Check for updates

#### OPEN ACCESS

EDITED BY Zhijie Zhao, Shanghai Jiao Tong University, China

REVIEWED BY Yuxuan Song, Peking University People's Hospital, China Fuhao Wang, Shandong University, China

\*CORRESPONDENCE Xin Song Songxin20250501@163.com

RECEIVED 13 May 2025 ACCEPTED 30 May 2025 PUBLISHED 12 JUNE 2025

#### CITATION

Ning W, Jia W, Ning J, Zhou L, Li Z, Zhang L and Song X (2025) Cross-cohort multi-omics analysis identifies novel clusters driven by EMT signatures in colorectal cancer. *Front. Immunol.* 16:1628005. doi: 10.3389/fimmu.2025.1628005

#### COPYRIGHT

© 2025 Ning, Jia, Ning, Zhou, Li, Zhang and Song. 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.

## Cross-cohort multi-omics analysis identifies novel clusters driven by EMT signatures in colorectal cancer

Wu Ning<sup>1</sup>, Wenqing Jia<sup>2</sup>, Jingyuan Ning<sup>3</sup>, Lei Zhou<sup>1</sup>, Zongze Li<sup>1</sup>, Lin Zhang<sup>1</sup> and Xin Song<sup>1\*</sup>

<sup>1</sup>Department of General Surgery, China-Japan Friendship Hospital, Beijing, China, <sup>2</sup>Institute of Clinical Medicine, China-Japan Friendship Hospital, Beijing, China, <sup>3</sup>Institute of Basic Medical Sciences, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China

**Introduction:** The pronounced heterogeneity of colorectal cancer (CRC) significantly impacts patient prognosis and therapeutic response, making elucidation of its molecular mechanisms critical for developing precision treatment strategies. This study aimed to systematically characterize tumor cell heterogeneity and explore its clinical implications.

**Methods:** Five single-cell RNA sequencing cohorts were integrated (comprising 70 CRC samples and 164,173 cells) to systematically analyze tumor cell heterogeneity. Unsupervised clustering analysis based on VEGFR+ tumor cell signature genes was used to stratify CRC patients. Key molecular mechanisms were validated through in vitro cellular experiments, in vivo animal models, molecular docking, and dynamics simulations.

**Results:** The analysis successfully identified five distinct tumor cell subtypes, with the VEGFR+ subtype exhibiting marked epithelial-mesenchymal transition (EMT) activation signatures and strong association with metastasis and poor clinical outcomes. Based on VEGFR+ signature genes, CRC patients were stratified into three subgroups: C1 (metabolically active), C2 (proliferative), and C3 (invasive), with the C3 subtype demonstrating high metastatic potential, stem-like properties, and an immunosuppressive microenvironment, along with a five-year survival rate below 50%. Mechanistic investigations identified HOXC6 as a key driver of the C3 subtype, with HOXC6 knockout significantly suppressing CRC cell proliferation, migration, and invasion. Furthermore, molecular docking revealed that the targeted agent abemaciclib effectively binds HOXC6, with both cellular and animal experiments confirming its ability to inhibit CRC cell functions and significantly reduce tumor burden in nude mice.

**Discussion:** This study establishes the first single-cell-resolution molecular classification system for CRC, delineates the mechanistic link between EMT subtypes and metastatic progression, and identifies HOXC6 as a novel therapeutic vulnerability. These findings provide a translational foundation for precision oncology and offer new rationale for precision diagnosis and treatment of colorectal cancer.

#### KEYWORDS

colorectal cancer, EMT, HOXC6, scRNA-seq, metastatic

#### 1 Introduction

Colorectal cancer ranks as the third leading cause of cancerrelated mortality worldwide, with five-year survival rates plummeting below 15% following distant metastasis (1). This dire clinical outcome stems from two pivotal biological hallmarks: persistent high recurrence rates and an inherent metastatic propensity observable even at primary tumor stages (2). Current clinical treatment for colorectal cancer relies on 5-FU-based chemotherapy (3), often combined with targeted therapies like anti-EGFR or anti-VEGF agents (4). While these regimens improve response rates, drug resistance and toxicity remain major limitations, especially in advanced-stage disease. Immunotherapy shows promise but is currently limited to a small subset of patients with MSI-H/dMMR tumors, highlighting the need for broader therapeutic strategies (5, 6).

Advances in tumor molecular biology have established tumor heterogeneity as the central driver of these clinical challenges (2, 7). Such heterogeneity manifests not only interpatient variability but, more critically, as spatiotemporal diversity among cellular subpopulations within individual tumors (8). This multilevel complexity fundamentally undermines the efficacy of conventional therapies. The recent advent of single-cell sequencing technologies has revolutionized our capacity to deconstruct tumor heterogeneity at cellular resolution, enabling precise molecular characterization of distinct cellular states within the tumor microenvironment (9).

Within the complex tumor microenvironment, distinct neoplastic subpopulations emerge as pivotal effectors of disease progression through their acquisition of stem-like properties and metastatic competence (10). These metastasis-prone cells are molecularly hallmarked by complete or partial epithelial-tomesenchymal transition (EMT) activation-a dynamic, reversible biological program that initiates the metastatic cascade through coordinated loss of epithelial polarity and intercellular junctions coupled with gain of mesenchymal traits, including enhanced migratory capacity and apoptosis resistance (11, 12). Notably, recent advances reveal EMT's pleiotropic role extends beyond physical motility, dually reprogramming cellular metabolism (e.g., augmented glycolysis and glutaminyls) and immune-related molecule expression (e.g., PD-L1 upregulation) to facilitate environmental adaptation during metastatic dissemination (13-15). This process is governed by a multilayered regulatory network: sequential activation of core transcription factors (16) (SNAIL, TWIST, ZEB families (17)) and microenvironmental signals (TGF- $\beta$  (18), WNT (19), inflammatory cytokines (20)) provide spatiotemporal control. Critically, EMT exhibits marked state plasticity, with tumor cells occupying full, partial, or hybrid EMT states-each demonstrating distinct metastatic potential and therapeutic vulnerability (21). This functional heterogeneity directly contributes to treatment resistance in colorectal cancer. Systematic dissection of EMT's molecular circuitry therefore offers dual utility: elucidating fundamental principles of metastatic biology while revealing actionable targets for precision therapeutics.

Here we systematically deconstruct colorectal cancer heterogeneity through single-cell transcriptomics, establishing the first EMT activity-based molecular classification that identifies the VEGFR+ EMT subtype—among five distinct tumor cell subpopulations—as clinically associated with metastatic progression and adverse outcomes. Our work mechanistically delineates HOXC6 as a master regulatory node governing this high-risk subtype, with both functional validation and therapeutic targeting experiments confirming its druggable potential. These findings crystallize the central role of EMT reprogramming in metastatic dissemination while providing a framework for precision oncology—from molecular stratification to targeted intervention—that advances personalized therapeutic paradigms in solid tumors.

#### 2 Result

# 2.1 Single-cell transcriptomic profiling reveals distinct tumor cell subpopulations in colorectal cancer

To systematically characterize tumor cell heterogeneity in colorectal cancer (CRC) patients, we analyzed single-cell RNA sequencing data from five independent cohorts (GSE132257, GSE132465, GSE144735, GSE188711, GSE205506; Figure 1A). After stringent quality control, we obtained high-quality transcriptomic profiles for 164,173 cells from 70 CRC samples. Unsupervised clustering based on canonical cell markers identified 11 major cell populations (Figure 1B), including plasma cells, B cells, macrophages, dendritic cells, fibroblasts, endothelial cells, mast cells, mural cells, monocytes, T/NK cells, and epithelial cells. The expression patterns of marker genes for each population were visualized using dot plots (Figure 1C). For in-depth analysis of tumor cell heterogeneity, epithelial cells were isolated and subjected to further clustering. Malignant cells were identified using the SCEVAN algorithm and subsequently classified into five molecularly distinct subpopulations (Figure 1D): VEGFA+TC, EDN1+TC, CDK1+TC, AKR1C3+TC, and GALNT3+TC. Dot plot visualization confirmed the expression of signature genes for each subpopulation (Figure 1E). Among these, the VEGFA+TC subpopulation exhibited marked upregulation of CDH1, MMP7, and VEGFA - key regulators of cell migration, invasion, and epithelial-mesenchymal transition (EMT). Functional enrichment analysis of the top 100 differentially expressed genes ( $log_2FC > 0.5$ ; adjusted p-value < 0.05) in each subpopulation revealed that VEGFA+TC cells were significantly enriched in pathways related to cell-cell adhesion, cell motility, and epithelial migration (Figure 1F), consistent with an EMT phenotype and metastatic potential. The other subpopulations displayed distinct functional characteristics: EDN1+TC was enriched in protein metabolic processes and endoplasmic reticulum stress responses; CDK1+TC showed strong cell cycle and DNA replication activity; AKR1C3 +TC was specialized in heavy metal detoxification and oxidative stress response; while GALNT3+TC participated in glycosylation



Single-cell sequencing analysis reveals CRC tumor cell heterogeneity (A) Dimensionality reduction clustering of five single-cell cohorts. (B)UMAP visualization of different cell types. (C) Dot plot showing marker gene expression across cell types. (D) UMAP visualization of tumor cells. (E) Dot plot displays marker genes of five tumor epithelial subtypes. (F) Bar plot of enriched pathways from differential gene analysis of five tumor epithelial subtypes. (G) Scatter plot showing correlation between VEGFA+TC subtype signature genes and metastatic potential. (H) Kaplan-Meier survival curves of patients with high/low VEGFA+TC subtype signature expression across three CRC cohorts.

and extracellular matrix remodeling. Validation in bulk transcriptomic datasets (Figure 1G) demonstrated a significant positive correlation between the VEGFA+TC signature and metastatic potential. Kaplan-Meier survival analysis (Figure 1H) further confirmed that patients with high VEGFA+TC signature expression had significantly worse clinical outcomes. In summary, our single-cell analysis identified a clinically relevant VEGFA+TC subpopulation characterized by EMT activation and heightened



Unsupervised clustering across cohorts identifies new CRC subtypes (A) Unsupervised clustering of CRC patients in TCGA dataset based on VEGFA +TC subtype signature scores. (B) Kaplan-Meier overall survival curves of three subtypes in TCGA dataset (log-rank test). (C, D) Subtype reproducibility across fifteen independent validation cohorts using NTP algorithm.

metastatic capacity, providing new insights into CRC tumor heterogeneity and progression.

#### 2.2 Cross-cohort multi-omics analysis identifies novel clusters driven by VEGFA +TC signatures in colorectal cancer

Using VEGFA+TC signature genes, we performed unsupervised clustering analysis on colorectal cancer patients from the TCGA database. We examined clustering results across different k values

(k=2-9). Optimal separation was achieved at k=3 (Figure 2A). Kaplan-Meier analysis revealed significant survival differences among the three subtypes, with C3 showing the worst prognosis (Figure 2B). The clustering results of the three subtypes remained stable in the TCGA database (Figure 2C). Furthermore, we extended the subtype classification using the NTP algorithm. Results demonstrated the stable existence of these three subtypes across multiple cohorts (GSE12945, GSE13067, GSE13294, GSE14333, GSE161158, GSE17536, GSE18088, GSE19862, GSE25071, GSE26682, GSE28702, GSE28722, GSE29621, GSE39582, GSE41258) (Figure 2D).



# 2.3 Significant differences in biological functions among the three subtypes

To characterize the biological behavior of these subtypes, we calculated subtype-specific biological processes using ssGSEA (Figure 3A). Subtype C1 exhibited suppression of Wnt target genes, along with enrichment in lipid metabolism (arachidonic acid production, sphingolipid catabolism) and immune regulation (negative regulation of Notch signaling, regulation of Th2 cell

differentiation). Subtype C2 was associated with mitochondrial transcription and metabolic processes (alditol phosphate metabolism, glycerol-3-phosphate metabolism), as well as meiotic cell cycle and germ cell development, suggesting connections to reproductive biology. Subtype C3 showed enrichment in cell junction organization and adhesion-dependent processes (matrix adhesion-dependent cell spreading, positive regulation of endothelial cell migration), indicating a role in cell motility. We next quantified transcription factor activity across the three



scores (E), and stromal scores (F) among three subtypes in TCGA-CRC dataset.

subtypes (Figure 3B). Results showed subtype C1 highly expressed transcription factors including ZNF639, SIX5, IRF3, and MAZ, suggesting associations with development and metabolism. Subtype C2 showed high expression of MNT, SNAI2, and MXI1, indicating associations with cell proliferation and cell cycle. Subtype C3 exhibited high expression of TEAD1, STAT2, HIF1A, MAFK, EBF1, and TCF12, suggesting involvement in cell migration. Additionally, PROGENy analysis (Figure 3C) revealed significant activation of MAPK, NF-KB and TNFa signaling pathways in subtype C3, while hypoxia and PI3K pathways were suppressed. This suggests subtype C3 may acquire metastatic advantages by remodeling signaling networks - utilizing inflammatory pathways  $(NF-\kappa B/TNF\alpha)$  to drive invasion while downregulating metabolism-adaptation pathways (PI3K) to maintain energy balance during metastasis. This unique "inflammatory activationmetabolic suppression" pattern further supports the highly metastatic nature of subtype C3.

# 2.4 Immunological landscape of the three subtypes

Our study systematically characterized the heterogeneity of immune microenvironments across the three CRC subtypes (C1-C3). Quantitative analysis of immune cell infiltration using ssGSEA revealed that the C3 subtype exhibited the most extensive immune cell infiltration, including key antitumor effector populations such as activated CD8+ T cells and CD4+ T cells (Figure 4A). Notably, this subtype concurrently overexpressed multiple immunomodulatory molecules, including antigen-presenting proteins (HLA-A/B/C), co-stimulatory molecules (CD28/CD80), and immune checkpoint proteins (PD-L1) (Figure 4B). However, the simultaneous enrichment of regulatory T cells, neutrophils, and monocyte-derived macrophages established a characteristic "hot but suppressed" immune microenvironment in C3. ESTIMATE algorithm analysis further confirmed these findings (Figures 4C-F), demonstrating that C3 possessed the highest immune and stromal scores but the lowest tumor purity. In contrast, the C2 subtype displayed a typical "immune-desert" phenotype with maximal tumor purity and minimal immune infiltration, while C1 exhibited transitional immune characteristics. These results not only validate the complexity of CRC immune microenvironments but, more importantly, reveal the unique coexistence of immune activation and suppression mechanisms in the C3 subtype. This finding provides novel insights into the biological basis of immunotherapy resistance in colorectal cancer.

## 2.5 HOXC6 serves as a marker molecule for C3 and correlates with poor prognosis

To elucidate the molecular characteristics and clinical significance of the C3 subtype with the worst prognosis, we identified three genes (HOXC6, LAMP5, and SPOCK1) through intersection analysis of C3 signature genes across TCGA, GSE12945, and GSE17536 cohorts (Figure 5A). Subsequent survival analyses consistently identified HOXC6, LAMP5, and SPOCK1 as significant prognostic risk factors. Kaplan-Meier analysis demonstrated significantly shorter median overall survival in patients with high expression of these genes (Figure 5B). The homeobox (HOX) genes encode a group of transcription factors that bind DNA and regulate gene expression, functioning as either activators or repressors (22, 23). In cancer biology, these genes frequently exhibit mutations or elevated expression levels, where they actively participate in driving malignant transformation and tumor advancement (24). Leveraging powerful platform, BEST, we conducted an in-depth analysis of HOXC6's functional expression patterns across diverse cancer types. The database's robust datasets enabled us to elucidate HOXC6's prognostic significance and potential biological roles in CRC (25).Clinical staging analysis showed a progressive increase in HOXC6 expression levels with disease advancement (Figure 5E). Analysis of anti-PD-1/PD-L1 treated patient cohorts revealed superior progression-free survival in the HOXC6 low-expression group (Figures 5F-G). Receiver operating characteristic (ROC) curve analysis demonstrated that HOXC6 expression levels achieved area under the curve values of 0.855 and 0.714 in two independent validation cohorts for predicting immunotherapy response (Figure 5H), indicating significant predictive value. Functional enrichment analysis showed HOXC6-associated genes significantly enriched in fundamental biological processes including DNA binding and secretory system regulation, particularly transcriptional activation and vesicular transport functions (Figure 5I). KEGG pathway analysis revealed HOXC6 primarily activates signal transduction pathways, including VEGF signaling and cytokine receptor interaction pathways (Figure 5J). GSEA pathway enrichment demonstrated HOXC6's association with cell junction assembly and cell adhesion (Figure 5K), findings highly consistent with C3 subtype characteristics and forming a comprehensive functional network from cell junction assembly to membrane skeleton formation. Notably, HOXC6 showed strong correlation with epithelial-mesenchymal transition (EMT) programs (Figure 5L). In colorectal cancer, the HOXC6 highexpression group exhibited typical EMT features, including downregulation of epithelial markers and upregulation of mesenchymal markers (Figure 5M). Genomic analysis revealed significant correlations between HOXC6 expression and specific mutation profiles, with APC and TP53 mutations more prevalent in the low-expression group (Figure 5N). This suggests HOXC6's potential role in regulating genomic stability, consistent with the loss of function of the APC tumor suppressor in CRC, which is associated with poor prognosis (26). AND the loss of function of TP53 was reported to lead to the formation of an NF-κB dependent inflammatory microenvironment and triggers EMT (27). These findings collectively establish HOXC6 as a key regulatory molecule for the C3 subtype, demonstrating not only prognostic value but also providing a potential precision therapeutic target for this patient subgroup.

# 2.6 HOXC6 knockdown suppresses malignant behaviors of colorectal cancer cells *in vitro* and *in vivo*

To investigate the effects of HOXC6 on colorectal cancer cell behaviors including proliferation, migration, and invasion, we established HOXC6-knockdown colorectal cancer cell lines. HOXC6 mRNA expression was significantly reduced in these cell lines (Figure 6A). CCK-8 assays performed 24 hours posttransfection demonstrated that the proliferation rate was significantly decreased in the sh-HOXC6 group compared to the sh-NC group (Figure 6C). Similarly, colony formation assays revealed a marked reduction in clonogenic potential in the sh-HOXC6 group relative to the sh-NC group (Figure 6D). These findings collectively indicate that HOXC6 knockdown suppresses colorectal cancer cell proliferation. Wound healing assays (Figure 6E) and Trans-well migration assays (Figure 6F) showed that the migration rate was significantly lower in sh-HOXC6 cells compared to sh-NC cells, demonstrating that HOXC6 downregulation inhibits colorectal cancer cell migratory capacity. Furthermore, Trans-well invasion assays conducted with these cell lines (Figure 6G) revealed a significant decrease in the number of cells penetrating through the membrane in the sh-HOXC6 group versus the sh-NC group, indicating that HOXC6 knockdown impairs the invasive capability of colorectal cancer cells. Quantitative analysis of tumor volumes at day 20 post subcutaneous implantation showed significantly reduced tumor sizes in the sh-HOXC6 group compared to the sh-NC group.

## 2.7 Abemaciclib targets HOXC6 and effectively inhibits tumor proliferation

Through an integrated approach combining computational drug screening and experimental validation, we systematically evaluated the therapeutic potential of HOXC6 in colorectal



#### FIGURE 5

HOXC6 promotes poor prognosis in CRC patients (A) Venn diagram of prognosis-associated genes across three cohorts. (B) Kaplan-Meier overall survival curves for HOXC6, LAMP5 and SPOCK1 in TCGA-CRC and two validation cohorts (log-rank test). (C) Association between HOXC6 expression and CRC clinical stages. (D) Progression-free survival in anti-PD-1/PD-L1 treated patients stratified by HOXC6 expression. (E) ROC analysis of HOXC6 expression for predicting immunotherapy response (AUC values shown). (F) GO functional analysis of HOXC6-associated biological processes. (G) KEGG pathway enrichment analysis. (H) GSEA pathway activation analysis. (I) Biological processes associated with HOXC6 in CRC. (J) Correlation between HOXC6 and EMT pathway in CRC. (K) Genomic analysis of HOXC6 expression and mutation profiles. (L) Pathway activity analysis of patients with high HOXC6 expression. (M) GSEA analysis of EMT pathway in patients with high HOXC6 expression. (N) Correlation analysis between HOXC6 expression level and gene mutation status.



cancer. Initial virtual screening of 2,110 FDA-approved compounds using molecular docking identified abemaciclib as exhibiting optimal binding characteristics based on free energy calculations (Figure 7A). Detailed docking analysis revealed that abemaciclib forms a stable interaction network within the HOXC6 active pocket, including hydrogen bonds with TYR-148 and GLN-146 residues (Figures 7B-C). In vitro functional assays confirmed abemaciclib's therapeutic potential. Treatment with abemaciclib significantly downregulated HOXC6 mRNA expression in colorectal cancer cell lines (Figure 7D). Systematic cellular functional experiments demonstrated abemaciclib's targeted inhibition of HOXC6 and its antitumor effects. CCK-8 assays showed significantly reduced cell proliferation 24 hours post-treatment (Figure 7E), while colony formation assays (Figure 7F) confirmed abemaciclib's ability to markedly decrease colony numbers, indicating effective inhibition of tumor cell proliferation through HOXC6 targeting. Migration capacity assessments, including wound healing (Figure 7G), Transwell migration (Figure 7H), and invasion assays (Figure 7I), consistently demonstrated significantly reduced cell migration rates and membrane-penetrating cell numbers in abemaciclib-treated groups, confirming the drug's ability to effectively block HOXC6mediated metastatic potential. In nude mouse subcutaneous tumor models (Figure 7J), abemaciclib-treated groups showed significantly smaller tumor volumes compared to controls. These multi-level results, from molecular interactions to cellular phenotypes, collectively demonstrate that abemaciclib specifically targets HOXC6, forms stable complexes, and significantly

inhibits colorectal cancer cell proliferation, migration, and invasion capabilities.

## **3** Discussion

Colorectal cancer, as one of the most heterogeneous malignancies in the digestive system, presents significant clinical challenges due to variations in treatment response and prognosis (28). CRC pathogenesis stems from the dysregulation of complex signaling networks that orchestrate key cellular functions such as proliferation, survival, and apoptosis (29). This aberrant signaling leads to the malignant transformation of the colonic epithelium. This study systematically characterized CRC tumor cell heterogeneity through single-cell RNA sequencing, leading to the first identification of the VEGFA+TC subpopulation with EMT activation and high metastatic potential. Based on its molecular features, we classified CRC patients into three stable subtypes. This classification demonstrated high reproducibility across multiple independent cohorts (TCGA and various GEO datasets), with the C3 subtype showing the highest EMT activity and poorest clinical outcomes, highlighting the clinical predictive value of this stratification system.

Multi-omics data and functional experiments systematically revealed the central regulatory role of HOXC6 in specific CRC subtypes, providing novel insights into tumor heterogeneity. Our findings demonstrate that HOXC6 not only serves as a marker molecule for the aggressive C3 subtype, but also drives disease



progression through multiple mechanisms including EMT program regulation, metabolic reprogramming, and immune microenvironment remodeling. Previous study identified TGF-B signaling exhibiting a dual role in CRC by suppressing early tumorigenesis but promoting EMT-driven metastasis in advanced stages through epigenetic dysregulation of its pathway components (30). Validation studies confirmed that HOXC6 is closely associated with EMT pathway activation and directly regulates CRC cell proliferation, migration, and invasion capabilities. Both in vitro and in vivo functional experiments showed that HOXC6 knockdown significantly inhibited tumor growth and metastasis, confirming its oncogenic role. Additionally, patients with high HOXC6 expression showed poorer response to PD-1 inhibitors, while those with low HOXC6 expression exhibited higher APC/ TP53 mutation burden, suggesting HOXC6 may promote tumor progression through mutation-independent transcriptional regulatory mechanisms. These findings not only fill the knowledge gap regarding molecular mechanisms of EMT and immune microenvironment crosstalk, but also provide novel biomarkers (e.g., HOXC6) and potential therapeutic targets for precision classification and targeted intervention in CRC.

Although the role of HOX genes in immunotherapy resistance has not been directly studied, their involvement in anti-angiogenic therapy resistance suggests a possible connection. HOX genes, such as HOXB9, promote tumor aggressiveness by upregulating proangiogenic and pro-inflammatory factors, which can also contribute to an immunosuppressive microenvironment (31). Since inflammation and myeloid cell recruitment are key mechanisms of resistance to both anti-angiogenic and immunotherapies, targeting HOX genes may help overcome treatment resistance. Functional studies further suggested that elevated expression of HOXB13 is related to ER downregulation and, consequently, to TAM-resistance in ER+ cancers (32).Further research should explore whether HOX family members influence immune evasion and response to checkpoint inhibitors.

Research across multiple cancer types has established that aberrant overexpression of transcription factors serves as a key driver of both primary tumor growth and distant metastasis (33– 35). The homeobox protein HOXC6 represents one such factor that exhibits cancer-specific upregulation and functionally contributes to malignant cell expansion. HOXC6 promotes cervical cancer progression by transcriptionally upregulating BCL2 to enhance anti-apoptotic effects and drive tumor cell proliferation (36) In prostate cancer, HOXC6 promotes tumor cell survival by transcriptionally repressing proapoptotic factors NEP and IGFBP-3, and its knockdown induces caspase-dependent apoptosis in both androgen-dependent and -independent cancer cells (37). However, in CRC, key questions remain regarding HOXC6's expression patterns, regulatory networks, and mechanistic contributions to EMT. Our single-cell resolution study not only confirms HOXC6's conserved oncogenic role across tumors but also establishes EMT's pivotal position in CRC metastasis. Elucidating HOXC6's EMTregulatory mechanisms will advance understanding of tumor heterogeneity and may yield novel molecular classification systems and therapeutic targets for precision medicine.

The discovery of abemaciclib's activity against HOXC6 represents a breakthrough in targeting transcription factors (TFs), traditionally considered "undruggable" due to their lack of defined binding pockets (38, 39). Our molecular dynamics simulations first identified an allosteric site within HOXC6's DNA-binding domain. Mechanistically, abemaciclib stabilizes HOXC6's inactive conformation, disrupting coactivator recruitment. Abemaciclib is an oral CDK4/6 inhibitor with preferential activity against CDK4 (40). Its favorable safety profile, including lower rates of myelosuppression, permits continuous dosing unlike other approved agents in this class (41, 42). Emerging preclinical synergy between cell cycle inhibitors (CDK4/6i) and either PI3K pathway blockers or immunotherapies is now being translated into clinical trial designs (43). Our studies revealed that while HOXC6high tumors respond poorly to PD-1 inhibitors, they exhibit marked sensitivity to abemaciclib-potentially through drugmediated immune microenvironment remodeling. Immune checkpoint inhibitors targeting the PD-1/PD-L1 axis have gained monotherapy approval for advanced colorectal cancers exhibiting mismatch repair deficiency (dMMR) and/or high microsatellite instability (MSI-H) (44-46). For high-risk C3-subtype patients, combination strategies incorporating HOXC6 targeting may be essential. Thus, combining Abemaciclib with immunotherapy could improve treatment sensitivity in HOXC6-high colorectal cancer patients.

In conclusion, the C3 subtype represents a clinically relevant EMT population whose characterization advances CRC heterogeneity understanding and precision therapy development. Future efforts should focus on translating these discoveries into clinical practice to improve outcomes. Our integrated computational-structural-preclinical approach not only establishes HOXC6 as a novel therapeutic target but also pioneers innovative TF-targeting paradigms, potentially ushering in a new era of molecular subtype-directed CRC treatment.

#### 4 Methods

#### 4.1 Data collection

For this study, we obtained 15 colorectal cancer cohort datasets from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases (GSE12945, GSE13067, GSE13294, GSE14333, GSE161158, GSE17536, GSE18088, GSE19862, GSE25071, GSE26682, GSE28702, GSE28722, GSE29621, GSE39582, and GSE41258). The RNA-seq count data from TCGA-CRC were converted to transcripts per million (TPM) and subsequently log2-transformed. Each dataset was matched with its respective sequencing platform. In cases where a gene was represented by multiple probes, the gene expression values were averaged. Subsequently, data were normalized in R (version 4.4.0) using the "normalizeBetweenArrays" function from the "limma" package (version 3.48.1). Each dataset was analyzed independently, thereby mitigating batch effects.

## 4.2 Single-cell RNA sequencing data analysis

Initial data processing was performed using Seurat (v4.0.4), where raw data from five CRC single-cell cohorts were imported via the Read10X function and converted to sparse matrix format (dgCMatrix). Datasets were integrated using the merge function with unique cell identifiers ensured by RenameCells. Quality control included: 1) Scrublet-based doublet removal; 2) filtration of low-quality cells (<100 genes detected); 3) exclusion of genes expressed in <3 cells. Data normalization employed LogNormalize scaling (factor=10,000) with identification of top 2,000 highly variable genes. Technical variation was regressed out using ScaleData (covariates: UMI counts and mitochondrial gene content). Dimensionality reduction (top 30 PCs) was followed by Harmony-based batch correction (47). UMAP visualization and Louvain clustering (resolution=0.6) were performed, with optimal clustering determined via clustree. Cell type annotation utilized established marker genes and literature references. Epithelial subpopulations were re-analyzed through repeated dimensionality reduction/clustering (resolution=0.2), with functional annotation based on differentially expressed genes.

#### 4.3 Functional annotation analysis

Functional annotation analysis was performed using ClusterProfiler (v4.4.4) (48). The analysis workflow comprised: (1) gene ID conversion from Ensembl IDs to standard gene names via the bitr function; (2) Gene Ontology (GO) enrichment analysis using enrichGO, covering Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) categories; (3) KEGG pathway analysis via enrichKEGG. Statistical significance was defined as P<0.05 with Benjamini-Hochberg multiple testing correction applied to all results.

#### 4.4 Unsupervised clustering

To elucidate tumor heterogeneity in gastric cancer patients, we employed a sample clustering method. The clustering was based on the signature gene of the VEGFR+TC subtypes. Specifically, sample clustering analysis was conducted using the ConsensusClusterPlus (49)package in R. The maximum number of clusters (K) was set to 9, with a step size of 1, to evaluate clustering results at different K values. To ensure the robustness of the clustering results, the clustering analysis was repeated 1,000 times. Pltem was used as a feature retention threshold, determining the number of features (genes) retained for clustering analysis in each bootstrap iteration. In our setup, Pltem was set to 0.8, meaning that 80% of the features were retained in each bootstrap iteration. The pFeature parameter indicated the probability of feature retention, determining whether each feature would be retained in each bootstrap iteration. In this case, pFeature was set to 1, implying that all features were retained without feature selection. This setup is suitable for scenarios where all features need to be considered, or when the number of features is relatively small. The clustering process employed the k-means clustering algorithm (clusterAlg = "km") and Euclidean distance (distance = "Euclidean"). The results were visualized using the "ggplot2" package to assess potential algorithmic biases in the ConsensusClusterPlus clustering method.

#### 4.5 Single-sample gene set enrichment analysis

We used the ssGSEA algorithm to study the enrichment of differential gene sets across different samples. ssGSEA is a computational method designed for single-sample gene set enrichment analysis, providing insights into the activity levels of differential genes in individual samples. The differential gene sets used in this study were derived from research by the Alexander Bagaev team. We implemented ssGSEA using the "GSVA" package in R and the "gsva()" function(Bagaev et al., 2021). In this analysis, the method parameter was set to "ssgsea," and enrichment analysis was performed using Gaussian kernel density estimation (kcdf = "Gaussian"). Additionally, we retained the absolute ranking information (abs.ranking = TRUE). To normalize the ssGSEA results, the ssGSEA scores were scaled to a range of 0 to 1, which facilitates better visualization and comparison.

#### 4.6 Stability of molecular feature investigation

The recently developed Nearest Template Prediction (NTP) (30) is a flexible, single-sample-based predictive method capable of cross-platform and multi-class prediction, along with confidence

assessment. In the test dataset, NTP was applied to identify the three defined subtypes by utilizing the expression profiles of the signature genes. Samples with a False Discovery Rate (FDR) < 0.05 were considered successfully classified.

#### 4.7 Molecular docking

Molecular docking was used to predict the binding patterns and affinities between Abemaciclib (ligand) and HOXC6(target protein). In brief, we obtained the structure of Abemaciclib from the PubChem database, which was then optimized using the MM2 force field in ChemBio3D and saved in PDB format. Based on scoring function optimization, molecular docking simulations were performed, and AutoDock Vina was used to predict the binding conformations between Savolitinib and HOXC6. The protein structure database AlphaFold was used to obtain the structure of HOXC6, which was also saved in PDB format. Subsequently, molecular docking simulations were conducted with AutoDock Vina, predicting the binding conformation of Abemaciclib and HOXC6 based on scoring function optimization. The scoring function serves as an indicator of the ligand-protein affinity, estimating the binding energy and thermodynamic stability of the protein-ligand complex. Finally, the results were visualized using PyMOL software, and the binding pose of the ligand at the protein's active site was determined.

#### 4.8 Western blot

Proteins were extracted, quantified, and separated on a 12% SDS-PAGE gel, then transferred to PVDF membranes. Membranes were blocked with skim milk and incubated with $\beta$ -actin and HOXC6 antibodies, followed by a secondary antibody. Detection was performed using an imaging system after exposure treatment.

## 4.9 Quantitative reverse transcription polymerase chain reaction

To extract total RNA of tissues, we used the RNAsimple Total RNA Kit (TIANGEN). And then the reverse-transcribed utilizing with HiScript III 1st Strand cDNA Synthesis Kit (+gDNA wiper) (Vazyme), and the NovoStart<sup>®</sup> SYBR qPCR SuperMix Plus (novoprotein) was arranged for qPCR were showed in Table 1. The relative expression of HOXC6 was evaluated using the 2– $\Delta\Delta$ Ct method, with GAPDH serving as the internal reference.

TABLE 1 PCR primer sequence information.

Gene	Forward	Reverse
HOXC6	ACAGACCTCAATCGCTCAGGA	AGGGGTAAATCTGGATACTGGC
GAPDH	CCAGCAAGAGCACAAGAGGAAGAG	GGTCTACATGGCAACTGTGAGGAG

## 4.10 Cell culture and HOXC6 knockout in CT26 CRC cell line

CT26 CRC cells were cultured in RPMI1640 medium supplemented with 10% fetal bovine serum (FBS) at 37°C and 5% CO2, with the medium replaced every 3 days. Cells from stable passages in the logarithmic growth phase were used for the experiments. In the knockout group, cells were infected with shRNA lentiviral particles targeting HOXC6, while the control group received lentiviral particles containing a scramble sequence.

#### 4.11 CCK-8 assay

In the CCK-8 assay, CT26 cells were seeded at a density of  $2 \times 10^3$  cells per well in a 96-well plate and cultured overnight. Subsequently, 20  $\mu$ L of CCK-8 solution (5 mg/mL) was added to each well and incubated for 4 hours. The optical density at 450 nm was measured using a microplate reader.

#### 4.12 Wound healing assay

The wound healing assay was used to evaluate cell migration in CT26 cells. After the cells reached 90% confluence, a wound was made, and the wells were washed three times with PBS. Fresh serum-free medium was added, and wound closure was monitored at 0 hours and 24 hours post-wounding using an Olympus X71 microscope. The distance between the cells was measured using ImageJ. For statistical accuracy, the assay was repeated three times.

## 4.13 Colony formation assay

CT26 cells were cultured in 6-well plates under standard conditions at 37°C for two weeks. After incubation, the plates were washed with cold phosphate-buffered saline (PBS) to remove non-adherent cells. The colonies were then fixed with 4% paraformaldehyde for 15 minutes and stained with 0.1% crystal violet at room temperature. Colony formation was imaged and quantified using an optical microscope.

#### 4.14 Transwell assay

The Transwell assay was used to assess cell migration. CT26 cells  $(4\times10^4)$  were placed in the upper chamber with medium containing 5% FBS, while 500  $\mu$ L of medium was added to the lower chamber. After 24 hours, non-invading cells were removed, and the remaining cells were fixed with paraformaldehyde and stained with crystal violet. The number of invading cells was quantified in five random fields under a microscope.

# 4.15 Subcutaneous tumor formation experiment in mice

We conducted a subcutaneous tumor formation experiment in nude mice using CT26 cell lines with sh-NC and sh-HOXC6. First, the cells were collected by centrifugation and suspended in PBS to prepare a cell suspension, which was counted and adjusted to a concentration of 107 cells/mL. The suspension was then kept on ice and mixed thoroughly before injection. Next, the right side of the rib cage of the nude mice was disinfected, and 0.2 mL of the cell suspension was injected using a 1 mL syringe. The nude mice were housed separately according to different experimental groups, and their initial body weights were recorded on the day of injection. From the day of injection, the body weight of the nude mice and the length and width of the subcutaneous tumor were recorded every 2 days for a total observation period of 15 days. On day 15 of the experiment, we recorded the body weight and tumor size again, after which the mice were euthanized, and photographs were taken for documentation. Finally, the nude mice were dissected, and the tumor tissues were extracted for weighing and photography for subsequent analysis.

## 4.16 *In vivo* antitumor experiment with savolitinib

We conducted a subcutaneous tumor formation experiment in nude mice using the CT26 cell line, following the same steps as described above. Three days prior to tumor implantation and every three days thereafter, Savolitinib and its control DMSO were injected around the tumor, with a dosage of 10 mg/kg.

## 4.17 Statistical analysis

All data processing, visualization, and statistical analyses were performed using R software. The correlation between two continuous variables was assessed using Spearman's correlation coefficient. Initially, normality tests were performed on the datasets. For normally distributed data with equal variances, Student's t-test and one-way analysis of variance (ANOVA) were used to compare differences between two or more groups. For nonnormally distributed data or data that did not meet the homogeneity of variance assumption, the Wilcoxon test and Kruskal-Wallis test were used for comparisons within two or more groups, respectively. Chi-square tests were used to analyze categorical variables. p-values were adjusted for false discovery rate (FDR), particularly when performing multiple pairwise comparisons. To validate the results, the experiments were performed in triplicate to confirm their reproducibility.

#### 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 author.

#### **Ethics statement**

The animal study was approved by Laboratory Animal Ethics Committee of the Academic Committee of the China-Japan Friendship Hospital. The study was conducted in accordance with the local legislation and institutional requirements.

#### Author contributions

WN: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. WJ: Conceptualization, Data curation, Investigation, Software, Writing – review & editing. JN: Investigation, Methodology, Writing – original draft, Writing – review & editing. LeZ: Conceptualization, Investigation, Methodology, Writing – review & editing. ZL: Formal Analysis, Investigation, Writing – review & editing. LiZ: Investigation, Methodology, Writing – review & editing. XS: Data curation, Investigation, Methodology, Project administration, Software, Supervision, Validation, Writing – original draft, Writing – review & editing.

#### References

1. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. Lancet. (2019) 394:1467-80. doi: 10.1016/s0140-6736(19)32319-0

2. Biller LH, Schrag D. Diagnosis and treatment of metastatic colorectal cancer: A review. Jama. (2021) 325:669–85. doi: 10.1001/jama.2021.0106

3. Vodenkova S, Buchler T, Cervena K, Veskrnova V, Vodicka P, Vymetalkova V, et al. 5-fluorouracil and other fluoropyrimidines in colorectal cancer: Past, present and future. *Pharmacol Ther.* (2020) 206:107447. doi: 10.1016/j.pharmthera.2019.107447

4. Xie YH, Chen YX, Fang JY. Comprehensive review of targeted therapy for colorectal cancer. *Signal Transduct Target Ther.* (2020) 5(1):22. doi: 10.1038/s41392-020-0116-z

5. Modest DP, Pant S, Sartore-Bianchi A. Treatment sequencing in metastatic colorectal cancer. Eur J Cancer. (2019) 109:70-83. doi: 10.1016/j.ejca.2018.12.019

6. Cervantes A, Adam R, Roselló S, Arnold D, Normanno N, Taïeb J, et al. Metastatic colorectal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol.* (2023) 34:10–32. doi: 10.1016/j.annonc.2022.10.003

7. Abedizadeh R, Majidi F, Khorasani HR, Abedi H, Sabour D. Colorectal cancer: a comprehensive review of carcinogenesis, diagnosis, and novel strategies for classified treatments. *Cancer Metastasis Rev.* (2024) 43:729–53. doi: 10.1007/s10555-023-10158-3

8. Cañellas-Socias A, Sancho E, Batlle E. Mechanisms of metastatic colorectal cancer. . Nat Rev Gastroenterol Hepatol. (2024) 21:609–25. doi: 10.1038/s41575-024-00934-z

9. Wen R, Zhou L, Peng Z, Fan H, Zhang T, Jia H, et al. Single-cell sequencing technology in colorectal cancer: a new technology to disclose the tumor heterogeneity and target precise treatment. *Front Immunol.* (2023) 14:1175343. doi: 10.3389/fimmu.2023.1175343

10. Mathonnet M, Perraud A, Christou N, Akil H, Melin C, Battu S, et al. Hallmarks in colorectal cancer: angiogenesis and cancer stem-like cells. *World J Gastroenterol.* (2014) 20:4189–96. doi: 10.3748/wjg.v20.i15.4189

### Funding

The author(s) declare that no financial support was received for the research and/or publication of 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.

11. Ang HL, Mohan CD, Shanmugam MK, Leong HC, Makvandi P, Rangappa KS, et al. Mechanism of epithelial-mesenchymal transition in cancer and its regulation by natural compounds. *Med Res Rev.* (2023) 43:1141–200. doi: 10.1002/med.21948

12. Chakraborty S, Mir KB, Seligson ND, Nayak D, Kumar R, Goswami A, et al. Integration of EMT and cellular survival instincts in reprogramming of programmed cell death to anastasis. *Cancer Metastasis Rev.* (2020) 39:553–66. doi: 10.1007/s10555-020-09866-x

13. Sánchez-Tilló E, Liu Y, de Barrios O, Siles L, Fanlo L, Cuatrecasas M, et al. EMTactivating transcription factors in cancer: beyond EMT and tumor invasiveness. *Cell Mol Life Sci.* (2012) 69:3429–56. doi: 10.1007/s00018-012-1122-2

14. Jiang Y, Zhan H. Communication between EMT and PD-L1 signaling: New insights into tumor immune evasion. *Cancer Lett.* (2020) 468:72-81. doi: 10.1016/j.canlet.2019.10.013

15. He Y, Zhu M, Lai X, Zhang H, Jiang W. The roles of PD-L1 in the various stages of tumor metastasis. *Cancer Metastasis Rev.* (2024) 43:1475–88. doi: 10.1007/s10555-024-10189-4

16. Radhakrishnan K, Truong L, Carmichael CL. An "unexpected" role for EMT transcription factors in hematological development and Malignancy. *Front Immunol.* (2023) 14:1207360. doi: 10.3389/fimmu.2023.1207360

17. Caramel J, Ligier M, Puisieux A. Pleiotropic roles for ZEB1 in cancer. *Cancer Res.* (2018) 78:30–5. doi: 10.1158/0008-5472.Can-17-2476

18. Morrison CD, Parvani JG, Schiemann WP. The relevance of the TGF- $\beta$  Paradox to EMT-MET programs. *Cancer Lett.* (2013) 341:30–40. doi: 10.1016/j.canlet.2013.02.048

19. Sun L, Xing J, Zhou X, Song X, Gao S. Wnt/β-catenin signalling, epithelialmesenchymal transition and crosslink signalling in colorectal cancer cells. *BioMed Pharmacother*. (2024) 175:116685. doi: 10.1016/j.biopha.2024.116685 20. Suarez-Carmona M, Lesage J, Cataldo D, Gilles C. EMT and inflammation: inseparable actors of cancer progression. *Mol Oncol.* (2017) 11:805-823 (2017). doi: 10.1002/1878-0261.12095

21. Yang J, Antin P, Berx G, Blanpain C, Brabletz T, Bronner M, et al. Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol.* (2020) 21:341–52. doi: 10.1038/s41580-020-0237-9

22. Shah N, Sukumar S. The Hox genes and their roles in oncogenesis. Nat Rev Cancer. (2010) 10:361-71. doi: 10.1038/nrc2826

23. Li B, Huang Q, Wei GH. The role of HOX transcription factors in cancer predisposition and progression. *Cancers (Basel)*. (2019) 11(4). doi: 10.3390/ cancers11040528

24. Malvi P, Chava S, Cai G, Hu K, Zhu LJ, Edwards YJK, et al. HOXC6 drives a therapeutically targeta ble pancreatic cancer growth and metastasis pathway by regulating MSK1 and PPP2R2B. *Cell Rep Med.* (2023) 4:101285. doi: 10.1016/j.xcrm.2023.101285

25. Liu Z, Liu L, Weng S, Xu H, Xing Z, Ren Y, et al. BEST: a web application for comprehensive biomarker exploration on large-scale data in solid tumors. *J Big Data*. (2023) 10:165. doi: 10.1186/s40537-023-00844-y

26. Janssen KP, Alberici P, Fsihi H, Gaspar C, Breukel C, Franken P, et al. APC and oncogenic KRAS are synergistic in enhancing Wnt signaling in intestinal tumor formation and progression. *Gastroenterology*. (2006) 131:1096–109. doi: 10.1053/j.gastro.2006.08.011

27. Schwitalla S, Ziegler PK, Horst D, Becker V, Kerle I, Begus-Nahrmann Y, et al. Loss of p53 in enterocytes generates an inflammatory microenvironment enabling invasion and lymph node metastasis of carcinogen-induced colorectal tumors. *Cancer Cell.* (2013) 23:93–106. doi: 10.1016/j.ccr.2012.11.014

28. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* (2024) 74:229–63. doi: 10.3322/caac.21834

29. Li Q, Geng S, Luo H, Wang W, Mo YQ, Luo Q, et al. Signaling pathways involved in colorectal cancer: pathogenesis and targeted therapy. *Signal Transduct Target Ther.* (2024) 9:266. doi: 10.1038/s41392-024-01953-7

30. Hao Y, Baker D, Ten Dijke P. TGF-β-mediated epithelial-mesenchymal transition and cancer metastasis. *Int J Mol Sci.* (2019) 20(11). doi: 10.3390/ ijms20112767

31. Contarelli S, Fedele V, Melisi D. HOX genes family and cancer: A novel role for homeobox B9 in the resistance to anti-angiogenic therapies. *Cancers (Basel).* (2020) 12 (11). doi: 10.3390/cancers12113299

32. Shah N, Jin K, Cruz LA, Park S, Sadik H, Cho S, et al. HOXB13 mediates tamoxifen resistance and invasiveness in human breast cancer by suppressing ER $\alpha$  and inducing IL-6 expression. *Cancer Res.* (2013) 73:5449–58. doi: 10.1158/0008-5472.Can-13-1178

33. Bradner JE, Hnisz D, Young RA. Transcriptional addiction in cancer. Cell. (2017) 168:629-43. doi: 10.1016/j.cell.2016.12.013

34. Saitoh M. Transcriptional regulation of EMT transcription factors in cancer. Semin Cancer Biol. (2023) 97:21–9. doi: 10.1016/j.semcancer.2023.10.001

35. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol.* (2014) 15:178–96. doi: 10.1038/nrm3758

36. Wang Y, Wang C, Liu N, Hou J, Xiao W, Wang H, et al. HOXC6 promotes cervical cancer progression via regulation of Bcl-2. *FASEB J.* (2019) 33:3901–11. doi: 10.1096/fj.201801099RR

37. Ramachandran S, Liu P, Young AN, Yin-Goen Q, Lim SD, Laycock N, et al. Loss of HOXC6 expression induces apoptosis in prostate cancer cells. *Oncogene*. (2005) 24:188–98. doi: 10.1038/sj.onc.1207906

38. Bushweller JH. Targeting transcription factors in cancer - from undruggable to reality. *Nat Rev Cancer.* (2019) 19:611–24. doi: 10.1038/s41568-019-0196-7

39. Henley MJ, Koehler AN. Advances in targeting 'undruggable' transcription factors with small molecules. *Nat Rev Drug Discov.* (2021) 20:669–88. doi: 10.1038/ s41573-021-00199-0

40. Ettl T, Schulz D, Bauer RJ. The renaissance of cyclin dependent kinase inhibitors. *Cancers (Basel).* (2022) 14(2). doi: 10.3390/cancers14020293

41. Wander SA, N, Litchfield LM, D, Morato Guimaraes C, Slamon DJ, et al. Targeting CDK4 and 6 in cancer therapy: emerging preclinical insights related to abemaciclib. *Oncologist.* (2022) 27:811–21. doi: 10.1093/oncolo/oyac138

42. Spring LM, Wander SA, Andre F, Moy B, Turner NC, Bardia A, et al. Cyclindependent kinase 4 and 6 inhibitors for hormone receptor-positive breast cancer: past, present, and future. *Lancet*. (2020) 395:817–27. doi: 10.1016/s0140-6736(20) 30165-3

43. Morrison L, Loibl S, Turner NC. The CDK4/6 inhibitor revolution — a gamechanging era for breast cancer treatment. *Nat Rev Clin Oncol.* (2024) 21:89–105. doi: 10.1038/s41571-023-00840-4

44. Lin KX, Istl AC, Quan D, Skaro A, Tang E, Zheng X, et al. PD-1 and PD-L1 inhibitors in cold colorectal cancer: challenges and strategies. *Cancer Immunol Immunother*. (2023) 72:3875–93. doi: 10.1007/s00262-023-03520-5

45. Weng J, Li S, Zhu Z, Liu Q, Zhang R, Yang Y, et al. Exploring immunotherapy in colorectal cancer. *J Hematol Oncol.* (2022) 15:95. doi: 10.1186/s13045-022-01294-4

46. Payandeh Z, Khalili S, Somi MH, Mard-Soltani M, Baghbanzadeh A, Hajiasgharzadeh K, et al. PD-1/PD-L1-dependent immune response in colorectal cancer. *J Cell Physiol*. (2020) 235:5461–75. doi: 10.1002/jcp.29494

47. 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

48. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*. (2021) 2:100141. doi: 10.1016/j.xinn.2021.100141

49. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. . *Bioinf.* (2010) 26:1572–3. doi: 10.1093/ bioinformatics/btq170