Deciphering colorectal carcinoma prognosis and therapy response via immunogenic cell death gene analysis and validation

Jinglu Yu [1], **Yongkang Cui** [2], **Lei Chen** [3], **Shuang Li** [2], **Yabin Gong** [Corresp., 4], **Zhenye Xu** [Corresp. 3]

[1] Department of Oncology, Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, Jingan District, China

[2] Department of Gastroenterology, Baoshan Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, Baoshan District, China

[3] Department of Oncology, Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, Xuhui District, China

[4] Department of Oncology, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, Xuhui District, China

Corresponding Authors: Yabin Gong, Zhenye Xu
Email address: gongyabin5791@126.com, ZhenYe_Xu@126.com

While the significance of Immunogenic Cell Death (ICD) in oncology is acknowledged, its specific impact on colorectal carcinoma remains underexplored. In this study, we delved into the role of ICD in colorectal carcinoma, a topic not yet comprehensively explored. A novel ICD quantification system was developed to forecast patient outcomes and the effectiveness of immunotherapy. Utilizing single-cell sequencing, we constructed an ICD score within the Tumor Immune Microenvironment (TIME) and examined Immunogenic Cell Death Related Genes (ICDRGs). Using data from TCGA and GEO, we discovered two separate molecular subcategories within 1184 patients diagnosed with Colon adenocarcinoma/Rectum adenocarcinoma Esophageal carcinoma (COADREAD). The ICD score was established by Principal Component Analysis (PCA), which classified patients into groups with low and high ICD scores. Further validation in three independent cohorts confirmed the model's accuracy in predicting immunotherapy success. Patients with higher ICD scores exhibited a "hot" immune phenotype and showed increased responsiveness to immunotherapy. Key genes in the model, such as AKAP12, CALB2, CYR61, and MEIS2, were found to enhance COADREAD cell proliferation, invasion, and PD-L1 expression. These insights offered a new avenue for anti-tumor strategies by targeting ICD, marking advances in colorectal carcinoma treatment.

Deciphering Colorectal Carcinoma Prognosis and Therapy Response via Immunogenic Cell Death Gene Analysis and Validation

Jinglu Yu [2, 5], Yongkang Cui [4, 5], Lei Chen [1], Shuang Li [4], Yabin Gong [3] * and Zhenye Xu [1] *

[1] Department of Oncology, Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China
[2] Department of Oncology, Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China.
[3] Department of Oncology, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China.
[4] Department of Gastroenterology, Baoshan Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China.
[5] The authors contributed equally to this work.

Corresponding Author:
Zhenye Xu [1]
Wanping South Road, Shanghai, 200032, China
Email address: Zhenye_Xu@126.com
Yabin Gong [3]
Ganhe Road, Shanghai, 200437, China
Email address: gongyabin5791@126.com

Abstract

While the significance of Immunogenic Cell Death (ICD) in oncology is acknowledged, its specific impact on colorectal carcinoma remains underexplored. In this study, we delved into the role of ICD in colorectal carcinoma, a topic not yet comprehensively explored. A novel ICD quantification system was developed to forecast patient outcomes and the effectiveness of immunotherapy. Utilizing single-cell sequencing, we constructed an ICD score within the Tumor Immune Microenvironment (TIME) and examined Immunogenic Cell Death Related Genes (ICDRGs). Using data from TCGA and GEO, we discovered two separate molecular subcategories within 1184 patients diagnosed with Colon adenocarcinoma/Rectum adenocarcinoma Esophageal carcinoma (COADREAD). The ICD score was established by Principal Component Analysis (PCA), which classified patients into groups with low and high ICD scores. Further validation in three independent cohorts confirmed the model's accuracy in predicting immunotherapy success. Patients with higher ICD scores exhibited a "hot" immune phenotype and showed increased responsiveness to immunotherapy. Key genes in the model, such as AKAP12, CALB2, CYR61, and MEIS2, were found to enhance COADREAD cell proliferation, invasion, and PD-L1 expression. These insights offered a new avenue for anti-tumor strategies by targeting ICD, marking advances in colorectal carcinoma treatment.

Introduction

Colorectal carcinoma ranks as a leading gastrointestinal cancer and holds the third position in global cancer mortality. It presents a 5-year survival rate near 65% [1]. Surgical resection remains the cornerstone for treating early-stage colorectal cancer, while advanced stages often benefit from a regimen combining chemotherapy and targeted therapies [2]. The advent of immune checkpoint inhibitors (ICI) has marked a revolutionary shift in oncological treatments, offering substantial clinical advantages [3]. Despite these advancements, the response to ICI is limited to a small patient subset [4,5], underscoring the urgency to better predict outcomes and responses in COADREAD, and to develop integrated approaches for overcoming immune resistance. Historically, biomarker research in oncology predominantly utilized RNA sequencing (RNA-Seq) of whole tumor tissues [6], capturing only a generalized genetic snapshot from a diverse cell population. This approach resulted in ICI biomarkers with suboptimal predictive accuracy. The introduction of scRNA-Seq has transformed this area by allowing for the dissection of gene expression on a cellular level, opening up possibilities for discovering improved new biomarkers [7].

ICD is a form of programmed cell death that enhances adaptive immunity by considering its physical, chemical, and operational aspects [8]. Tumor cell death, induced by external stimuli, releases various danger signals including high mobility group protein B1, surface-exposed calreticulin, type I interferon. These molecules convert non-immunogenic cells into immunogenic ones and serve as ligands for pattern recognition receptors. This interaction triggers the recruitment and activation of antigen-presenting cells, thereby initiating tumor-specific immune responses [9,10]. Primary stressors inducing ICD include chemotherapy drugs, targeted agents, radiotherapy, and intracellular pathogens [11]. Researches indicated that the synergistic application of chemotherapy and immunotherapy can enhance tumor cell immunogenicity, boost anti-cancer immune reactions [12]. This interplay positions ICD as a promising biomarker, intricately linked to patient prognosis and immunotherapy responsiveness. Previous studies have begun exploring ICD's clinical utility, particularly as a prognostic biomarker in cancers like clear cell renal cell carcinoma, melanoma[13,14]. However, there remains unknown in leveraging ICD for predicting responses to immunotherapy and chemotherapy in colorectal cancer. Addressing this gap was the focal point of our study, marking a step forward in this research domain.

In this investigation, our objective was to amalgamate single-cell and bulk RNA sequencing data to decipher the nuances of immune cell death in COADREAD patients, assessing its biomarker potential for predicting ICI outcomes. Additionally, we sook to elucidate the molecular and immune profiles associated with immune cell death and their implications for COADREAD prognosis. By analyzing ICDRGs, we have distinguished two subtypes and developed a robust predictive model, the ICD Score. This scoring system not only quantifies immune cell death but also holds promise as an instrumental marker in the realm of personalized precision medicine.

Materials & Methods

Data Sources and Acquisition

The single cell data for this research was acquired from the GEO Database (http //www.ncbi.nlm.nih.gov/geo/) [15], datasets GSE166555, which consisted of 12 samples of colorectal carcinoma. The TCGA database at https //portal.gdc.cancer.gov [16] provided bulk RNA-seq data and clinical details for 380 COADREAD patients, GEO database also provided additional datasets, GSE17538 (n=238) and GSE39582 (n=566).

Examining RNA sequencing at the individual cell level

The scRNA-seq data in R was analyzed using Seurat (v4.1.1) [17]. Initial quality control was performed using the “Seurat” package to filter out cells that did not meet specific criteria. The criteria for selecting the dataset included nFeature_RNA between 200 and 7500, nCount_RNA between 200 and 35,000, and a mitochondrial gene percentage below 10%. After these criteria, the dataset consisted of 66,050 cells and 21,753 genes. The dataset was then batch corrected using the 'harmony' package. Normalization of the data was achieved by the 'ScaleData' function. PCA was performed to extract the top 28 principal components using the top 2,000 genes with high variability. Unsupervised clustering and visualization of cell subpopulations on a two-dimensional map was carried out using Uniform Manifold Approximation and Projection (UMAP) [18].

To compare gene expression differences between clusters, the 'FindAllMarkers' function was used. Marker genes for each cluster were determined based on a log2 (fold change) value greater than 1, an adjusted *P*-value lower than 0.05, and a resolution of 0.8. Finally, the cell subpopulations in the various clusters were annotated using the 'SingleR' package along with the CellMarker database [19] and PanglaoDB database [20].

Using the R function 'FindAllMarkers', we examined genes that were expressed differently in various cell types. In order to evaluate the diversity in sets of genes, we employed Gene Set Variation Analysis (GSVA) (version 1.40.1) [21]. Scores were computed for 50 hallmark gene sets acquired from MSigDB (https://www.gsea-msigdb.org/gsea/msigdb) [22] and 34 ICD genes from the publication referenced by PMID 27057433 [23]. Cell clusters were classified into high and low ICD score groups based on the median scores. Heatmaps were used to visualize the correlations between ICD and Hallmark scores, generated using the R packages 'pheatmap' and 'corrplot' with default thresholds.

Bulk RNA-Sequencing Data Integration

For the bulk RNA-seq datasets, we integrated the data and corrected batch effects using the ‘limma’ and ‘sva’ packages. Subsequently, a log2 transformation was applied to all datasets. The integration process's effectiveness, before and after, was visualized using PCA plots, created with the ‘FactoMineR’ and ‘factoextra’ packages in R.

Development of ICDRGs Prognostic Signature

We initially examined correlations between ICDRGs in COADREAD patients. Genes associated with prognosis were identified by univariate Cox regression analysis using the survival package. The analysis was conducted with a significance threshold of $p < 0.05$, as stated in the source https://rdocumentation.org/packages/survival/versions/2.42-3.

Molecular Subtyping via Consensus Clustering

To categorize molecular subtypes based on ICD prognostic signature genes, we employed the 'ConsensusClusterPlus' package (1.59.0) in R. This involved utilizing the K-means algorithm for clustering with a range of k = 2-10 to determine the optimal cluster number. To ensure stability, this clustering process was iteratively conducted 1,000 times.

ICDRGs Prognostic Signature Development

By conducting univariate Cox regression analysis using the survival package at a significance level of $p < 0.05$, we identified genes linked to prognosis in COADREAD patients, while examining the correlations among ICDRGs.

Molecular Subtypes and Clinicopathological Correlation

We conducted an analysis to assess the clinical significance of molecular subtypes in COADREAD by examining their correlation with clinicopathological characteristics such as age, gender, grade, and TNM stage. The correlation was visualized through heatmaps using the 'pheatmap' package in R. Additionally, we performed survival analysis using the 'survminer' package to evaluate the prognostic implications of these subtypes.

Pathway Enrichment and Immune Infiltration Analysis

In order to understand the TIME in different molecular subtypes, we employed GSVA using hallmark, KEGG, and Reactome gene sets obtained from the MSigDB database. By utilizing the 'estimate' package in R, the ESTIMATE algorithm (https://sourceforge.net/projects/estimateproject/)[24] computed stromal, immune, and ESTIMATE scores in tumor tissues. In addition, we utilized single-sample gene set enrichment analysis (ssGSEA) to investigate the immune infiltration pattern in COADREAD, specifically targeting 23 types of immune cells. The distribution of these immune cells in the samples was quantified and visualized using box plots.

Molecular Analysis and DEGs Identification

The GSCA datasets [25] were used to analyze prognostic genes, including mutation, SNV, CNV, and Methylation. DEGs between clusters were identified using the 'limma' package, with a threshold of |log2 (Fold Change)| > 1 and an adjusted p-value < 0.05. Subsequently, the R package 'clusterProfiler' was employed for conducting GO and KEGG analyses. Significant DEGs with prognostic value were detected using univariate Cox regression and presented in a

forest plot.

ICD Score System Development

The ICD Score system was created through the use of PCA [26]. This score represents the level of individual ICD regulation and is derived by summing the values of PC1 and PC2, which capture the most significant variances in the expression data. The ICD score was calculated for each sample, and an optimal threshold value was determined to maximize the accuracy of prognostic predictions. Based on this threshold, patients from the datasets were categorized into either low or high ICD score groups. Survival analyses were conducted to evaluate the prognostic significance of the ICD score in COADREAD. Box and bar diagrams were used to visualize the correlation between the ICD score and clinical features.

Immune landscape analysis of ICD score

To demonstrate the connections between the ICD score and the infiltration of immune cells, as well as the expression of cytokines, chemokines, and their corresponding receptors, we utilized heatmaps. The relationship between the ICD score and hallmark pathways was evaluated using GSVA. Furthermore, we analyzed the mRNA expression of immune checkpoints and gene mutations in COADREAD patients from the TCGA database. We identified and visualized somatic mutations in different ICD score groups using waterfall plots and forest plots, providing insights into the genetic variations associated with ICD scores.

Immunotherapy prediction and chemotherapy sensitivity analysis

The predictive value of the ICD score for immunotherapy response and chemotherapy sensitivity was assessed in our study. We analyzed the IMvigor210 cohort, which included 298 patients with urothelial carcinoma, to predict the immunotherapy response based on the ICD score. The data was processed using the R package 'IMvigor210CoreBiologies'. Additionally, we included the GSE61676 group, which consisted of transcriptomic data from advanced non-squamous non-small cell lung tumors treated with erlotinib and bevacizumab, to evaluate the predictive ability of the ICD score. Furthermore, information on the response of patients with Kidney Renal Clear Cell Carcinoma (KIRC) to nivolumab (anti-PD-1) was obtained from Checkmate-KIRC (PMID 32472114) [27]. Kaplan-Meier analysis was used to examine the prognostic significance of the ICD score in COADREAD. The ICD score was also employed to estimate the IC50 of commonly targeted therapeutic drugs using the 'pRRophetic' R package and data from the Genomics of Drug Sensitivity in Cancer (GDSC) database [28].

RNA Extraction and PCR Analysis

Total RNA was extracted from colorectal cells using the Trizol method (Invitrogen), followed by cDNA synthesis using the Prime Script RT reagent Kit (Takara). Gene expression was quantified using TB Green Premix Ex Taq (Takara) and normalized to GAPDH. Supplementary Table 1 provides primer details.

199
200 ### Cell Culture
201 Human colorectal cell lines including NCM460, HCT116, HT29, and LOVO were cultured in
202 RPMI 1640 or DMEM (Gibco) supplemented with 10% FBS. The cells were incubated at 37°C
203 in a 5% CO2 environment. To validate their authenticity, the cell lines obtained from the Chinese
204 Academy of Sciences underwent STR profiling and mycoplasma testing, which can be verified at
205 http://www.cellbank.org.cn/.
206
207 ### siRNA Transfection
208 siRNAs were used to target AKAP12, CALB2, CYR61, MEIS2, or a control, which were then
209 transfected into HCT116 and HT29 cells. The cells were seeded at a density of 3×10^5 cells per
210 dish (60 mm). Transfection was performed using Lipofectamine RNAiMax, followed by a 24-
211 hour culture in RPMI 1640 or DMEM supplemented with 10% FBS. The siRNA sequences can
212 be found in Supplementary Table 1.
213
214 ### Proliferation Assay
215 A total of 3000 cells were placed in 96-well plates. The cells' viability was assessed every 24
216 hours using a CCK-8 kit manufactured by Dojindo. The microplate reader (Thermo Fisher, USA)
217 was used to measure the absorbance at 450 nm. The obtained data was then analyzed using
218 GraphPad Prism 9.5.0.
219
220 ### Transwell Migration Assay
221 A total of 100,000 cells were positioned in the top compartment on a membrane coated with
222 Matrigel. RPMI 1640 or DMEM (without FBS) was added above, while a medium containing 20%
223 FBS was placed below. Post 24-hour incubation, cells were fixed, stained with Giemsa crystal
224 violet, and imaged. Non-migrated cells from the upper chamber were removed.
225
226 ### Statistical Analysis
227 The R software (version 4.1.2) was utilized for data analysis. Correlations were assessed using
228 Pearson or Spearman methods, and the Wilcoxon test was employed to compare two groups.
229 Survival outcomes among different subgroups were analyzed using Kaplan–Meier and log-rank
230 tests. The prognostic significance of ICDRGs and clinicopathologic attributes was evaluated
231 through univariate Cox regression. The Student's t-test was used to analyze the qRT-PCR results.
232 A p-value less than 0.05 indicated statistical significance, denoted as $*p < 0.05$, $**p < 0.01$, and
233 $***p < 0.001$.
234
235 ### Ethical Considerations
236 The research did not involve human participants or animals as per the authors' confirmation.
237
238 ## Results

Single Cell Sequencing Data Interpretation

Reducing Complexity

Our analysis commenced with the single-cell sequencing dataset GSE166555, focusing on COADREAD. We integrated data from 12 distinct samples, ensuring minimal batch effect. Employing the Uniform Manifold Approximation and Projection (UMAP) algorithm, we segregated the cells into 28 identifiable clusters. By scrutinizing the expression of surface marker genes across these clusters, we pinpointed eight cell types: B cells, dendritic cells, mono-macrophages, natural killer (NK) cells, T cells, malignant cells, endothelial cells, and fibroblasts. These cell types were distinctively represented in different clusters, as illustrated in Figure 2 A.

Transcriptomic Landscape Analysis

Our analysis revealed the most influential genes, with the top five contributors detailed in Figure 2 B. In malignant cells, the highest expressed marker genes were KRT8, LGALS4, PIGR, ELF3, and CLDN4, while the least expressed were SRGN, VIM, TSC22D3, CXCR4, and RGS1. For monocyte-macrophages, the most abundantly expressed markers included AIF1, TYROBP, IL1-B, CXCL8, and S100A9, contrasting with the lowest expressed markers KRT8, LGALS4, PHGR1, PIGR, and JCHAIN. In NK cells, the highest expressed genes were TPSAB1, CPA3, MS4A2, HPGDS, and KIT, with the least expressed being KRT8, PHGR1, FXYD3, CLDN4, and IFI27. T cells exhibited high expression of CD3D, IL7R, CD2, CD3E, and TRAC, and low expression of LGALS4, KRT8, PIGR, IFI27, and CLDN3. Gene Set Variation Analysis (GSVA) further delineated the Hallmark pathways enriched in each cell type, as shown in Figure 2 C. B cells and T cells demonstrated similar pathway enrichment patterns, while fibroblasts and monocyte-macrophages shared a different, yet comparable pattern. B cells and T cells demonstrated similar pathway enrichment patterns, while fibroblasts and monocyte-macrophages shared a different, yet comparable pattern.

Expression Profiles of ICDRGs

In our analysis, 34 genes associated with ICD were identified using the "FindAllMarkers" function, enabling us to discern the expression patterns of ICDRGs across various cell types. These expression patterns were effectively visualized through bubble plots (Fig. 2 D). Notably, HSP90AA1 exhibited high expression in B cells, T cells, and NK cells, while showing lower expression in dendritic cells, monocyte-macrophages, malignant cells, endothelial cells, and fibroblasts. Conversely, HMGB1 was predominantly expressed in NK cells, T cells, and fibroblasts, but less so in B cells, dendritic cells, monocyte-macrophages, malignant cells, and endothelial cells. CALR expression was significantly elevated in B cells compared to other cell types. PDIA3 showed a similar pattern, with heightened expression in B cells and NK cells, but lower levels in other cell types. Additionally, IFNGR1, CD4, LY96, and IL1B were predominantly expressed in monocyte-macrophages, while IFNGR1 and CD4, along with MYD88, were highly expressed in dendritic cells.

ICD Score Elucidation and Implications

In our analysis, cells were scored for expression of ICDRGs using the GSVA method, as shown in Figure 2 E. This led to categorization into low and high ICD score clusters (Figure 2 F), with the score distribution visualized in box plots (Figure 2 G). Monocyte-macrophages exhibited the highest ICD scores, in contrast to tumor cells, which scored the lowest. Composition charts displayed the cell type distribution in these clusters, based on count and proportion (Figure 2 H), highlighting a prevalence of tumor cells in low-score and T cells in high-score clusters. Correlations between ICD scores and Hallmark pathway scores (Figure 2 I) revealed positive associations with allograft rejection, complement, and inflammatory response pathways in monocyte-macrophages, NK cells, and endothelial cells. B cells, dendritic cells, and T cells showed negative correlations with pathways like KRAS signaling DN and WNT beta-catenin signaling. High TLS score group exhibited elevated expression of immune activation pathways (Figure 2 J), with inflammatory response associated pathways markedly enriched.

Comprehensive analysis of Bulk-Seq Data

Data from 1184 patients were aggregated from the TCGA (TCGA-COADREAD) and GEO databases (GSE39582, GSE17538). We addressed batch effects using the "limma" and "sva" methods. Figure 3 A-B showed a notable reduction in the dimensions of components 1 and 2 from 76.3% to 15.6%.

Delineation of ICD Molecular Subtypes and Their Clinical Relevance

Our survival analysis identified 12 ICDRGs - PIK3CA, P2RX7, NT5E, ENTPD1, IL1R1, LY96, BAX, CASP1, CASP8, IL17A, FOXP3, CXCR3 - significantly correlated with overall survival (OS) in COADREAD patients (Figures 3 D, $p < 0.05$). Univariate Cox regression highlighted that PIK3CA, P2RX7, NT5E, ENTPD1, IL1R1, LY96 positively impacted survival, whereas BAX, CASP1, CASP8, IL17A, FOXP3, CXCR3 had a negative association. We constructed an ICD network to elucidate these genes' interrelationships and prognostic significance (Figure 3 C). COADREAD patients were divided into two molecular subtypes, A and B, based on ICDRG expression profiles using a consensus clustering algorithm. Optimal separation (k=2) indicated distinct survival probabilities between these subtypes, with subtype A showing higher survival (log-rank test, $p = 0.001$; Figure 3 F). Most TLRGs were highly expressed in B cluster (Figure 3 G). Additionally, the variances between ICDRG expression and clinical features like age, stage, gender, recurrence, metastasis, fustat, and futime across the two molecular subtypes had no significant differences (Figure 3 H).

Analyzing ICD Subtypes Through GSVA

GSVA enrichment analyses on molecular subtypes using Hallmark, KEGG, and Reactome pathways showed subtype B's significant enrichment in inflammation-related pathways, including interferon gamma and alpha, IL2-STAT5, IL-6-JAK-STAT3, and Toll-like receptor and B cell receptor signaling (Figure 3 I-K), further establishing the link between subtype B and

inflammation, a key aspect of immunogenic cell death.

ICDRGs Analysis via GSCA Database

We analyzed ICDRGs through the GSCA database (http://bioinfo.life.hust.edu.cn/GSCA/#/). The analysis covered SNV percentages (Sup-Figure 1 A), mutation frequencies (Sup-Figure 1 B), CNVs (Sup-Figure 1 C), and correlations between CNV, methylation, and mRNA expression (Sup-Figures 1 D-F) in COADREAD. This provided insights into genetic and epigenetic variations of ICD signature genes.

Investigating Immune Infiltration in ICD Subtypes

PCA analysis demonstrated distinct ICD transcription profiles between subtypes A and B (Figure 4 A). Investigating ICDRGs' roles in the TIME of COADREAD, we found subtype B showed higher stromal, immune, and ESTIMATE scores (Figure 4 B). The estimate algorithm was used to calculate human immune cell subsets for each COADREAD sample in immune cell infiltration between the subtypes via ssGSEA (Figure 4 C). Specifically, cells like activated B and CD4 T cells, dendritic cells, MDSC, macrophages, and various T-helper cells showed lower infiltration levels in subtype A compared to subtype B. These findings suggested that subtype B is enriched in immune pathways and closely associated with TIME.

Exploring ICD-DEGs: Functional and Pathway Enrichment

We identified 448 differentially expressed genes (DEGs) between ICD molecular subtypes A and B, visualized in a volcano plot (Figures 4 D, $|\log FC|>1$, $P < 0.05$). GO and KEGG analyses were conducted to identify associated biological pathways. GO enrichment revealed DEGs predominantly involved in cytokine production regulation, leukocyte migration and mediated immunity, cell and leukocyte chemotaxis, immune receptor and chemokine activity, and pattern recognition receptor pathways (Figures 4 E). KEGG enrichment highlighted DEGs in cytokine-cytokine receptor interaction, chemokine signaling, viral protein interaction with cytokine and cytokine receptor (Figures 4 F). A cnetplot elucidated specific gene networks in these pathways, with the top 5 pathways shown in Figure 4 G. This enrichment analysis suggested ICD's significant role in immune activation and cytokine-chemokine interactions in COADREAD.

Development and Implications of ICD Score

Univariate Cox regression analysis was employed on the 448 DEGs to assess their prognostic significance, leading to the identification of 25 OS-related genes. These genes were represented in a forest plot (Figure 4 H). Subsequently, PCA was conducted on these 25 prognostic ICD-DEGs, dividing patients into low and high ICD score genomic subtypes (score=PCA1-PCA2).

Examining the Association with Somatic Mutations

The somatic mutations in the two ICD score clusters were further analyzed. The top ten mutated genes in the high ICD score cluster included TP53, APC, TTN, and KRAS (Figures 4 I upper),

while the low ICD score cluster showed a different gene mutation profile (Figures 4 I under), indicating a higher overall mutation frequency in the high score cluster. A forest plot highlighted the top 12 genes with the most significant variation in mutation frequency between the clusters (Figures 4 J). This analysis suggested a positive correlation between higher gene mutation frequency and variation in the high ICD score cluster.

ICD Score Correlation with TIME and Pathways

The relationship between the ICD score and hallmark pathway scores, analyzed using GSVA, revealed a positive correlation with immune activation pathways, including interferon gamma and alpha responses, inflammatory response, IL-2-STAT5 signaling, and TNFA signaling via NFKB. Conversely, a negative correlation was observed with pathways related to malignancy like peroxisome, MYC targets, oxidative phosphorylation, and the reactive oxygen species pathway (Fig. 4 K). This suggests the ICD score's alignment with immune-activated pathways. Further analysis of cytokine-chemokine networks highlighted significant enrichment of chemokines, interleukins, and interferons along with their receptors in the high-ICD cluster (Fig. 4 L). Spearman correlation analysis indicated a positive correlation of the ICD score with the infiltration of 23 immune cell types, suggesting a more robust immune component in the TIME of the high-score cluster and potentially better immune prognosis (Fig. 5 B).

ICD Score's Impact on COADREAD Prognosis

Survival analysis revealed that COADREAD patients in the low-score ICD cluster exhibited higher survival rates than those in the high-score cluster (Figure 5 A, $p < 0.001$). Significant variations in recurrence, metastasis, and disease stage were observed between the clusters, with advanced stages showing higher ICD scores (Fig. 5 C-J). Further analysis demonstrated elevated expression of immune checkpoints like PDCD1, CTLA4, TIGIT, and LAG3 in the high-score group (Figures 5 K-N), suggesting a potential for better response to immune checkpoint inhibitors in this group. This underlined the ICD score's relevance in predicting immunotherapy efficacy and regulating the tumor immune microenvironment in COADREAD.

Efficacy of ICD Score in Predicting Response to Immunotherapy

We assessed the prognostic value of the ICD score for immunotherapies, including PD-L1 blockade, in various cohorts. In the GSE61676 cohort, a higher response rate to bevacizumab combined with erlotinib was observed in the high-ICD cluster (57%) compared to the low cluster (36%) (Figure 5 P). This cluster also showed significantly longer overall survival (OS) (Figure 5 O, p=0.03). In the IMvigor210 cohort, patients' responses to anti-PD-L1 therapy varied. Patients with stable or progressive disease had lower ICD scores compared to those with complete or partial responses (Figure 5 R). The high-ICD cluster demonstrated significant clinical benefits and longer OS (Figure 5 Q, p=0.0036). Similarly, in the Checkmate cohort, the high-ICD cluster had significantly longer OS (Figure 5 S, p=0.0061) and a higher percentage of complete/partial responses (Figure 5 T). These findings indicated that the ICD score was a potential predictor of

immunotherapy efficacy in COADREAD.

Drug Sensitivity Analysis Based on ICD Score

Patients in the high-score cluster exhibited lower IC50 values for drugs like AP.24534, A.770041, ABT.263, AG.014699, AICAR, and AMG.706 (Figures 5 U-Z), indicating a higher sensitivity to these treatments. These findings suggested the utility of the ICD score as a predictor for selecting effective anticancer drugs.

ICDRGs' Role in COADREAD Cellular Behaviors

Our research involved RT-qPCR analysis of AKAP12, CALB2, CYR61, and MEIS2 in normal colon and COADREAD cell lines, revealing significant overexpression in tumor cells (Figure 6 A-D). We further investigated these genes' roles in COADREAD using siRNA knockdown in HCT116 and HT29 cells, with siRNA-1 and siRNA-2 selected for their high transfection efficiency (Sup-Figure 2 A-H). The impact on cell proliferation and invasion was assessed through CCK8 and Transwell assays, respectively. Knockdown of these four genes markedly inhibited proliferation and invasion in both cell lines (Figures 6 E-N). Additionally, their knockdown significantly reduced PD-L1 expression, suggesting potential in combining ICD inducers with PD-L1 inhibitors (Figure 6 P-Q). Immunohistochemical analysis from the HPA database further confirmed higher protein expression levels of these genes in COADREAD stroma (Figure 6 O).

In summary, AKAP12, CALB2, CYR61, and MEIS2 were key regulators significantly influencing the biological behaviors of COADREAD cells.

Discussion

Colorectal adenocarcinoma, a leading malignancy globally, is witnessing a surge in incidence across various regions. Despite recent advancements, its heterogeneous nature and aggressive behavior pose challenges in prognostic assessment[2]. Identifying novel biomarkers for personalized therapy development and prognosis improvement is thus a pressing need.

The pro-inflammatory immune microenvironment and neoantigen development in COADREAD have been linked to heightened immunotherapy response [29,30]. ICD involves pathways triggering intense inflammatory responses, potentially reshaping the tumor microenvironment and releasing tumor neoantigens [8]. Given the high heterogeneity of colorectal cancers and specific molecular subtypes' sensitivity to immunotherapy [31], ICD offers insights into predicting and identifying patient groups suited for immunotherapy, including those with low response rates.

Therefore, exploring ICD in COADREAD presents a novel approach to understanding its microenvironmental impacts, prognostic significance, and predictive value. Addressing the unclear role of ICD in COADREAD research is vital for advancing therapeutic strategies and patient management.

scRNA-seq stands out from traditional bulk RNA-seq by providing detailed insights into

individual gene expression levels across cell subpopulations. This technique is instrumental in identifying specific biomarkers and understanding cellular heterogeneity in various cancers, including COADREAD [32].

In our study, we merged the insights gained from both bulk RNA-seq and scRNA-seq to perform a thorough analysis of COADREAD. This included examining ICD expression profiles, conducting clustering analyses, assessing immune infiltration, exploring the mutation landscape, and screening for prognostic signature genes. The culmination of these analyses was the development of the ICD score, which has demonstrated considerable prognostic and predictive utility for immunotherapy responses in COADREAD.

In COADREAD, scRNA-seq analysis identified two ICD score clusters. The high-score cluster showed enriched inflammatory and TNFA signaling via NF-κB, indicating enhanced immune activation. Contrarily, these pathways were suppressed in the low-score cluster. Predominant in the high-score group were monocyte-macrophages and CD4+/CD8+ T cells, suggesting a robust anti-tumor immune response. These findings align with the concept that higher ICD scores correlate with more immunologically active "hot tumors."

We identified 12 ICD-related prognostic genes and divided COADREAD patients into two clusters based on these genes. Our analysis highlighted significant differences in prognosis, clinicopathological characteristics, immune infiltration, and pathway enrichment between the clusters, emphasizing ICD's role in regulating immune interactions in COADREAD.

In our study, we identified 448 differentially expressed genes (DEGs) between ICD score clusters in COADREAD, using univariate Cox regression analysis to pinpoint 25 OS-related genes. These genes, integral to the ICD score model developed via Principal Component Analysis, included SCG2, CALB2, HOXC6, CRYAB, ZNF532, and others, each substantiating their roles in tumor progression, metastasis, and response to therapy.

For instance, SCG2, known for its prognostic value in immune infiltration and macrophage polarization, aligns with its role in chemotherapy and immunotherapy in colorectal cancer [33]. CALB2's modulation of 5-fluorouracil sensitivity [34], HOXC6's association with poor OS and immunogenicity [35], and CRYAB's involvement in CSC formation via the Wnt/β-catenin pathway [36] all highlighted their significance. Similarly, ZNF532's role in oncogenic chromatin propagation in NUT midline carcinoma [37] and AKAP12's influence on colon cancer metastasis [38] were noteworthy. Additionally, FABP4's influence on colon cancer invasion [39], MEIS2's role in cell migration [40], and SLC2A3's correlation with prognosis and immune signature [41] were pivotal. Other genes like SPARCL1, CYR61, CYP1B1, and COLEC12 have been implicated in various cancer dynamics, from tumor microenvironment modulation to disease progression and treatment response [42-45].

Our study, for the first time, reported the comprehensive effect of these ICD-regulated genes in COADREAD, offering new insights into their potential as therapeutic targets and prognostic markers in this complex disease.

Our study revealed stark differences in intrinsic properties between high and low ICD score groups in COADREAD. The high-score group, despite a poorer prognosis, exhibited more

immune cell infiltration, a higher tumor mutation load, and elevated levels of immune checkpoint expression, including increased cytokines, chemokines, and their receptors. Furthermore, a positive correlation was observed between the ICD score and immune-activated pathways.

Interestingly, the response to immunotherapy varied based on the ICD score. Patients in the high-ICD score cluster showed a potential benefit from ICIs and targeted therapies. This aligns with their higher immune mutational load and more active immune environment, as confirmed by previous studies highlighting the role of both immunotherapy and chemotherapy in inducing ICD and bolstering anti-tumor immunity [8,9,12].

Our findings underscored that a high ICD score, indicative of increased immune cell infiltration and activated anti-tumor pathways, was predictive of a favorable response to immunotherapy. This suggested the close relationship between ICD status and the effectiveness of immunotherapy and other immunogenic treatments in COADREAD.

Our study faced limitations such as its retrospective nature and reliance on public sequencing data. Future studies should validate our findings in larger, prospective clinical trials, and include proteomic cross-validation for clinical applicability. Further basic research is needed to understand how ICD-related genes affect prognosis and response to therapy. Additionally, examining single-cell changes during anti-PD1 treatment could provide insights into treatment response heterogeneity in colorectal cancer [46].

Conclusions

In summary, our analysis of ICDRGs offers insights into the TIME and could guide future COADREAD research, particularly in drug development and personalized immunotherapy therapy.

Acknowledgements

We are grateful to Yuanyuan Han, HuaShao for technical assistance and advice.

References

1 Siegel, R. L., Miller, K. D., Wagle, N. S. & Jemal, A. Cancer statistics, 2023. *CA: a cancer journal for clinicians* **73**, 17-48 (2023). https://doi.org/10.3322/caac.21763

2 Miller, K. D. *et al.* Cancer treatment and survivorship statistics, 2022. *CA: a cancer journal for clinicians* **72**, 409-436 (2022). https://doi.org/10.3322/caac.21731

3 Ciardiello, F. *et al.* Clinical management of metastatic colorectal cancer in the era of precision medicine. *CA: a cancer journal for clinicians* **72**, 372-401 (2022). https://doi.org/10.3322/caac.21728

4 Lenz, H. J. *et al.* First-Line Nivolumab Plus Low-Dose Ipilimumab for Microsatellite Instability-High/Mismatch Repair-Deficient Metastatic Colorectal Cancer: The Phase II CheckMate 142 Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **40**, 161-170 (2022). https://doi.org/10.1200/jco.21.01015

5 Gurjao, C. *et al.* Intrinsic Resistance to Immune Checkpoint Blockade in a Mismatch Repair-Deficient Colorectal Cancer. *Cancer immunology research* **7**, 1230-1236 (2019). https://doi.org/10.1158/2326-6066.Cir-18-0683

6 Simoneau, J., Dumontier, S., Gosselin, R. & Scott, M. S. Current RNA-seq methodology reporting limits reproducibility. *Briefings in bioinformatics* **22**, 140-145 (2021). https://doi.org/10.1093/bib/bbz124

7 Wu, Y. *et al.* Spatiotemporal Immune Landscape of Colorectal Cancer Liver Metastasis at Single-Cell Level. *Cancer discovery* **12**, 134-153 (2022). https://doi.org/10.1158/2159-8290.Cd-21-0316

8 Kroemer, G., Galassi, C., Zitvogel, L. & Galluzzi, L. Immunogenic cell stress and death. *Nature immunology* **23**, 487-500 (2022). https://doi.org/10.1038/s41590-022-01132-2

9 Hayashi, K. *et al.* Tipping the immunostimulatory and inhibitory DAMP balance to harness immunogenic cell death. *Nature communications* **11**, 6299 (2020). https://doi.org/10.1038/s41467-020-19970-9

10 Haen, S. P., Löffler, M. W., Rammensee, H. G. & Brossart, P. Towards new horizons: characterization, classification and implications of the tumour antigenic repertoire. *Nature reviews. Clinical oncology* **17**, 595-610 (2020). https://doi.org/10.1038/s41571-020-0387-x

11 Salas-Benito, D. *et al.* Paradigms on Immunotherapy Combinations with Chemotherapy. *Cancer discovery* **11**, 1353-1367 (2021). https://doi.org/10.1158/2159-8290.Cd-20-1312

12 Galluzzi, L., Humeau, J., Buqué, A., Zitvogel, L. & Kroemer, G. Immunostimulation with chemotherapy in the era of immune checkpoint inhibitors. *Nature reviews. Clinical oncology* **17**, 725-741 (2020). https://doi.org/10.1038/s41571-020-0413-z

13 Liu, J., Shi, Y. & Zhang, Y. Multi-omics identification of an immunogenic cell death-related signature for clear cell renal cell carcinoma in the context of 3P medicine and based on a 101-combination machine learning computational framework. *The EPMA journal* **14**, 275-305 (2023). https://doi.org/10.1007/s13167-023-00327-3

14 Ren, J. *et al.* Comprehensive characterisation of immunogenic cell death in melanoma revealing the association with prognosis and tumor immune microenvironment. *Frontiers in immunology* **13**, 998653 (2022). https://doi.org/10.3389/fimmu.2022.998653

15 Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets--update. *Nucleic acids research* **41**, D991-995 (2013). https://doi.org/10.1093/nar/gks1193

16 Blum, A., Wang, P. & Zenklusen, J. C. SnapShot: TCGA-Analyzed Tumors. *Cell* **173**, 530 (2018). https://doi.org/10.1016/j.cell.2018.03.059

17 Mangiola, S., Doyle, M. A. & Papenfuss, A. T. Interfacing Seurat with the R tidy universe. *Bioinformatics (Oxford, England)* **37**, 4100-4107 (2021). https://doi.org/10.1093/bioinformatics/btab404

18 Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nature biotechnology* (2018). https://doi.org/10.1038/nbt.4314

19 Hu, C. *et al.* CellMarker 2.0: an updated database of manually curated cell markers in

human/mouse and web tools based on scRNA-seq data. *Nucleic acids research* **51**, D870-d876 (2023). https://doi.org/10.1093/nar/gkac947

20 Franzén, O., Gan, L. M. & Björkegren, J. L. M. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database : the journal of biological databases and curation* **2019** (2019). https://doi.org/10.1093/database/baz046

21 Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC bioinformatics* **14**, 7 (2013). https://doi.org/10.1186/1471-2105-14-7

22 Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell systems* **1**, 417-425 (2015). https://doi.org/10.1016/j.cels.2015.12.004

23 Garg, A. D., De Ruysscher, D. & Agostinis, P. Immunological metagene signatures derived from immunogenic cancer cell death associate with improved survival of patients with lung, breast or ovarian malignancies: A large-scale meta-analysis. *Oncoimmunology* **5**, e1069938 (2016). https://doi.org/10.1080/2162402x.2015.1069938

24 Scire, J. *et al.* estimateR: an R package to estimate and monitor the effective reproductive number. *BMC bioinformatics* **24**, 310 (2023). https://doi.org/10.1186/s12859-023-05428-4

25 Liu, C. J. *et al.* GSCA: an integrated platform for gene set cancer analysis at genomic, pharmacogenomic and immunogenomic levels. *Briefings in bioinformatics* **24** (2023). https://doi.org/10.1093/bib/bbac558

26 Ringnér, M. What is principal component analysis? *Nature biotechnology* **26**, 303-304 (2008). https://doi.org/10.1038/nbt0308-303

27 Braun, D. A. *et al.* Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nature medicine* **26**, 909-918 (2020). https://doi.org/10.1038/s41591-020-0839-y

28 Yang, W. *et al.* Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic acids research* **41**, D955-961 (2013). https://doi.org/10.1093/nar/gks1111

29 Zhang, J., Huang, D., Saw, P. E. & Song, E. Turning cold tumors hot: from molecular mechanisms to clinical applications. *Trends in immunology* **43**, 523-545 (2022). https://doi.org/10.1016/j.it.2022.04.010

30 Blass, E. & Ott, P. A. Advances in the development of personalized neoantigen-based therapeutic cancer vaccines. *Nature reviews. Clinical oncology* **18**, 215-229 (2021). https://doi.org/10.1038/s41571-020-00460-2

31 Weng, J. *et al.* Exploring immunotherapy in colorectal cancer. *Journal of hematology & oncology* **15**, 95 (2022). https://doi.org/10.1186/s13045-022-01294-4

32 Qi, J. *et al.* Single-cell and spatial analysis reveal interaction of FAP(+) fibroblasts and SPP1(+) macrophages in colorectal cancer. *Nature communications* **13**, 1742 (2022). https://doi.org/10.1038/s41467-022-29366-6

33 Weng, S. *et al.* SCG2: A Prognostic Marker That Pinpoints Chemotherapy and

Immunotherapy in Colorectal Cancer. *Frontiers in immunology* **13**, 873871 (2022). https://doi.org/10.3389/fimmu.2022.873871

34 Stevenson, L. *et al.* Calbindin 2 (CALB2) regulates 5-fluorouracil sensitivity in colorectal cancer by modulating the intrinsic apoptotic pathway. *PloS one* **6**, e20276 (2011). https://doi.org/10.1371/journal.pone.0020276

35 Qi, L. *et al.* The Effects of Differentially-Expressed Homeobox Family Genes on the Prognosis and HOXC6 on Immune Microenvironment Orchestration in Colorectal Cancer. *Frontiers in immunology* **12**, 781221 (2021). https://doi.org/10.3389/fimmu.2021.781221

36 Dai, A. *et al.* Effects of the CRYAB gene on stem cell-like properties of colorectal cancer and its mechanism. *Journal of cancer research and therapeutics* **18**, 1328-1337 (2022). https://doi.org/10.4103/jcrt.jcrt_212_22

37 Alekseyenko, A. A. *et al.* Ectopic protein interactions within BRD4-chromatin complexes drive oncogenic megadomain formation in NUT midline carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E4184-e4192 (2017). https://doi.org/10.1073/pnas.1702086114

38 Deng, Y. *et al.* HDAC6-dependent deacetylation of AKAP12 dictates its ubiquitination and promotes colon cancer metastasis. *Cancer letters* **549**, 215911 (2022). https://doi.org/10.1016/j.canlet.2022.215911

39 Tian, W. *et al.* FABP4 promotes invasion and metastasis of colon cancer by regulating fatty acid transport. *Cancer cell international* **20**, 512 (2020). https://doi.org/10.1186/s12935-020-01582-4

40 Wan, Z. *et al.* MEIS2 promotes cell migration and invasion in colorectal cancer. *Oncology reports* **42**, 213-223 (2019). https://doi.org/10.3892/or.2019.7161

41 Gao, H. *et al.* A Prognosis Marker SLC2A3 Correlates With EMT and Immune Signature in Colorectal Cancer. *Frontiers in oncology* **11**, 638099 (2021). https://doi.org/10.3389/fonc.2021.638099

42 Shen, C. *et al.* The KDM6A-SPARCL1 axis blocks metastasis and regulates the tumour microenvironment of gastrointestinal stromal tumours by inhibiting the nuclear translocation of p65. *British journal of cancer* **126**, 1457-1469 (2022). https://doi.org/10.1038/s41416-022-01728-3

43 Xie, L. *et al.* Aberrant activation of CYR61 enhancers in colorectal cancer development. *Journal of experimental & clinical cancer research : CR* **38**, 213 (2019). https://doi.org/10.1186/s13046-019-1217-9

44 Chen, C. *et al.* CYP1B1 inhibits ferroptosis and induces anti-PD-1 resistance by degrading ACSL4 in colorectal cancer. *Cell death & disease* **14**, 271 (2023). https://doi.org/10.1038/s41419-023-05803-2

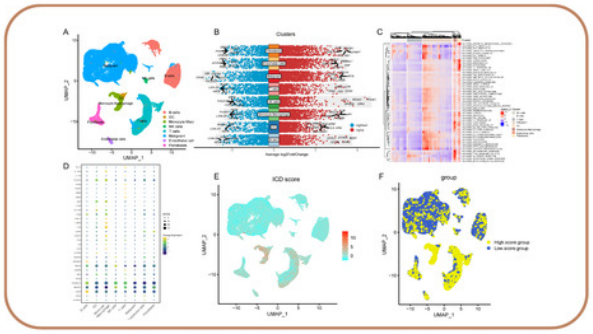
45 Tong, W. *et al.* Exploration of shared TF-miRNA–mRNA and mRNA-RBP-pseudogene networks in type 2 diabetes mellitus and breast cancer. *Frontiers in immunology* **13**, 915017 (2022). https://doi.org/10.3389/fimmu.2022.915017

46 Bassez, A. *et al.* A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. *Nature medicine* **27**, 820-832 (2021). https://doi.org/10.1038/s41591-021-01323-8

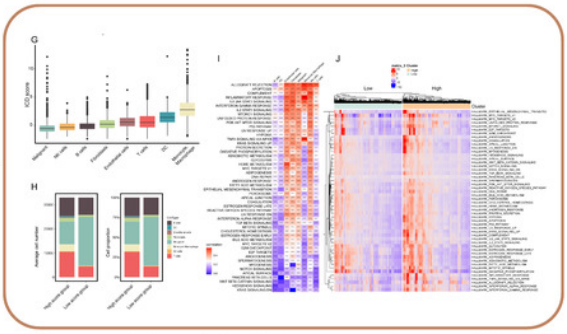
Figure 1

Flowchart of this study.

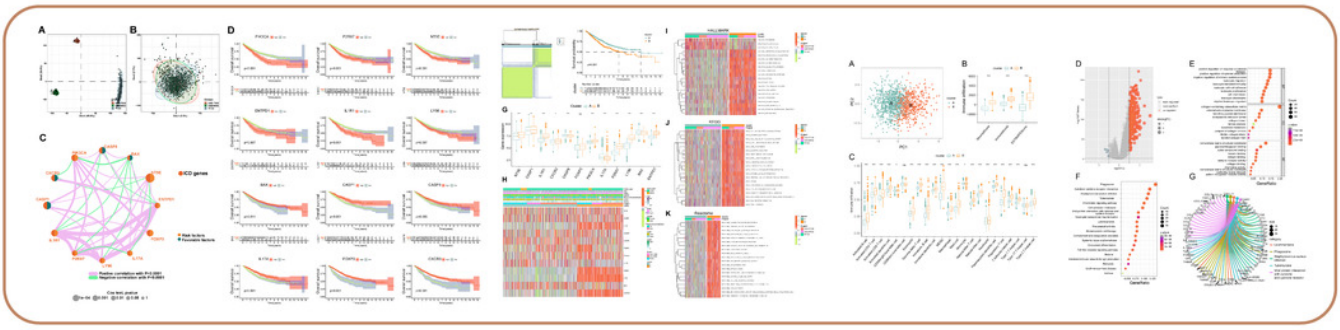
Flowchart of this study.



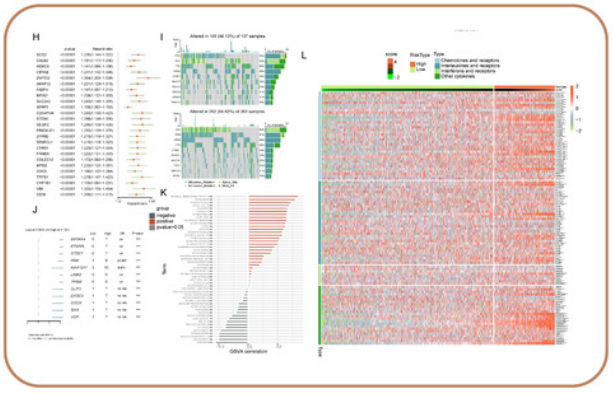
Construction of the ICD score at single cell level



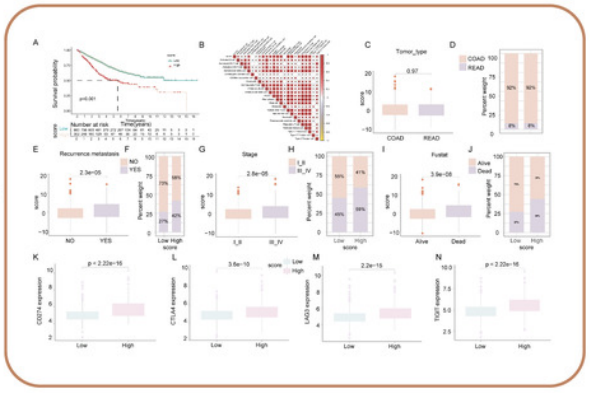
Characteristics of the ICD score at single cell level



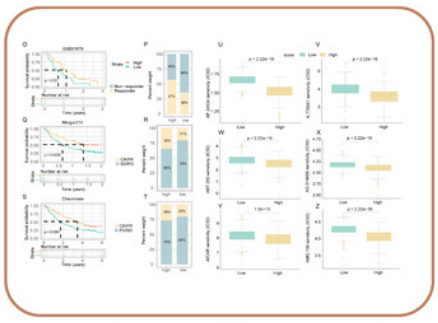
Construction and landscape of the ICD molecular clusters



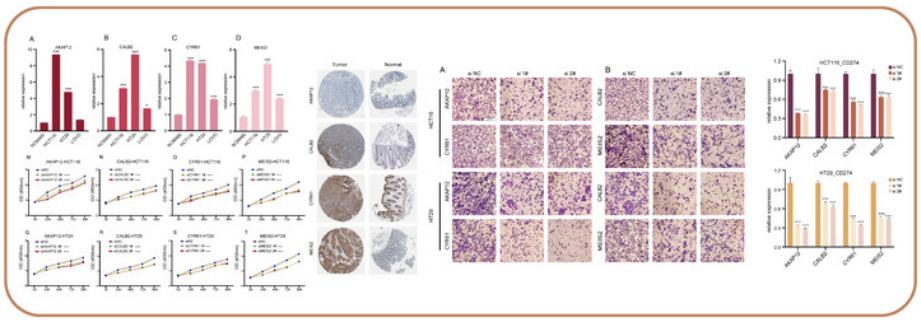
Construction of the ICD score



landscape of the ICD score



Immunotherapy and drug susceptibility



Experiment verification

Figure 2

Elaboration of ICD score and cellular diversity in COADREAD via scRNA-seq.

(A) The COADREAD cell dataset was delineated into 8 distinct clusters via UMAP, featuring varied cell types: B cells, dendritic cells, endothelial cells, fibroblasts, malignant cells, mono-macrophages, NK cells, and T cells. (B) A volcano plot delineated differentially expressed genes among these cell types, highlighting the top five. (C) Displayed through a heatmap, this illustrated the enrichment levels of Hallmark pathways in each tumor-infiltrating cell type, with a bar indicating the relative pathway enrichment. (D) A dot plot revealed the expression levels of 34 ICDRGs across these varied cell types. (E) GSVA was used to compute the ICD score for each cell type, visualized through a UMAP plot indicating the expression of ICD gene clusters. (F) Cell clusters were categorized into high and low ICD score groups, illustrated via UMAP plot. (G) Box plots represented the ICD scores across cell clusters, organized in ascending order. (H) A composition chart compared the number and proportion of cells in high versus low ICD score groups. (I) A heatmap demonstrated the relationship between ICD score and Hallmark pathway scores across all clusters. (J) This heatmap contrasted Hallmark pathway scores between high and low ICD score groups.

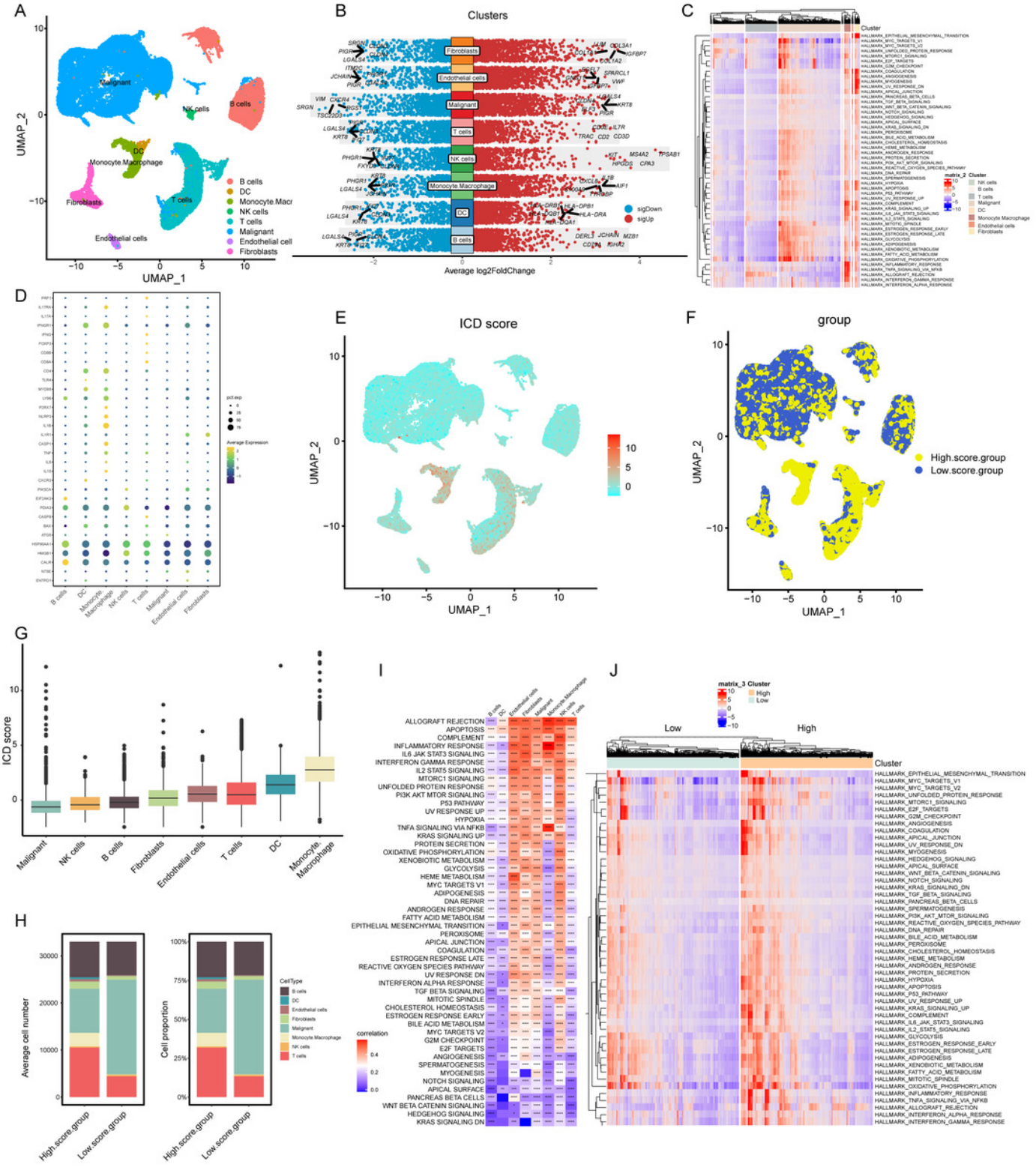

A
UMAP_2
UMAP_1
Malignant
NK cells
B cells
DC
Monocyte Macrophage
T cells
Fibroblasts
Endothelial cells
B cells
DC
Monocyte Macr
NK cells
T cells
Malignant
Endothelial cell
Fibroblasts
B
Clusters
Fibroblasts
Endothelial cells
Malignant
T cells
NK cells
Monocyte Macrophage
DC
B cells
sigDown
sigUp
Average log2FoldChange
C
Cluster
matrix_2 Cluster
NK cells
B cells
T cells
Malignant
DC
Monocyte Macrophage
Endothelial cells
Fibroblasts
D
Average Expression
B cells
DC
Monocyte Macrophage
NK cells
T cells
Malignant
Endothelial cells
Fibroblasts
E
ICD score
UMAP_2
UMAP_1
F
group
High.score.group
Low.score.group
UMAP_2
UMAP_1
G
ICD score
Malignant
NK cells
B cells
Fibroblasts
Endothelial cells
T cells
DC
Monocyte Macrophage
H
Average cell number
Cell proportion
High.score.group
Low.score.group
CellType
B cells
DC
Endothelial cells
Fibroblasts
Malignant
Monocyte Macrophage
NK cells
T cells
I
ALLOGRAFT REJECTION
APOPTOSIS
COMPLEMENT
INFLAMMATORY RESPONSE
IL6 JAK STAT3 SIGNALING
INTERFERON GAMMA RESPONSE
IL2 STAT5 SIGNALING
MTORC1 SIGNALING
UNFOLDED PROTEIN RESPONSE
PI3K AKT MTOR SIGNALING
P53 PATHWAY
UV RESPONSE UP
HYPOXIA
TNFA SIGNALING VIA NFKB
KRAS SIGNALING UP
PROTEIN SECRETION
OXIDATIVE PHOSPHORYLATION
XENOBIOTIC METABOLISM
GLYCOLYSIS
HEME METABOLISM
MYC TARGETS V1
ADIPOGENESIS
DNA REPAIR
ANDROGEN RESPONSE
FATTY ACID METABOLISM
EPITHELIAL MESENCHYMAL TRANSITION
PEROXISOME
APICAL JUNCTION
COAGULATION
ESTROGEN RESPONSE LATE
REACTIVE OXYGEN SPECIES PATHWAY
UV RESPONSE DN
INTERFERON ALPHA RESPONSE
TGF BETA SIGNALING
MITOTIC SPINDLE
CHOLESTEROL HOMEOSTASIS
ESTROGEN RESPONSE EARLY
BILE ACID METABOLISM
MYC TARGETS V2
G2M CHECKPOINT
E2F TARGETS
ANGIOGENESIS
SPERMATOGENESIS
MYOGENESIS
NOTCH SIGNALING
APICAL SURFACE
PANCREAS BETA CELLS
WNT BETA CATENIN SIGNALING
HEDGEHOG SIGNALING
KRAS SIGNALING DN
correlation
J
matrix_3 Cluster
High
Low
Low
High
Cluster

Figure 3

Delineation and biological characterization of ICD molecular subtypes in colorectal carcinoma.

(A) Illustrating the distribution of transcriptional expressions of 34 ICDRGs in three colorectal adenocarcinoma cohorts. (B) Enhanced visualization post-adjustment, clarifying transcriptional diversity. (C) Spearman correlation analyses depicted the interactions among ICDRGs in 1184 colorectal cancer samples, with line thickness indicating correlation strength-pink for positive and green for negative. (D) Log-rank tests evaluated the survival impact of specific ICDRGs (PIK3CA, P2RX7, NT5E, ENTPD1, IL1R1, LY96, BAX, CASP1, CASP8, IL17A, FOXP3, CXCR3). (E) Consensus clustering analysis revealed two distinct subtypes (k = 2) in colorectal cancer samples. (F) Highlights significant survival differences between molecular subtypes A and B. (G) Analyzed the abundance of ICD signature genes within the two identified clusters. (H) A heat-map demonstrated the association of clinicopathologic characteristics with each subtype, using red to indicate up-regulation and green for down-regulation. (I-K) GSVA analysis of Hallmark (I), KEGG (J), and Reactome (K) pathways between molecular subtypes A and B, with red denoting more enriched and green denoting less enriched pathways.

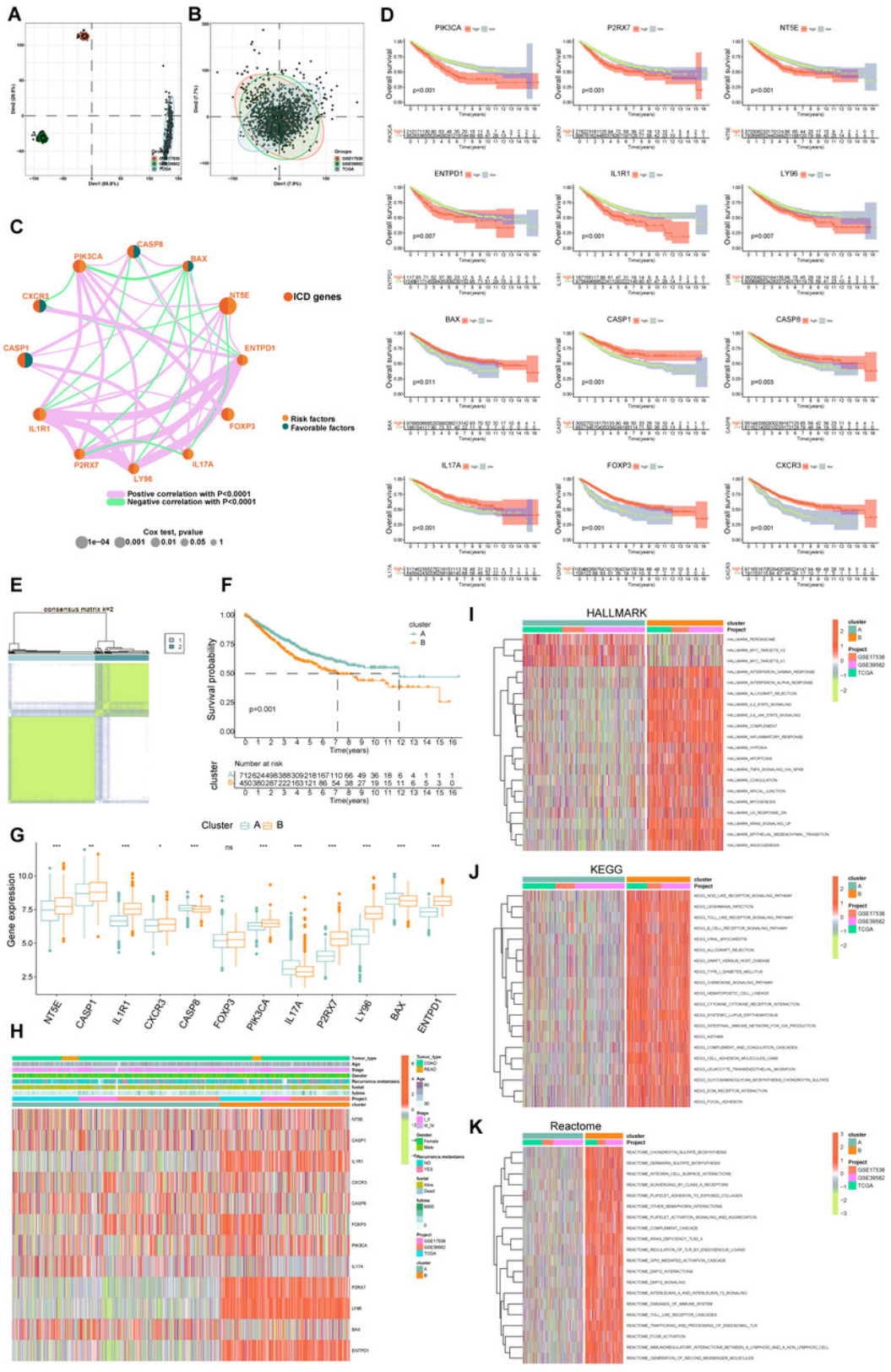

A
B
C
D
PIK3CA
P2RX7
NT5E
ENTPD1
IL1R1
LY96
BAX
CASP1
CASP8
IL17A
FOXP3
CXCR3
ICD genes
Risk factors
Favorable factors
Postive correlation with P<0.0001
Negative correlation with P<0.0001
Cox test, pvalue
E
consensus matrix k=2
F
Survival probability
p=0.001
Number at risk
G
Cluster A B
Gene expression
H
I
HALLMARK
J
KEGG
K
Reactome

Figure 4

Comprehensive ICD score analysis in colorectal carcinoma.

(A) PCA illustrating notable distribution differences among ICD molecular subtypes. (B) Comparison of stromal, immune, and ESTIMATE scores in different ICD clusters. (C) Analysis of immune cell infiltration in clusters A and B. (D) Volcano plot presenting 448 ICD-DEGs from varied molecular subtypes. (E) GO enrichment analysis for ICD-DEGs, with plot size indicating gene counts, categorized by BP, CC, and MF. (F) KEGG enrichment analysis of ICD-DEGs, detailing gene enrichment. (G) Chord graph highlighting five key KEGG pathways and associated genes. (H) PCA-based ICD score construction from 25 ICD-DEGs and subsequent survival analysis of two score clusters. (I) Waterfall plots of somatic mutations in high and low ICD score groups. (J) Forest plot comparing mutational variation and HRs in high and low-score clusters. (K) GSVA correlation of ICD score with 50 hallmark pathways. (L) GSVA for cytokines, chemokines, and their receptors in high and low ICD score clusters.

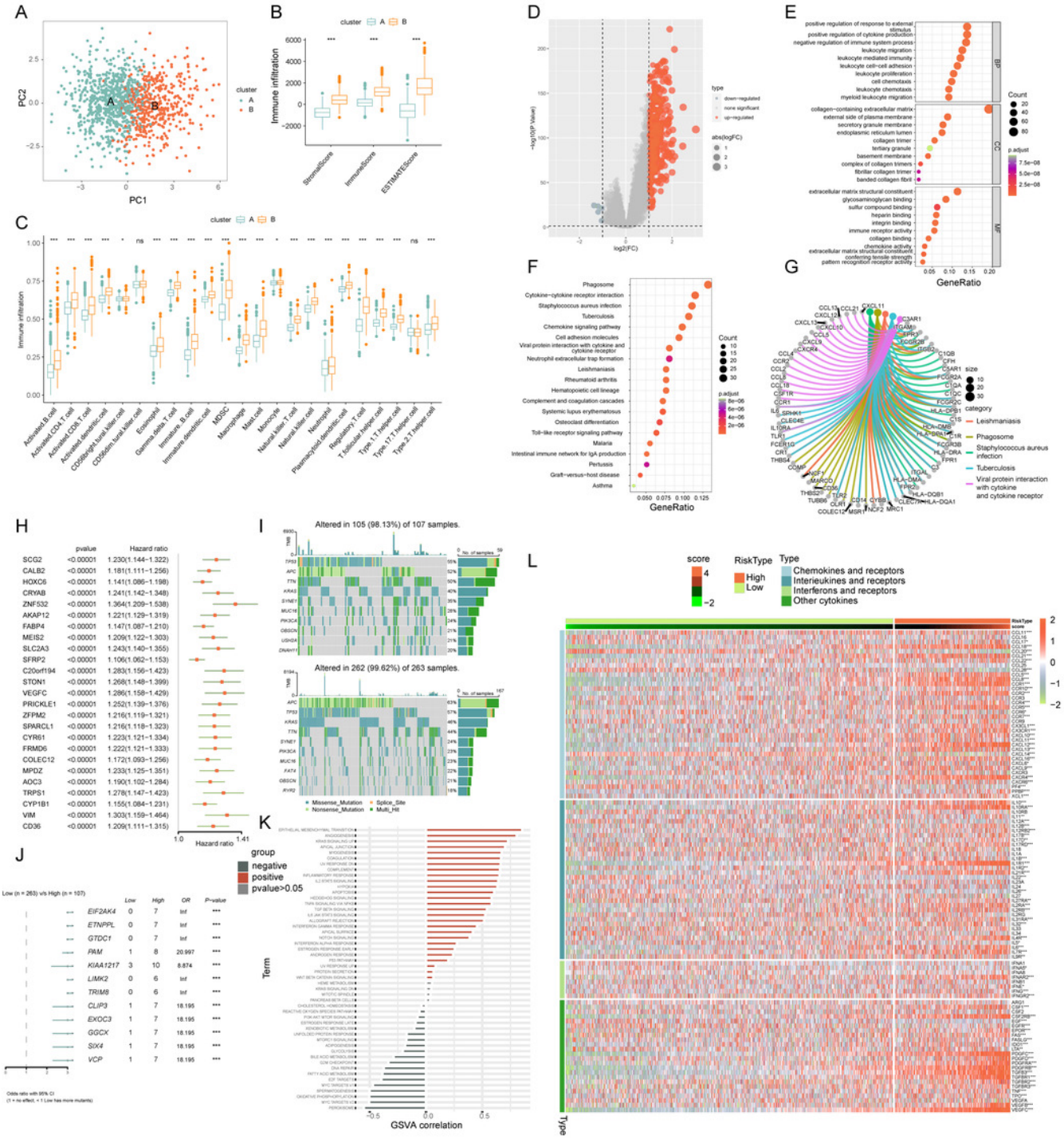

A
B
C
D
E
F
G
H
I
J
K
L
cluster
Immune infiltration
GeneRatio
Count
Hazard ratio
Altered in 105 (98.13%) of 107 samples.
Altered in 262 (99.62%) of 263 samples.
group
negative
positive
pvalue>0.05
GSVA correlation
RiskType
High
Low
Type
Chemokines and receptors
Interleukins and receptors
Interferons and receptors
Other cytokines

Figure 5

Assessment of clinical impact, immunotherapy response, and drug sensitivity in relation to ICD score.

(A) Survival analysis examined the prognostic differences between high and low ICD score groups. (B) Analysis of the positive relationship between ICD score and tumor-infiltrating immune cells. (C-J) Investigates the differences in tumor type (C, D), recurrence and metastasis (E, F), stage (G, H), and survival status (I, J) across ICD score clusters. (K-N) Examined the variance in immune checkpoint levels between high and low score groups. (O-P) Targeted therapy response in lung cancer (GSE61676). Comparison of responses and survival outcomes between low and high ICD score clusters. (Q-R) Immunotherapy efficacy in urothelial carcinoma (IMvigor210). Analysis of response and survival differences in low versus high ICD score clusters. (S-T) Anti-PD-1 therapy response in KIRC cohort (Checkmate). Evaluated therapeutic response and survival rates between low and high ICD score groups. (U-Z) Box diagrams depicting the differences in drug sensitivity to targeted therapy between high and low ICD score clusters.

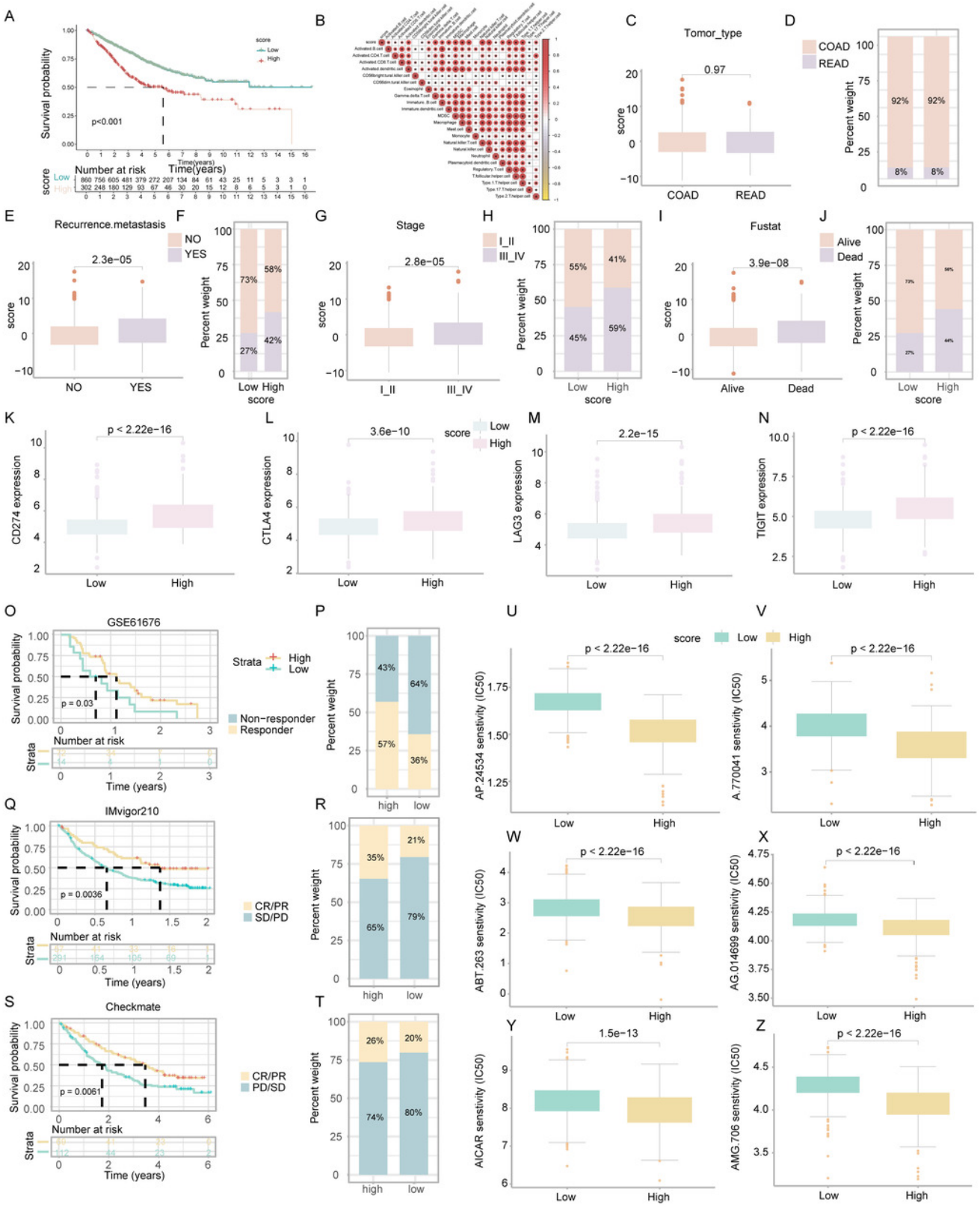
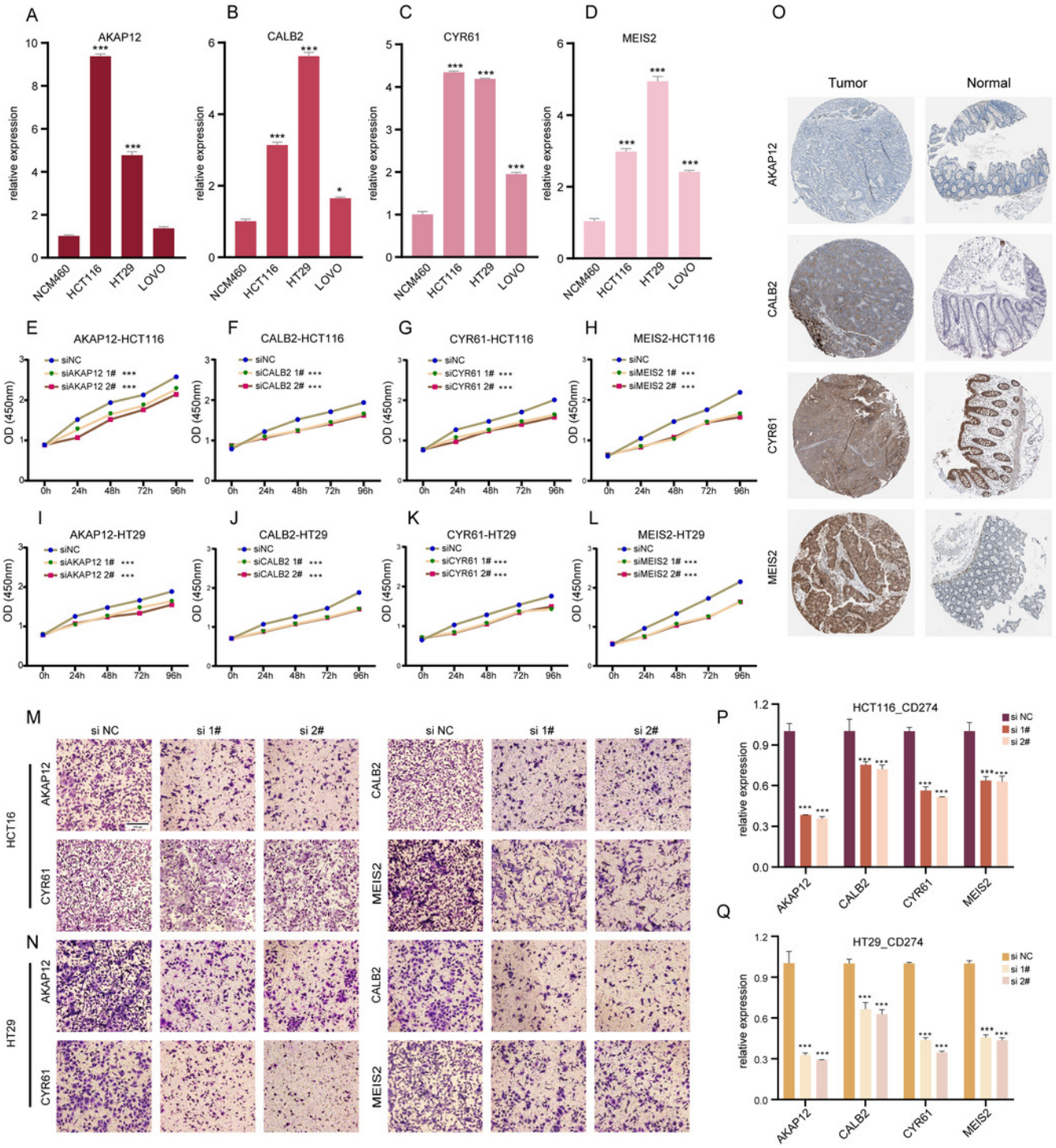

A
Survival probability
p<0.001
score
Low
High
Number at risk
Time(years)
B
C
Tomor_type
0.97
score
COAD
READ
D
COAD
READ
Percent weight
92%
8%
E
Recurrence.metastasis
2.3e−05
NO
YES
F
73%
27%
58%
42%
Low High
score
G
Stage
2.8e−05
I_II
III_IV
H
55%
45%
41%
59%
I
Fustat
3.9e−08
Alive
Dead
J
K
CD274 expression
p < 2.22e−16
L
CTLA4 expression
3.6e−10
M
LAG3 expression
2.2e−15
N
TIGIT expression
p < 2.22e−16
O
GSE61676
Survival probability
p = 0.03
Strata
High
Low
Number at risk
Time (years)
Non−responder
Responder
P
43%
57%
64%
36%
high
low
Percent weight
Q
IMvigor210
p = 0.0036
CR/PR
SD/PD
R
35%
65%
21%
79%
S
Checkmate
p = 0.0061
T
26%
74%
20%
80%
U
AP.24534 senstivity (IC50)
p < 2.22e−16
V
A.770041 senstivity (IC50)
W
ABT.263 senstivity (IC50)
X
AG.014699 senstivity (IC50)
Y
AICAR senstivity (IC50)
1.5e−13
Z
AMG.706 senstivity (IC50)

Figure 6

Experimental validation of gene expression and functional assays.

(A-D) Assessing expression levels of AKAP12 (A), CALB2 (B), CYR61 (C), and MEIS2 (D) in normal colon cells (NCM460) and COADREAD cell lines (HCT116, HCT29, VOLO). (E-L) Post-knockdown viability reduction in HCT116 and HT29 cells were measured for AKAP12 (E, I), CALB2 (F, J), CYR61 (G, K), and MEIS2 (H, L). (M-N) Demonstrating decreased invasion capacity in HCT116 and HCT29 cells following AKAP12, CALB2, CYR61, and MEIS2 knockdown. (O) Immunohistochemistry data from HPA database showed protein expressions of AKAP12, CALB2, CYR61, and MEIS2. (P-Q) Significant reduction in PD-L1 expression in HCT116 (P) and HT29 (Q) cells after knockdown of the respective genes.


A
AKAP12
relative expression
NCM460
HCT116
HT29
LOVO
B
CALB2
C
CYR61
D
MEIS2
E
AKAP12-HCT116
siNC
siAKAP12 1#
siAKAP12 2#
OD (450nm)
0h
24h
48h
72h
96h
F
CALB2-HCT116
siCALB2 1#
siCALB2 2#
G
CYR61-HCT116
siCYR61 1#
siCYR61 2#
H
MEIS2-HCT116
siMEIS2 1#
siMEIS2 2#
I
AKAP12-HT29
J
CALB2-HT29
K
CYR61-HT29
L
MEIS2-HT29
O
Tumor
Normal
AKAP12
CALB2
CYR61
MEIS2
M
si NC
si 1#
si 2#
HCT16
N
HT29
P
HCT116_CD274
Q
HT29_CD274