Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Fibroblast-Associated Stromal–Immune Exclusion in Colorectal Cancer: Prognostic Modeling, Spatial Characterization, and Functional Investigation of the KLF2–CD73 Axis

Yinyin Pang
Hainan Medical University
Qianyuan Zhang
Hainan Medical University
Longxi Xia
Hainan Medical University
Xiang Wang
Hainan Medical University
Yunfu Lv
Hainan Medical University
Fan Jiang
Hainan Medical University
Ning Liu
liuning@muhn.edu.cn
Hainan Medical University

Research Article

Keywords: Colorectal cancer, Cancer-associated fibroblasts, Stromal-immune exclusion, Prognostic risk model, KLF2–CD73 axis

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11670662/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Fibroblast-Associated Stromal–Immune Exclusion in Colorectal Cancer: Prognostic Modeling, Spatial Characterization, and Functional Investigation of the KLF2–CD73 Axis

Yinyin Pang[1], Qianyuan Zhang[1], Longxi Xia[1], Xiang Wang[1], Yunfu Lv[1], Fan Jiang[2 *]

Ning Liu[1 *]

[1] Department of Gastrointestinal Surgery, Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University, Haikou, Hainan 570311, P.R. China

[2] Department of the Center of Gerontology, Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University, Haikou, Hainan 570311, P.R. China

* Corresponding authors: Department of Gastrointestinal Surgery, Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University, Haikou 570311, P.R. China. Email: liuning@muhn.edu.cn; yunfu_lv@126.com;

Department of the Center of Gerontology, Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University, Haikou, Hainan 570311, P. R. China. Email: jiangfan719@hainmc.edu.cn

Abstract

Objective To establish a translational framework linking computational risk modeling, spatial immune exclusion, and therapeutic targets in colorectal cancer (CRC), and to evaluate how the cancer-associated fibroblast (CAF) Krüppel-like factor 2 (KLF2)- cluster of differentiation 73 (CD73) pathway may contribute to an immunosuppressive microenvironment through the spatial and functional suppression of $CD8^+$ T cells.

Methods Five independent transcriptomic cohorts were integrated. Using the GSE39582 discovery cohort, 45 machine learning algorithms were compared via nested cross-validation to construct the CAF stromal-immune exclusion risk (CMIR) score, which was validated for overall survival (OS) prediction in independent external cohorts (GSE17536, TCGA-COAD). Single-cell transcriptomic data (GSE178341) and multiplex immunofluorescence (mIF) staining (evaluating a 4-protein mIF-based CMIR proxy score representing key model components) were combined to quantify in situ cell infiltration density and spatial physical distances. In primary fibroblasts, lentivirus-mediated KLF2 knockdown and overexpression, pharmacological CD73 inhibition, and ternary co-cultures were used to examine $CD8^+$ T-cell cytotoxic activity.

Results A 17-gene CMIR score was constructed with the optimal LASSO algorithm and simplified into an 8-gene parsimonious model. In all OS cohorts, patients with high CMIR scores had significantly shorter overall survival than those with low scores (all $P < 0.05$), and combining the CMIR score with tumor-node-metastasis (TNM) stage significantly improved the concordance index and clinical net benefit of

survival prediction. Furthermore, multi-omics profiling revealed that a high CMIR score was significantly correlated with immune-excluded microenvironment features commonly observed in immunotherapy-unresponsive tumors. In situ multiplex immunofluorescence demonstrated that high CMIR scores were associated with reduced $CD8^+$ T-cell infiltration, increased physical distance between T cells and tumor parenchyma, and elevated KLF2/CD73 expression. Mechanistically, KLF2 promoted CD73 expression in CAFs to suppress $CD8^+$ T-cell cytotoxicity, which was effectively reversed by a small-molecule CD73 inhibitor.

Conclusion This study establishes a translational framework connecting prognostic risk stratification, histological spatial immune exclusion, and actionable molecular targets, while an mIF-based 4-protein proxy score recapitulates these spatial features at the tissue level. The KLF2-CD73-CAF axis represents a rational target to modulate stromal immunosuppression and enhance the efficacy of immunotherapy in colorectal cancer.

Keywords: Colorectal cancer; Cancer-associated fibroblasts; Stromal-immune exclusion; Prognostic risk model; KLF2–CD73 axis.

1 Introduction

Colorectal cancer (CRC) is among the leading digestive tract malignancies worldwide in both incidence and mortality [1,2]. Despite advances in surgery combined with chemotherapy and targeted therapy, the prognosis of patients with advanced, recurrent, or metastatic disease remains poor [3,4]. The cellular and stromal components of the tumor microenvironment (TME) play critical roles in regulating

tumor progression and immune evasion [5,6]. In particular, impaired infiltration of CD8-positive cytotoxic T lymphocytes ($CD8^+$ T cells) into the tumor parenchyma, termed the immune exclusion phenotype, is a key mechanism underlying the failure of antitumor immune responses [7,8].

Cancer-associated fibroblasts (CAFs), an abundant stromal cell population within the tumor stroma, establish physical and chemical barriers by depositing extracellular matrix, chemokines, and metabolites, thereby restricting the spatial localization and cytotoxic function of $CD8^+$ T cells [9-11]. Previous prognostic models have mostly relied on computational transcriptomic profiling without functional or spatial validation, remaining limited in clinical translation [12,13]. Specifically, existing risk signatures rarely capture the spatial architecture of immune exclusion or address the microenvironmental features associated with primary non-response to immune checkpoint blockade in stroma-rich tumors. In addition, while the transcription factor Krüppel-like factor 2 (KLF2) plays important regulatory roles in cell differentiation and homeostasis, its potential involvement in CAF activation and metabolic immunosuppression via the cluster of differentiation 73 (CD73) pathway warrants further investigation [14-16].

To establish a translational framework connecting computational risk modeling, in situ spatial histology, and therapeutic target identification, this study integrated multi-center transcriptomic data and applied nested cross-validation to construct and freeze the CAF stromal-immune exclusion risk (CMIR) score. Beyond prognostic evaluation, we assessed whether the CMIR score could indicate primary resistance to immune

checkpoint blockade. Single-cell RNA sequencing (scRNA-seq), multiplex immunofluorescence (mIF), and in vitro cell experiments were further combined to systematically characterize its microenvironment exclusion features at the levels of cellular origin, spatial distribution, and biological function, providing a candidate biomarker for risk stratification in CRC supported by spatial mapping and experimental evidence.

2 Methods

2.1 Data cohorts and clinical samples

Four independent CRC transcriptomic datasets were integrated, comprising one discovery study (partitioned into a training set and an internal validation set) and three external validation cohorts. The GSE39582 cohort served as the discovery cohort; it comprised 566 CRC tumor tissue samples, of which 556 met the inclusion criteria for prognostic analysis. After requiring histologically confirmed primary CRC with complete overall survival (OS) follow-up and clinical variables, patients were assigned according to the prespecified discovery/validation split of the original study to a training set (n = 439, 147 OS events) and an internal validation set V1 (n = 117, 40 OS events). Three independent external validation cohorts were included: GSE17536 (n = 177, OS endpoint, 73 events), GSE14333 (n = 226, disease-free survival [DFS] endpoint, 176 events), and the Cancer Genome Atlas colon adenocarcinoma (TCGA-COAD) cohort (RNA-seq HTSeq-FPKM-UQ format; 248 patients with 67 OS events were included after excluding those with incomplete OS follow-up or key clinical variables). Single-cell validation used the GSE178341 dataset (371,223 cells from 62

patients with CRC, including 34 mismatch repair-deficient and 28 mismatch repair-proficient tumors). Sample screening procedures, exclusion criteria, endpoint definitions, platform information, censoring rates, and median follow-up times for each cohort are provided in Supplementary Table 1.

The mIF validation cohort consisted of 60 pairs of formalin-fixed paraffin-embedded (FFPE) CRC tissues and paired adjacent non-tumor tissues (≥5 cm from the tumor margin) surgically resected at Hainan General Hospital (Hainan Affiliated Hospital of Hainan Medical University) between January 2024 and December 2025. Pathological diagnoses were confirmed by two independent pathologists, with disagreements arbitrated by a third senior pathologist. Inclusion criteria were histologically confirmed primary CRC, complete clinicopathological data, FFPE tissue with a necrosis fraction <30%, and age ≥18 years; exclusion criteria were severe systemic disease, preoperative neoadjuvant therapy, incomplete medical records, or a history of other malignancies. This study was approved by the Ethics Committee of Hainan General Hospital (Hainan Affiliated Hospital of Hainan Medical University) (approval no. EC-YLY-2025-103-01), conducted in accordance with the Declaration of Helsinki, and all patients provided written informed consent. Public cohort data were obtained from published databases and required no additional ethical approval.

2.2 Preprocessing of multi-cohort transcriptomic data

Microarray cohorts (GSE39582, GSE17536, and GSE14333) were downloaded as series matrix files from the Gene Expression Omnibus (GEO), and probes were mapped to gene symbols using the platform-specific annotation files, with multiple probes for

the same gene collapsed by the median. TCGA-COAD HTSeq-FPKM-UQ data were log2(x+1) transformed. To remove cross-platform batch effects, z-score standardization was performed independently within each cohort; for cross-cohort portability analysis, an alternative standardization using training-set-frozen means and standard deviations (method B) was applied and compared with within-cohort standardization (method A). Expression matrices retained only the 17 model genes and the genes required for related analyses.

2.3 CMIR score construction and nested cross-validation

The CMIR score was constructed exclusively within the training set to avoid information leakage. Candidate genes were first screened by univariate Cox proportional hazards regression ($P < 0.05$), and 45 machine learning algorithm combinations (including least absolute shrinkage and selection operator (LASSO), Ridge, elastic net, CoxBoost, survival support vector machine, and their combinations) were then systematically compared by nested cross-validation (outer 5-fold for unbiased performance estimation, inner 5-fold for hyperparameter tuning, 10 repetitions, random seed 42). The algorithm combination with the highest nested cross-validation C-index (LASSO + LASSO) was selected, and the final model was fitted on the full training set, freezing the regression coefficients of the 17 genes, the standardization parameters (mean and standard deviation), and the risk threshold (-0.023753). The CMIR score was calculated as the weighted sum of standardized gene expression and the corresponding coefficients. Model simplification used the selection frequency of 500-bootstrap LASSO combined with repeated 5-fold cross-validation,

and the minimum gene subset (8-gene parsimonious model) was determined by the 1-SE rule. Algorithm selection, gene screening, model size, and threshold determination were all completed within the training set; external cohorts were used only for independent validation of the frozen model.

2.4 Model validation and performance evaluation

The frozen model was validated in the internal validation set V1 and three independent external validation cohorts (GSE17536, GSE14333 with DFS endpoint, and TCGA-COAD). Kaplan-Meier curves and the log-rank test compared survival between high- and low-CMIR groups dichotomized by the cohort-specific median; univariate and multivariate Cox regression assessed the independent prognostic value of the CMIR score (multivariate models adjusted for age, sex, TNM stage, and other clinical variables); the C-index, 12/36/60-month time-dependent area under the curve (AUC) of receiver operating characteristic curves, calibration curves, and decision curve analysis (DCA) evaluated discrimination, calibration, and clinical net benefit, in comparison with TNM stage and the combined CMIR + TNM model. Threshold portability was assessed by comparing the high-risk proportions and stratification ability of cohort-median versus fixed training-set-threshold grouping; cross-platform portability was assessed by comparing the C-index and score correlation between within-cohort standardization (method A) and training-set-frozen standardization (method B); external calibration was evaluated using the calibration slope.

2.5 Immune microenvironment analysis of bulk transcriptomes

To elucidate the relationship between the CMIR score and the tumor immune

microenvironment, immune microenvironment analysis was performed on the complete expression matrix of all 566 CRC tumor tissue samples in the GSE39582 cohort. This analysis did not require complete OS follow-up or clinical variables and therefore included 10 patients in addition to the 556 included in the prognostic analysis. CMIR scores were calculated after within-cohort z-score standardization. Single-sample gene set enrichment analysis (ssGSEA, GSVA package) was used to score 14 immune and stromal signatures, including $CD8^+$ T cells, T cells, cytotoxicity, natural killer (NK) cells, B cells, macrophages, M2 macrophages, regulatory T cells (Tregs), CAFs, endothelial cells, extracellular matrix (ECM), transforming growth factor-β (TGF-β), adenosine signaling, and the CXCL9/10/11 chemokine signature. Expression of 10 immune checkpoint genes (*PDCD1*, *CD274*, *PDCD1LG2*, *CTLA4*, *LAG3*, *TIGIT*, *HAVCR2*, *IDO1*, *VSIR*, and *CD276*) was also evaluated. The immunedeconv package integrated ESTIMATE, MCP-counter, and xCell to infer immune cell infiltration and stromal/immune scores. Signature scores and checkpoint expression were compared between high- and low-CMIR groups.

2.6 Molecular subtyping and confounder-stratified analyses

To evaluate the relationship between the CMIR score and CRC molecular subtypes and to exclude confounding by molecular subtypes, consensus molecular subtype (CMS) classification was performed in the GSE39582 cohort using the CMScaller package, and mismatch repair (MMR) status, CpG island methylator phenotype (CIMP), chromosomal instability (CIN), and *KRAS* and *BRAF* mutation status were extracted from the clinical annotations. Differences in CMIR scores among molecular

subtype/feature groups were compared with the Wilcoxon rank-sum test, and the prognostic value of the CMIR score was assessed within each molecular subtype (Cox regression) to determine whether its prognostic effect was independent of molecular subtypes.

2.7 Single-cell transcriptomic analysis

GSE178341 single-cell data (371,223 cells in the full dataset) underwent quality control, normalization, highly variable gene selection, principal component analysis, and uniform manifold approximation and projection (UMAP), followed by graph-based clustering for cell population identification and annotation of cell types based on canonical markers. A total of 109,806 cells from the tumor microenvironment compartment were extracted for subsequent analysis and annotated into 15 cell types; CAFs were further subdivided into 11 subtypes and $CD8^+$ T cells into 7 subtypes. Bulk CMIR scores were mapped onto single cells to resolve their cellular origin. Association analyses between patient-level CAF-CMIR scores and $CD8^+$ T cell functional signatures (exclusion, cytotoxicity, and exhaustion) included only tumor tissue (SPECIMEN_TYPE = "T"), used per-patient medians, and required ≥10 cells per patient, yielding 42 patients.

2.8 Cell-cell communication analysis

The LIANA framework was used to integrate multiple methods for inferring ligand-receptor interactions between CAFs and $CD8^+$ T cells, ranked by specificity rank and magnitude rank. To ensure the reliability of database-curated ligand-receptor pairs, the TGF-β family ligands were specifically verified: whether *TGFB1* pairs with its

canonical receptors *TGFBR1*/*TGFBR2* was checked pair by pair, and non-canonical receptor pairings such as CXCR4 were flagged as pseudo-pairs and removed, retaining only canonical ligand-receptor axes for mechanistic interpretation.

2.9 Multiplex immunofluorescence staining and quantification

mIF staining was performed on 60 FFPE CRC tissues and paired adjacent tissues. Sections of 4 μm thickness were deparaffinized, rehydrated through graded alcohols, and subjected to pressure-cooker antigen retrieval in citrate buffer (pH 6.0), followed by 5 rounds of sequential antibody incubation and Opal fluorophore staining (Opal 7-Color Manual Immunohistochemistry (IHC) Kit, Akoya Biosciences); antigen retrieval was repeated after each round to completely strip antibodies from the previous round, and nuclei were counterstained with 4',6-diamidino-2-phenylindole (DAPI). Target antibodies and dilutions were: CD8 (Abcam, ab237709, 1:200), fibroblast activation protein (FAP) (Abcam, ab207178, 1:100), alpha-smooth muscle actin (α-SMA) (Abcam, ab7817, 1:200), CD73 (CST, #13160, 1:100), and KLF2 (Abcam, ab203591, 1:100).

Because the mIF panel did not cover all genes in the transcriptomic CMIR model, an mIF-based CMIR proxy score was constructed by equally weighting 4 protein markers (FAP, α-SMA, CD73, and KLF2; i.e., the sum of the z-scores of each protein), and patients were dichotomized at the median into high and low groups (30 each). The panel did not include epithelial markers such as panCK/EpCAM, so tumor regions were delineated by morphological criteria (nuclear atypia, glandular architecture, and stromal boundaries) under the guidance of pathologists. Images were acquired with a Vectra Polaris multispectral imaging system; 5 random non-overlapping fields covering

the tumor center, invasive front, and tumor-stroma interface were selected per sample, excluding necrotic and artifact regions. After spectral unmixing with inForm software (v2.4.10), cell segmentation based on DAPI nuclear detection was performed with QuPath software (v0.4.3) (minimum cell area 20 μm^2; fluorescence thresholds determined from negative controls). Quantified metrics included $CD8^+$ T cell density, CD8-FAP spatial distance, CD8-tumor epithelium distance, FAP- and α-SMA-positive area fractions, the CD73/FAP colocalization coefficient (Pearson correlation), and the mean fluorescence intensity of CD73 and KLF2. Measurements from the 5 fields per patient were averaged into one patient-level value, and the patient (rather than the individual field) was used as the statistical unit to avoid pseudoreplication; image acquisition and data analysis were performed blinded throughout.

2.10 Spatial transcriptomic analysis

Spatial transcriptomic analysis was performed as a single-sample exploratory supplementary analysis based on 1 CRC sample (SN84_A120838_Rep1, 328 spots). CMIR, CAF, and CD8 signature scores were computed for each spot; Spearman correlation was used to assess spatial colocalization between the CMIR and CAF signatures, and Moran's I was used to test spatial autocorrelation.

2.11 Cell experiments and functional validation

2.11.1 Isolation and culture of primary CAFs

Primary CAFs were isolated from freshly resected tumor tissues of three independent patients with CRC (n = 3 distinct patient donors) to account for patient-to-patient biological variability. Tissue samples were thoroughly washed with sterile

phosphate-buffered saline containing penicillin and streptomycin, minced into approximately 1 mm^3 fragments, and digested with agitation at 37 ℃ for 1 h in digestion medium containing collagenase (1 mg/mL) and DNase I. The cell suspension was filtered through a 70 μm strainer and centrifuged, and the cells were resuspended in DMEM/F12 complete medium containing 10% fetal bovine serum and 1% penicillin-streptomycin and incubated at 37 ℃ with 5% CO_2. Cells at passages 3 to 8 were used in all experiments. After morphological observation and immunofluorescence verification of markers (FAP-positive, α-SMA-positive, and EpCAM-negative), cells were used in subsequent experiments.

2.11.2 Construction of KLF2 knockdown and overexpression cell lines

KLF2 knockdown (KLF2-KD) was achieved with lentiviral vector-mediated short hairpin RNA (shRNA). An shRNA targeting human *KLF2* (target sequence 5'-GACCUACACCAAGAGUUCG-3'; reference sequence NM_016270) was packaged into the GV112 lentiviral vector driven by the human U6 promoter (AgeI and EcoRI cloning sites, with a puromycin selection marker). For overexpression (KLF2-OE), the full-length human *KLF2* coding sequence was cloned into the CMV promoter-driven GV341 lentiviral vector (3×FLAG tag, AgeI and NheI cloning sites, with a puromycin selection marker). Recombinant plasmids were verified by colony PCR, restriction enzyme digestion, and Sanger sequencing. Primary CAFs were transduced with lentivirus at the optimal multiplicity of infection; 72 h after transduction, puromycin (Sigma-Aldrich, P8833) was added to select stable cell lines, and expression levels were verified by RT-qPCR and Western blot. Control CAFs were transduced with empty

vector lentivirus. Four groups were established: Control CAF, KLF2-KD, KLF2-OE, and KLF2-OE plus CD73 inhibitor (KLF2-OE^{+}CD73i). CD73 was inhibited with the small-molecule compound AB680 (quemliclustat, MedChemExpress, HY-125286).

2.11.3 Real-time quantitative PCR

Total RNA was extracted from CAFs of each group using TRIzol reagent; the A260/A280 ratio measured on a NanoDrop 2000 spectrophotometer was required to be between 1.8 and 2.0, and RNA integrity was assessed by agarose gel electrophoresis. cDNA was synthesized from 1 μg of total RNA using the PrimeScript RT kit. Amplification was performed with TB Green Premix Ex Taq II on an Applied Biosystems 7500 Fast system in a total volume of 20 μL. The amplification program consisted of initial denaturation at 95 ℃ for 3 min, followed by 40 cycles of 95 ℃ for 10 s and 60 ℃ for 30 s, and melting curve analysis was performed after amplification. Relative expression was calculated by the $2^{-\Delta\Delta Ct}$ method with *GAPDH* as the reference gene. Primer sequences were as follows: *GAPDH* forward 5'-ACAGTCAGCCGCATCTTCTT-3' and reverse 5'-TTGATTTTGGAGGGATCTCG-3'; *KLF2* forward 5'-CCACCTGTTGTGTGCATTGG-3' and reverse 5'-CTCCAGGTGGGTTTCCAGAC-3'.

2.11.4 Western blot

CAF lysates were prepared with RIPA lysis buffer, protein concentrations were determined with a BCA protein assay kit, and equal amounts of protein were separated by SDS-PAGE and wet-transferred onto PVDF membranes. Membranes were blocked with 5% non-fat milk at room temperature for 1 h and then incubated with primary

antibodies overnight at 4 ℃ on a shaker. Primary antibodies included KLF2 (Proteintech, 23384-1-AP, 1:1000), CD73 (Proteintech, 12231-1-AP, 1:1000), FAP (Proteintech, 1:1000), α-SMA (Proteintech, 23081-1-AP, 1:1000), and GAPDH (Biosharp, BL006B, 1:2000). After washing, membranes were incubated with HRP-conjugated secondary antibodies at room temperature for 1 h, ECL chemiluminescence substrate was added, and images were acquired with a ChemiDoc MP system. Semi-quantitative analysis of target band gray values was performed with GAPDH as the loading control, and each group included 3 independent biological replicates (derived from CAFs isolated from the three distinct patient donors).

2.11.5 Ternary co-culture of CAFs, tumor cells, and $CD8^+$ T cells and functional assays

Human $CD8^+$ T cells were isolated from peripheral blood mononuclear cells of healthy donors by negative immunomagnetic bead sorting (purity >95%) and activated with anti-CD3 and anti-CD28 antibodies combined with IL-2 for 48 h. CAFs subjected to different genetic interventions were co-cultured with CRC cells at a 1:1 ratio for 24 h, and activated $CD8^+$ T cells were then added at an effector-to-target ratio of 10:1 for continued co-culture. Before testing, cells were incubated with PMA (50 ng/mL), ionomycin (1 μg/mL), brefeldin A (10 μg/mL), and anti-CD107a antibody for 4 to 6 h. Cell surfaces were stained with anti-CD3 (clone SK7) and anti-CD8 (clone SK1) antibodies at 4 ℃ in the dark for 30 min, and after fixation and permeabilization, intracellular IFN-γ (clone 4S.B3) and granzyme B (clone GB11) staining was performed. Data were acquired on a BD FACSCanto II flow cytometer with >100,000 events per sample, and the percentages of IFN-γ^+, granzyme B^+, and $CD107a^+$ cells

within the $CD3^+CD8^+$ population were analyzed by gating.

2.11.6 Tumor cell-specific killing assay

The cytotoxic activity of $CD8^+$ T cells against target cells was measured by the lactate dehydrogenase (LDH) release assay. After 4 h of ternary co-culture, culture supernatants were collected by centrifugation, with spontaneous release control wells, maximum release control wells, and medium blank wells set up in parallel. LDH cytotoxicity detection reagent was added and incubated in the dark for color development, absorbance at 490 nm was measured on a microplate reader, and the percentage of specific killing of tumor cells was calculated using the standard formula provided with the kit. All experiments were performed in triplicate wells per group and independently repeated at least 3 times.

2.12 Statistical analysis

Normality of continuous variables was assessed with the Shapiro-Wilk test; comparisons between two groups used the Mann-Whitney U test, and comparisons among multiple groups used the Kruskal-Wallis test with Dunn's post hoc test (Benjamini-Hochberg false discovery rate correction); comparisons among multiple groups in cell experiments used one-way analysis of variance (ANOVA) with Tukey's post hoc test. Correlation analysis used Spearman rank correlation. Survival analyses used the Kaplan-Meier method, log-rank test, and Cox proportional hazards regression, reporting hazard ratios (HRs) with 95% confidence intervals; the Cox proportional hazards assumption was verified with Schoenfeld residual tests (cox.zph, survival package). Multivariate Cox models adjusted for clinical covariates; tumor purity-

adjusted analyses added purity scores to the Cox models. Calibration curves used 500 bootstrap resamples, and the 95% confidence interval of the C-index was estimated with 1000 bootstrap replicates; time-dependent AUCs were calculated by the inverse probability of censoring weighting (IPCW) method (timeROC package). The multiple comparison correction strategy was as follows: exploratory analyses (immune cell infiltration, immune checkpoint expression, single-cell CMIR score comparisons, and mIF spatial quantification) were corrected with BH-FDR; high-dimensional pairwise comparisons among cell subtypes were corrected with Bonferroni; and prespecified primary endpoints (survival analysis and multivariate Cox regression) and single-hypothesis exploratory analyses were not corrected. Missing data were handled by complete-case analysis. Two-sided $P < 0.05$ was considered statistically significant. Statistical analyses were performed with R 4.4.3 (R Foundation for Statistical Computing, Vienna, Austria) and Python 3.11 (Python Software Foundation, Wilmington, DE, USA); LASSO-Cox modeling used glmnet 5.1, single-cell analysis used scanpy 1.9, cell-cell communication analysis used LIANA 1.7.1, and plotting used ggplot2 and matplotlib.

3 Results

3.1 Construction of the CMIR score by nested cross-validation and multi-cohort prognostic validation

In the GSE39582 training set of 439 patients with CRC, 45 machine learning algorithm combinations were compared for prognostic discrimination by nested cross-validation (Fig. 1A). The LASSO + LASSO combination achieved the highest

concordance index (C-index = 0.624) (Fig. 1B). The final model was fitted and frozen on this algorithm, and 17 feature genes were selected, with positive regression coefficients for *SLC2A1*, *NT5E*, *JUNB*, *KLF2*, *EGR3*, *CD36*, and *PIM1* and negative coefficients for *CCL22*, *CXCL11*, *ILVBL*, *IL17RB*, *FUT4*, *C2*, *FZD5*, *CCR7*, *CASP6*, and *CXCL9* (Fig. 1C). In 500 bootstrap resampling analyses, the selection frequencies of *CCL22*, *ILVBL*, *IL17RB*, *CXCL11*, *SLC2A1*, and *NT5E* exceeded 0.8 (Fig. 1D).

In survival analyses, after dichotomization at the median CMIR score, the high-CMIR group had shorter OS than the low-CMIR group in the training set (Fig. 2A), internal validation set V1 (Fig. 2B), GSE17536 cohort (Fig. 2C), and TCGA-COAD cohort (Fig. 2D) (all $P < 0.05$). In the GSE14333 cohort, DFS did not differ significantly between the high- and low-CMIR groups ($P = 0.97$) (Fig. 2E). Univariate Cox proportional hazards regression showed that a one-standard-deviation increase in the CMIR score was associated with a higher OS risk in all four OS cohorts, whereas the hazard ratio was not statistically significant in the GSE14333 DFS cohort (Fig. 2F).

3.2 Incremental value of the CMIR score over TNM stage, threshold portability analysis, and model simplification

In the four OS cohorts (Training, V1, GSE17536, and TCGA), combining the CMIR score with TNM stage improved the C-index over TNM stage alone, with increments of +0.111 ($P < 0.001$), +0.069 ($P < 0.001$), +0.022 ($P < 0.001$), and +0.025 ($P < 0.05$), respectively (Fig. 3A). Time-dependent AUC analysis showed that the CMIR + TNM combined model yielded higher AUC values than TNM stage alone at 12, 36, and 60 months (Fig. 3B). Calibration curves for 3-year mortality showed good

agreement between predicted and observed mortality in each cohort (Fig. 3C). Five-year decision curve analysis showed that the CMIR + TNM combined model provided higher net benefit than the CMIR score or TNM stage alone in the training set (Fig. 3D).

In the risk threshold portability assessment, log-rank tests were statistically significant in all OS cohorts when dichotomized by the cohort-specific median ($P < 0.05$), whereas with the fixed training-set threshold, log-rank tests did not reach significance in the GSE17536 and TCGA cohorts ($P > 0.05$) (Fig. 4A). Statistics of the high-risk patient proportion showed that the fixed training-set threshold classified 0% of patients as high risk in GSE17536, 25.0% in TCGA, and 99.6% in GSE14333 (Fig. 4B). In the cross-platform standardization portability assessment, the C-index and calibration slope of within-cohort standardization (method A) and training-set-frozen standardization (method B) were generally similar across cohorts (Suppl Fig. 1A, 1B), and the CMIR scores calculated by the two methods were positively correlated across cohorts (Spearman rho between 0.83 and 1.00) (Suppl Fig. 1D); under method B, the high-risk proportion was 25.0% in TCGA and 99.6% in GSE14333 (Suppl Fig. 1C).

In model simplification, an 8-gene parsimonious model was distilled from the 17 feature genes using the 1-SE rule (Fig. 4C). In the training set and three independent external cohorts, the C-index of the 8-gene parsimonious model was close to that of the full 17-gene model in each cohort (Fig. 4D).

3.3 Association of the CMIR score with bulk transcriptomic tumor microenvironment features and molecular feature subgroup analyses

In the GSE39582 cohort, the CMIR score was positively correlated with stromal

and immunosuppressive features such as TGF-β, ECM, adenosine, and the CAF signature, and negatively correlated with $CD8^+$ T cells, cytotoxicity, CXCL9/10/11, and immune checkpoint molecules (Fig. 5A). The CAF score and stromal score were higher in the high-CMIR group than in the low-CMIR group (both $P < 0.001$, Fig. 5B and Fig. 5C), whereas the $CD8^+$ T cell score and cytotoxicity score were lower in the high-CMIR group (both $P < 0.001$, Fig. 5D and Fig. 5E).

After tumor purity adjustment, the positive correlations of the CMIR score with the CAF and stromal scores and the negative correlations with the cytotoxicity and immune scores persisted, with increased correlation coefficients (Suppl Fig. 2A). Across molecular feature subgroups including MMR, *KRAS*, CIMP, CIN, and *BRAF*, a one-standard-deviation increment in the CMIR score was consistently associated with a higher OS risk (Suppl Fig. 2B). Between-group comparisons showed that CMIR scores were lower in patients with dMMR tumors than in those with pMMR tumors ($P = 5.8 \times 10^{-4}$, Suppl Fig. 2C), and higher in patients with *KRAS*-mutant tumors than in those with wild-type tumors ($P = 4.4 \times 10^{-5}$, Suppl Fig. 2D).

3.4 Single-cell transcriptomic tracing of the CAF cellular origin and molecular features of the CMIR score

In the GSE178341 single-cell transcriptomic dataset, tumor microenvironment cells were reduced, clustered, and annotated by UMAP into 15 major cell types (Fig. 6A). After mapping the CMIR score onto CAFs, the score showed a heterogeneous continuous distribution across the CAF population (Fig. 6B). Further subclustering of CAFs identified 11 subtypes with distinct distribution patterns in UMAP space (Fig.

6C). $CD8^+$ T cells were classified into 7 subtypes (Fig. 6D).

Gene set enrichment analysis (GSEA) of CAF pseudo-bulk transcriptomes identified positive correlations between high CMIR scores and pathways including FGFR downstream signaling, *TGFBR3* expression, *NFE2L2*-regulated genes, laminin interactions, and NOTCH4 intracellular domain regulation, whereas negatively correlated pathways were mainly enriched in ribosomal and protein translation processes (Suppl Fig. 3A). Transcription factor activity analysis revealed differential expression activity of RELA, NFKB1, HIF1A, JUN, FOS, STAT3, and EGR1 across CAF subtypes (Suppl Fig. 3B). Among microenvironment-related genes, *KLF2* expression was positively correlated with the CAF CMIR score (Spearman rho > 0.5, $P < 0.001$), and the expression of *CXCL9*, *NT5E*, *CXCL12*, and *ENTPD1* was also positively correlated with the CMIR score (Suppl Fig. 3C).

3.5 Ligand-receptor interaction analysis between CAFs and $CD8^+$ T cells and validation by spatial transcriptomic colocalization

In single-cell ligand-receptor interaction analysis, CAF subpopulations signaled to $CD8^+$ T cells mainly through the *CXCL14-CXCR4* and *CXCL12-CXCR4* chemokine axes, with CXCL14-CXCR4 showing the highest specificity rank and mean expression in the $CXCL14^+$, BMP-producing, and $CCL8^+$ CAF subpopulations (Fig. 7A). Expression distribution analysis revealed that *CXCL9*, *CXCL10*, and *CXCL11* were mainly expressed in macrophages, monocytes, and dendritic cells and were absent or barely detectable in CAF subpopulations, whereas their receptor *CXCR3* was specifically expressed across $CD8^+$ T cell subpopulations (Fig. 7B). Among canonical

non-chemokine interactions, the *LGALS1* (galectin-1)-CD69 axis was detected in most CAF subpopulations (Fig. 7C). Among extracellular matrix-related interactions, *COL1A1*, *COL6A3*, *FN1*, and *THBS2* showed extensive interactions with their receptors *CD44*, *ITGA1-ITGB1*, and *CD47* in CAF subpopulations (Fig. 7D).

In the single-sample CRC spatial transcriptomic analysis, spot-level CMIR scores (Suppl Fig. 4A), CAF signature scores (Suppl Fig. 4B), and CD8 signature scores (Suppl Fig. 4C) exhibited distinct spatial distribution patterns. Spot-level correlation analysis showed a positive correlation between the CMIR score and the CAF signature score (Spearman rho = 0.38, $P = 5.1 \times 10^{-13}$) (Suppl Fig. 4D).

3.6 In situ multiplex immunofluorescence quantification and spatial distribution features of the microenvironment

In mIF staining of 60 CRC tissues and paired adjacent tissues, representative images show the expression distribution of DAPI counterstaining and of CD8, FAP, α-SMA, CD73, and KLF2, together with the stromal-tumor boundary demarcation, across the four groups of high-CMIR tumor, low-CMIR tumor, high-CMIR adjacent, and low-CMIR adjacent tissues (Fig. 8).

Quantitative analysis revealed a significantly lower $CD8^+$ T cell density in the high-CMIR tumor group than in the low-CMIR tumor group ($P < 0.001$), which was also lower than in both adjacent tissue groups (Fig. 9A). Both the physical distance between $CD8^+$ T cells and FAP^+ CAFs and the distance between $CD8^+$ T cells and the tumor epithelium were significantly greater in the high-CMIR tumor group than in the low-CMIR tumor group and the adjacent groups (all $P < 0.001$, Fig. 9A). The FAP-

positive and α-SMA-positive area fractions were significantly higher in the high-CMIR tumor group than in the low-CMIR tumor group and the adjacent groups (all $P < 0.001$, Fig. 9B). The mIF-based CMIR proxy score showed a clear high-low stratification across the large sample groups (Fig. 9C). The CD73/FAP Pearson colocalization coefficient, the mean fluorescence intensity of CD73, and the mean fluorescence intensity of KLF2 were all significantly higher in the high-CMIR tumor group than in the low-CMIR tumor group and both adjacent tissue groups (all $P < 0.001$, Fig. 9D).

3.7 Functional validation of KLF2 regulation of CAF activation, CD73 expression, and $CD8^+$ T cell effector function

In primary CAFs, RT-qPCR revealed significantly lower relative mRNA expression of KLF2 in the KLF2-KD group than in the Control CAF group ($P < 0.001$), whereas expression in the KLF2-OE and KLF2-OE+CD73i groups was significantly higher than in the Control CAF group (both $P < 0.001$) (Fig. 10A). Western blot analysis revealed significant downregulation of the relative protein expression of FAP ($P < 0.05$), CD73 ($P < 0.001$), α-SMA ($P < 0.001$), and KLF2 ($P < 0.001$) in CAFs, whereas KLF2 overexpression significantly upregulated the protein expression of FAP, CD73, α-SMA, and KLF2 (all $P < 0.001$) (Fig. 10B). Combined treatment with the CD73 inhibitor on the KLF2 overexpression background (KLF2-OE+CD73i) significantly downregulated the protein expression of FAP ($P < 0.05$), CD73 ($P < 0.001$), and α-SMA ($P < 0.01$), whereas the protein expression of KLF2 itself showed no statistically significant difference (ns) (Fig. 10B).

In the co-culture system of CAFs and $CD8^+$ T cells, flow cytometry demonstrated

that the expression levels of IFN-γ, GZMB, and CD107a in $CD8^+$ T cells were markedly suppressed in the CD8 plus Control CAF group compared with the CD8 T alone group (Fig. 11A). Quantitative analysis revealed that compared with the CD8 plus Control CAF group, the CD8 plus KLF2-KD CAF group showed significantly higher percentages of IFN-γ^+ ($P < 0.001$), $GZMB^+$ ($P < 0.05$), and $CD107a^+$ ($P < 0.05$) $CD8^+$ T cells, whereas the CD8 plus KLF2-OE CAF group exhibited further profound reductions in all three effector molecules (all $P < 0.001$) (Fig. 11B). Notably, pharmacological blockade of CD73 (KLF2-OE CAF plus CD73i) significantly restored the percentages of IFN-γ^+, $GZMB^+$, and $CD107a^+$ $CD8^+$ T cells relative to the KLF2-OE group (all $P < 0.001$) (Fig. 11B). Consistent with these functional markers, the tumor cell-specific killing assay showed that the cytotoxic activity of $CD8^+$ T cells was impaired by Control CAFs ($P < 0.001$), rescued by KLF2 knockdown ($P < 0.001$), exacerbated by KLF2 overexpression ($P < 0.001$), and reversed by the CD73 inhibitor ($P < 0.001$) (Fig. 11C).

4 Discussion

Nested cross-validation effectively removes overfitting and evaluation bias during feature selection [12,17]. In this study, nested cross-validation was used to systematically evaluate 45 algorithms and construct the CMIR model. The LASSO algorithm has been shown to be well suited for handling high-dimensional gene data with collinear features [18]. The model was distilled into an 8-gene combination while maintaining excellent discrimination. A minimal gene set helps reduce testing costs and facilitates clinical translation based on qPCR or NanoString technology [19]. The

training-set-frozen standardization strategy maintained cross-platform consistency; this strategy has been widely used to overcome batch effects across sequencing platforms [20,21]. The model showed independent prognostic value in OS cohorts but not in the DFS cohort. DFS mainly reflects local recurrence, whereas OS is more susceptible to the influence of long-term stromal remodeling and systemic immune exhaustion [3,4]. High CMIR scores in this study were characterized by TGF-β activation, extracellular matrix deposition, and adenosine metabolism enrichment; the enrichment of these pathways is a typical feature of immunosuppressive tumor stroma [22,23].

In the single-cell and spatial transcriptomic analyses, the CMIR signature was mainly anchored in CAF subpopulations, with frequent CXCL12-CXCR4 and CXCL14-CXCR4 interactions between CAFs and $CD8^+$ T cells. The CXCR4 signaling axis has been shown to mediate lymphocyte retention in the peritumoral stroma [24,25]. mIF quantification of 60 tissue samples demonstrated a significant reduction in CD8-positive T cell density in high-CMIR tumors, with significantly increased physical distances from FAP-positive CAFs and the tumor epithelium. This spatial physical segregation is consistent with the stromal immune exclusion phenotype previously reported in tumors [7,10]. Dense collagenous stroma combined with disrupted chemokine gradients hinders effector T cell infiltration into the tumor core [9,26], which constitutes an important anatomical basis for the primary resistance of microsatellite-stable colorectal cancer to immune checkpoint blockade [27-30]. Accordingly, the CMIR score reflects these spatial and stromal properties, serving as a functional indicator of tumors that may exhibit intrinsic resistance to immune

checkpoint blockade due to the physical exclusion of effector T cells.

Previous studies have reported that KLF2 mainly maintains cell quiescence and homeostasis in vascular endothelium and immature T cells [15,16]. In this study, KLF2 intervention in primary CAFs coordinately regulated the expression of FAP, α-SMA, and CD73. As the rate-limiting enzyme that converts extracellular AMP to adenosine, CD73 is a core molecule of the chemical metabolic barrier in the tumor microenvironment [31]. Our co-culture system showed that CAFs overexpressing KLF2 significantly reduced interferon-γ, granzyme B, and CD107a positivity in $CD8^+$ T cells and decreased cytotoxic killing. High concentrations of extracellular adenosine can activate purinergic receptors on T cells, thereby blocking T cell receptor signaling and effector molecule release [32,33].

The small-molecule CD73 inhibitor AB680 (quemliclustat) prevents adenosine production by binding the catalytic site of CD73 [34,35]. In vitro intervention in this study showed that combined use of AB680 significantly reversed the reduction in T cell effector molecules and cytotoxicity caused by KLF2-overexpressing CAFs. Blockade of the CD73-adenosine axis has been shown to improve immune infiltration in the tumor microenvironment and to produce synergistic antitumor effects with PD-1/PD-L1 inhibitors [36,37]. Targeting the KLF2-CD73-CAF axis therefore provides an actionable pharmacological strategy to modulate stromal-mediated immunosuppression and potentially sensitize immune-excluded colorectal tumors to immune checkpoint blockade.

This study has several limitations. First, the CMIR prognostic model was mainly

constructed and validated on retrospective high-throughput genomic datasets, and its independent clinical discrimination requires validation in large-scale prospective clinical cohorts. Additionally, technical batch variations among retrospective platforms led to extreme risk proportions under a fixed cutoff (0% in GSE17536 and 99.6% in GSE14333), highlighting the necessity of cohort-specific medians for retrospective datasets and the requirement for standardized clinical assays to establish an absolute threshold in prospective practice. Second, the regulatory mechanism of the KLF2-CD73-CAF axis on $CD8^+$ T cells was mainly based on an in vitro co-culture system of primary CAFs and T cells; future studies should establish immunocompetent mouse tumor models with CAF-specific knockdown or overexpression to evaluate in vivo the role of this pathway in spatial immune cell infiltration and in sensitivity to immune checkpoint blockade.

In summary, this study establishes a translational pipeline linking computational risk modeling, in situ spatial exclusion phenotypes, and targeted therapeutic intervention. We constructed a cross-platform-stable CMIR prognostic model, revealed the features of CAF-mediated spatial immune exclusion associated with primary immunotherapy resistance, and suggested a potential mechanism by which the KLF2-CD73-CAF axis modulates $CD8^+$ T cell killing function, partially through adenosine production, providing an actionable basis for clinical risk stratification and stroma-targeting combination immunotherapy in colorectal cancer.

Abbreviations

AUC, area under the curve; CAF, cancer-associated fibroblast; CD73, cluster of

differentiation 73; CD107a, cluster of differentiation 107a; CMIR, CAF stromal-immune exclusion risk; CRC, colorectal cancer; DFS, disease-free survival; FAP, fibroblast activation protein; GZMB, granzyme B; ICB, immune checkpoint blockade; IFN-γ, interferon-gamma; KLF2, Krüppel-like factor 2; LASSO, least absolute shrinkage and selection operator; mIF, multiplex immunofluorescence; OS, overall survival; scRNA-seq, single-cell RNA sequencing; ST, spatial transcriptomics; TCGA, The Cancer Genome Atlas; TME, tumor microenvironment; α-SMA, alpha-smooth muscle actin.

Declarations

Ethics approval and consent to participate

The studies involving human participants were reviewed and approved by the Ethics Committee of Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University (Approval No.: EC-YLY-2025-103-01). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants or their legal guardians/next of kin.

Consent for publication

Not applicable.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

We gratefully acknowledge the contributions of the patients who participated in this study. We also thank the data contributors of the NCBI GEO and TCGA databases for sharing the public transcriptomic and single-cell RNA-seq datasets.

Author contributions

YYP: Data curation, Investigation, Validation. QYZ: Software, Visualization, Data curation. LXX: Investigation, Clinical sample organization, Validation. XW: Software, Formal analysis, Visualization. YFL: Resources, Supervision, Writing – review & editing. FJ: Conceptualization, Methodology, Supervision, Funding acquisition, Writing – review & editing. All authors contributed to the article and approved the submitted version. NL: Conceptualization, Methodology, Formal analysis, Writing – original draft, Project administration.

Funding

This study was supported by the National Natural Science Foundation of China (82560470), Academic Enhancement Support Program of Hainan Medical University (XSTS2025040, XSTS2025063), the Innovation Platform for Academicians of Hainan Province (YSPTZX202005), Undergraduate Research and Innovation Training Program of Hainan Medical University (RZ2600001107, RZ2500002184), Hainan Medical University College Students' Innovation and Entrepreneurship Training Program (X202611810014), and the Open Research Fund of the Institute of Oncology, Hainan Medical University (hnmu-hams-IC-2026-14).

Data availability statement

The transcriptomic profiles, single-cell RNA sequencing data, and clinical information utilized in this study are publicly available in the NCBI GEO database (https://www.ncbi.nlm.nih.gov/geo/) under the accession numbers GSE39582, GSE17536, GSE14333, and GSE178341. The TCGA-COAD transcriptomic dataset (HTSeq-FPKM-UQ-format) and corresponding clinical data were accessed via the official Genomic Data Commons (GDC) Data Portal of The Cancer Genome Atlas (TCGA) database (GDC Legacy Archive / Data Release 18.0; https://portal.gdc.cancer.gov/). All other original data are included in this article and available from the corresponding authors upon request.

References

[1] Bray F, Laversanne M, Sung H, 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(3): 229-263. https://doi.org/10.3322/caac.21834.

[2] Sung H, Filho A M, Laversanne M, et al. Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 34 cancers in 186 countries. CA Cancer J Clin. 2026; 76(4): e70090. https://doi.org/10.3322/caac.70090.

[3] Garcia-Aguilar J, Patil S, Gollub M J, et al. Organ Preservation in Patients With Rectal Adenocarcinoma Treated With Total Neoadjuvant Therapy. J Clin Oncol. 2022; 40(23): 2546-2556. https://doi.org/10.1200/jco.22.00032.

[4] Canellas-Socias A, Sancho E, Batlle E. Mechanisms of metastatic colorectal cancer. Nat Rev Gastroenterol Hepatol. 2024; 21(9): 609-625.

https://doi.org/10.1038/s41575-024-00934-z.

[5] Chen D S, Mellman I. Elements of cancer immunity and the cancer-immune set point. Nature. 2017; 541(7637): 321-330. https://doi.org/10.1038/nature21349.

[6] Bagaev A, Kotlov N, Nomie K, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. Cancer Cell. 2021; 39(6): 845-865.e7. https://doi.org/10.1016/j.ccell.2021.04.014.

[7] Mariathasan S, Turley S J, Nickles D, et al. TGFbeta attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature. 2018; 554(7693): 544-548. https://doi.org/10.1038/nature25501.

[8] Xia Z, Chen H, Shen Y, et al. Multicellular immunosuppressive networks in microsatellite-stable colorectal cancer: CD8+ T cell exclusion, dysfunction, exhaustion, and therapeutic implications. Front Immunol. 2026; 17: 1945393. https://doi.org/10.3389/fimmu.2026.1945393.

[9] Ohlund D, Handly-Santana A, Biffi G, et al. Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. J Exp Med. 2017; 214(3): 579-596. https://doi.org/10.1084/jem.20162024.

[10] Pelka K, Hofree M, Chen J H, et al. Spatially organized multicellular immune hubs in human colorectal cancer. Cell. 2021; 184(18): 4734-4752.e20. https://doi.org/10.1016/j.cell.2021.08.003.

[11] Ding Y, Zhang S, Guo Y, et al. Recent advances in cancer-associated fibroblasts in colorectal cancer: functional subtypes, spatiotemporal dynamics and therapeutic targets. Neoplasia. 2026; 82: 101382.

https://doi.org/10.1016/j.neo.2026.101382.

[12] Chen X, Zhang D, Ou H, et al. Bulk and single-cell RNA sequencing analyses coupled with multiple machine learning to develop a glycosyltransferase associated signature in colorectal cancer. Transl Oncol. 2024; 49: 102093. https://doi.org/10.1016/j.tranon.2024.102093.

[13] Lyu A, Fan Z, Clark M, et al. Evolution of myeloid-mediated immunotherapy resistance in prostate cancer. Nature. 2025; 637(8048): 1207-1217. https://doi.org/10.1038/s41586-024-08290-3.

[14] Allard B, Longhi M S, Robson S C, et al. The ectonucleotidases CD39 and CD73: Novel checkpoint inhibitor targets. Immunol Rev. 2017; 276(1): 121-144. https://doi.org/10.1111/imr.12528.

[15] Fagerberg E, Attanasio J, Dien C, et al. KLF2 maintains lineage fidelity and suppresses CD8 T cell exhaustion during acute LCMV infection. Science. 2025; 387(6735): eadn2337. https://doi.org/10.1126/science.adn2337.

[16] Wang Y, Yu J, Huang W, et al. KLF2 as a prognostic and potential therapeutic target associated with ferroptosis in serous ovarian cancer: An integrative bioinformatics and clinical validation study. Oncol Lett. 2026; 32(5): 494. https://doi.org/10.3892/ol.2026.15849.

[17] Lyu H, Wang S, Guo G, et al. A machine learning-derived immune-related prognostic model identifies PLXNA3 as a functional risk gene in colorectal cancer. Front Immunol. 2025; 16: 1653794. https://doi.org/10.3389/fimmu.2025.1653794.

[18] Khaliq A M, Erdogan C, Kurt Z, et al. Refining colorectal cancer classification and clinical stratification through a single-cell atlas. Genome Biol. 2022; 23(1): 113. https://doi.org/10.1186/s13059-022-02677-z.

[19] Haque S, Mathkor D M, Wahid M, et al. Machine learning-enabled multi-omics discovery of prognostic biomarkers and signaling targets in pancreatic cancer. Comput Biol Chem. 2026; 125: 109320. https://doi.org/10.1016/j.compbiolchem.2026.109320.

[20] Dunne P D, Arends M J. Molecular pathological classification of colorectal cancer-an update. Virchows Arch. 2024; 484(2): 273-285. https://doi.org/10.1007/s00428-024-03746-3.

[21] Xia P, Qu Y, Liu Q, et al. Identification of programmed cell death-related subtypes reveals immune heterogeneity and therapeutic divergence in colon cancer. Theranostics. 2026; 16(9): 4821-4840. https://doi.org/10.7150/thno.126314.

[22] Tauriello D V F, Palomo-Ponce S, Stork D, et al. TGFbeta drives immune evasion in genetically reconstituted colon cancer metastasis. Nature. 2018; 554(7693): 538-543. https://doi.org/10.1038/nature25492.

[23] Liu X, Ding Q, Zhang H, et al. The CD39-CD73-adenosine axis: Master regulator of immune evasion and therapeutic target in pancreatic ductal adenocarcinoma. Biochim Biophys Acta Rev Cancer. 2025; 1880(5): 189443. https://doi.org/10.1016/j.bbcan.2025.189443.

[24] Sun R, Sun Y, Wu C, et al. CXCR4-modified CAR-T cells suppresses MDSCs

recruitment via STAT3/NF-kappaB/SDF-1alpha axis to enhance efficacy against pancreatic cancer. Mol Ther. 2023; 31(11): 3193-3209. https://doi.org/10.1016/j.ymthe.2023.09.010.

[25] Liu S, Zhang Z, Wang Z, et al. SPP1 Drives Colorectal Cancer Liver Metastasis and Immunotherapy Resistance by Stimulating CXCL12 Production in Cancer-Associated Fibroblasts. Cancer Res. 2026; 86(1): 58-79. https://doi.org/10.1158/0008-5472.can-24-4916.

[26] Ji X, Wu X, Sun W, et al. Fibroblasts in the tumor microenvironment: heterogeneity and dynamic interactions in tumor progression revealed by spatial transcriptomics. Cell Oncol (Dordr). 2025; 48(6): 1615-1629. https://doi.org/10.1007/s13402-025-01108-y.

[27] Zhu H, Roelands J, Ahmed E I, et al. Location matters: spatial dynamics of tumor-infiltrating T cell subsets is prognostic in colon cancer. Front Immunol. 2024; 15: 1293618. https://doi.org/10.3389/fimmu.2024.1293618.

[28] Niu D, Liu P, Chen X, et al. Mechanisms of immune resistance in microsatellite-stable colorectal cancer. Front Oncol. 2026; 16: 1893907. https://doi.org/10.3389/fonc.2026.1893907.

[29] Zhou Y, Yu J H, Sui Q Q, et al. POSTN(+) cancer-associated fibroblasts forming fortress-like barrier to mediate immunotherapy resistance in colorectal cancer. Cancer Lett. 2026; 657: 218683. https://doi.org/10.1016/j.canlet.2026.218683.

[30] Ziyi G, Zhou W, Zou K. Mapping the physical topology of the tumor immune microenvironment: macroscopic image fractals as surrogate markers for immune

749 exclusion phenotype and treatment resistance in colorectal cancer liver
750 metastases. Front Immunol. 2026; 17: 1889571.
751 https://doi.org/10.3389/fimmu.2026.1889571.
752 [31] Sun P, Zheng X, Li X. The Effects of CD73 on Gastrointestinal Cancer
753 Progression and Treatment. J Oncol. 2022; 2022: 4330329.
754 https://doi.org/10.1155/2022/4330329.
755 [32] Chi J J, Xie P, Cheng M H, et al. MGAT1-Guided complex N-Glycans on CD73
756 regulate immune evasion in triple-negative breast cancer. Nat Commun. 2025;
757 16(1): 3552. https://doi.org/10.1038/s41467-025-58524-9.
758 [33] Sambucci M, Fenoglio D, Ferrera F, et al. CD39 on CD8(+) Tumour-Infiltrating
759 Lymphocytes: A Paradoxical Marker of Tumour Reactivity and
760 Immunosuppression. Eur J Immunol. 2026; 56(10): e70288.
761 https://doi.org/10.1002/eji.70288.
762 [34] Lawson K V, Kalisiak J, Lindsey E A, et al. Discovery of AB680: A Potent and
763 Selective Inhibitor of CD73. J Med Chem. 2020; 63(20): 11448-11468.
764 https://doi.org/10.1021/acs.jmedchem.0c00525.
765 [35] Piovesan D, Tan J B L, Becker A, et al. Targeting CD73 with AB680
766 (Quemliclustat), a Novel and Potent Small-Molecule CD73 Inhibitor, Restores
767 Immune Functionality and Facilitates Antitumor Immunity. Mol Cancer Ther.
768 2022; 21(6): 948-959. https://doi.org/10.1158/1535-7163.mct-21-0802.
769 [36] Iwane K, Nakanishi Y, Muta Y, et al. Targeting fibroblast derived
770 thrombospondin 2 disrupts an immune-exclusionary environment at the tumor

front in colorectal cancer. Nat Commun. 2025; 16(1): 11590. https://doi.org/10.1038/s41467-025-66485-2.

[37] Wainberg Z A, Manji G A, Bahary N, et al. Quemliclustat and chemotherapy with or without zimberelimab in metastatic pancreatic adenocarcinoma: a randomized phase 1 trial. Nat Med. 2026; 32(4): 1267-1277. https://doi.org/10.1038/s41591-026-04283-z.

Legends

Fig. 1 Construction of the 17-gene CMIR prognostic model based on nested cross-validation and machine learning algorithm combinations

A Workflow of machine learning-based CMIR score modeling and multi-cohort independent validation; **B** Comparison of the prognostic discrimination of 45 machine learning algorithm combinations based on nested cross-validation C-index; **C** Regression coefficients of the 17 CMIR feature genes fitted by the LASSO-Cox algorithm; **D** Selection frequencies of candidate genes evaluated by 500-bootstrap LASSO resampling. Data are presented as C-index values or regression coefficients; algorithm evaluation was based on outer 5-fold and inner 5-fold nested cross-validation repeated 10 times; dashed lines indicate the optimal C-index baseline (B) and the 0.8 selection frequency threshold (D). Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; CV, cross-validation; C-index, concordance index; LASSO, least absolute shrinkage and selection operator; Enet, elastic net; CoxBoost, Cox boosting; SurvivalSVM, survival support vector machine; OS, overall survival; DFS, disease-free survival; TCGA, The Cancer Genome Atlas.

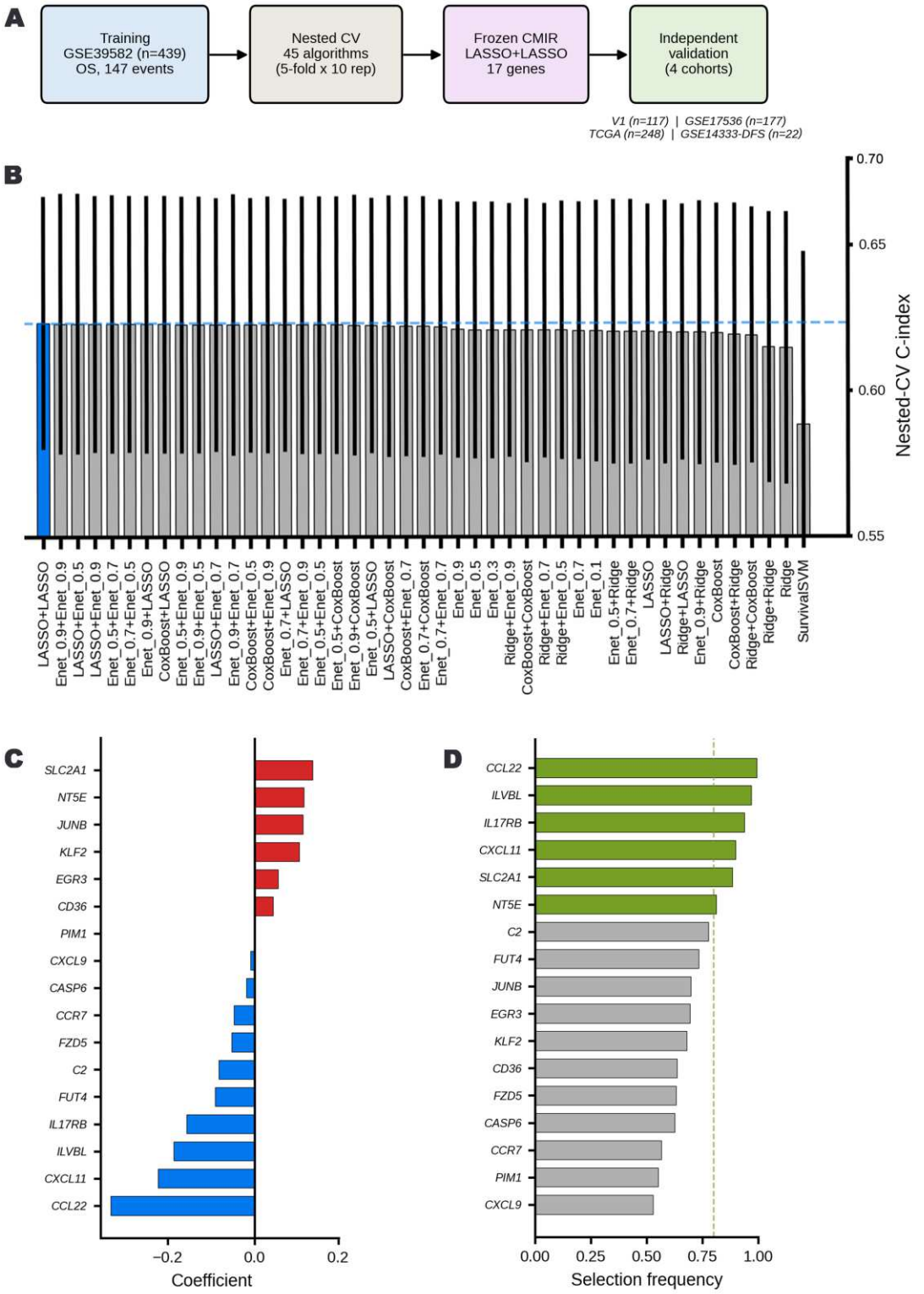


Fig. 2 Prognostic stratification of the CMIR score across multiple overall survival cohorts and endpoint-specific evaluation

A-E: Kaplan-Meier survival curves and the log-rank test were used to evaluate

differences between high- and low-CMIR groups in overall survival (OS) in the training set (A), internal validation set V1 (B), GSE17536 cohort (C), and TCGA-COAD cohort (D), and in disease-free survival (DFS) in the GSE14333 cohort (E); **F** Univariate Cox proportional hazards regression was used to estimate hazard ratios for patient prognosis per one-standard-deviation increase in the CMIR score in each cohort. Shading on survival curves represents 95% confidence intervals; forest plot data are presented as hazard ratios (HRs) with 95% confidence intervals; *** indicates $P < 0.001$, ** indicates $P < 0.01$, * indicates $P < 0.05$, and ns indicates $P \geq 0.05$. Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; OS, overall survival; DFS, disease-free survival; HR, hazard ratio; SD, standard deviation; TCGA, The Cancer Genome Atlas.

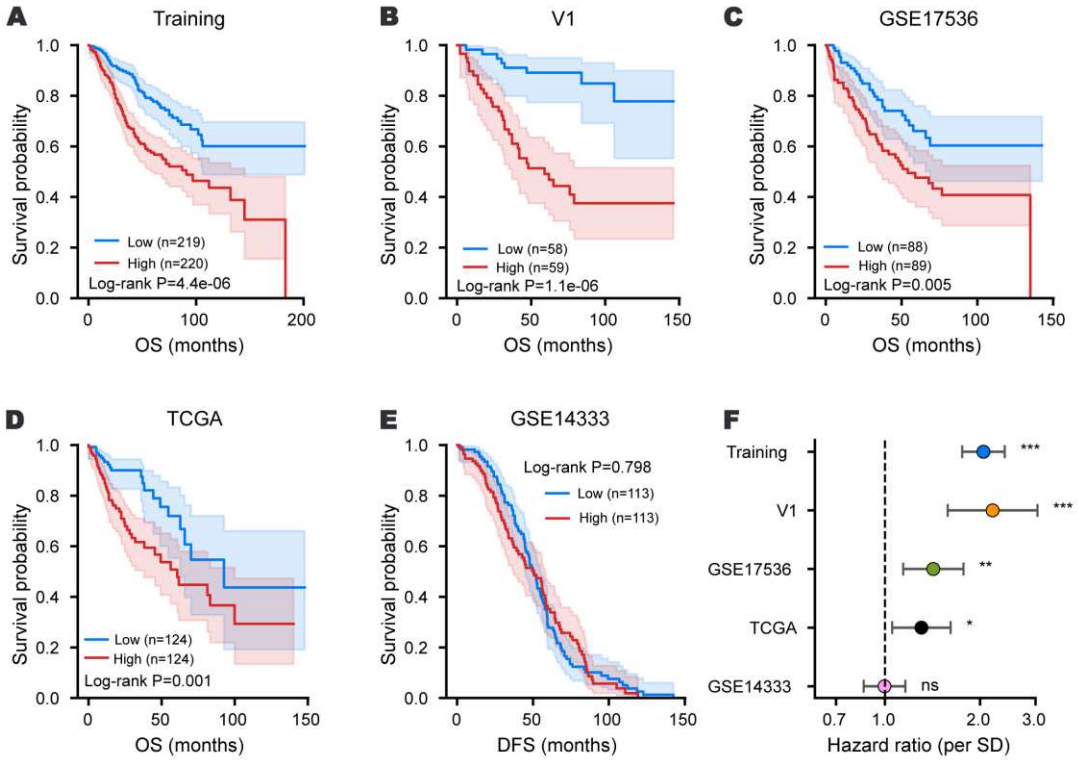


Fig. 3 Incremental prognostic value of the CMIR score combined with TNM stage and evaluation of clinical net benefit

A The C-index was used to evaluate the discrimination of TNM stage, the CMIR score, and the CMIR + TNM combined model in the Training, V1, GSE17536, and TCGA cohorts; **B** Time-dependent receiver operating characteristic curves were used to evaluate AUC values at 12, 36, and 60 months in each cohort; **C** Calibration curves were used to evaluate the agreement between predicted and observed 3-year mortality in each cohort; **D** Decision curve analysis was used to evaluate the net benefit of 5-year survival prediction in the training set. C-index increments are annotated above the bars; *** indicates $P < 0.001$ and * indicates $P < 0.05$. Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; TNM, tumor-node-metastasis; C-index, concordance index; AUC, area under the curve; TCGA, The Cancer Genome Atlas.

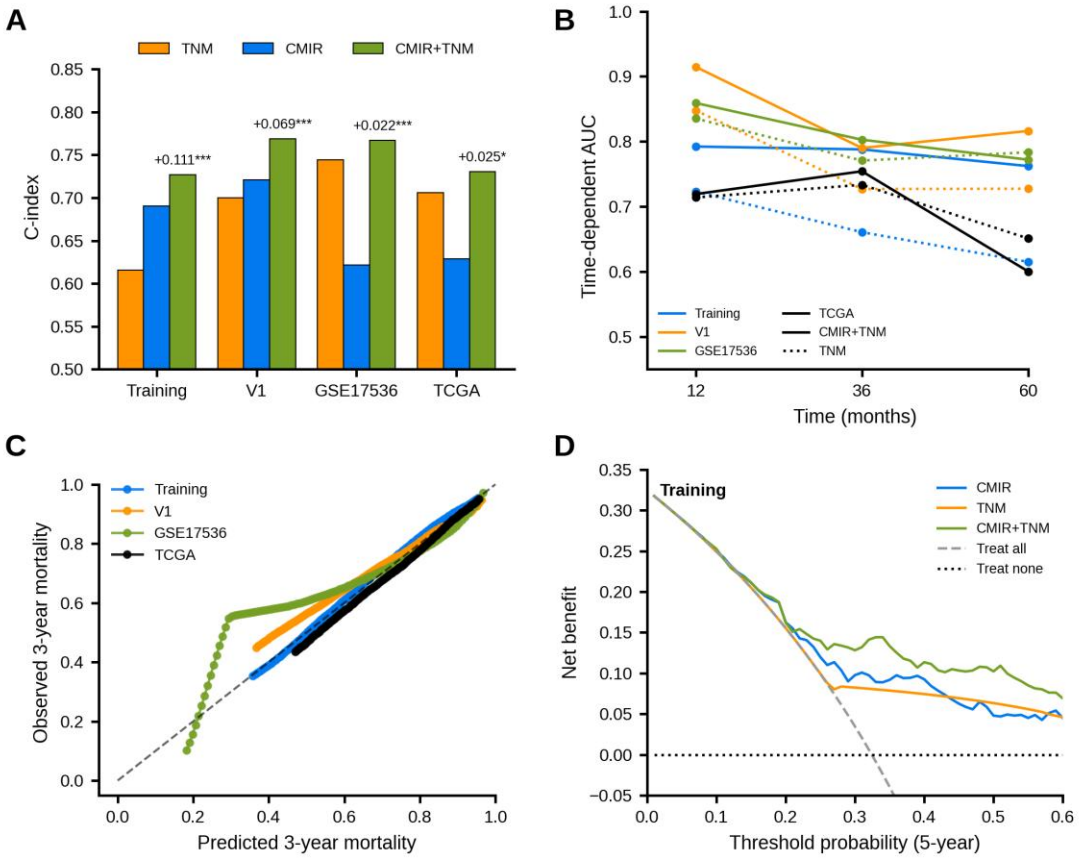


Fig. 4 Comparison of CMIR score threshold portability and simplification to an 8-gene parsimonious model

A The log-rank test was used to evaluate -log10(*P*) values for cohort-median grouping versus fixed training-set-threshold grouping in each cohort; **B** The percentage of patients classified as high risk under the two threshold strategies was tabulated for each cohort; **C** Nested cross-validation was used to evaluate the C-index of models with different gene numbers, and the 8-gene parsimonious model was selected by the 1-SE rule; **D** The C-index of the full 17-gene model and the 8-gene parsimonious model was compared across overall survival (OS) cohorts. Dashed lines indicate the significance threshold of $P = 0.05$ (A) and the 50% high-risk proportion reference line (B); orange dots represent the full 17-gene model and green dots represent the 8-gene parsimonious model (C). Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; C-index, concordance index; SE, standard error; TCGA, The Cancer Genome Atlas.

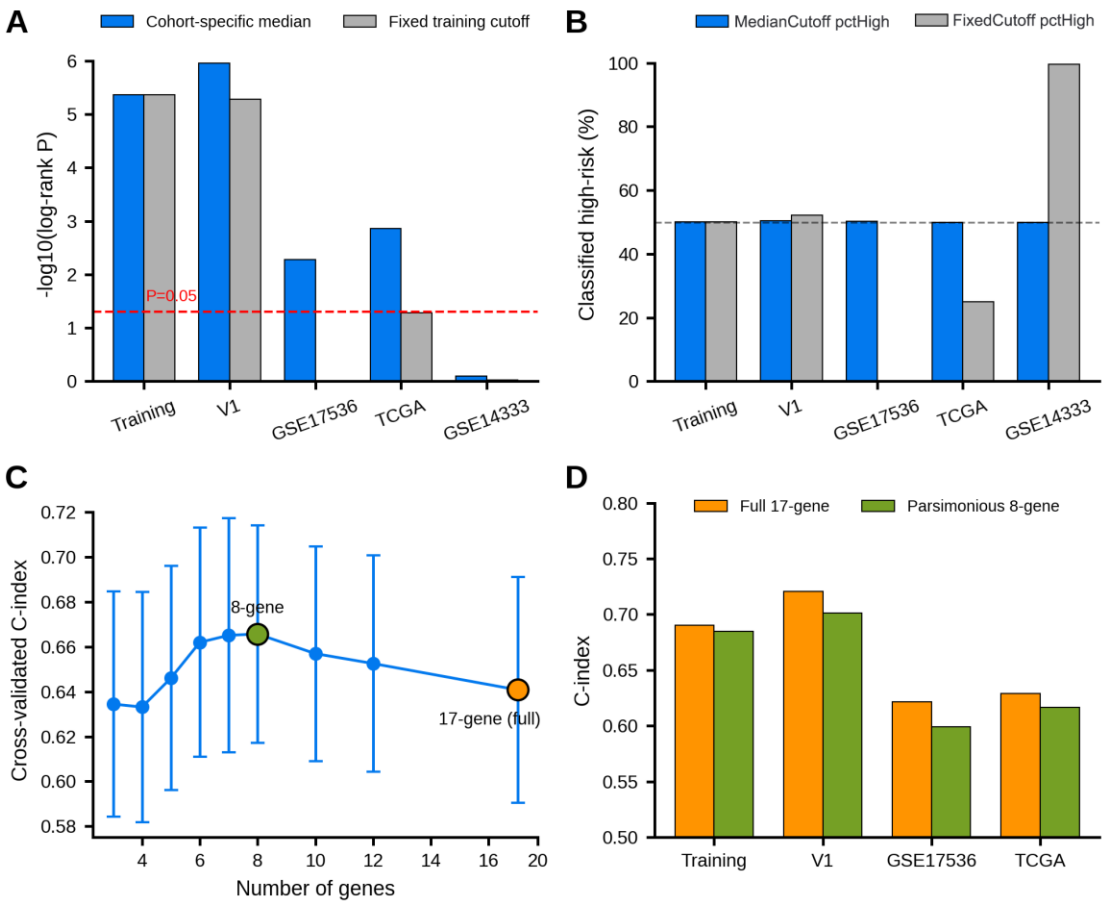


Fig. 5 Correlation of the CMIR score with bulk transcriptomic immune microenvironment features and between-group comparisons

A Spearman correlation analysis was used to evaluate correlations between the CMIR score and the expression of immune and stromal signature scores in the GSE39582 cohort; **B-E** Wilcoxon rank-sum tests were used to compare the CAF score (MCP-counter) (B), stromal score (ESTIMATE) (C), $CD8^+$ T cell score (ssGSEA) (D), and cytotoxicity score (ssGSEA) (E) between high- and low-CMIR groups. *** indicates $P < 0.001$ and * indicates $P < 0.05$. Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; CAF, cancer-associated fibroblast; ECM, extracellular matrix; MCP, MCP-counter algorithm; GSVA, gene set variation analysis; ssGSEA, single-sample gene set enrichment analysis; ESTIMATE, estimation of stromal and

immune cells in malignant tumors using expression data.

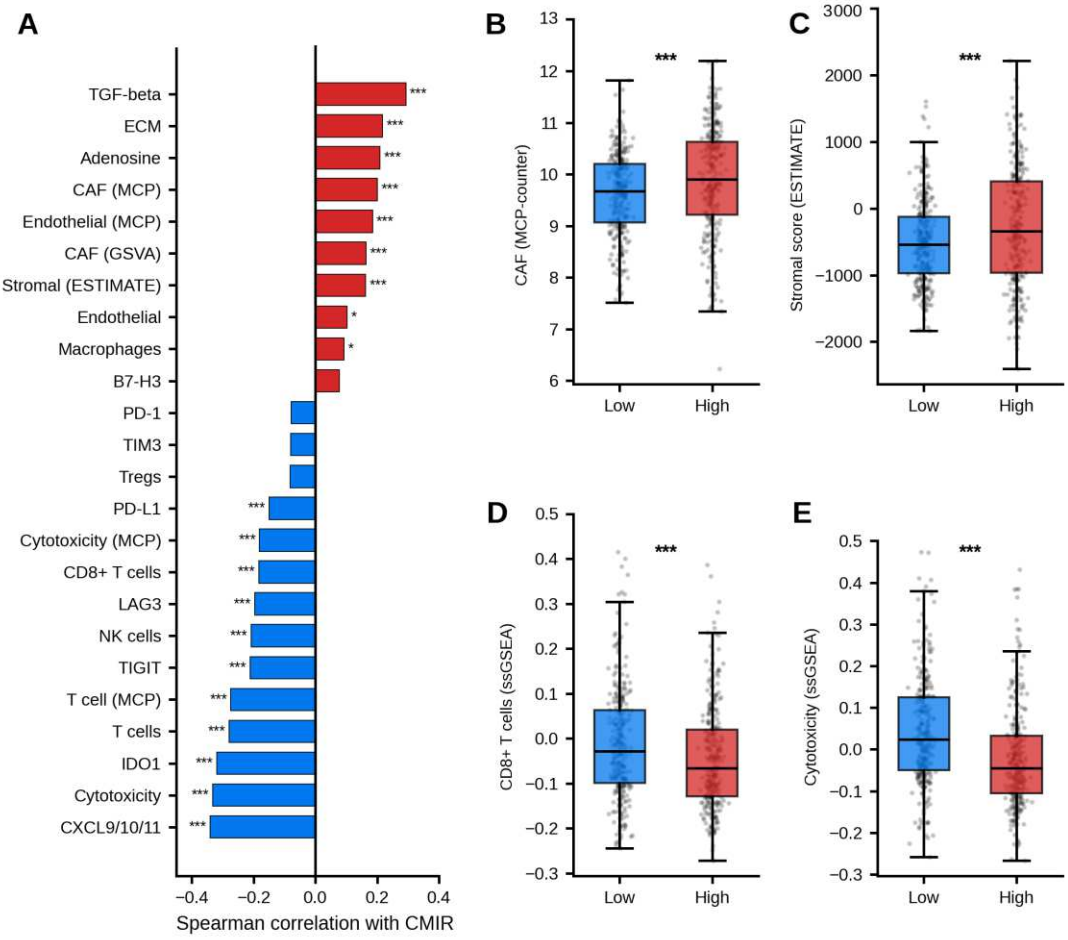


Fig. 6 Single-cell transcriptomic atlas annotation and distribution of the CMIR score across CAF subpopulations

A-D UMAP was used to display the clustering of 15 major cell types in the tumor microenvironment (A), the continuous distribution of the CMIR score in CAFs (B), the clustering of 11 CAF subtypes (C), and the clustering of 7 $CD8^+$ T cell subtypes (D). Data are presented as UMAP two-dimensional coordinates with expression color gradients. Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; CAF, cancer-associated fibroblast; UMAP, uniform manifold approximation and projection; DC, dendritic cell; Endo, endothelial cell; Fibro,

fibroblast; Granulo, granulocyte; ILC, innate lymphoid cell; Macro, macrophage; Mono, monocyte; NK, natural killer cell; Peri, pericyte; TCD4, $CD4^+$ T cell; TCD8, $CD8^+$ T cell.

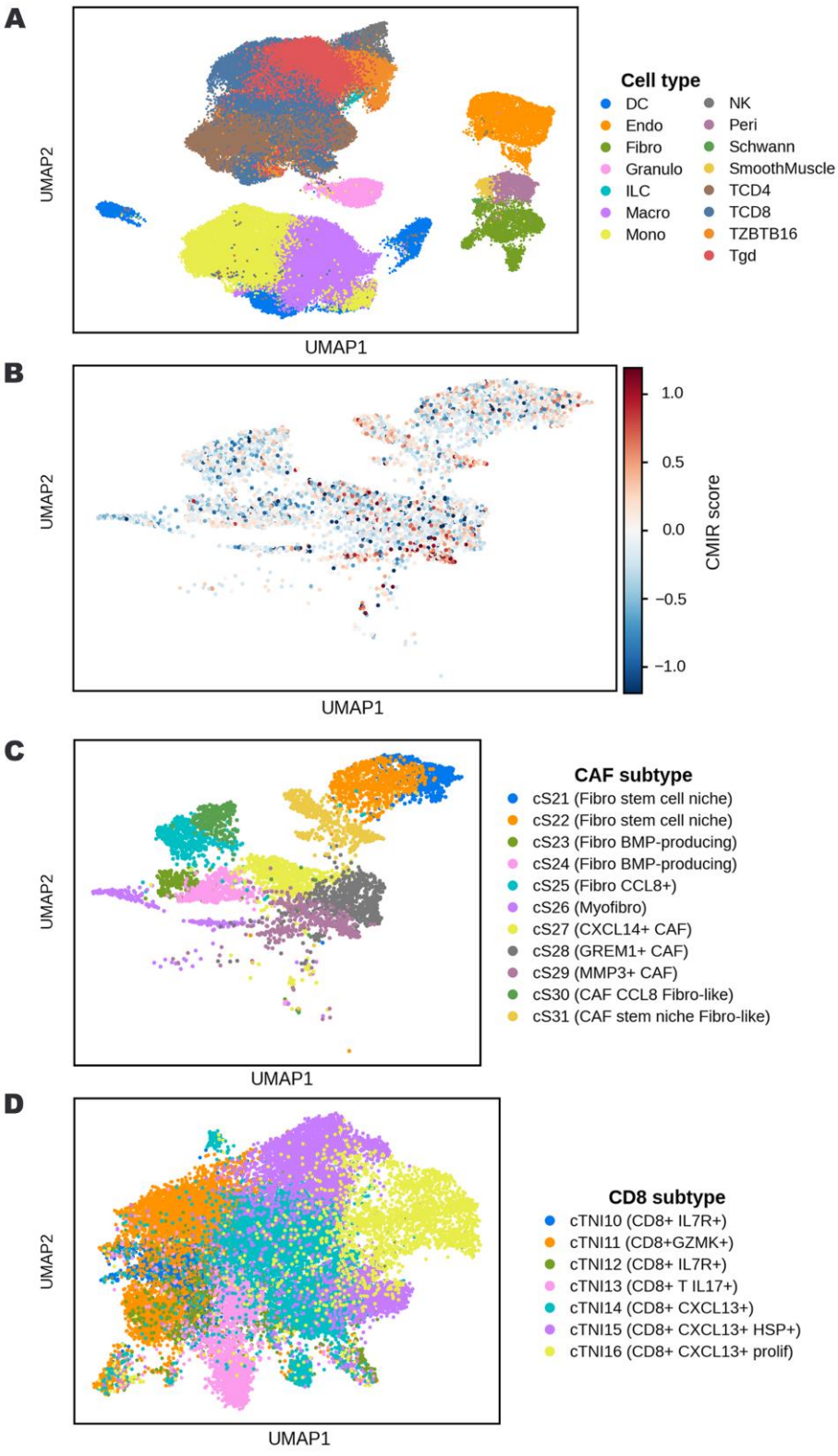


Fig. 7 Chemokine, effector molecule, and stromal ligand-receptor interactions between CAFs and CD8$^+$ T cells

A The LIANA tool was used to evaluate the interaction specificity and magnitude of chemokine ligand-receptor pairs sent by each CAF subpopulation; **B** The proportion of expressing cells and the mean expression of *CXCL9*, *CXCL10*, *CXCL11*, and their receptor *CXCR3* were evaluated across CAF, myeloid, and CD8$^+$ T cell subpopulations; **C** Canonical non-chemokine ligand-receptor pairs between CAF subpopulations and CD8$^+$ T cells were evaluated; **D** Extracellular matrix-related ligand-receptor pairs between CAF subpopulations and CD8$^+$ T cells were evaluated. Bubble size represents the mean ligand-receptor expression or the proportion of expressing cells, and color intensity represents -log10(specificity rank) or mean expression. Abbreviations: CAF, cancer-associated fibroblast; LR, ligand-receptor; DC, dendritic cell; Mono, monocyte; Macro, macrophage.

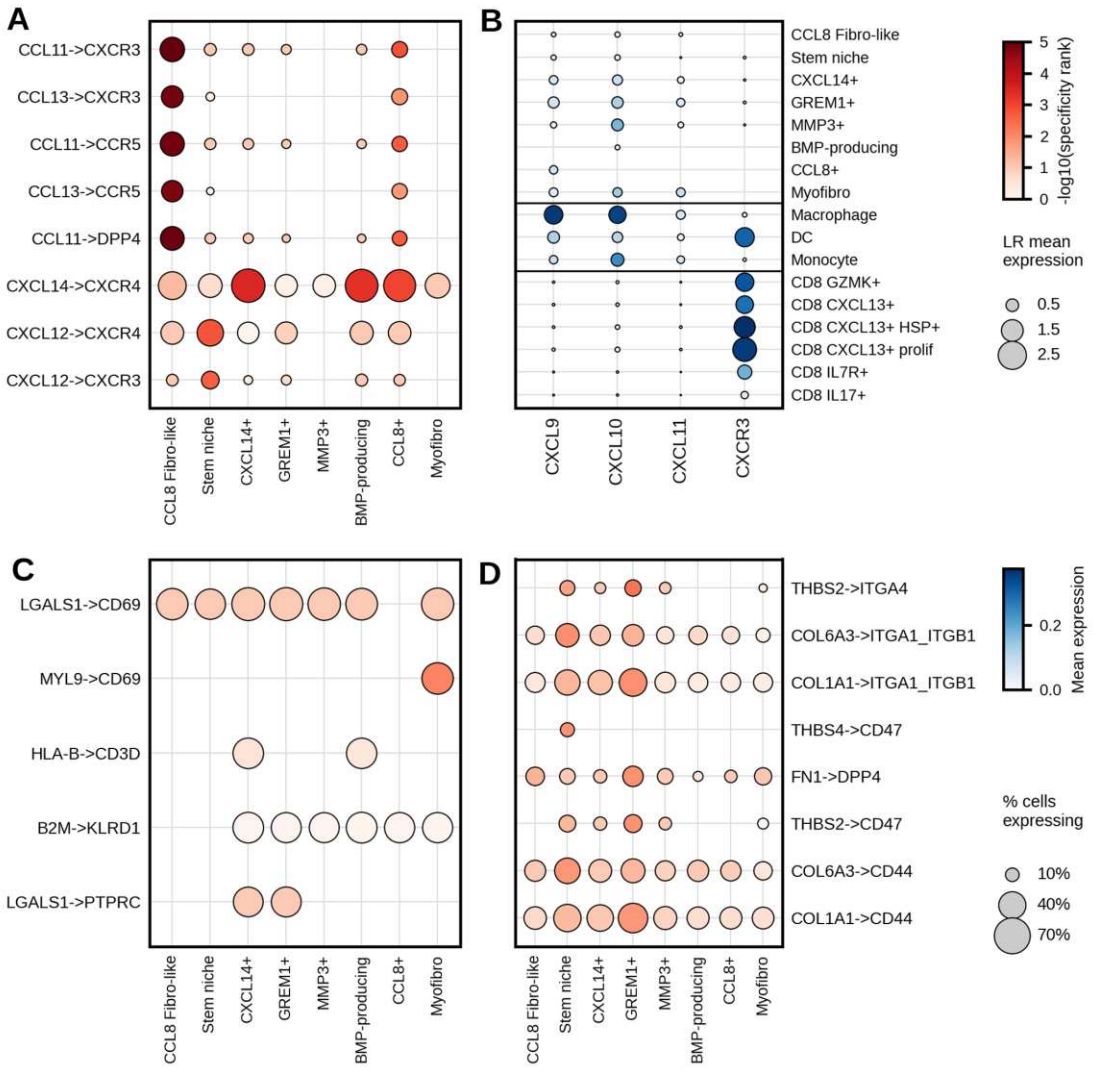
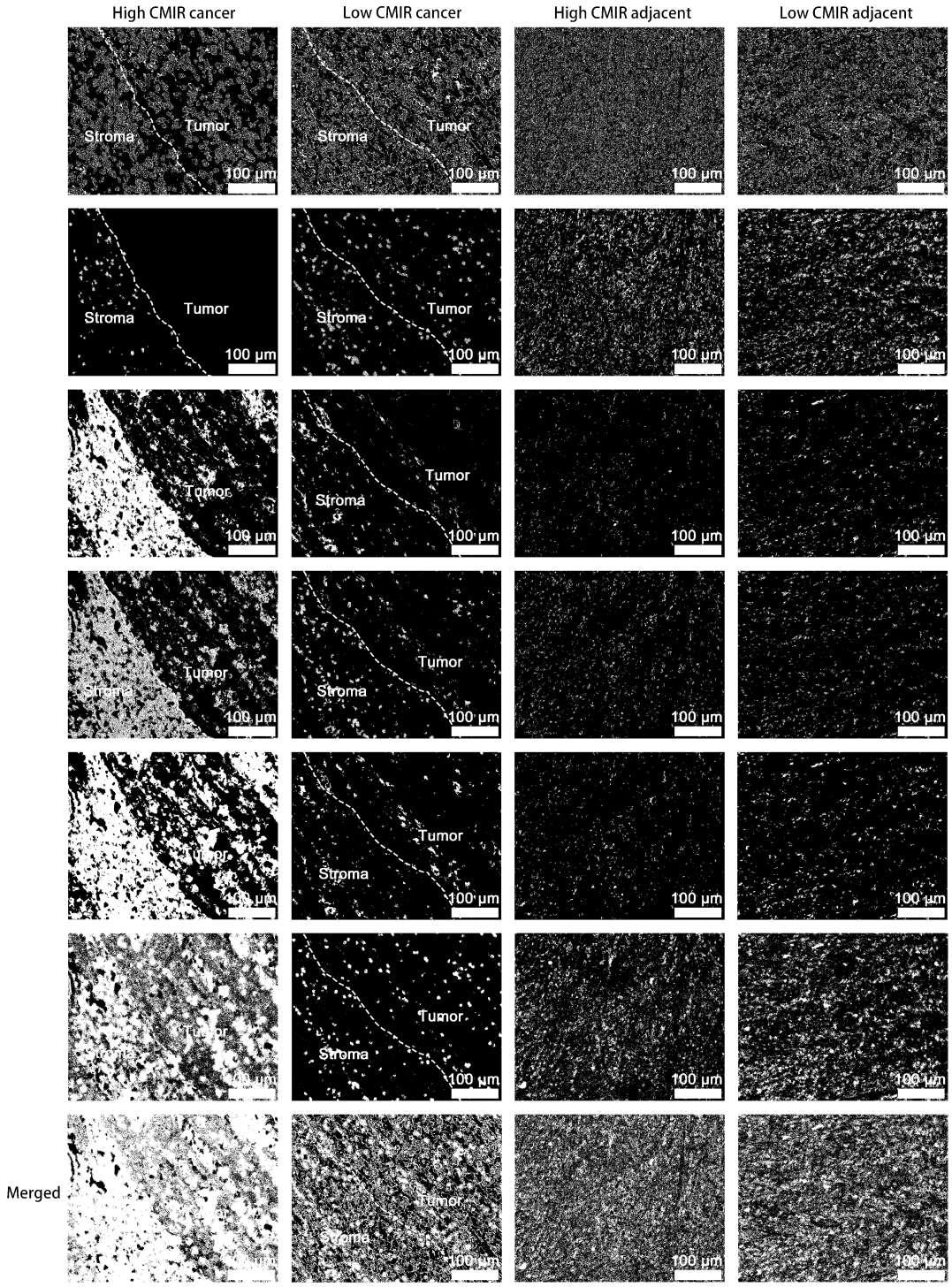


Fig. 8 Expression and spatial localization of microenvironmental markers in colorectal cancer and adjacent tissues detected by multiplex immunofluorescence

The panels show DAPI-counterstained nuclear signals and the single-channel and merged (Merged) expression of CD8, FAP, α-SMA, CD73, and KLF2 in tissue sections from the four groups of high-CMIR tumor, low-CMIR tumor, high-CMIR adjacent, and low-CMIR adjacent tissues, with dashed lines demarcating the stroma and tumor regions. White scale bars represent 100 μm. Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; FAP, fibroblast activation protein; α-SMA, α-smooth muscle actin; DAPI, 4',6-diamidino-2-phenylindole.



887

888 **Fig. 9 Multiplex immunofluorescence quantification of tumor microenvironment**

889 **protein expression, spatial distances, and colocalization metrics**

890 **A** Dunn's test was used to compare $CD8^+$ T cell density, CD8-FAP spatial distance, and

891 CD8-tumor epithelium physical distance among groups; **B** Differences in the FAP-

positive and α-SMA-positive area fractions among groups were evaluated; **C** The distribution of the proxy CMIR score constructed from the four mIF proteins was evaluated; **D** Differences in the CD73/FAP colocalization coefficient (Pearson), the mean fluorescence intensity of CD73, and the mean fluorescence intensity of KLF2 among groups were evaluated. Data are presented as violin plots combined with box plots (medians and interquartile ranges); multiple comparisons were corrected with BH-FDR; *** indicates $P < 0.001$, ** indicates $P < 0.01$, * indicates $P < 0.05$, and ns indicates $P \geq 0.05$. Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; FAP, fibroblast activation protein; α-SMA, α-smooth muscle actin; mIF, multiplex immunofluorescence; BH-FDR, Benjamini-Hochberg false discovery rate; AU, arbitrary units; CC, correlation coefficient; High-Ca, high-CMIR cancer; Low-Ca, low-CMIR cancer; High-Adj, high-CMIR adjacent; Low-Adj, low-CMIR adjacent.

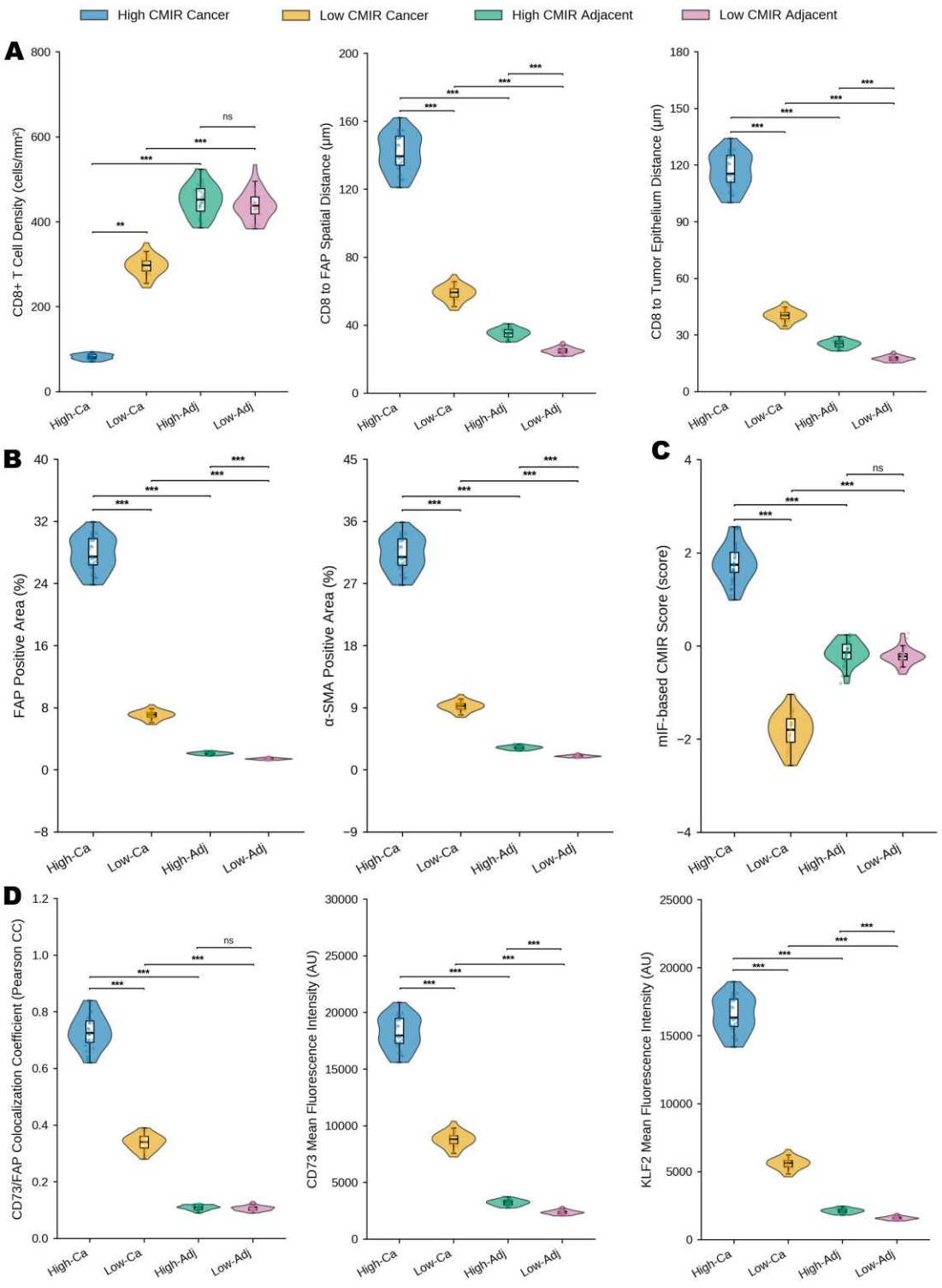


Fig. 10 Regulatory effects of KLF2 intervention on CAF activation markers and CD73 expression

A The relative KLF2 mRNA expression was measured by RT-qPCR in the four groups of Control CAF, KLF2-KD CAF, KLF2-OE CAF, and KLF2-OE^{+}CD73i CAF; **B**

Protein bands of FAP, CD73, α-SMA, KLF2, and the loading control GAPDH were detected by Western blot in CAFs of each group, followed by densitometric quantification. Data are presented as mean ± standard deviation; *** indicates $P < 0.001$, ** indicates $P < 0.01$, * indicates $P < 0.05$, and ns indicates $P \geq 0.05$. Abbreviations: CAF, cancer-associated fibroblast; KD, knockdown; OE, overexpression; CD73i, CD73 inhibitor; FAP, fibroblast activation protein; α-SMA, α-smooth muscle actin; GAPDH, glyceraldehyde-3-phosphate dehydrogenase.

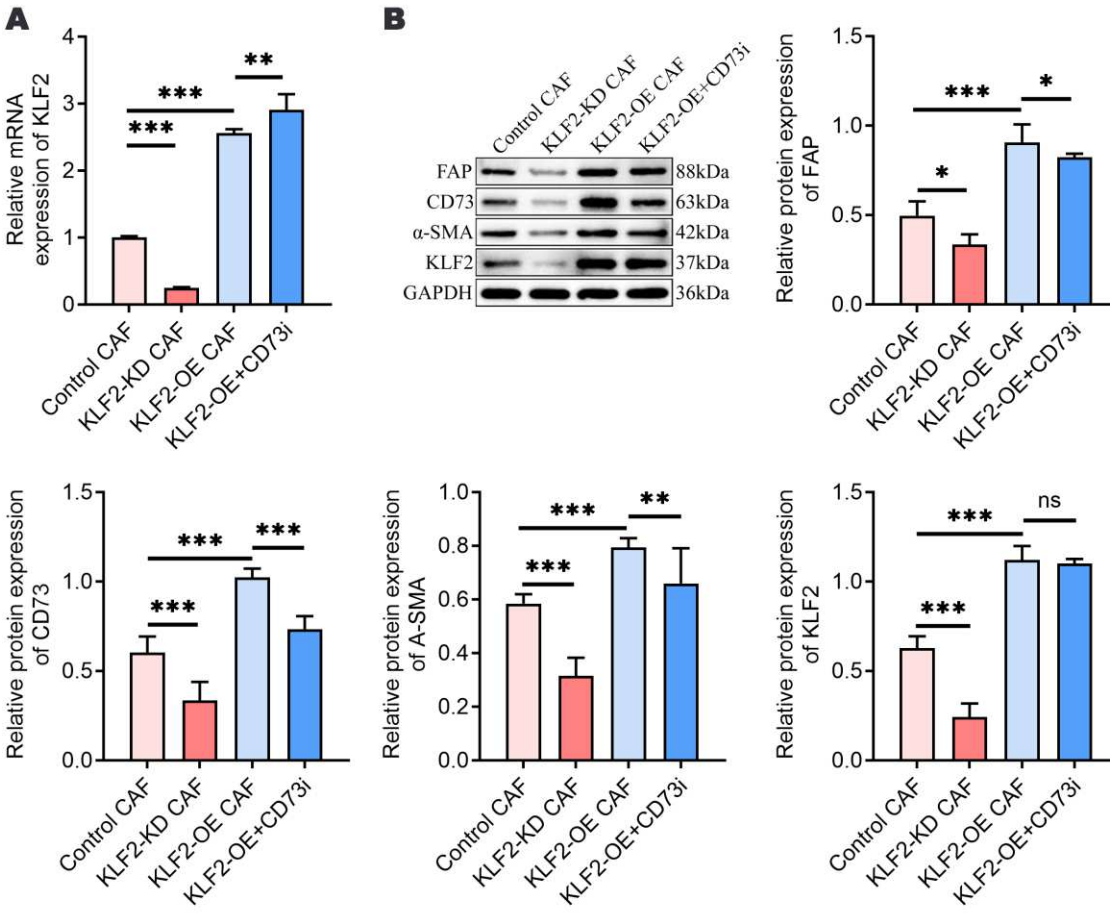
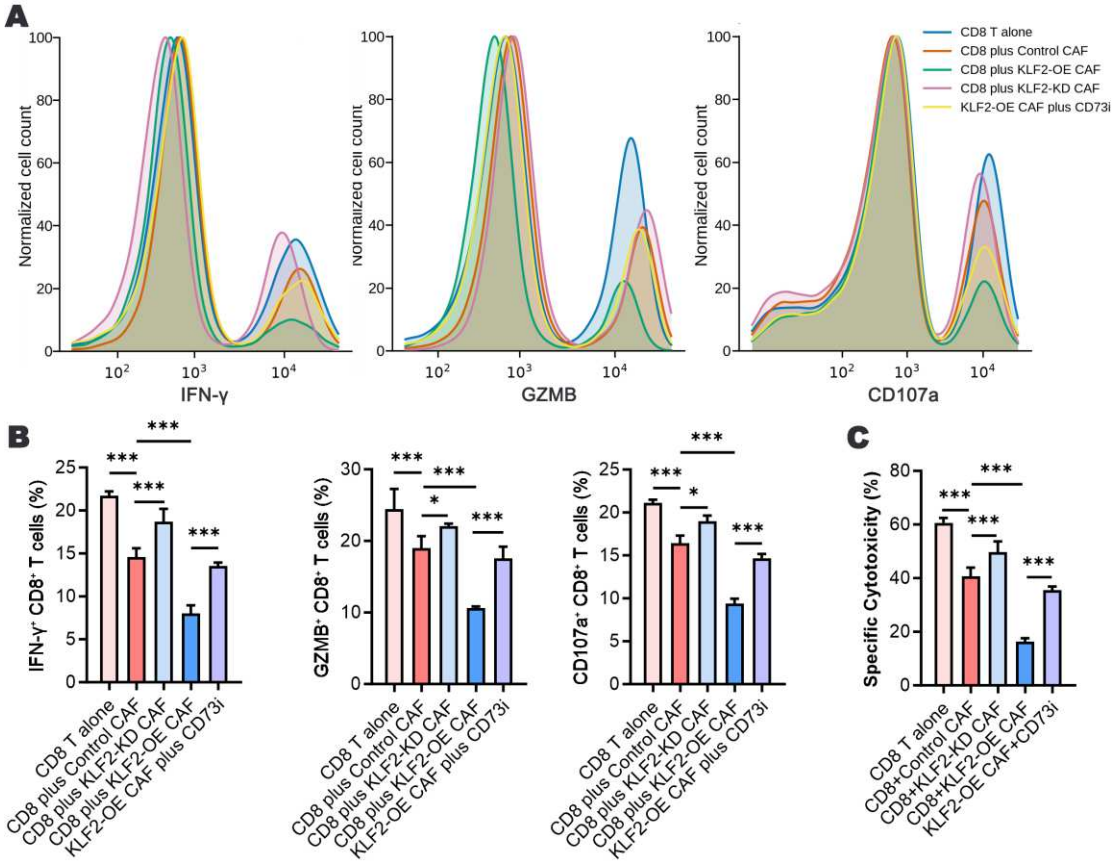


Fig. 11 Regulatory role of the KLF2-CD73-CAF axis in CD8$^+$ T cell effector function and tumor cell killing activity

A Representative flow cytometry histograms demonstrating the fluorescence intensity and expression distribution of IFN-γ, GZMB, and CD107a in CD8$^+$ T cells across the

experimental groups. **B** Flow cytometric quantification comparing the percentages of IFN-γ$^+$, GZMB$^+$, and CD107a$^+$ CD8$^+$ T cells among groups. **C** The specific cytotoxic killing rate (%) of colorectal cancer cells by CD8$^+$ T cells evaluated by LDH release assay across groups. Data are presented as mean ± standard deviation; *** indicates $P < 0.001$ and * indicates $P < 0.05$. Abbreviations: CAF, cancer-associated fibroblast; KD, knockdown; OE, overexpression; CD73i, CD73 inhibitor; IFN-γ, interferon-γ; GZMB, granzyme B.



Suppl Fig 1 Evaluation of cross-platform standardization strategies for CMIR score portability and performance comparison

A The C-index was used to evaluate within-cohort standardization (method A: within-z), training-set-frozen standardization (method B: frozen-z), and rank-based transformation (method C: rank) in each cohort; **B** Calibration slopes of methods A and

B were evaluated in each cohort; **C** The percentage of patients classified as high CMIR risk under methods A and B was tabulated for each cohort; **D** Spearman correlation analysis was used to evaluate the correlation coefficient rho of CMIR scores calculated by methods A and B in each cohort. Dashed lines indicate the standard calibration slope value of 1.0 (B) and the 50% high-risk proportion reference line (C). Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; C-index, concordance index; TCGA, The Cancer Genome Atlas.

Suppl Fig. 2 Purity adjustment, prognostic analysis across molecular feature subgroups, and comparison of CMIR scores among molecular subtypes

A Changes in Spearman correlation coefficients between the CMIR score and microenvironment cell scores before and after purity adjustment were evaluated; **B** Univariate Cox regression was used to estimate hazard ratios for overall survival (OS) per one-standard-deviation increment in the CMIR score across molecular feature subgroups; **C-D** Wilcoxon rank-sum tests were used to compare CMIR scores between the pMMR and dMMR groups (C) and between the *KRAS* wild-type (WT) and mutant (M) groups (D). Data are presented as correlation coefficients, hazard ratios, and 95% confidence intervals. Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; CAF, cancer-associated fibroblast; MMR, mismatch repair; dMMR, mismatch repair-deficient; pMMR, mismatch repair-proficient; KRAS, KRAS proto-oncogene; CIMP, CpG island methylator phenotype; CIN, chromosomal instability; BRAF, BRAF proto-oncogene; HR, hazard ratio; SD, standard deviation; EST, ESTIMATE algorithm.

Suppl Fig. 3 CMIR-related pathway enrichment, transcription factor activity, and key gene correlations in CAFs

A GSEA was used to evaluate normalized enrichment scores (NESs) of CMIR-related pathways in CAF pseudo-bulk transcriptomes; **B** A heatmap was used to display the activity distribution of key transcription factors across CAF subtypes; **C** Spearman correlation analysis was used to evaluate the correlation coefficients between the expression of *KLF2*, *CXCL9*, *NT5E*, *CXCL12*, and other genes and the CMIR score in CAFs. *** indicates $P < 0.001$. Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; CAF, cancer-associated fibroblast; GSEA, gene set enrichment analysis; NES, normalized enrichment score.

Suppl Fig 4 Spatial colocalization of the CMIR score and the CAF signature in spatial transcriptomics

A-C Spatial distribution maps of the CMIR score (A), CAF signature score (B), and CD8 signature score (C) across spatial transcriptomic spots; **D** Spearman correlation analysis was used to evaluate the correlation between the CMIR score and the CAF signature score across spots. Data are presented as expression color gradients or scatter plots with fitted lines. Abbreviations: CMIR, colorectal cancer fibroblast stromal-immune exclusion risk; CAF, cancer-associated fibroblast; spot, spatial transcriptomic capture site.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- CMIRSupplementaryCode.zip
- SupplementaryTable1.docx
- Supplementarywesternblotrawdata.docx
- FigureS1.tif
- FigureS2.tif
- FigureS4.tif