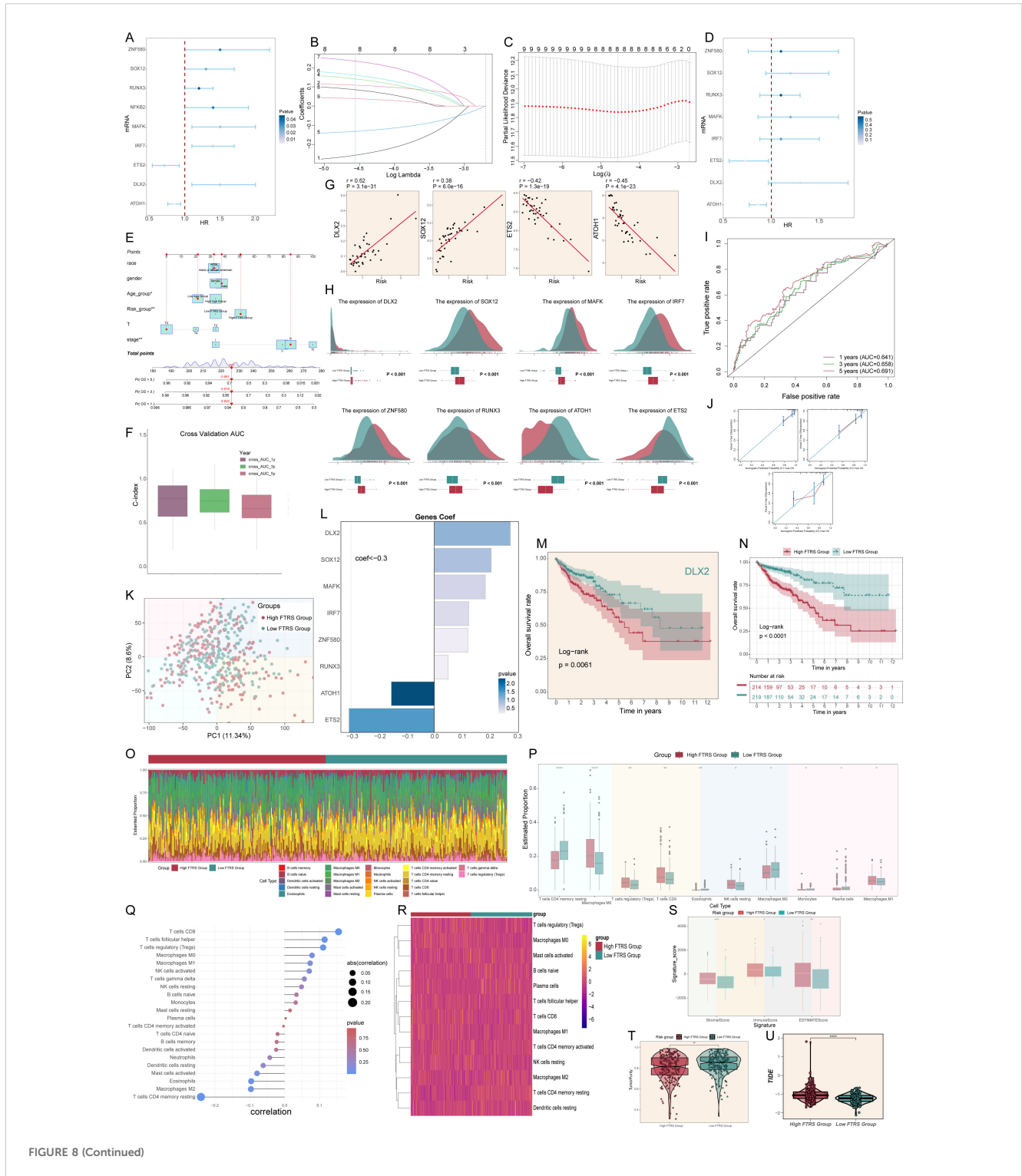


FIGURE 8 (Continued)



FIGURE 8 (Continued)

Constructed risk score profiles and immune infiltration and survival prognostic analyses for CRC. **(A)** Forest plot from univariate Cox regression analysis showcased statistically significant genes ( $P < 0.05$ ) with  $HR < 1$  indicating protective factors and  $HR > 1$  indicating risk factors. **(B, C)** Selection of eight prognostic-related genes (non-zero regression coefficients) was made via LASSO regression analysis. LASSO coefficient curve determined by optimal lambda **(B)**, with optimal parameter (lambda) determined through cross-validation **(C)**. **(D)** Forest plot of eight prognosis-related genes. **(E)** Column line graphs were used to predict patients' prognosis at 1 year, 3 years and 5 years. For categorical variables, the importance of each variable was ranked according to the standard deviation of the column-line graph scale. To use a column line plot, individual patient-specific points are located on each variable axis. Red lines and dots are plotted upward to identify the point at which each variable is received; the sum of these points (280) is located on the Total Points axis, and a line is plotted downward to the Survival axis to identify the probability of overall survival. **(F)** Box line plots depict the C-index of the AUC values of the risk scores for predicting 1-, 3-, and 5-year survival. **(G)** Four genes that showed a significant correlation with the risk scores (*DLX2*, *SOX12*, *ETS2* and *ATOH1*). **(H)** The ridge and box line plots showed that gene expression of these four genes (*DLX2*, *SOX12*, *MAFK*, *IRF7*) was higher in the high-risk group compared to the low-risk group, while gene expression of these four genes (*ZNF580*, *RUNX3*, *ATOH1*, *ETS2*) was lower in the high-risk group. **(I)** ROC curves depict the sensitivity and specificity of the risk scores for predicting 1-, 3-, and 5-year survival. **(J)** Calibration curves for column charts predicting 1-, 3-, and 5-year overall survival. The OS predicted by the line plot model is plotted on the x-axis and the actual OS is plotted on the y-axis. **(K)** The PCA plot demonstrated the difference in the distribution of prognosis-related genes in the high FTRS and low FTRS groups. **(L)** Bar chart showed the coefficient (Coef) values of genes utilized for model construction. **(M)** OS curve of *DLX2*, a highly expressed gene screened by LASSO. **(N)** OS curves for different scoring subgroups in a cohort (high FTRS group and low FTRS group). **(O)** Proportion of each infiltrating immune cell type in the high- and low-FTRS groups were shown using CIBERSOFT. **(P)** Statistically different infiltrating immune cell type in the high- and low-FTRS groups were demonstrated using CIBERSOFT. \* stands for  $P \leq 0.05$ , \*\* stands for  $P \leq 0.01$  and \*\*\*\* represents  $P \leq 0.0001$ . **(Q)** Lollipop chart of immune cell versus risk score. **(R)** Heatmap demonstrated the difference in expression of different immune cells in the two risk groups. **(S)** Stromal score, immune score, and estimate score were calculated for the high- and low-FTRS groups, respectively, using ESTIMATE. \* stands for  $P \leq 0.05$ , \*\* stands for  $P \leq 0.01$  and \*\*\* represents  $P \leq 0.001$ . **(T)** TumorPurity was calculated using ESTIMATE for the high- and low-FTRS groups, respectively. \*\* stands for  $P \leq 0.01$ . **(U)** Violin plot demonstrating the difference in Tumor Immune Dysfunction and Exclusion (TIDE) scores in the two risk groups. \*\*\*\* represents  $P \leq 0.0001$ , a statistically significant difference.

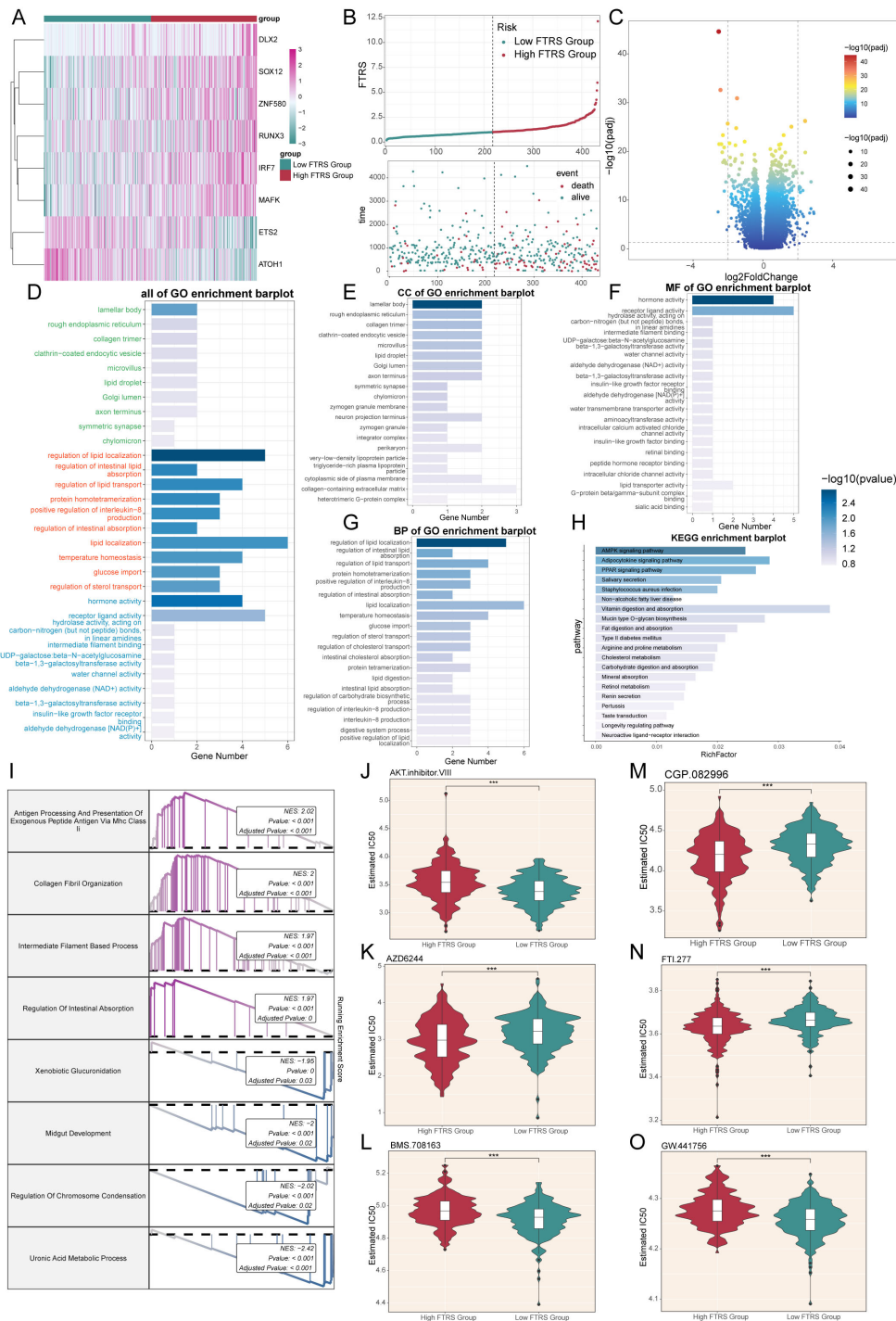
with higher TIDE scores correlating with poorer immune checkpoint inhibition therapy (105), and violin plot of high FTRS group scores being significantly greater than lower groups (Figure 8U).

### 3.9 Identification of differentially expressed genes and their enrichment analysis results and drug sensitivity analysis

Immediately following the previous section, we proceeded to show the differential expression patterns of the genes used to construct the model using heatmap (Figure 9A). In addition, the curve and scatter plots showed the differences in risk scores and survival results between the two risk groups, suggesting that the high FTRS group was associated with a poorer prognosis, corresponding to the previous section (Figure 9B).

Next, volcano plot showed up- and down-regulated DEGs (Figure 9C). The filtering criteria for fold change were based on a  $\log_2$  fold change threshold of 2, which meant that only genes with a fold change greater than or equal to 4 (or less than or equal to -4) were considered as differentially expressed. In addition to this fold change criterion, a p-value cutoff of 0.05 was applied to further refine the selection of significant DEGs. We then used these DEGs to gain insight into the biological processes involved by employing various enrichment methods. All GO enrichment results and their terms were shown, with different colors representing different subclasses (Figure 9D). GO-CC enrichment analysis showed that DEGs were predominantly enriched in lamellar body and rough endoplasmic reticulum, the latter of which suggested that there may be abundant protein synthesis in high-risk groups (Figure 9E). In GO-MF enrichment analysis, we observed enrichment in hormone and receptor ligand activity, suggesting that there may be an abundance of communication (Figure 9F). As for GO-BP enrichment analysis, the items mainly including regulation of lipid

localization, intestinal lipid absorption and lipid transport (Figure 9G). All of the above enriched items are associated with lipid metabolism, and the former has been richly studied for its association with cancer. Dysregulated lipid metabolism is one of the most prominent metabolic alterations in cancer, and cancer cells utilize lipid metabolism for energy, biofilm components, and signaling molecules required for proliferation, survival, invasion, metastasis, and response to tumor microenvironmental influences and cancer therapy (106, 107). For another enrichment method analysis KEGG, the pathways were mainly AMPK, adipocytokine and PPAR signaling pathway (Figure 9H). Among them, AMPK plays a crucial role in maintaining energy homeostasis and metabolism, and its aberrant activation can lead to pro-carcinogenic effects, which is a double-edged sword (108); recent studies also show that AMPK signaling can also reversibly regulate highly active MAPK signaling in cancer cells by phosphorylating its key component, the RAF/KSR family of kinases, which not only affects carcinogenesis, but also the outcome of targeted cancer therapy against MAPK signaling. And there was an experiment showing that adipocytokines were involved in the carcinogenesis process (109), and an adipocytokine, LCN-2, has been reported to have the ability to destroy the extracellular matrix, which may lead to cancer progression and metastatic spread (110). At the same time, previous article had shown that PPAR-d upregulation increases susceptibility to colon tumorigenesis, which may impact the development of strategies to molecularly target PPAR-d in cancer and non-cancer diseases (111). Through the GSEA results we found that up-regulated DEGs were mainly enriched in the processes of antigen processing and presentation of exogenous peptide antigen, collagen fibril organization, intermediate filament-based process and regulation of intestinal absorption, while down-regulated DEGs were enriched in xenobiotic glucuronidation, midgut development, regulation of chromosome condensation and uronic acid metabolic process (Figure 9I).



**FIGURE 9**

Enrichment analysis and drug sensitivity analysis in different risk groups. **(A)** Heatmap displayed differential expression of model genes, with color scale based on normalized data. **(B)** Curve chart illustrated the risk scores of high and low FTRS groups, and scatter plot depicted survival/death events over time for both groups. **(C)** Volcano plot showed significantly differentially expressed genes. Each spot represents a gene. **(D-H)** Bar charts separately presented the enrichment analysis results of differential genes in all of GO, GO-CC, GO-MF, GO-BP and KEGG pathways for high and low FTRS groups. **(I)** Results of GSEA of differentially expressed genes. **(J-O)** Violin plots demonstrated the sensitivity of tumor cells in the high and low FTRS risk groups to different drugs. \*\*\* represents  $P \leq 0.001$ .

Finally, in order to study the sensitivity of the high and low FTRS groups to different drugs, we predicted the drug sensitivity of each patient based on the drug sensitivity data in the GDSC database using the “pRRophetic” R software package. The results

showed that the drugs AZD6244, CGP.082996, and FTI.277 had lower IC50 values and higher drug sensitivities in tumor cells of the high FTRS group, whereas the other drugs had lower IC50 values in tumor cells of the low FTRS group (Figures 9J–O).

### 3.10 *In vitro* experimental verifications of *DLX2*

Through the previous construction of CRC prognostic model, we obtained the prognostically strongly related gene *DLX2*. To further investigate its role in CRC, we performed *in vitro* experiments using NCI-H716 and SW837 cell lines. First, we knocked down *DLX2* and measured mRNA and protein expression levels before and after knockdown. Compared with the control group, we observed a significant decrease in both mRNA and protein expression levels in both cell lines (Figure 10A). Subsequently, CCK-8 assay showed that tumor cell viability was significantly decreased after *DLX2* knockdown (Figure 10B). Colony formation assay confirmed that *DLX2* gene knockdown inhibited tumor cell aggregation (Figure 10C). In addition, after *DLX2* knockdown, the migratory and invasive abilities of tumor cells were assessed using the scratch and the transwell method, and the results showed that the levels of migration and invasion were significantly reduced (Figures 10D, E). Finally, EDU staining method also confirmed these results (Figure 10F). Together, these results indicated that knockdown of *DLX2* inhibited tumor cell activity, proliferation, migration, and invasion, thereby suppressing tumor growth.

## 4 Discussion

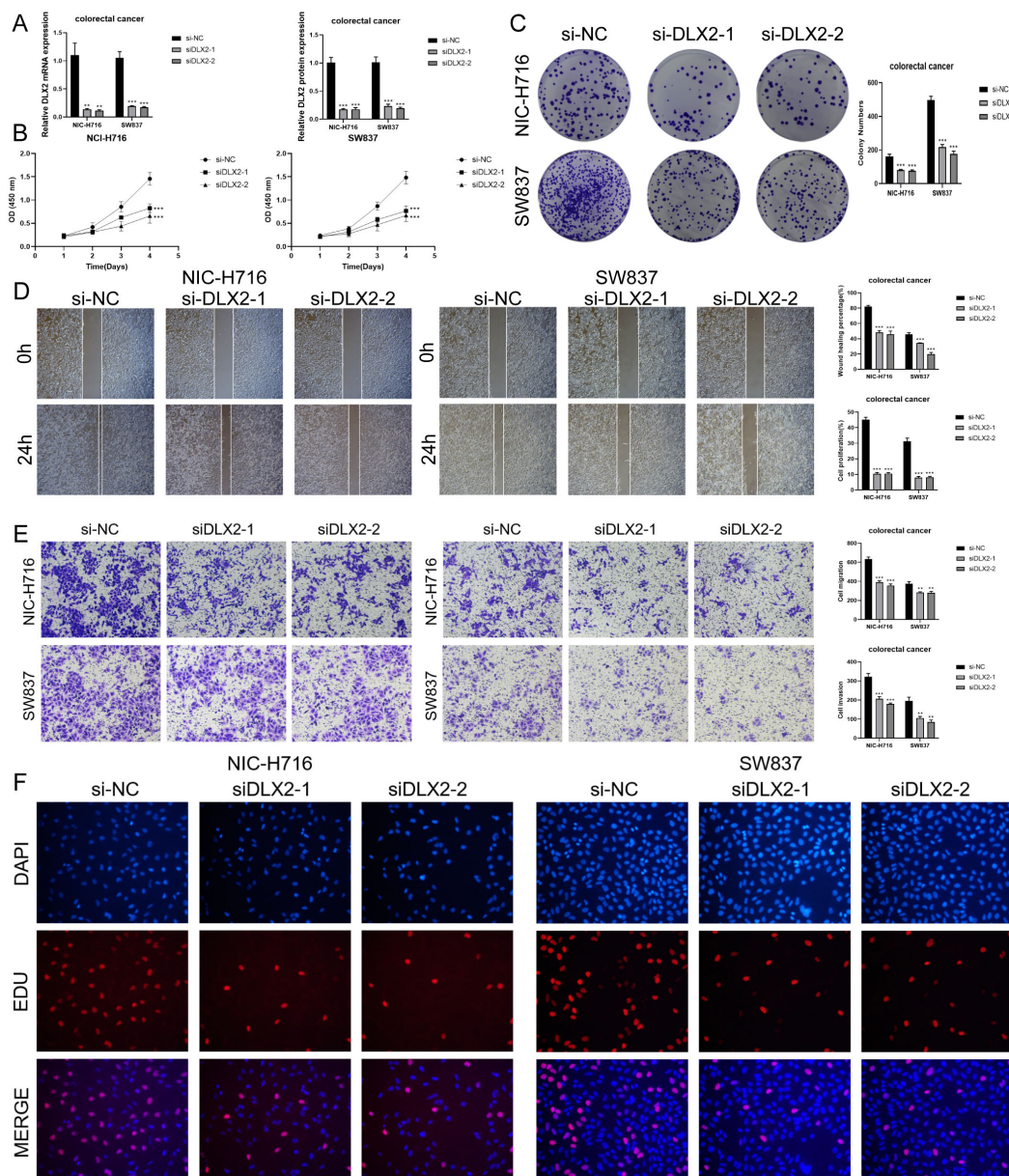
CRC is a highly heterogeneous disease with complex interactions between tumor cells and the surrounding microenvironment, making it difficult to fully understand the mechanisms driving progression and drug resistance (112–114). scRNA-seq has emerged as a powerful tool for revealing cellular diversity and gene expression profiles within tumors, allowing a more nuanced understanding of the cellular states and signaling pathways that contribute to CRC pathogenesis (115). Recent studies have shown that different tumor cell subtypes and the tumor microenvironment play a role in influencing CRC progression, immune evasion, and treatment response (116, 117). By utilizing scRNA-seq data, we can identify key cellular subtypes and regulatory networks that are critical for tumor growth, metastasis, and treatment response. This study builds on existing research by analyzing the cellular composition of CRC, identifying novel tumor cell subtypes, and mapping developmental trajectories and constructing prognostic models, ultimately providing new insights into the molecular and cellular dynamics of CRC.

First, we identified 8 distinct cell types, highlighting the complexity of the TME in CRC. Notably, as CRC progressed from normal to adenoma and carcinoma, the abundance of epithelial cells (EPCs), which have been considered a major driver of tumorigenesis due to their origin in the intestinal epithelium, increased markedly, and were transformed into tumor cells via EMT (118), an observation that is consistent with previous findings that epithelial expansion and transformation is associated with an increase in colorectal tumors with increased malignancy (119). The immunosuppressive nature of TME in CRC is highlighted by the progressive increase in proliferating cells and decrease in immune cells such as T and NK

cells and B cells as the cancer progresses. The decrease in immunoreactivity at the stage of carcinogenesis suggests the presence of an underlying immune evasion mechanism, a phenomenon well documented in many types of solid tumors (120–122). This immunosuppression may also explain why immunotherapy has had limited success in patients with CRC, especially those without high microsatellite instability (MSI) (123). Subsequently, the identification of 7 TCs subtypes provided important information for understanding the cellular basis of CRC progression. The enrichment of C0 *FXYD5*+ TCs in cancer tissues and their association with poor prognosis confirmed previous studies that *FXYD5* expression is associated with increased epithelial-mesenchymal transition (EMT) and metastatic potential (89, 90). The different tissue distribution patterns of these subtypes reinforce the concept of CRC tumor heterogeneity. For example, C0 *FXYD5*+ TCs are abundantly present in cancerous tissues, whereas C2 *MUC2*+ TCs and C4 *OTOP2*+ TCs are predominantly present in normal and para-cancer tissues. This spatial and molecular heterogeneity highlights the need for a precision medicine approach to treating CRC, as different subtypes may respond differently to therapeutic interventions.

Studies of stemness genes in various tumor subtypes have revealed that *MYC* is significantly expressed in C0 *FXYD5*+ TCs, suggesting that this subtype has a high proliferative potential and is critical for tumor growth. Stemness has long been considered one of the hallmarks of cancer, and higher levels of stem-like features are associated with tumor aggressiveness and drug resistance (124–126). Pseudotime analyses using Monocle and Slingshot suggest that TCs follow distinct developmental trajectories. Lineage 1 appears to represent the key lineage associated with CRC tumorigenesis. Our findings also suggest that targeting key regulators on this trajectory, such as those driving the transition from C1 *APCDD1*+ to C0 *FXYD5*+ TCs, may provide novel strategies to prevent CRC progression. GSEA showed that C0 *FXYD5*+ TCs subtypes were highly involved in metabolic processes such as oxidative phosphorylation and aerobic respiration. This is consistent with the established concept of metabolic reprogramming in cancer cells, which enhances energy production through processes such as glycolysis and oxidative phosphorylation to support rapid cell proliferation (127–129). The enrichment of ribosome biogenesis and protein synthesis pathways in this subtype further emphasizes its role in maintaining the high metabolic demands of cancer cells (130). The up-regulation of genes such as *TMSB10*, *S100A10*, and *GSTP1* in the C0 subtype suggests that they are closely associated with the progression and poor prognosis of CRC. These genes are associated with tumor metastasis and chemotherapy resistance, and thus are potential targets for therapeutic intervention (94, 96, 97). SCENIC analysis identified several key TFs, including *ETV4*, *MYC*, and *XBP1*, which are major regulators of C0 *FXYD5*+ TCs subtypes. In particular, the role of *ETV4* in promoting CRC invasion and metastasis through activation of the EMT program highlights its potential as a therapeutic target (100).

In the end, our analysis highlights several key genes and pathways associated with poor CRC prognosis. A risk-scoring model based on the expression of eight genes, including *DLX2*



**FIGURE 10**

*In vitro* experiments confirmed the effects of *DLX2* knockdown. (A) *DLX2* knockdown significantly reduced the expression levels of mRNA and protein in both experimental groups. \*\* stands for  $P \leq 0.01$  and \*\*\* represents  $P \leq 0.001$ . (B) CCK-8 assay showed that tumor cell viability was significantly decreased after *DLX2* knockdown compared with the control group. \*\*\* represents  $P \leq 0.001$ . (C) Colony formation assays revealed a significant reduction in colony numbers after *DLX2* knockdown. \*\*\* represents  $P \leq 0.001$ . (D) The scratch assay indicated that *DLX2* knockdown inhibited tumor cell migration. Bar graphs showed a significant reduction in both tumor cell migration and proliferation capabilities post-*DLX2* knockdown. \*\*\* represents  $P \leq 0.001$ . (E) Transwell experiments indicated that *DLX2* knockdown inhibited the migration and invasion capabilities of tumor cells in the NIC-H716 and SW837 cell lines. \*\* stands for  $P \leq 0.01$  and \*\*\* represents  $P \leq 0.001$ . (F) EDU staining assay confirmed that *DLX2* knockdown inhibited the proliferation of tumor cells by DAPI staining for comparison.

and *SOX12*, provides a reliable framework for predicting the overall survival of CRC patients. Validation of this model in the TCGA cohort emphasizes its potential clinical utility in stratifying patients according to risk and adjusting treatment strategies accordingly. In addition, immune infiltration analysis showed that higher infiltration of regulatory T cells (Tregs) was associated with a poorer prognosis, which is consistent with previous findings that Tregs promote immune evasion in CRC by suppressing anti-tumor immune responses (104, 131, 132). Tumors with high expression of

*FXYD5*+ TCs have higher immune cell infiltration and lower tumor purity, an observation that suggests that these tumors may be more resistant to immune checkpoint blockade therapy. This finding has important implications for the design of immunotherapies for the treatment of CRC, especially for high-risk patients with *FXYD5*+ TCs subtypes.

In conclusion, the present study describes in detail the cellular and molecular characteristics of CRC using scRNA-seq technology. The identification of different tumor subtypes, their developmental

trajectories and key regulators provides new insights into the progression of CRC. The findings highlight potential prognostic markers and therapeutic targets, including *FXYD5*+ TCs, stem cell-associated genes, and key TFs, which could guide future interventions in CRC, and *in vitro* experiments were performed to validate the strongly prognostic-related gene *DLX2*. Given the complexity and heterogeneity of CRC, our study highlights the importance of a personalized approach in CRC treatment with the ultimate goal of improving patient prognosis.

However, our experiments still have shortcomings. First, although the study identified several key tumor cell subtypes and related markers, these findings were based only on bioinformatics analysis and *in vitro* experiments. Experimental validation, such as *in vivo* functional testing, is needed to confirm the role of these markers and subtypes in CRC progression and their potential as therapeutic targets. Second, although this study described immune cell types such as T and NK cells, B cells, and Tregs, there is limited understanding of the mechanisms of how these immune cells interact with tumor cells. A more detailed analysis of immune signaling pathways, immune cell function, and the potential for immunotherapeutic interventions could enhance the impact of this study, especially given the increasingly important role that immunotherapy is playing in the treatment of CRC. In addition, although the study identified several prognostic markers and developed a risk score model, it lacked sufficient correlation with clinical outcomes such as patient survival, response to treatment, or relapse rate. The inclusion of longitudinal clinical data would help validate the prognostic value of these markers and support their potential application in clinical practice. In future studies, we will address these limitations, improve the robustness of the findings, and expand their application in the diagnosis and treatment of CRC.

## Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

## Author contributions

QL: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Methodology, Software, Visualization. ZW: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Methodology, Software, Visualization. JW: Writing – review & editing, Data curation, Software, Validation, Visualization. MX: Writing – review & editing, Methodology, Software, Validation. XZ: Conceptualization, Data curation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. PS: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Methodology, Software, Visualization. YY: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Methodology, Software, Visualization.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was qualified by Shanghai Municipal Science and Technology Commission (20ZR1451700), and Shanghai Collaborative Innovation Center for Translational Medicine (TM201731). Research Fund for Key Subject Construction Project of Shanghai Municipal Health Commission (Gastrointestinal Surgery, No. ZK2019A15), and the Research Fund of Shanghai Tongren Hospital, Shanghai Jiaotong University School of Medicine.

## Acknowledgments

Our Graphical Abstract was drawn using Figdraw. Image ID: ARSTOfee30. We would like to thank Figdraw for its contribution to this article.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2024.1509658/full#supplementary-material>

### SUPPLEMENTARY FIGURE 1

The heatmap demonstrated the relative expression intensity of each cell in each chromosomal region, the upper heatmap represents the demonstrated results of the reference cells (T and NK cells), and the lower heatmap represents the demonstrated results of the target cells (EPCs). Each row of the graph represents one cell and each column represents one gene. Red colour indicates CNV amplification, blue colour indicates CNV deletion, darker colour indicates more obvious CNV variation.

## References

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin.* (2015) 65:5–29. doi: 10.3322/caac.21254
- Lyu N, Pedersen B, Shklovskaya E, Rizos H, Molloy MP, Wang Y. SERS characterization of colorectal cancer cell surface markers upon anti-EGFR treatment. *Exploration.* (2022) 2:20210176. doi: 10.1002/EXP.20210176
- Visvader JE. Cells of origin in cancer. *Nature.* (2011) 469:314–22. doi: 10.1038/nature09781
- Vaiopoulos AG, Kostakis ID, Koutsilieris M, Papavassiliou AG. Colorectal cancer stem cells. *Stem Cells.* (2012) 30:363–71. doi: 10.1002/stem.1031
- Fessler E, Medema JP. Colorectal cancer subtypes: developmental origin and microenvironmental regulation. *Trends Cancer.* (2016) 2:505–18. doi: 10.1016/j.trecan.2016.07.008
- Ahmad R, Singh JK, Wunnavu A, Al-Obeed O, Abdulla M, Srivastava SK. Emerging trends in colorectal cancer: Dysregulated signaling pathways (Review). *Int J Mol Med.* (2021) 47(3):14. doi: 10.3892/ijmm.2021.4847
- Marmol I, Sanchez-de-Diego C, Pradilla DA, Cerrada E, Rodriguez YM. Colorectal carcinoma: A general overview and future perspectives in colorectal cancer. *Int J Mol Sci.* (2017) 18(1):197. doi: 10.3390/ijms18010197
- Requena DO, Garcia-Buitrago M. Molecular insights into colorectal carcinoma. *Arch Med Res.* (2020) 51:839–44. doi: 10.1016/j.arcmed.2020.09.014
- Huang J, Liu Y, Lin J, Su J, Redshaw C, Feng X, et al. Novel pyrene-based aggregation-induced emission luminogen (AIEgen) composite phase change fibers with satisfactory fluorescence anti-counterfeiting, temperature sensing, and high-temperature warning functions for solar-thermal energy storage. *Adv Compos Hybrid Mater.* (2023) 6:126. doi: 10.1007/s42114-023-00706-4
- Hou K, Gu H, Yang Y, Lam SS, Li H, Sonne C, et al. Recent progress in advanced covalent organic framework composites for environmental remediation. *Adv Compos Hybrid Mater.* (2023) 6:199. doi: 10.1007/s42114-023-00776-4
- Vernia F, Longo S, Stefanelli G, Viscido A, Latella G. Dietary factors modulating colorectal carcinogenesis. *Nutrients.* (2021) 13(1):143. doi: 10.3390/nu13010143
- Montalban-Arques A, Scharl M. Intestinal microbiota and colorectal carcinoma: Implications for pathogenesis, diagnosis, and therapy. *Ebiomedicine.* (2019) 48:648–55. doi: 10.1016/j.ebiom.2019.09.050
- Gu Y, Jiang L, Shui M, Luo H, Zhou X, Zhang S, et al. Revealing the association between East Asian oral microbiome and colorectal cancer through Mendelian randomization and multi-omics analysis. *Front Cell Infect Microbiol.* (2024) 14:1452392. doi: 10.3389/fcimb.2024.1452392
- Cramer L, Hildebrandt B, Kung T, Wichmann K, Springer J, Doehner W, et al. Cardiovascular function and predictors of exercise capacity in patients with colorectal cancer. *J Am Coll Cardiol.* (2014) 64:1310–19. doi: 10.1016/j.jacc.2014.07.948
- Hao Z, Zhang Y. Dapagliflozin in heart failure with reduced ejection fraction: A real-world study. *Cardiovasc Innov Appl.* (2022) 6(4):219–23. doi: 10.15212/CVIA.2022.0005
- Zhou J, Lin H, Lv T, Hao J, Zhang H, Sun S, et al. Inappropriate activation of TLR4/NF- $\kappa$ B is a cause of heart failure. *Cardiovasc Innov Appl.* (2022) 7(1). doi: 10.15212/CVIA.2022.0020
- Ma S, He L, Zuo Q, Zhang G, Guo Y. Canagliflozin regulates ferroptosis, potentially via activating AMPK/PGC-1 $\alpha$ /nrf2 signaling in HFpEF rats. *Cardiovasc Innov Appl.* (2023) 7(1). doi: 10.15212/CVIA.2022.0024
- Rohani S, Suhada N, Mohammad Z, Azhar S. Traditional herbal medicine as adjunctive therapy for colorectal cancer: a scoping review. *Tradit Med Res.* (2022) 7:15. doi: 10.53388/TMR20220127260
- Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, et al. Immunotherapy in colorectal cancer: rationale, challenges and potential. *Nat Rev Gastroenterol Hepatol.* (2019) 16:361–75. doi: 10.1038/s41575-019-0126-x
- Shi J, Sun Z, Gao Z, Huang D, Hong H, Gu J. Radioimmunotherapy in colorectal cancer treatment: present and future. *Front Immunol.* (2023) 14:1105180. doi: 10.3389/fimmu.2023.1105180
- Zhao K, Qi L, Li Q, Wang Y, Qian C, Shi Z. Self-absorbing multilayer skin-like composite with Phyllostachys nigra polysaccharides promotes wound healing. *Adv Compos Hybrid Mater.* (2024) 7:225. doi: 10.1007/s42114-024-01018-x
- Liang H, Lu Q, Yang J, Yu G. Supramolecular biomaterials for cancer immunotherapy. *Res (Wash D C).* (2023) 6:211. doi: 10.34133/research.0211
- Xie S, Cai Y, Chen D, Xiang Y, Cai W, Mao J, et al. Single-cell transcriptome analysis reveals heterogeneity and convergence of the tumor microenvironment in colorectal cancer. *Front Immunol.* (2022) 13:1003419. doi: 10.3389/fimmu.2022.1003419
- He R, Huang S, Lu J, Su L, Gao X, Chi H. Unveiling the immune symphony: decoding colorectal cancer metastasis through immune interactions. *Front Immunol.* (2024) 15:1362709. doi: 10.3389/fimmu.2024.1362709
- Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* (2012) 487:330–37. doi: 10.1038/nature11252
- Wang H, Tian T, Zhang J. Tumor-associated macrophages (TAMs) in colorectal cancer (CRC): from mechanism to therapy and prognosis. *Int J Mol Sci.* (2021) 22(16):8470. doi: 10.3390/ijms22168470
- Mei Y, Xiao W, Hu H, Lu G, Chen L, Sun Z, et al. Single-cell analyses reveal suppressive tumor microenvironment of human colorectal cancer. *Clin Transl Med.* (2021) 11:e422. doi: 10.1002/ctm2.422
- Qi J, Crinier A, Escaliere B, Ye Y, Wang Z, Zhang T, et al. Single-cell transcriptomic landscape reveals tumor specific innate lymphoid cells associated with colorectal cancer progression. *Cell Rep Med.* (2021) 2:100353. doi: 10.1016/j.xcrm.2021.100353
- Khaliq AM, Erdogan C, Kurt Z, Turgut SS, Grunvald MW, Rand T, et al. Refining colorectal cancer classification and clinical stratification through a single-cell atlas. *Genome Biol.* (2022) 23:113. doi: 10.1186/s13059-022-02677-z
- Valdeolivas A, Amberg B, Giroud N, Richardson M, Galvez E, Badillo S, et al. Profiling the heterogeneity of colorectal cancer consensus molecular subtypes using spatial transcriptomics. *NPJ Precis Oncol.* (2024) 8:10. doi: 10.1038/s41698-023-00488-4
- Roelands J, van der Ploeg M, Ijsselsteijn ME, Dang H, Boonstra JJ, Hardwick J, et al. Transcriptomic and immunophenotypic profiling reveals molecular and immunological hallmarks of colorectal cancer tumorigenesis. *Gut.* (2023) 72:326–39. doi: 10.1136/gutjnl-2022-327608
- Liebl MC, Hofmann TG. The role of p53 signaling in colorectal cancer. *Cancers (Basel).* (2021) 13(9):2125. doi: 10.3390/cancers13092125
- Zhao H, Ming T, Tang S, Ren S, Yang H, Liu M, et al. Wnt signaling in colorectal cancer: pathogenic role and therapeutic target. *Mol Cancer.* (2022) 21:144. doi: 10.1186/s12943-022-01616-7
- Jackstadt R, van Hooff SR, Leach JD, Cortes-Lavaud X, Lohuis JO, Ridgway RA, et al. Epithelial NOTCH signaling rewires the tumor microenvironment of colorectal cancer to drive poor-prognosis subtypes and metastasis. *Cancer Cell.* (2019) 36:319–36. doi: 10.1016/j.ccell.2019.08.003
- Tang Q, Chen J, Di Z, Yuan W, Zhou Z, Liu Z, et al. TM4SF1 promotes EMT and cancer stemness via the Wnt/beta-catenin/SOX2 pathway in colorectal cancer. *J Exp Clin Cancer Res.* (2020) 39:232. doi: 10.1186/s13046-020-01690-z
- Yang C, Dou R, Wei C, Liu K, Shi D, Zhang C, et al. Tumor-derived exosomal microRNA-106b-5p activates EMT-cancer cell and M2-subtype TAM interaction to facilitate CRC metastasis. *Mol Ther.* (2021) 29:2088–107. doi: 10.1016/j.jymthe.2021.02.006
- Al ZM, Boukhoud M, Shammaa H, Mouslem H, El AL, Iratni R, et al. Immunotherapy and immunoevasion of colorectal cancer. *Drug Discovery Today.* (2023) 28:103669. doi: 10.1016/j.drudis.2023.103669
- Zhong F, Lin Y, Jing X, Ye Y, Wang S, Shen Z. Innate tumor killers in colorectal cancer. *Cancer Lett.* (2022) 527:115–26. doi: 10.1016/j.canlet.2021.12.022
- Zhang S, Zhang G, Wang M, Guo SB, Wang F, Li Y, et al. Artificial intelligence hybrid survival assessment system for robot-assisted proctectomy: A retrospective cohort study. *Jco Precis Oncol.* (2024) 8:e2400089. doi: 10.1200/PO.24.00089
- Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol.* (2018) 36:411–20. doi: 10.1038/nbt.4096
- Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods.* (2019) 16:1289–96. doi: 10.1038/s41592-019-0619-0
- Zhou Y, Yang D, Yang Q, Lv X, Huang W, Zhou Z, et al. Single-cell RNA landscape of intratumoral heterogeneity and immunosuppressive microenvironment in advanced osteosarcoma. *Nat Commun.* (2020) 11:6322. doi: 10.1038/s41467-020-20059-6
- Li X, Lin Z, Zhao F, Huang T, Fan W, Cen L, et al. Unveiling the cellular landscape: insights from single-cell RNA sequencing in multiple myeloma. *Front Immunol.* (2024) 15:1458638. doi: 10.3389/fimmu.2024.1458638
- Sun L, Shao W, Lin Z, Lin J, Zhao F, Yu J. Single-cell RNA sequencing explored potential therapeutic targets by revealing the tumor microenvironment of neuroblastoma and its expression in cell death. *Discovery Oncol.* (2024) 15:409. doi: 10.1007/s12672-024-01286-5
- Zhao F, Hong J, Zhou G, Huang T, Lin Z, Zhang Y, et al. Elucidating the role of tumor-associated ALOX5+ mast cells with transformative function in cervical cancer progression via single-cell RNA sequencing. *Front Immunol.* (2024) 15:1434450. doi: 10.3389/fimmu.2024.1434450
- Jiang H, Yu D, Yang P, Guo R, Kong M, Gao Y, et al. Revealing the transcriptional heterogeneity of organ-specific metastasis in human gastric cancer using single-cell RNA Sequencing. *Clin Transl Med.* (2022) 12:e730. doi: 10.1002/ctm2.730
- Wu F, Fan J, He Y, Xiong A, Yu J, Li Y, et al. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat Commun.* (2021) 12:2540. doi: 10.1038/s41467-021-22801-0

48. Nie W, Zhao Z, Liu Y, Wang Y, Zhang J, Hu Y, et al. Integrative single-cell analysis of cardiomyopathy identifies differences in cell stemness and transcriptional regulatory networks among fibroblast subpopulations. *Cardiol Res Pract.* (2024) 2024:3131633. doi: 10.1155/2024/3131633
49. Jin W, Zhang Y, Zhao Z, Gao M. Developing targeted therapies for neuroblastoma by dissecting the effects of metabolic reprogramming on tumor microenvironments and progression. *Theranostics.* (2024) 14:3439–69. doi: 10.7150/thno.93962
50. Zhang Y, Zhao Z, Huang W, Kim BS, Lin L, Li X, et al. Pan-cancer single-cell analysis revealing the heterogeneity of cancer-associated fibroblasts in skin tumors. *Curr Gene Ther.* (2024) 25:e15665232331353. doi: 10.2174/0115665232331353240911080642
51. Huang W, Kim BS, Zhang Y, Lin L, Chai G, Zhao Z. Regulatory T cells subgroups in the tumor microenvironment cannot be overlooked: Their involvement in prognosis and treatment strategy in melanoma. *Environ Toxicol.* (2024) 39:4512–30. doi: 10.1002/tox.24247
52. Ding Y, Zhao Z, Cai H, Zhou Y, Chen H, Bai Y, et al. Single-cell sequencing analysis related to sphingolipid metabolism guides immunotherapy and prognosis of skin cutaneous melanoma. *Front Immunol.* (2023) 14:1304466. doi: 10.3389/fimmu.2023.1304466
53. Ge Q, Zhao Z, Li X, Yang F, Zhang M, Hao Z, et al. Deciphering the suppressive immune microenvironment of prostate cancer based on CD4+ regulatory T cells: Implications for prognosis and therapy prediction. *Clin Transl Med.* (2024) 14:e1552. doi: 10.1002/ctm2.1552
54. Zhang X, Lan Y, Xu J, Quan F, Zhao E, Deng C, et al. CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res.* (2019) 47:D721–28. doi: 10.1093/nar/gky900
55. Zhou W, Lin Z, Tan W. Deciphering the molecular landscape: integrating single-cell transcriptomics to unravel myofibroblast dynamics and therapeutic targets in clear cell renal cell carcinomas. *Front Immunol.* (2024) 15:1374931. doi: 10.3389/fimmu.2024.1374931
56. Shao W, Lin Z, Xiahou Z, Zhao F, Xu J, Liu X, et al. Single-cell RNA sequencing reveals that MYBL2 in Malignant epithelial cells is involved in the development and progression of ovarian cancer. *Front Immunol.* (2024) 15:1438198. doi: 10.3389/fimmu.2024.1438198
57. Zheng L, Qin S, Si W, Wang A, Xing B, Gao R, et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science.* (2021) 374:abe6474. doi: 10.1126/science.abe6474
58. Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods.* (2017) 14:1083–86. doi: 10.1038/nmeth.4463
59. Gulati GS, Sikandar S, Wesche DJ, Manjunath A, Bharadwaj A, Berger MJ, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science.* (2020) 367:405–11. doi: 10.1126/science.aax0249
60. Liu P, Xing N, Xiahou Z, Yan J, Lin Z, Zhang J. Unraveling the intricacies of glioblastoma progression and recurrence: insights into the role of NFYB and oxidative phosphorylation at the single-cell level. *Front Immunol.* (2024) 15:1368685. doi: 10.3389/fimmu.2024.1368685
61. Lin Z, Li X, Shi H, Cao R, Zhu L, Dang C, et al. Decoding the tumor microenvironment and molecular mechanism: unraveling cervical cancer subpopulations and prognostic signatures through scRNA-Seq and bulk RNA-seq analyses. *Front Immunol.* (2024) 15:1351287. doi: 10.3389/fimmu.2024.1351287
62. Zhao Z, Ding Y, Tran LJ, Chai G, Lin L. Innovative breakthroughs facilitated by single-cell multi-omics: manipulating natural killer cell functionality correlates with a novel subcategory of melanoma cells. *Front Immunol.* (2023) 14:1196892. doi: 10.3389/fimmu.2023.1196892
63. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods.* (2017) 14:979–82. doi: 10.1038/nmeth.4402
64. Xing J, Cai H, Lin Z, Zhao L, Xu H, Song Y, et al. Examining the function of macrophage oxidative stress response and immune system in glioblastoma multiforme through analysis of single-cell transcriptomics. *Front Immunol.* (2023) 14:1288137. doi: 10.3389/fimmu.2023.1288137
65. Zhao ZJ, Wei DP, Zheng RZ, Peng T, Xiao X, Li FS. The gene coexpression analysis identifies functional modules dynamically changed after traumatic brain injury. *Comput Math Methods Med.* (2021) 2021:5511598. doi: 10.1155/2021/5511598
66. Li XY, Zhao ZJ, Wang JB, Shao YH, Hui-Liu, You JX, et al. m7G methylation-related genes as biomarkers for predicting overall survival outcomes for hepatocellular carcinoma. *Front Bioeng Biotechnol.* (2022) 10:849756. doi: 10.3389/fbioe.2022.849756
67. Zhao ZJ, Zheng RZ, Wang XJ, Li TQ, Dong XH, Zhao CY, et al. Integrating lipidomics and transcriptomics reveals the crosstalk between oxidative stress and neuroinflammation in central nervous system demyelination. *Front Aging Neurosci.* (2022) 14:870957. doi: 10.3389/fnagi.2022.870957
68. Zhao Z, Li T, Dong X, Wang X, Zhang Z, Zhao C, et al. Untargeted metabolomic profiling of cuprizone-induced demyelination in mouse corpus callosum by UPLC-orbitrap/MS reveals potential metabolic biomarkers of CNS demyelination disorders. *Oxid Med Cell Longev.* (2021) 2021:7093844. doi: 10.1155/2021/7093844
69. Lin Z, Fan W, Yu X, Liu J, Liu P. Research into the mechanism of intervention of SanQi in endometriosis based on network pharmacology and molecular docking technology. *Med (Baltimore).* (2022) 101:e30021. doi: 10.1097/MD.00000000000030021
70. Zhao J, Jiao W, Sui X, Zou J, Wang J, Lin Z. Construction of a prognostic model of luteolin for endometrial carcinoma. *Am J Transl Res.* (2023) 15:2122–39.
71. Suo S, Zhu Q, Saadatpour A, Fei L, Guo G, Yuan GC. Revealing the critical regulators of cell identity in the mouse cell atlas. *Cell Rep.* (2018) 25:1436–45. doi: 10.1016/j.celrep.2018.10.045
72. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun.* (2021) 12:1088. doi: 10.1038/s41467-021-21246-9
73. Wang Y, Zhao ZJ, Kang XR, Bian T, Shen ZM, Jiang Y, et al. lncRNA DLEU2 acts as a miR-181a sponge to regulate SEPP1 and inhibit skeletal muscle differentiation and regeneration. *Aging (Albany Ny).* (2020) 12:24033–56. doi: 10.18632/aging.104095
74. Zheng R, Zhuang Z, Zhao C, Zhao Z, Yang X, Zhou Y, et al. Chinese admission warning strategy for predicting the hospital discharge outcome in patients with traumatic brain injury. *J Clin Med.* (2022) 11(4):974. doi: 10.3390/jcm11040974
75. Zheng RZ, Zhao ZJ, Yang XT, Jiang SW, Li YD, Li WJ, et al. Initial CT-based radiomics nomogram for predicting in-hospital mortality in patients with traumatic brain injury: a multicenter development and validation study. *Neurol Sci.* (2022) 43:4363–72. doi: 10.1007/s10072-022-05954-8
76. Lin Z, Zou J, Sui X, Yao S, Lin L, Wang J, et al. Necroptosis-related lncRNA signature predicts prognosis and immune response for cervical squamous cell carcinoma and endocervical adenocarcinomas. *Sci Rep.* (2022) 12:16285. doi: 10.1038/s41598-022-20858-5
77. Zou J, Lin Z, Jiao W, Chen J, Lin L, Zhang F, et al. A multi-omics-based investigation of the prognostic and immunological impact of necroptosis-related mRNA in patients with cervical squamous carcinoma and adenocarcinoma. *Sci Rep.* (2022) 12:16773. doi: 10.1038/s41598-022-20566-0
78. Zhao J, Zou J, Jiao W, Lin L, Wang J, Lin Z. Construction of N-7 methylguanine-related mRNA prognostic model in uterine corpus endometrial carcinoma based on multi-omics data and immune-related analysis. *Sci Rep.* (2022) 12:18813. doi: 10.1038/s41598-022-22879-6
79. Zhao ZJ, Chen D, Zhou LY, Sun ZL, Wang BC, Feng DF. Prognostic value of different computed tomography scoring systems in patients with severe traumatic brain injury undergoing decompressive craniectomy. *J Comput Assist Tomogr.* (2022) 46:800–07. doi: 10.1097/RCT.0000000000001343
80. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol.* (2019) 37:773–82. doi: 10.1038/s41587-019-0114-2
81. Hu H, Zhang W, Zhou Y, Zhao K, Kuang J, Liu X, et al. Engineered mitochondrial ROS scavenger nanocapsule to enhance lung biodistribution and reduce inflammation for the treatment of ARDS. *Adv Compos Hybrid Mater.* (2024) 7:194. doi: 10.1007/s42114-024-00989-1
82. Zeng X, Wang S, Peng Z, Wang M, Zhao K, Xu BB, et al. Rapid screening and sensing of stearoyl-CoA desaturase 1 (SCD1) inhibitors from ginger and their efficacy in ameliorating non-alcoholic fatty liver disease. *J Food Meas Charact.* (2024) 18:6843–57. doi: 10.1007/s11694-024-02697-2
83. Uzbekov R, Prigent C. A journey through time on the discovery of cell cycle regulation. *Cells.* (2022) 11(4):704. doi: 10.3390/cells11040704
84. Lichtenstern CR, Ngu RK, Shalpour S, Karin M. Immunotherapy, inflammation and colorectal cancer. *Cells.* (2020) 9(3):618. doi: 10.3390/cells9030618
85. Jung Y, Lee S, Choi HS, Kim SN, Lee E, Shin Y, et al. Clinical validation of colorectal cancer biomarkers identified from bioinformatics analysis of public expression data. *Clin Cancer Res.* (2011) 17:700–09. doi: 10.1158/1078-0432.CCR-10-1300
86. Niu Y, Shao Z, Wang H, Yang J, Zhang F, Luo Y, et al. LASP1-S100A11 axis promotes colorectal cancer aggressiveness by modulating TGFbeta/Smad signaling. *Sci Rep.* (2016) 6:26112. doi: 10.1038/srep26112
87. Asghari AF, Goliaei B, Minuchehr Z. Analyzing large scale gene expression data in colorectal cancer reveals important clues; CLCA1 and SELENBP1 downregulated in CRC not in normal and not in adenoma. *Am J Cancer Res.* (2022) 12:371–80.
88. Zhang X, Hong R, Bei L, Hu Z, Yang X, Song T, et al. SELENBP1 inhibits progression of colorectal cancer by suppressing epithelial-mesenchymal transition. *Open Med (Wars).* (2022) 17:1390–404. doi: 10.1515/med-2022-0532
89. Bai Y, Li LD, Li J, Chen RF, Yu HL, Sun HF, et al. A FXYD5/TGF-beta/SMAD positive feedback loop drives epithelial-to-mesenchymal transition and promotes tumor growth and metastasis in ovarian cancer. *Int J Oncol.* (2020) 56:301–14. doi: 10.3892/ijo.2019.4911
90. Bai Y, Li L, Li J, Lu X. Association analysis of FXYD5 with prognosis and immunological characteristics across pan-cancer. *Heliyon.* (2024) 10:e30727. doi: 10.1016/j.heliyon.2024.e30727
91. Click B, Pinsky PF, Hickey T, Doroudi M, Schoen RE. Association of colonoscopy adenoma findings with long-term colorectal cancer incidence. *Jama.* (2018) 319:2021–31. doi: 10.1001/jama.2018.5809
92. Paunovic V, Kosic M, Misirkic-Marjanovic M, Trajkovic V, Harhaji-Trajkovic L. Dual targeting of tumor cell energy metabolism and lysosomes as an anticancer strategy. *Biochim Biophys Acta Mol Cell Res.* (2021) 1868:118944. doi: 10.1016/j.bbamer.2020.118944
93. Ma J, Sun L, Gao W, Li Y, Dong D. RNA binding protein: coordinated expression between the nuclear and mitochondrial genomes in tumors. *J Transl Med.* (2023) 21:512. doi: 10.1186/s12967-023-04373-3

94. FeiFei W, HongHai X, YongRong Y, PingXiang W, JianHua W, XiaoHui Z, et al. FBX8 degrades GSTP1 through ubiquitination to suppress colorectal cancer progression. *Cell Death Dis.* (2019) 10:351. doi: 10.1038/s41419-019-1588-z
95. Petrova DT, Asif AR, Armstrong VW, Dimova I, Toshev S, Yaramov N, et al. Expression of chloride intracellular channel protein 1 (CLIC1) and tumor protein D52 (TPD52) as potential biomarkers for colorectal cancer. *Clin Biochem.* (2008) 41:1224–36. doi: 10.1016/j.clinbiochem.2008.07.012
96. Yang J, Yang X, Guo T, Wang L, Zhao Z, Hu Z, et al. Can thymosin beta 10 function both as a non-invasive biomarker and chemotherapeutic target in human colorectal cancer? *Transl Oncol.* (2024) 46:102026. doi: 10.1016/j.tranon.2024.102026
97. Arai K, Ishimatsu H, Iwasaki T, Tsuchiya C, Sonoda A, Ohata K. Membranous S100A10 involvement in the tumor budding of colorectal cancer during oncogenesis: report of two cases with immunohistochemical analysis. *World J Surg Oncol.* (2020) 18:289. doi: 10.1186/s12957-020-02075-4
98. Bernacki RJ, Niedbala MJ, Korytnyk W. Glycosidases in cancer and invasion. *Cancer Metastasis Rev.* (1985) 4:81–101. doi: 10.1007/BF00047738
99. Wang Z, Pang J, Wang L, Dong Q, Jin D. CEBPB regulates the bile acid receptor FXR to accelerate colon cancer progression by modulating aerobic glycolysis. *J Clin Lab Anal.* (2022) 36:e24703. doi: 10.1002/jcla.24703
100. Horiuchi S, Yamamoto H, Min Y, Adachi Y, Itoh F, Imai K. Association of ets-related transcriptional factor E1AF expression with tumour progression and overexpression of MMP-1 and matrilysin in human colorectal cancer. *J Pathol.* (2003) 200:568–76. doi: 10.1002/path.1387
101. Liu S, Gao Q, Li Y, Lun J, Yu M, Zhang H, et al. XBP1s acts as a transcription factor of IRE1alpha and promotes proliferation of colon cancer cells. *Arch Biochem Biophys.* (2023) 737:109552. doi: 10.1016/j.abb.2023.109552
102. Zhang B, Du X, Fan Y, Qu G, Pang LK, Zhao R, et al. DLX2 promotes osteosarcoma epithelial-mesenchymal transition and doxorubicin resistance by enhancing HOXC8-CDH2 axis. *Iscience.* (2023) 26:108272. doi: 10.1016/j.isci.2023.108272
103. Tang P, Huang H, Chang J, Zhao GF, Lu ML, Wang Y. Increased expression of DLX2 correlates with advanced stage of gastric adenocarcinoma. *World J Gastroenterol.* (2013) 19:2697–703. doi: 10.3748/wjg.v19.i17.2697
104. Takeuchi Y, Nishikawa H. Roles of regulatory T cells in cancer immunity. *Int Immunol.* (2016) 28:401–09. doi: 10.1093/intimm/dxw025
105. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med.* (2018) 24:1550–58. doi: 10.1038/s41591-018-0136-1
106. Bian X, Liu R, Meng Y, Xing D, Xu D, Lu Z. Lipid metabolism and cancer. *J Exp Med.* (2021) 218(1):e20201606. doi: 10.1084/jem.20201606
107. Cheng C, Geng F, Cheng X, Guo D. Lipid metabolism reprogramming and its potential targets in cancer. *Cancer Commun (Lond).* (2018) 38:27. doi: 10.1186/s40880-018-0301-4
108. Hsu CC, Peng D, Cai Z, Lin HK. AMPK signaling and its targeting in cancer progression and treatment. *Semin Cancer Biol.* (2022) 85:52–68. doi: 10.1016/j.semcancer.2021.04.006
109. Kosova F, Coskun T, Kaya Y, Kara E, Ari Z. Adipocytokine levels of colon cancer patients before and after treatment. *Bratisl Lek Listy.* (2013) 114:394–97. doi: 10.4149/bll\_2013\_083
110. Jaber SA, Cohen A, D'Souza C, Abdulrazzaq YM, Ojha S, Bastaki S, et al. Lipocalin-2: Structure, function, distribution and role in metabolic disorders. *BioMed Pharmacother.* (2021) 142:112002. doi: 10.1016/j.biopha.2021.112002
111. Zuo X, Xu M, Yu J, Wu Y, Moussalli MJ, Manyam GC, et al. Potentiation of colon cancer susceptibility in mice by colonic epithelial PPAR-delta/beta overexpression. *J Natl Cancer Inst.* (2014) 106:dju052. doi: 10.1093/jnci/dju052
112. Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat Rev Gastroenterol Hepatol.* (2019) 16:713–32. doi: 10.1038/s41575-019-0189-8
113. Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. *Transl Oncol.* (2021) 14:101174. doi: 10.1016/j.tranon.2021.101174
114. Sudhi M, Shukla VK, Shetty DK, Gupta V, Desai AS, Naik N, et al. Advancements in bladder cancer management: A comprehensive review of artificial intelligence and machine learning applications. *Engineered Sci.* (2023) 26:1003. doi: 10.30919/es1003
115. Zhang L, Li Z, Skrzypczynska KM, Fang Q, Zhang W, O'Brien SA, et al. Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. *Cell.* (2020) 181:442–59. doi: 10.1016/j.cell.2020.03.048
116. Song S, Feng L, Xi K, Sun Z, Kong D, Luo Z, et al. Single-cell profiling of the copy-number heterogeneity in colorectal cancer. *Chin Med J (Engl).* (2023) 136:707–18. doi: 10.1097/CM9.0000000000002469
117. Chi H, Huang J, Yan Y, Jiang C, Zhang S, Chen H, et al. Unraveling the role of disulfidptosis-related LncRNAs in colon cancer: a prognostic indicator for immunotherapy response, chemotherapy sensitivity, and insights into cell death mechanisms. *Front Mol Biosci.* (2023) 10:1254232. doi: 10.3389/fmolb.2023.1254232
118. Wang M, Deng C, Yang C, Yan M, Lu H, Zhang Y, et al. Unraveling temporal and spatial biomarkers of epithelial-mesenchymal transition in colorectal cancer: insights into the crucial role of immunosuppressive cells. *J Transl Med.* (2023) 21:794. doi: 10.1186/s12967-023-04600-x
119. Briede I, Balodis D, Gardovskis J, Strumfa I. Stemness, inflammation and epithelial-mesenchymal transition in colorectal carcinoma: the intricate network. *Int J Mol Sci.* (2021) 22(23):12891. doi: 10.3390/ijms222312891
120. Deng S, Zhang Y, Wang H, Liang W, Xie L, Li N, et al. ITPRIPL1 binds CD3epsilon to impede T cell activation and enable tumor immune evasion. *Cell.* (2024) 187:2305–23. doi: 10.1016/j.cell.2024.03.019
121. Zhang J, Huang D, Saw PE, Song E. Turning cold tumors hot: from molecular mechanisms to clinical applications. *Trends Immunol.* (2022) 43:523–45. doi: 10.1016/j.it.2022.04.010
122. Zeng Q, Saghafinia S, Chryplewicz A, Fournier N, Christe L, Xie YQ, et al. Aberrant hyperexpression of the RNA binding protein FMRP in tumors mediates immune evasion. *Science.* (2022) 378:eabl7207. doi: 10.1126/science.abl7207
123. Picard E, Verschoor CP, Ma GW, Pawelec G. Relationships between immune landscapes, genetic subtypes and responses to immunotherapy in colorectal cancer. *Front Immunol.* (2020) 11:369. doi: 10.3389/fimmu.2020.00369
124. Mamun MA, Mannoor K, Cao J, Qadri F, Song X. SOX2 in cancer stemness: tumor Malignancy and therapeutic potentials. *J Mol Cell Biol.* (2020) 12:85–98. doi: 10.1093/jmcb/mjy080
125. Lambert AW, Pattabiraman DR, Weinberg RA. Emerging biological principles of metastasis. *Cell.* (2017) 168:670–91. doi: 10.1016/j.cell.2016.11.037
126. Battle E, Clevers H. Cancer stem cells revisited. *Nat Med.* (2017) 23:1124–34. doi: 10.1038/nm.4409
127. Faubert B, Solmonson A, DeBerardinis RJ. Metabolic reprogramming and cancer progression. *Science.* (2020) 368:eaaw5473. doi: 10.1126/science.aaw5473
128. Vaupel P, Schmidberger H, Mayer A. The Warburg effect: essential part of metabolic reprogramming and central contributor to cancer progression. *Int J Radiat Biol.* (2019) 95:912–19. doi: 10.1080/09553002.2019.1589653
129. Jiang L, Liu J, Zhang S, Jiang C, Huang J, Chen H, et al. Role of glycosylation-related gene MGAT1 in pancreatic ductal adenocarcinoma. *Front Immunol.* (2024) 15:1438935. doi: 10.3389/fimmu.2024.1438935
130. Smith RL, Soeters MR, Wust R, Houtkooper RH. Metabolic flexibility as an adaptation to energy resources and requirements in health and disease. *Endocr Rev.* (2018) 39:489–517. doi: 10.1210/er.2017-00211
131. Sasidharan NV, Elkord E. Immune checkpoint inhibitors in cancer therapy: a focus on T-regulatory cells. *Immunol Cell Biol.* (2018) 96:21–33. doi: 10.1111/imcb.1003
132. Chen H, Zuo H, Huang J, Liu J, Jiang L, Jiang C, et al. Unravelling infiltrating T-cell heterogeneity in kidney renal clear cell carcinoma: Integrative single-cell and spatial transcriptomic profiling. *J Cell Mol Med.* (2024) 28:e18403. doi: 10.1111/jcmm.18403