



Chaperonin containing TCP1 subunit 5 as a novel pan-cancer prognostic biomarker for tumor stemness and immunotherapy response: insights from multi-omics data, integrated machine learning, and experimental validation

Jiajun Li^{1,2} · Nuo Xu² · Leyin Hu⁴ · Jiayue Xu¹ · Yifan Huang¹ · Deqi Wang³ · Feng Chen¹ · Yi Wang³ · Jiani Jiang³ · Yanggang Hong¹ · Huajun Ye³

Received: 23 March 2025 / Accepted: 28 April 2025
© The Author(s) 2025

Abstract

Background Chaperonin containing TCP1 subunit 5 (CCT5), a vital component of the molecular chaperonin complex, has been implicated in tumorigenesis, cancer stemness maintenance, and therapeutic resistance. Nevertheless, its comprehensive roles in pan-cancer progression, underlying biological functions, and potential as a predictor of immunotherapy response remains poorly understood.

Methods We performed a comprehensive multi-omics pan-cancer analysis of CCT5 across 33 cancer types, integrating bulk RNA-seq, single-cell RNA-seq (scRNA-seq), and spatial transcriptomics data. CCT5 expression patterns, prognostic relevance, stemness association, and immune microenvironment relationships were evaluated. A novel CCT5-based signature (CCT5.Sig) was developed using machine learning on 23 immune checkpoint blockade (ICB) cohorts ($n = 1394$) spanning eight cancer types. Model performance was assessed using AUC metrics and survival analyses.

Results CCT5 was significantly overexpressed in tumor tissues and primarily localized to malignant and cycling cells. High CCT5 expression correlated with poor prognosis in multiple cancers and was enriched in oncogenic, cell cycle, and DNA damage repair pathways. CCT5 expression was positively associated with mRNasi, mDNasi, and CytotTRACE scores, indicating a role in stemness maintenance. Furthermore, CCT5-high tumors exhibited immune-cold phenotypes, with reduced TILs and CD8⁺ T cell activity. The CCT5.Sig model, based on genes co-expressed with CCT5, achieved superior predictive accuracy for ICB response (AUC = 0.82 in validation and 0.76 in independent testing), outperforming existing pan-cancer signatures.

Jiajun Li, Nuo Xu, Leyin Hu, Jiayue Xu and Yifan Huang contributed equally to this work.

✉ Yanggang Hong
sun160414@icloud.com

✉ Huajun Ye
yehujun2020@163.com

¹ The Second School of Clinical Medicine, Wenzhou Medical University, Wenzhou 325000, Zhejiang, China

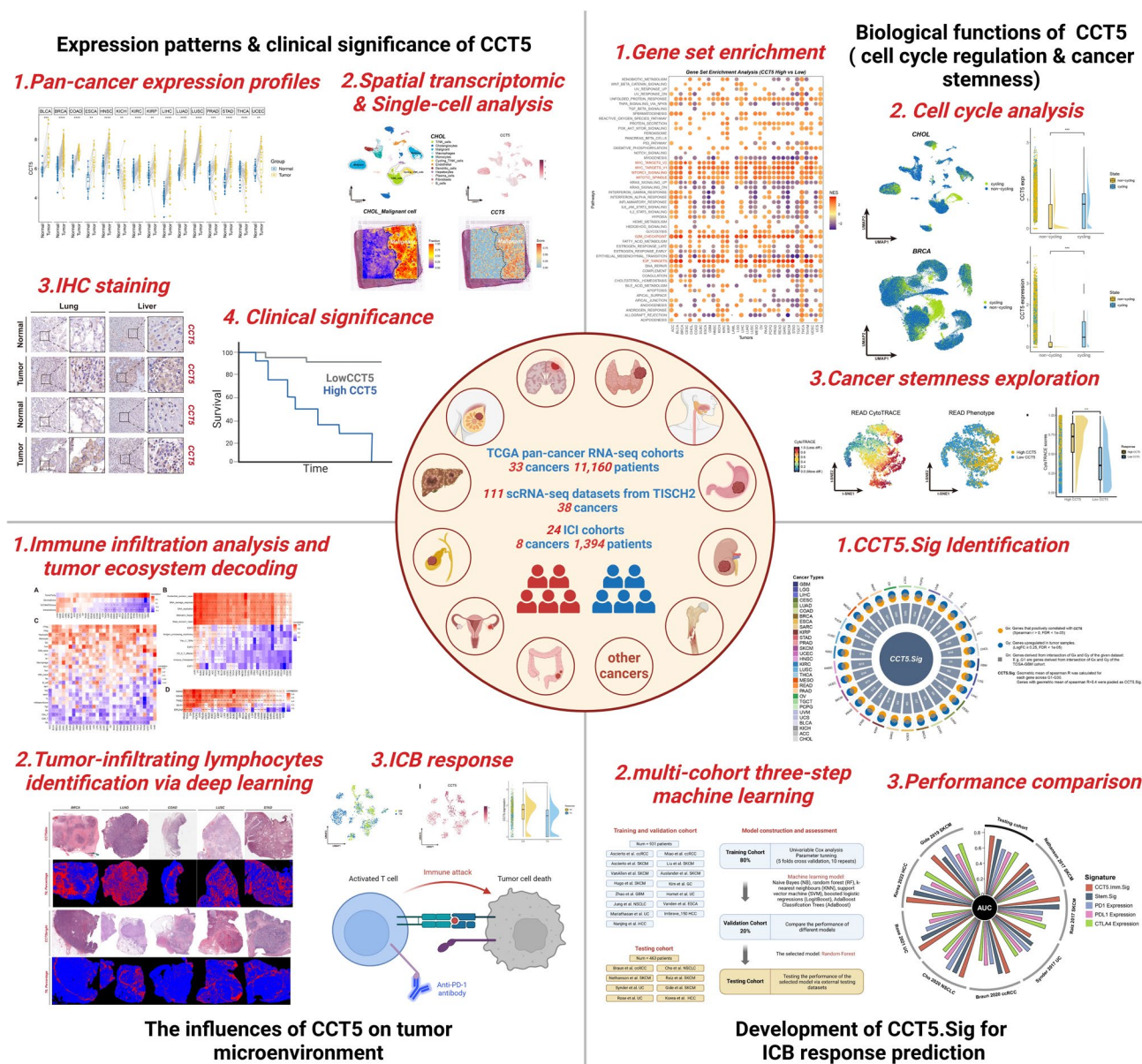
² State Key Laboratory of Systems Medicine for Cancer, Shanghai Cancer Institute and Department of Liver Surgery, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200001, China

³ Department of Gastroenterology, The First Affiliated Hospital, Wenzhou Medical University, Wenzhou 325000, Zhejiang, China

⁴ Department of Pathology, Zhejiang Cancer Hospital, Hangzhou 305000, Zhejiang, China

Conclusion This study reveals the multifaceted oncogenic roles of CCT5 and highlights its potential as a pan-cancer biomarker for prognosis and immunotherapy response. The machine learning-derived CCT5.Sig model provides a robust tool for patient stratification and may inform personalized immunotherapy strategies.

Graphical abstract



Keywords CCT5 · Pan-cancer · Cancer stemness · Tumor microenvironment · Immune checkpoint blockade · Single-cell analysis · Machine learning

Introduction

Cancer is a leading cause of death globally. According to Global Cancer Statistics 2022, GLOBOCAN estimates suggest that approximately one in five men or women

develop cancer during their lifetime, whereas approximately one in nine men and one in 12 women would die from it [1]. Conventional therapy methods, including surgical resection, chemotherapy, radiotherapy, hormone therapy, and immunotherapy based on immune checkpoint

blockade (ICB), have revolutionized the treatment landscape across multiple tumor types [2, 3] for instance, Pembrolizumab, a PD-1 inhibitor, has been approved as a first-line and second-line treatment for non-small cell lung cancer (NSCLC) since 2015. However, not all patients respond well to ICB therapy. Hence, it is of great significance to identify novel biomarkers or establish new models for immunotherapy efficacy prediction to develop personalized immunotherapy strategies.

Chaperonin containing TCP1 subunit 5 (CCT5), a member of the chaperonin family, plays a critical role in maintaining protein folding and cell cycle regulation [4, 5]. Recently, high expression of CCT5 has been observed in a variety of cancers, including liver hepatocellular carcinoma (LIHC) [6], rectal cancer (READ) [7], stomach adenocarcinoma (STAD) [8] and lung adenocarcinoma (LUAD) [9]. Studies have revealed that CCT5 induces epithelial-mesenchymal transition (EMT) to promote STAD lymph node metastasis by activating the Wnt/ β -catenin signaling pathway [8]. Another study has highlighted the role of the CCT5/miR-139-5p axis in promoting LIHC progression. CCT5 high expression was also found to be associated with poor treatment response in patients with locally advanced READ undergoing neoadjuvant chemoradiotherapy [7]. Additionally, CCT5, RGS3, and YKT6 promoted docetaxel resistance in breast cancer (BRCA) [10]. Guo et al. reported that CCT5 has the potential to maintain LIHC cell stemness [6]. Wang et al. focused on CCT5's role in enhancing the stemness of high-grade serous ovarian cancer (OV) [11]. The aforementioned insights suggest that CCT5 may play complicated roles in cancer progression, including the regulation of tumor cell proliferation, invasion, migration, maintenance of stemness, and therapeutic resistance. Additionally, given the close association between stemness and ICB response, CCT5 has the potential to be developed as a novel biomarker for ICB efficacy prediction [12]. Therefore, a comprehensive pan-cancer analysis for CCT5 is warranted to explore its predictive value for prognosis, tumor stemness, and ICB response.

In this study, a systematic investigation of CCT5 across 33 different cancer types was conducted using multiomic databases. First, we examined CCT5 expression patterns using bulk RNA-seq, single-cell RNA-seq (scRNA-seq), and spatial transcriptomics (stRNA-seq). Immunohistochemistry (IHC) staining was performed for *in vitro* validation. Next, we assessed CCT5's clinical significance by analyzing its impact on overall survival (OS), disease-specific survival (DSS), disease-free interval (DFI), and progression-free interval (PFI). To explore CCT5's functional role in cancer, we categorized samples into CCT5-high and CCT5-low groups based on average expression levels and performed a pan-cancer Gene Set Enrichment Analysis (GSEA) to identify common biological pathways. Given its strong

association with cell cycle regulation, we conducted a detailed cell cycle analysis using scRNA-seq datasets. Since cancer stem cells (CSCs) drive tumor initiation, progression, and metastasis, we examined whether CCT5 functions as a stemness driver. The mRNA stemness index (mRNasi) and methylation-based DNA stemness index (mDNasi) were used to assess stemness in bulk RNA-seq data, while CytoTRACE was applied to scRNA-seq data. Considering the well-established connection between stemness, immune evasion, and therapy resistance, we further explored CCT5's role in the tumor immune microenvironment (TIME). Using the ESTIMATE algorithm, we compared tumor purity, stromal, and immune cell abundance between CCT5-high and CCT5-low groups. The CIBERSORT algorithm was used to analyze immune cell composition within the TIME. We also characterized TIME using prognostic and immunotherapy-relevant gene signatures from previous studies. Since DNA damage and repair pathways were significantly enriched, we examined the relationship between CCT5 and mismatch repair (MMR)-related genes. Using deep learning algorithms, we compared tumor-infiltrating lymphocyte (TIL) abundance on HE-stained images between CCT5-high and CCT5-low groups. Additionally, we assessed CCT5 expression differences in ICB responders and non-responders in both bulk RNA-seq and scRNA-seq datasets. Given that the aforementioned analyses indicated CCT5's association with ICB resistance through multiple mechanisms, we integrated 23 ICB bulk RNA-seq datasets (1,394 patients, 8 cancers) and developed a CCT5-based signature (CCT5.Sig) for ICB response prediction using a multi-step machine learning strategy. The CCT5.Sig genes were selected based on the expression patterns across 30 cancer types. These genes underwent univariate Cox regression, machine learning model training, validation, and independent testing. The CCT5.Sig model demonstrated superior AUC performance in predicting ICB response compared to several existing signatures. Collectively, this study provides comprehensive insights into CCT5's role in cancer progression and establishes a machine learning-based model with high accuracy and robustness for ICB response prediction.

Materials and methods

Pan-cancer transcriptomic data

We downloaded the expression matrix of 33 cancers and clinical traits of the patients from UCSC Xena data portal (<https://xena.ucsc.edu/>) to investigate the expression patterns of CCT5, its prognostic values, its biological and pathological functions and to pool the candidate genes for CCT5.Sig. The HE staining images of BRCA, LUAD, colon

adenocarcinoma (COAD), LUSC, and STAD were collected from TCGA data portal (<https://portal.gdc.cancer.gov/>).

Pan-cancer scRNA-seq cohorts

To explore the expression patterns of CCT5 in single-cell resolution, 111 scRNA-Seq datasets containing both malignant and stromal/immune cell data, as well as detailed annotations, were obtained from the TISCH2 portal (<http://tisch.compbio.cn/home/>) [13]. These datasets include a total of 38 cancer types (Additional file 1: Table S1). The included cancers were acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), bladder urothelial carcinoma (BLCA), BRCA, cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), cholangiocarcinoma (CHOL), chronic lymphoblastic leukemia (CLL), colorectal cancer (CRC), esophageal carcinoma, glioma, hepatoblastoma (HB), head and neck cancer (HNSC), kidney chromophobe carcinoma (KICH), pan-kidney cohort (KIPAN), kidney renal clear cell carcinoma (KIRC), LIHC, Laryngeal squamous cell carcinoma (LSCC), medulloblastoma (MB), cutaneous B and T cell lymphoma (MF), neuroblastoma (NB), multiple myeloma (MM), neuroendocrine tumor (NET), cutaneous T cell lymphoma (CTCL, NHL), nasopharyngeal carcinoma (NPC), NSCLC, osteosarcoma (OS), oral squamous cell carcinoma (OSCC), OV, pancreatic adenocarcinoma (PAAD), pleuropulmonary blastoma (PPB), prostate adenocarcinoma (PRAD), retinoblastoma (RB), squamous cell carcinoma (SCC), small cell lung cancer (SCLC) skin cutaneous melanoma (SKCM), STAD, thyroid carcinoma (THCA), uterine corpus endometrial carcinoma (UCEC), and uveal melanoma (UVM).

The scRNA-seq cohorts for LIHC, BRCA, COAD, and CHOL were downloaded from GEO data portal to explore the CCT5 expression patterns in single-cell resolution (ID: GSE149614, GSE246613, GSE205506, and GSE138709, respectively) [14–17], while the other three scRNA-seq datasets for liver cancer (LIHC and CHOL), BRCA, and READ were retrieved for cancer stemness exploration, which were also downloaded from GEO data portal (ID: GSE125449, GSE246613 and GSE188711, respectively) [18–20].

Spatial transcriptomic datasets

Spatial transcriptomics data were utilized to validate CCT5 expression patterns observed in bulk and single-cell RNA sequencing analyses. The raw spatial transcriptomic datasets for GBM and UCEC were retrieved from GEO data portal (ID: GSE194329 and GSE203612, respectively) [21, 22]. While the BRCA dataset was obtained from Zenodo data repository [23], and the CHOL dataset was downloaded from <https://www.10xgenomics.com/products/spatial-gene-expression> [24].

Pan-cancer bulk RNA-seq and scRNA-seq ICB cohorts

One SKCM ICB scRNA-seq cohort (Rodman 2018 [25]) was downloaded from GEO data portal (ID: GSE115978) to explore the association between CCT5 expression levels in malignant cells and ICB response.

To train, validate and test the CCT5.Sig machine learning model, 23 ICB cohorts including eight cancer types were collected, including eight SKCM cohorts (Van Allen [26], Auslander [27], Hugo [28], Ascierto [29], Nathanson [30], Raiz [31], Liu [32], Gide [33]), four UC cohorts (Snyder [34], Mariathasan [35], Rose [36], Homet 2019), three ccRCC cohorts (Ascierto [37], Braun [38], Miao [39]), three HCC cohorts (Korea 2022, Nanjing 2023, imbrave150 [40]), two NSCLC cohorts (Jung [41], Cho [42]), one GC cohort (Kim [43]), one ESCA cohort (Van den [44]), and one GBM cohort (Zhao [45]). Detailed information of these cohorts are summarized in Additional file 1: Table S2.

scRNA-seq data processing

A total of 111 scRNA-seq datasets were obtained from TISCH2 data portal and the CCT5 expression levels in various cell types were extracted and presented in one single heatmap via “pheatmap” R package.

The other five scRNA-seq datasets were explored via the Seurat pipeline. For scRNA-seq data processing, we retained high-quality cells that had fewer than 20% mitochondrial gene expression and expressed at least 200 genes. As for genes, we focused on ones that were expressed at levels between 200 and 6000 and present in a minimum of three cells. The scRNA-seq data were normalized and scaled using a linear regression model with the “Log-normalization” method. The top 2000 highly variable genes were identified through the “FindVariableFeatures” function. We then reduced dimensionality using principal component analysis (PCA) and applied the “Harmony” package to correct for batch effects. Clustering of cells was performed with the “FindClusters” function, setting the resolution at 0.6. To annotate cell clusters, we focused on genes with high expression, unique patterns, and known canonical markers [46, 47].

Differential analysis

We first retrieved the expression levels of CCT5 in normal tissues from the Genotype-Tissue Expression (GTEx) database (<https://www.gtexportal.org/home/>). The expression data were processed and compared across different normal tissue types to evaluate baseline expression patterns. For cancer types that contained both tumor and adjacent normal tissue samples, we performed an unpaired differential expression analysis to determine the overall differences in CCT5 expression. Considering that only limited normal

samples were included in TCGA cohorts, we combined the GTEx and TCGA cohorts to validate the findings. Additionally, paired differential expression analysis was applied to cohorts with matched tumor and adjacent normal samples to detect intra-patient expression differences. Differential expression analysis was performed using the “DESeq2” and “TCGAplot” R packages.

scRNA-seq analysis

Malignant cells were annotated by both known canonical markers and chromosomal copy number variation (CNV). The canonical markers were retrieved from CellMarker 2.0 scRNA-seq data portal (<http://117.50.127.228/CellMarker/>) [48]. To identify malignant cells with extensive chromosomal CNV, the CNV profiles were determined using the “inferCNV” package [49]. The CNV score was calculated as the mean of the squared CNV values for each chromosome. Malignant and non-malignant cell labels were assigned based on the distribution of CNV scores relative to a reference, with malignancy being determined by the identification of bimodal patterns in the scores.

Cell cycle annotation was conducted via “CellCycleScoring” algorithm, “Seurat” R package. Cells within the “G1” phase were annotated as “non-cycling,” while the remaining cells as “cycling” [50].

The CytoTRACE algorithm, created by Gulati et al., functions to identify, smooth, and compute the expression levels of genes most correlated with single-cell gene counts from scRNA-seq data. Following the calculation, CytoTRACE assigns each single cell a score that reflects its stemness within the dataset. This algorithm offers a reliable computational framework for predicting differentiation states based on scRNA-seq data and has shown superior performance to existing techniques for stemness analysis validated on large-scale datasets. Using the “CytoTRACE” R package, CytoTRACE scores were calculated specifically for malignant cells only within the scRNA-seq data. These scores, ranging from 0 to 1, represent stemness, where higher scores indicate greater stemness (or lesser differentiation) and lower scores indicate increased differentiation [51].

stRNA-seq analysis

For spatial transcriptomics analysis, we utilized the “Seurat” R package to preprocess, normalize, and scale the raw spatial transcriptomics data. Tumor cell identification, ecosystem characterization, and cell type deconvolution were performed using “SpaCET” R package [52]. SpaCET determines cancer cell fractions using a gene pattern dictionary that captures copy number alterations and expression variations across different tumor types. This approach outperforms the conventional inferCNV-based method in

both predictive accuracy and computational efficiency. For non-malignant cells, SpaCET employs a hierarchical two-tier model to break down their composition. At the first tier, the proportions of primary cell lineages are estimated, while at the second tier, the sublineage fractions are inferred based on their corresponding major lineage fractions. These estimations utilize reference profiles derived from scRNA-seq datasets spanning multiple cancer types [53]. The CCT5 expression patterns were visualized via “SpaCET.GenSetScore” algorithm.

Prognostic value evaluation

The prognostic significance of CCT5 in pan-cancer was evaluated using univariate Cox regression analysis for four survival outcomes: OS, DSS, PFI, and DFI. The analysis was conducted using the “survival” package in R. The hazard ratios (HRs) and corresponding 95% confidence intervals (CIs) were calculated to assess the association between CCT5 expression and patient prognosis in different cancer types. The results were visualized using forest plots, which were generated via “ggplot2” R package.

Gene set enrichment

Patients were stratified into CCT5-high and CCT5-low groups based on CCT5 expression levels. Gene Set Enrichment Analysis (GSEA) was performed to investigate differences in signaling pathways between the two groups [54]. Version 7.0 of the hallmark gene sets, obtained from the MSigDB database (<https://www.gsea-msigdb.org/gsea/msigdb>), was used as the reference [55]. Differential pathway analysis was conducted across 33 cancer types, and significantly enriched gene sets were identified using an adjusted *p* value threshold of < 0.05. Results were visualized using the R package “ggplot2”.

mRNAsi and mDNAsi calculation

Based on mean-centered RNA-seq data from PSCs in the PCBC database (<https://www.synapse.org/Synapse:syn2701943>), the stemness signatures for each cancer type in TCGA cohort were identified using the one-class logistic regression (OCLR) machine learning algorithm and validated through leave-one-out cross-validation. Subsequently, the stemness index was determined by scaling Spearman correlation coefficients to a range of 0 to 1, where higher mRNAsi and/or mDNAsi values indicated more significant tumor dedifferentiation and greater stemness characteristics [56]. The correlation analysis between CCT5 expression levels in 33 cancers and their matched figures of mRNAsi and mDNAsi were then conducted, and the results were visualized by R package “ggradar”.

Immune infiltration analyses and tumor microenvironment characterization

Tumor purity, stromal scores, ESTIMATE scores, and immune scores were calculated via “ESTIMATE” R package. The fractions of infiltrated non-cancerous cells were evaluated through “CIBERSORT” R package. We then used signatures summarized by Zeng et al. for tumor microenvironment characterization [57]. Detailed information of the signatures could be seen in Additional file 1: Table S3. Considering that signatures regarding DNA damage and repairing were significantly enriched, we further analyzed the association between CCT5 expression and MMR genes which were reported to be correlated with ICB response.

Tumor-infiltrating lymphocyte identification via deep learning

Saltz et al. developed deep learning algorithms to identify tumor-infiltrating lymphocytes (TILs) from HE-stained images. Here, HE-stained images of both CCT5-high and CCT5-low expression groups were downloaded from the TCGA data portal, and the corresponding deep learning-based TIL quantification was obtained using the online tools developed by Saltz et al. (https://cancerimagingarchive.net/datascope/TCGA_TilMap/) [58]. These results provided a valuable approach to validate the in silico findings, offering additional insights into the immune landscape of tumors.

ICB response exploration

The Raiz SKCM ICB cohort was utilized to explore the association between CCT5 expression and response to ICB therapy. CCT5 expression levels were compared between responders (R) and non-responders (NR) to assess its potential role in predicting ICB efficacy. Additionally, the Rodmen SKCM scRNA-seq ICB cohort was employed for further investigation of the relationship between CCT5 expression and ICB response at the single-cell level. Given that CCT5 was predominantly expressed in malignant cells, we first isolated malignant cells from the dataset. Subsequently, we compared CCT5 expression levels between ICB-resistant malignant cells (NR) and treatment-naïve malignant cells (TN) to assess its potential role in therapy resistance.

CCT5.Sig machine learning model construction

Our investigation into the association between CCT5, cancer stemness, and TIME modulation suggested that CCT5 may influence ICB response through multiple mechanisms. Therefore, constructing a prediction model based on CCT5 is a worthwhile endeavor to further explore its potential as a biomarker for immunotherapy response.

Details of the model construction are as follows:

Step1: Selecting candidate genes for CCT5.Sig

Genes positively correlated with CCT5 (Spearman $R > 0$, $FDR < 1e-05$) were classified as Gx, while those significantly upregulated in tumor samples were designated as Gy. To identify tumor-specific genes linked to CCT5, we intersected Gx and Gy, defining the resulting gene set as Gn for all 33 bulk RNA-seq datasets within the TCGA data portal. For cohorts lacking normal tissue samples, Gn is identical to Gx. To ensure robustness across datasets, we calculated the geometric mean of Spearman R for each gene across G1 to G33. Finally, genes with a geometric mean Spearman $R > 0.4$, indicating a moderate to strong correlation, were compiled into the CCT5.Sig signature.

Step2: Datasets integration

We applied the ComBat method to remove batch effects across different ICB cohorts. The 23 ICB cohorts were stratified into two groups: one for training and validation, and the other for independent testing. Both cohorts included all eight cancer types, except for GBM, ESCA, and GC, as each had only one available cohort, which were all assigned to the first cohort for model training and validation. The first cohort included 15 ICB cohorts consist of 931 patients from the following datasets: Ascierto et al. (ccRCC, SKCM), Miao et al. (ccRCC), Liu et al. (SKCM), VanAllen et al. (SKCM), Auslander et al. (SKCM), Hugo et al. (SKCM), Zhao et al. (GBM), Kim et al. (GC), Jung et al. (NSCLC), Homet et al. (UC), Mariathasan et al. (UC), Vanden et al. (ESCA), Nanjing et al. (HCC), and Imbrave_150 (HCC). We randomly divided the first cohort into two datasets: a training set comprising 80% of the patients ($n = 745$) and a validation set comprising 20% of the patients ($n = 186$). This division ensures a balanced distribution for model training and evaluation. The remaining eight cohorts were consolidated as an independent testing set.

Step3: Univariate Cox regression

All genes within the CCT5.Sig signature underwent univariate Cox regression analysis across multiple cohorts. The genes that consistently demonstrated significance (threshold: p value < 0.2) across all cohorts were identified through intersection. The remaining genes were then selected for subsequent machine learning analyses.

Step4: Model training and parameter tuning

We developed an ICB response classification model using CCT5.Sig and trained it with six widely used machine learning (ML) algorithms: support vector machine (SVM), Naïve Bayes (NB), random forest (RF), k-nearest neighbors (KNN), AdaBoost classification trees (AdaBoost), and boosted logistic regression (LogitBoost). For all ML algorithms requiring parameter tuning, fivefold cross-validation (CV) was employed for hyperparameter optimization, ensuring optimal model

performance. To enhance robustness, the optimization process was repeated 10 times with different random seeds for each resampling. Final genes involved in the CCT5.Sig are presented in Additional File 1: Table S4.

Step5: Model validation and testing

We developed six models from the training set, each utilizing a different ML algorithm. These models were then applied to the validation set, and their performance was compared. The model with the best performance was selected as the final CCT5.Sig model.

To assess its predictive value, the final model was further applied to the testing set, ensuring its robustness and generalizability.

Step6: Performance comparison

To further assess the predictive value of CCT5.Sig, we compared it with previously reported ICI response signatures, including four pan-cancer signatures, Stem.Sig [12], CD274 (PDL1) expression [59], CD279 (PD1) expression [32, 60], and CTLA4 expression [61]. Details of the Stem.Sig signature and its corresponding algorithms are provided in Additional File 1: Table S5.

Immunohistochemical staining (IHC staining)

After embedding, tumor and adjacent paracancerous tissues were sectioned into 4 μm paraffin slices, followed by deparaffinization and hydration. Antigen retrieval was performed using citrate buffer, and endogenous peroxidase activity was blocked with 3% H_2O_2 , followed by blocking with 5% BSA. The sections were then incubated overnight at 4 $^\circ\text{C}$ with primary antibodies, CCT5 (11603-1-AP, Proteintech). The following day, after incubation with secondary antibodies, the sections were stained with 3,3'-diaminobenzidine (DAB). Subsequently, hematoxylin counterstaining was performed, followed by air-drying and mounting with neutral resin. Finally, all slides were observed and imaged.

Statistical analysis

R software (version 4.2.1) was used for data processing, statistical analysis, and visualization. To explore relationships between continuous variables, we calculated Spearman's correlation coefficients. For comparisons involving categorical data, the χ^2 test was applied. Continuous variables were assessed using either T-tests or Wilcoxon rank-sum tests, depending on their distribution. A significance level of $p < 0.05$ was set for all statistical tests, including GSEA (based on adjusted p-values), except for the univariable Cox regression in the third step of ML, where the threshold was set at 0.2.

Results

Expression patterns of CCT5 in normal and cancer tissue

We analyzed CCT5 expression across normal human tissues using the GTEx database (Fig. 1A). CCT5 was highly expressed in tissues such as the testis, esophagus, uterus, adipose tissue, cervix, thyroid, vagina, breast, bladder, adrenal gland, and lung, which are prone to tumorigenesis. In contrast, non-proliferative tissues such as the heart and brain exhibited lower CCT5 expression levels. To compare CCT5 expression between normal and tumor tissues, we conducted differential expression analysis using TCGA cohorts (Fig. 1B). Given the limited normal tissue samples in TCGA, we also integrated data from TCGA and GTEx to obtain a more comprehensive comparison (Fig. 1C). The results showed significantly higher CCT5 expression in tumor tissues, except for LAML and THCA, where expression was lower than in normal tissues. Further analysis of paired tumor and adjacent normal tissues validated these findings (Fig. 1D), indicating the potential roles of CCT5 in tumorigenesis.

To further validate the differential expression of CCT5, we performed IHC staining on tumor and adjacent normal tissues from various cancer types (Fig. 1E). The results demonstrated increased CCT5 protein expression in tumor tissues compared to normal tissues across multiple cancers, including LUAD, LIHC, COAD, BRCA, STAD, and READ. These findings further support the upregulation of CCT5 in tumor samples at both the transcriptomic and protein levels.

Expression patterns of CCT5 in scRNA-seq and stRNA-seq data

To investigate CCT5 expression at the single-cell level, we analyzed scRNA-seq datasets from various cancer types. CCT5 was predominantly expressed in malignant cells, with lower but notable expression in proliferative T cells, monocytes, and macrophages (Fig. 2A). Furthermore, in order to further elucidate CCT5 expression across different cell types, we visualized single-cell transcriptomic data from LIHC, BRCA, COAD, and CHOL, and mapped CCT5 expression on UMAP plots. The results revealed higher CCT5 expression in malignant cell and cycling cells (Fig. 2B–I). Violin plots for CCT5 expression in scRNA-seq data are presented in Additional File2: Fig. S1. Finally, to validate the preferential expression of CCT5 in malignant cells, we performed spatial transcriptomic analysis across multiple cancer types, including GBM,

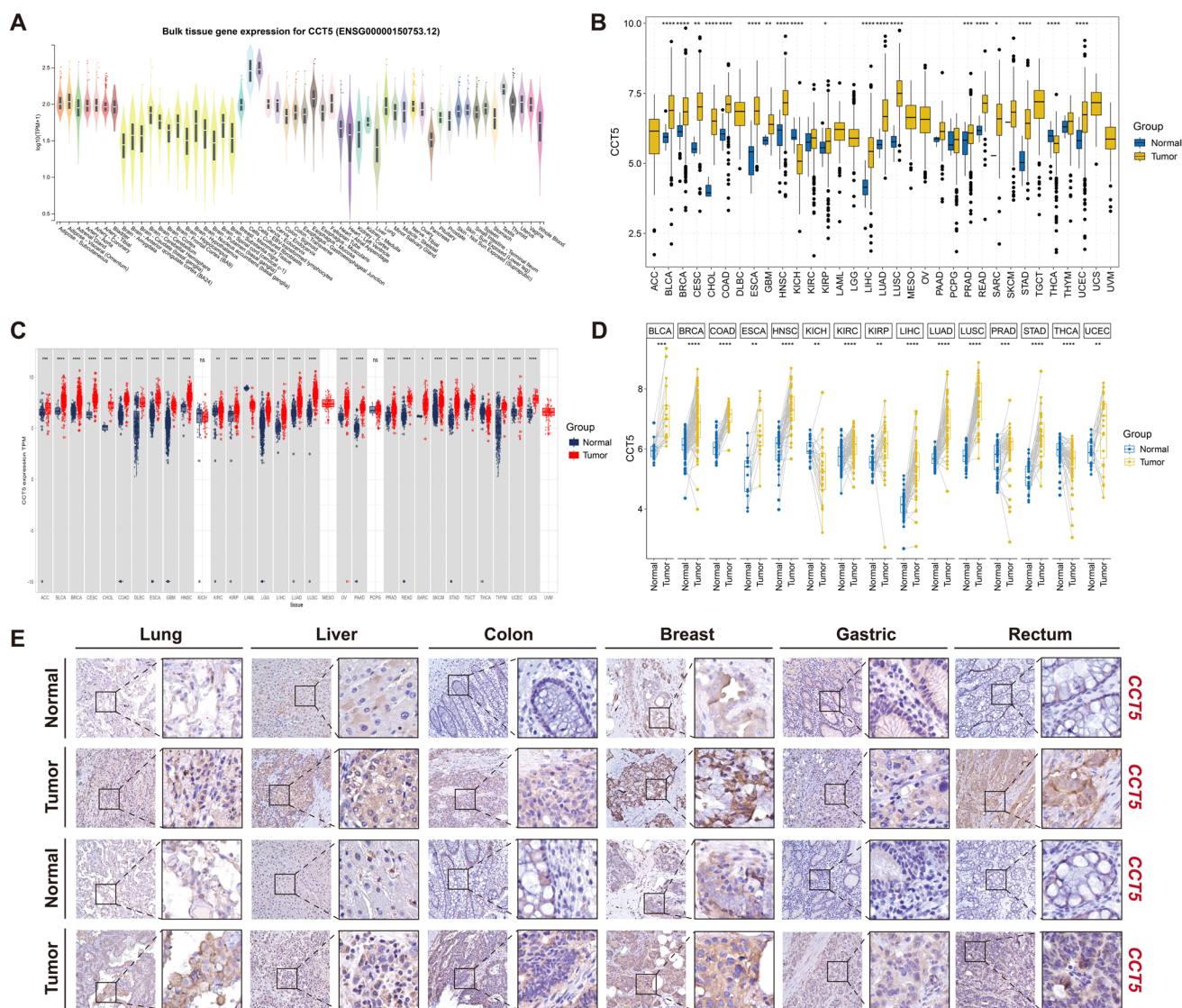


Fig. 1 Analysis of CCT5 expression patterns using bulk RNA-seq and IHC. **(A)** Violin plot showing CCT5 expression across various human tissues. **(B)** Boxplot comparing CCT5 expression in normal (blue) and tumor (yellow) tissues across cancer types (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). **(C)** Scatter plot of CCT5

expression in normal (blue) and tumor (red) tissues. **(D)** Paired dot plot of CCT5 expression in matched normal and tumor samples. **(E)** Immunohistochemical staining of CCT5 in normal and tumor tissues from lung, liver, colon, breast, gastric, and rectal samples

CHOL, BRCA, and UCEC (Fig. 2J–Q). The malignant cell regions exhibited a higher CCT5 expression level, further supporting its tumor-specific upregulation.

Clinical significance of CCT5 in pan-cancer

The prognostic significance of CCT5 expression was assessed across 33 cancer types using univariate Cox regression, focusing on OS, DSS, DFI, and PFI. High CCT5 expression was identified as a risk factor for OS in 11 cancers, including ACC, BRCA, KICH, KIRP, LGG, LIHC, LUAD, MESO, PAAD, SARC, and UVM (Fig. 3A).

Similarly, CCT5 overexpression correlated with shorter DSS in LIHC, KIRP, PRAD, MESO, SARC, UVM, KICH, LGG, HNSC, ACC, LUAD, and PAAD, and with shorter DFI in LIHC and KIRP (Fig. 3B, C). For PFI, high CCT5 expression was associated with poor prognosis in LIHC, UVM, KIRP, SARC, ACC, BLCA, PAAD, and MESO (Fig. 3D). Notably, in OV, high CCT5 expression was linked to a better prognosis across all four survival indicators. These findings highlighted CCT5 as a potential prognostic biomarker, with its overexpression generally indicating worse clinical outcomes in multiple cancers.

Gene set enrichment of CCT5

Samples within each bulk RNA-seq cohort were divided into CCT5-low and CCT5-high groups, and differential analysis was conducted. The signaling pathways associated with CCT5 and its potential molecular mechanisms in the progression of 33 cancer types were subsequently investigated using GSEA. A bubble diagram was employed to visualize pan-cancer enrichment of GSEA pathways based on high and low CCT5 expression. The analysis revealed that high CCT5 expression was significantly enriched in key oncogenic and cell cycle regulation related pathways, including MYC_TARGETS_V1, MTORC1_SIGNALING, MITOTIC_SPINDLE, G2M_CHECKPOINT, and E2F_TARGETS (Fig. 4A).

Given that many of the enriched pathways are involved in cell cycle regulation, we further investigated the relationship between CCT5 expression and proliferative cell states in specific cancers, including LIHC, SKCM, CHOL, and BRCA, given their relatively higher NES scores of the aforementioned signaling. UMAP clustering revealed distinct populations of cycling and non-cycling cells (Fig. 4B, D, F, H), and the box plots demonstrated that CCT5 expression was significantly higher in cycling cells compared to non-cycling cells (Fig. 4C, E, G, I), which were in line with the GSEA findings.

CCT5's association with cancer stemness

Firstly, we analyzed mRNAsi and mDNAsi across various cancer types in TCGA cohorts. The radar plots demonstrated a significant positive correlation between CCT5 expression and mRNAsi in multiple cancers (Fig. 5A), indicating a potential role of CCT5 in maintaining tumor stemness at the transcriptomic level. Similarly, CCT5 expression correlated with mDNAsi in several cancers, further supporting its association with cancer stem-like properties (Fig. 5B).

To study the impact of CCT5 on the stemness of malignant cells at the single-cell level, we selected primary liver cancer (LIHC and CHOL), READ, and BRCA datasets based on the results of mRNAsi and mDNAsi results. To avoid the influence of normal epithelial cells on the results, we first extracted the malignant cells from the datasets and then performed CytoTRACE analysis to assess their stemness. The T-SNE and box plots revealed that malignant cells with high CCT5 expression were poorly differentiated, as indicated by higher CytoTRACE scores, in the primary liver cancer, READ, and BRCA datasets (Fig. 5C–H).

Association between CCT5 and tumor immune microenvironment

To explore the role of CCT5 in the tumor immune microenvironment (TIME), we analyzed its correlation with immune

infiltration and immune-related pathways across multiple cancer types. Figure 6A shows that high CCT5 expression was associated with lower immune infiltration levels, as indicated by higher tumor purity, and lower ESTIMATE, Immune, and Stromal scores. Additionally, CCT5 expression was positively correlated with DNA damage, repair and replication pathways, while showing a negative correlation with immune activation pathways including antigen processing and CD8⁺T cell activation (Fig. 6B). According to the CIBERSORT results, CCT5 expression was positively associated with nTreg, iTreg, and neutrophil infiltration levels in a majority of cancer types (Fig. 6C). Conversely, CCT5 expression was negatively associated with the abundance of CD4⁺T, CD8⁺T, Tfh, Tgd, and NK cells within the TME in the majority of cancers. Notably, CCT5 expression was positively correlated with CD8⁺T abundance in UVM (Fig. 6C). CCT5 expression showed either positive or negative correlation with the infiltration of 16 other immune cell types across multiple cancers (Fig. 6C). Given that MMR status is a well-established biomarker for ICB response, and its associated signaling pathways are highlighted in Fig. 6B, we further examined the relationship between CCT5 expression and MMR genes. The analysis revealed a strong positive correlation between CCT5 and MMR gene expression, suggesting a potential link between CCT5 and genomic instability, which may influence ICB response (Fig. 6D). We also performed TIL mapping using HE-stained images across multiple cancer types. CCT5-high tumor expression exhibited lower TIL infiltration, suggesting CCT5 contributes to an immunosuppressive tumor microenvironment (Fig. 6E).

We next investigated the association between CCT5 expression and immune checkpoint blockade (ICB) response in SKCM patients. In the 2017_Riaz_SKCM cohort, patients with ICB resistance (NR) showed significantly higher CCT5 expression than responders (R) (Fig. 6F). Subsequently, we managed to validate this association in scRNA-seq data. Since data for responders (R) were not available in this cohort, we compared cancer stemness between non-responders (NR) and treatment-naïve (TN) patients instead. Ideally, a direct comparison between responders and non-responders would provide more precise insights. However, given that treatment-naïve patients (TN) likely include both potential responders and non-responders, we followed a previously established approach to assess differences of CCT5 expression between NR and TN patients. We first annotated the SKCM scRNA-seq data and extracted the malignant cells from the cellular landscape, as CCT5 was predominantly expressed in these malignant cells (Fig. 6G, H). We subsequently compared the CCT5 expression level between malignant cells from NR and TN patients. The results shown that CCT5 was predominantly expressed in ICB-resistant malignant cells (NR) compared to treatment-naïve (TN) cells (Fig. 6I, J).

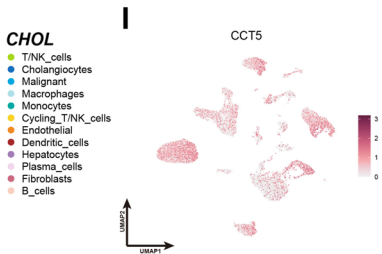
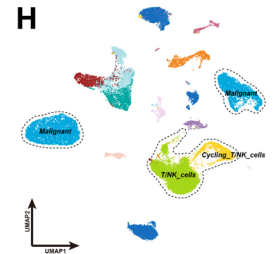
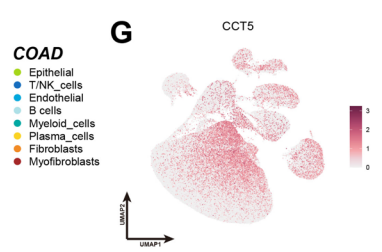
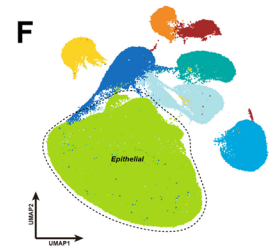
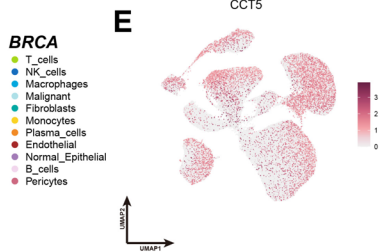
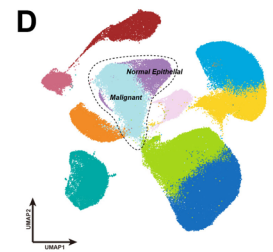
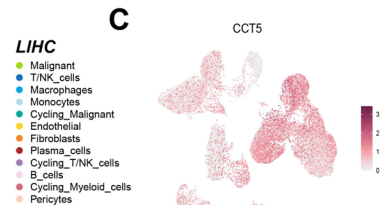
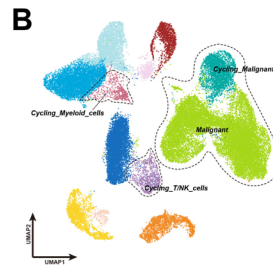
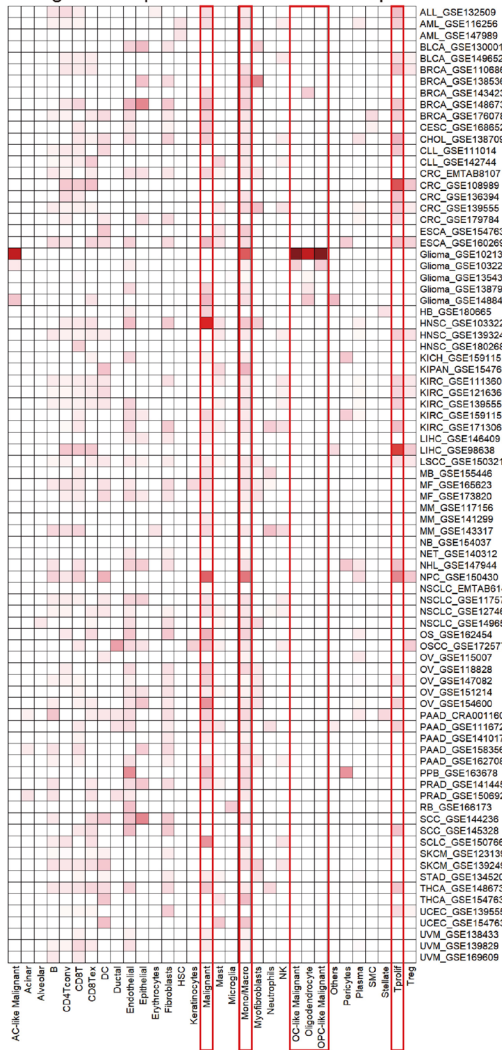
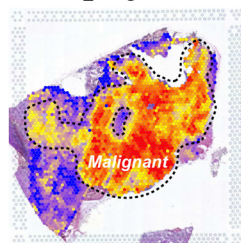
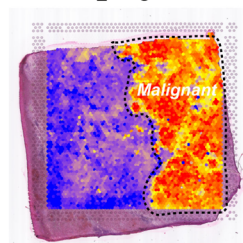
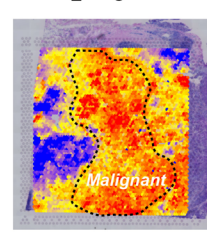
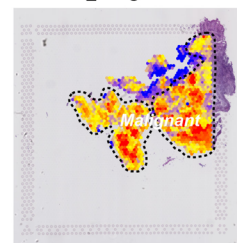
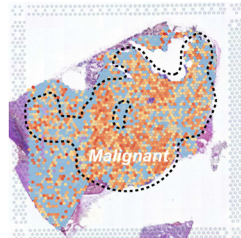
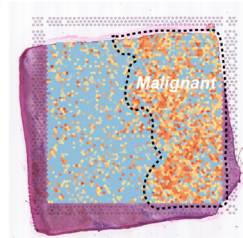
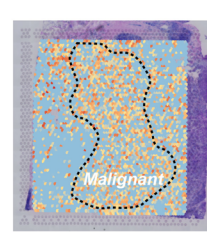
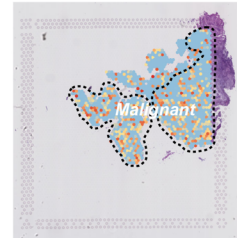
A Single cell expression data of CCT5 in pan-cancer**J** GBM_Malignant cell**L** CHOL_Malignant cell**N** BRCA_Malignant cell**P** UCEC_Malignant cell**K** CCT5**M** CCT5**O** CCT5**Q** CCT5

Fig. 2 Analysis of CCT5 expression patterns using scRNA-seq and stRNA-seq data. **(A)** Heatmap showing CCT5 single-cell expression across various cancer types. **(B–C)** UMAP plots of LIHC: **(B)** cell type distribution, **(C)** CCT5 expression. **(D–E)** UMAP plots of BRCA: **(D)** cell type distribution, **(E)** CCT5 expression. **(F–G)** UMAP plots of COAD: **(F)** cell type distribution, **(G)** CCT5 expression. **(H–I)** UMAP plots of CHOL: **(H)** cell type distribution, **(I)** CCT5 expression. **(J–K)** Spatial transcriptomics of GBM: **(J)** malignant cell distribution, **(K)** CCT5 expression. **(L–M)** Spatial transcriptomics of CHOL: **(L)** malignant cell distribution, **(M)** CCT5 expression. **(N–O)** Spatial transcriptomics of BRCA: **(N)** malignant cell distribution, **(O)** CCT5 expression. **(P–Q)** Spatial transcriptomics of UCEC: **(P)** malignant cell distribution, **(Q)** CCT5 expression

Immunotherapy outcome prediction by CCT5.Sig

The flow chart of CCT5.Sig genes selection and machine learning process is presented in Fig. 7. Firstly, we pooled a total of 605 CCT5.Sig genes from pan-cancer TCGA cohort. After the univariable Cox regression, 82 (13.5%) genes were kept for further ML progress. Subsequently, we trained the model using six different machine learning algorithms and optimized the parameters of each model through 10-time repeated fivefold cross-validation. This process resulted in the development of six distinct models. After training, we assessed and compared their AUC performance in the internal validation cohort. Among them, the random forest model achieved the highest AUC of 0.74 (95% CI 0.69–0.78), making it the optimal choice for the CCT5.Sig model (Fig. 8A). To further evaluate the CCT5.Sig model, we applied it to an independent testing set to predict ICB response. The model demonstrated consistent performance, achieving the AUC of 0.70 (95% CI 0.65–0.75) (Fig. 8B–C).

To assess whether the CCT5.Sig model can predict OS, we stratified ICB-treated patients into low-risk and high-risk subgroups based on their predicted response status. The Kaplan–Meier analysis of OS demonstrated a significantly longer OS in the low-risk group across the training, validation, and independent testing sets (all log-rank $p < 0.01$).

In the validation set, high-risk patients identified by the CCT5.Sig model had a median OS of 8.43 months, significantly shorter than the 32.67 months observed in low-risk patients (HR: 2.48; 95% CI 1.23–4.99) (Fig. 8D). A similar trend was seen in the independent testing cohort, where high-risk patients had a median OS of 35.77 months, while the figure for high-risk patients was 22.87 months (HR: 1.48, 95% CI 1.06–2.05) (Fig. 8E).

We then conducted a subgroup analysis across five individual cohorts that contributed to the testing set. For ICI response prediction, the AUC ranged from 0.52 to 0.84 across these cohorts (Additional File2: Fig. S2). Among them, the Korea HCC cohort achieved the highest performance with an AUC of 0.84 (95% CI 0.70–0.99), closely followed by the Rose 2021 UC cohort (AUC: 0.82; 95% CI 0.73–0.92). In contrast, the Synder 2017 UC cohort

exhibited the lowest predictive performance, with an AUC of 0.52 (95% CI 0.26–0.79).

We further compared the performance of CCT5.Sig with previously established predictive gene signatures. Among the pan-cancer signatures, including Stem.Sig, PD1 expression, PDL1 expression, and CTLA4 expression, CCT5.Sig demonstrated the best performance in the testing set, achieving an AUC of 0.76, followed by Stem.Sig, which attained an AUC of 0.73 (Fig. 8F). Other pan-cancer signatures exhibited strong predictive performance in only three or four cohorts (Fig. 8G).

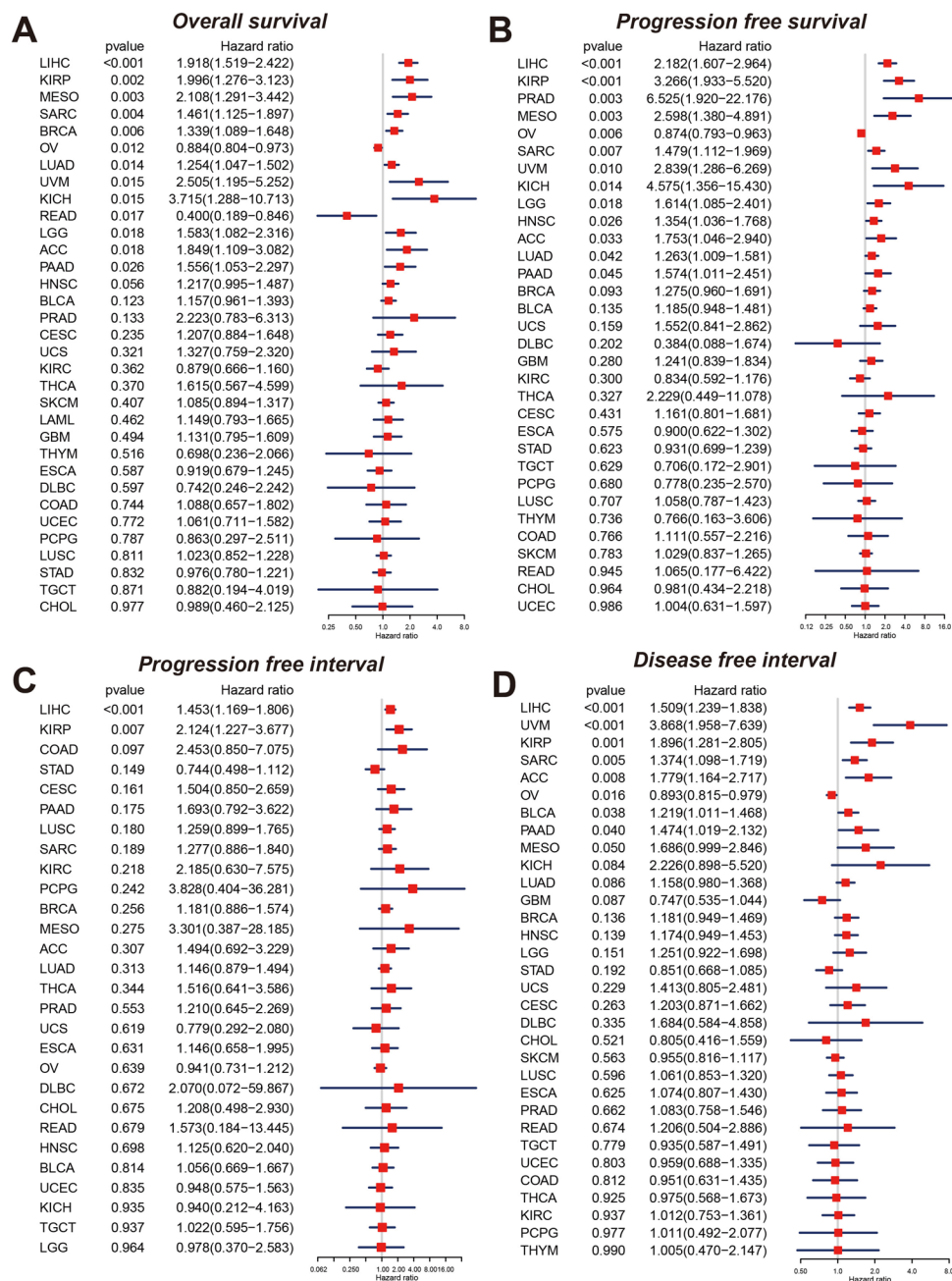
Discussion

Significant advancements in immunotherapy have enabled a great number of cancer patients to extend their lives. However, its efficacy varies among individuals due to the heterogeneity of the tumor microenvironment (TME) [62]. Identifying key predictive factors or developing models to assess immunotherapy response holds great promise for achieving precision and personalized medicine [63]. Despite the discovery of numerous biomarkers and the development of predictive models for immunotherapy response, many patients still experience resistance. Notably, most existing models have been tailored to single tumor types, leading to inconsistent predictive performance across different cancers and limiting their broader clinical applicability. Furthermore, even models developed using pan-cancer immunotherapy cohorts often suffer from other limitations, such as the inclusion of excessive gene markers or modest overall AUC performance, which may hinder their clinical utility [64].

In this study, we conducted a comprehensive pan-cancer analysis of CCT5, revealing its critical role in cancer progression, tumor stemness, and immunotherapy response. By integrating bulk RNA-seq, scRNA-seq, and stRNA-seq data, we provided a multi-dimensional perspective on CCT5 expression and its biological implications across multiple tumor types. Furthermore, through machine learning, we established CCT5.Sig, a robust pan-cancer predictive model for ICB response, which demonstrated superior performance compared to existing signatures.

Previous studies have highlighted CCT5 as a molecular chaperone involved in protein folding and cell cycle regulation, with elevated expression observed in multiple cancers such as LIHC, STAD, and READ [8, 10, 65]. However, these studies primarily focused on individual cancer types, lacking a comprehensive pan-cancer perspective. Our study expanded on these findings by demonstrating that CCT5 was consistently upregulated across multiple tumor types and correlates with poor prognosis, establishing it as a potential universal prognostic biomarker. Moreover, we revealed CCT5's role as a key regulator of cell cycle, DNA

Fig. 3 Prognostic value of CCT5 in pan-cancer. (A–D) CCT5's association with OS (A), PFS (B), PFI (C), DFI (D)



damage repairing, tumor dedifferentiation and stemness maintenance, further supporting its contribution to aggressive tumor phenotypes. A particularly novel aspect of our study is the identification of CCT5's involvement in immune evasion and immunotherapy resistance. We found that high CCT5 expression is associated with increased regulatory T cells (Tregs) and neutrophil infiltration, as well as reduced recruitment of tumor-infiltrating lymphocytes, creating an immunosuppressive tumor microenvironment that contributes to poor response to immune checkpoint blockade (ICB) therapy. To translate these findings into clinical applications, we developed CCT5.Sig, a machine learning-based model

trained on 23 ICB cohorts across eight cancer types, which demonstrated high predictive accuracy (AUC = 0.82 in validation set, 0.76 in independent testing set), outperforming several previously established pan-cancer immunotherapy biomarkers [66, 67].

By integrating bulk RNA-seq, scRNA-seq, and stRNA-seq, our study provides a comprehensive, multi-dimensional analysis of CCT5 expression patterns. Bulk RNA-seq data reveal that CCT5 was significantly upregulated in tumor tissues across various cancer types, emphasizing its potential as a tumor-associated biomarker. Single-cell analysis further demonstrated that CCT5 was predominantly expressed in

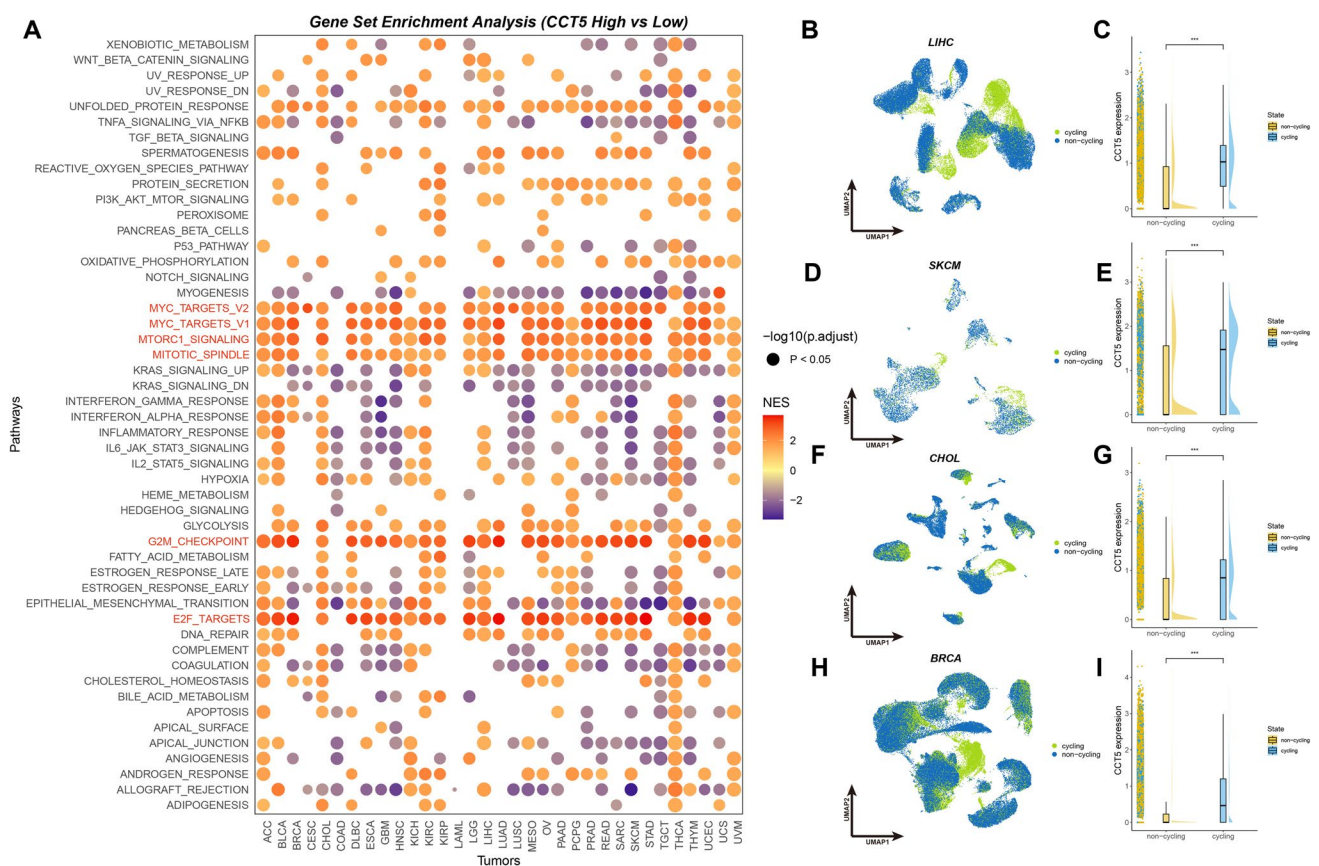


Fig. 4 Gene set enrichment and cell cycle analysis of CCT5 in pan-cancer. **(A)** GSEA results of CCT5. **(B–I)** Cell cycle analyses regarding CCT5 in LIHC **(B–C)**, SKCM **(D–E)**, CHOL **(F–G)** and BRCA **(H–I)**

malignant cells, particularly those undergoing active cell cycling, indicating its role in tumor proliferation and progression. Spatial transcriptomics supported these findings as CCT5 expression was enriched in tumor regions rather than adjacent normal tissues, reinforcing its tumor-specific oncogenic function. Together, these results indicated that CCT5 might be a key factor in cancer biology [68].

Our findings also revealed that CCT5 could serve as a significant prognostic biomarker across multiple cancers, where its high expression consistently correlates with poor patient outcomes. This strong association between CCT5 overexpression and poor prognosis likely stemmed from its involvement in cell proliferation, stemness maintenance, and immune evasion, which contributed to tumor aggressiveness and therapy resistance. Notably, while elevated CCT5 expression generally predicted worse survival, its prognostic significance appeared to be cancer-type-specific, as seen in ovarian cancer, where it was linked to better outcomes.

Through Gene Set Enrichment Analysis, our study identified CCT5 as a key player in cell cycle regulation, with its high expression enriched in major oncogenic pathways. Specifically, CCT5 was strongly associated with pathways driving cell cycle progression, including MYC targets, MTORC1

signaling, the G2/M checkpoint, E2F targets, and mitotic spindle regulation [69]. Further validation at the single-cell level confirmed that CCT5 was predominantly expressed in cycling malignant cells, especially those in the G2/M phase, reinforcing its role in mitotic regulation. These findings suggested that CCT5 promoted cell division and may contribute to aggressive tumor behavior [70].

In addition to its role in cell cycle regulation, CCT5 expression showed a strong correlation with DNA damage response (DDR) pathways, linking it to genomic instability and tumor survival mechanisms [71]. Given its association with tumor proliferation, CCT5 may enhance DNA repair dependency, allowing tumor cells to withstand replication stress and maintain survival. High CCT5 expression was positively correlated with MMR genes, suggesting a role in modulating DNA repair efficiency [3]. This interplay between cell cycle dysregulation and DDR activation could explain why CCT5-overexpressing tumors exhibit resistance to chemotherapy, aligning with findings from Ooe A and colleagues on CCT5's role in BRCA chemo-resistance [10]. These results highlighted CCT5 as a potential therapeutic target, where disrupting its function may enhance cancer treatment efficacy.

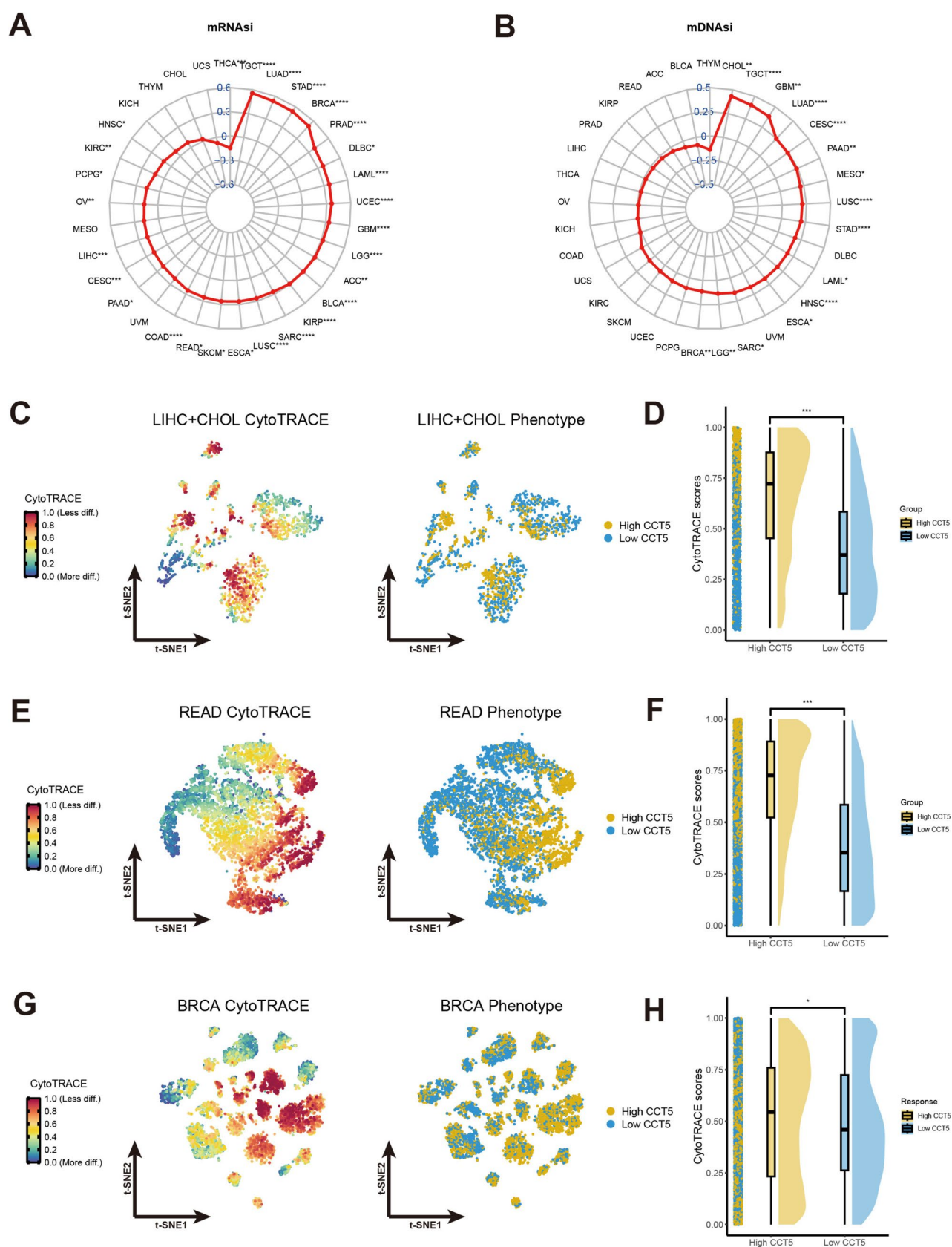


Fig. 5 CCT5 expression is correlated with cancer stemness. (A) CCT5 expression and mRNAi. (B) CCT5 expression with mDNAi. (C–H) CCT5 expression levels and CytoTRACE scores in primary liver cancer (C–D), READ (E–F) and BRCA (G–H)

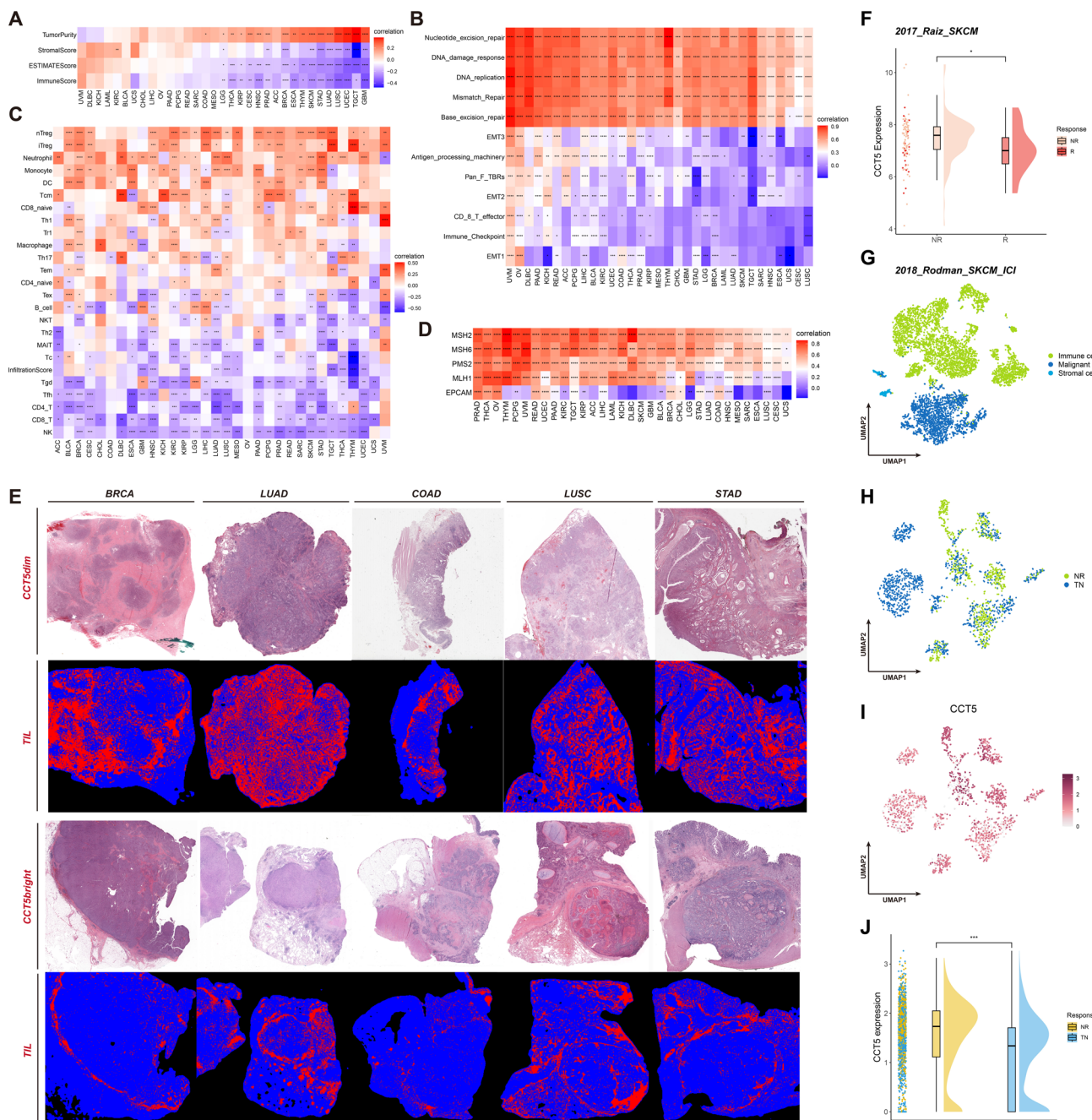


Fig. 6 Correlation of CCT5 expression with tumor microenvironment and immunotherapy response. **(A)** Heatmap showing the correlation of CCT5 expression with stromal and immune scores across different cancer types. **(B)** Correlation between CCT5 expression and key biological pathways, including DNA replication, immune activation, and epithelial-mesenchymal transition (EMT). **(C)** Heatmap displaying the relationship between CCT5 expression and various immune cell types in different cancers. **(D)** Correlation of CCT5 expression with markers of mesenchymal stem cells (MSCs) and epithelial cell adhesion molecule (EPCAM). **(E)** Histopathological analysis of tumor

samples from BRCA, LUAD, COAD, LUSC, and STAD, showing CCT5-stained slides (top) and corresponding tumor-infiltrating lymphocyte (TIL) distributions (bottom). **(F)** Violin plot comparing CCT5 expression between responders (R) and non-responders (NR) in the 2017_Riaz_SKCM cohort. **(G–I)** UMAP plots of 2018_Rodman_SKCM_ICI dataset: **(G)** immune cell clustering, **(H)** immunotherapy response groups, and **(I)** CCT5 expression levels. **(J)** Violin plot comparing CCT5 expression in treat-naïve patients (TN) and ICB non-responders (NR)

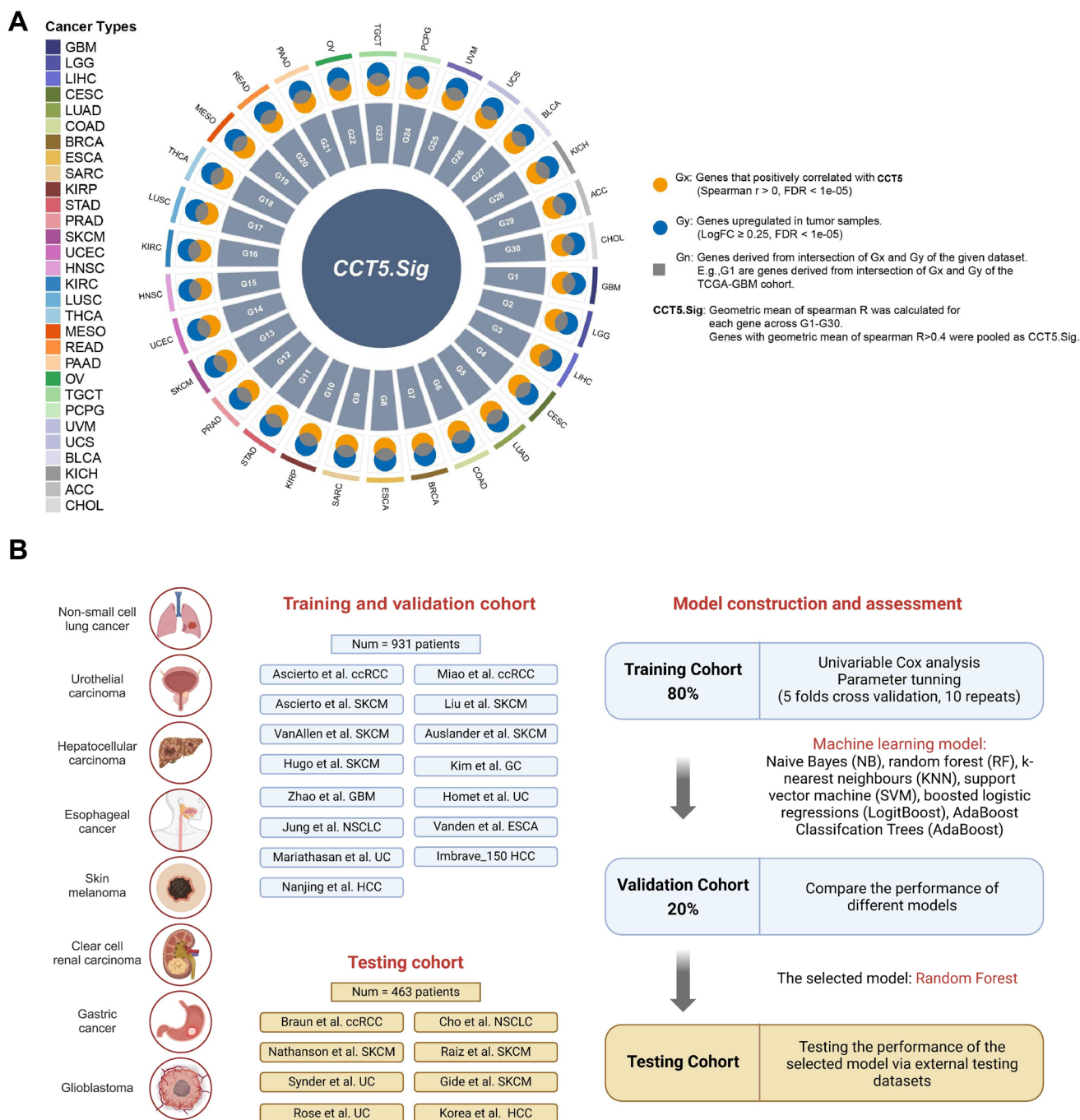


Fig. 7 Flow charts of the machine learning process. **(A)** Selection of CCT5.Sig genes. **(B)** Datasets integration and machine learning process

At the same time, our study found a strong correlation between CCT5 expression and cancer stemness, with high CCT5 levels positively associated with mRNasi and mDNasi across multiple cancer types. Single-cell analysis further illustrated its enrichment in poorly differentiated malignant cells, suggesting a role in maintaining stem-like properties.

Given that microsatellite instability (MSI) and cancer stemness were associated with ICB resistance, we investigated CCT5's role in TIME regulation [72–75]. In this study, the relationships between CCT5 expression and TIME characteristics across various cancer types use a combination of gene set scoring and Spearman correlation analysis. Gene set scoring, based on predefined immune-related gene sets validated in multiple studies, enabled us to systematically assess immune pathway activities. This approach provided valuable insights into how CCT5 expression may influence immune responses within the tumor microenvironment. Spearman correlation analysis, a nonparametric method, was particularly effective in detecting monotonic relationships between CCT5 expression and immune-related pathways. By combining these methods, our analysis revealed that high CCT5 expression correlated with an immunosuppressive (“cold”) TIME, characterized by lower immune infiltration and reduced CD8⁺ T cell activation, potentially contributing to ICB resistance [76]. To translate these findings into clinical application, we developed CCT5.Sig, a machine learning-based predictive model that demonstrated high accuracy in ICB response prediction, offering a promising tool for personalized immunotherapy strategies. Unlike using CCT5 alone, CCT5.Sig provided enhanced predictive power and broader applicability. That is because a single-gene signature like CCT5 may not fully capture the complex interactions between tumor microenvironment, stemness, and immune modulation, whereas CCT5.Sig integrated multiple CCT5-associated genes, resulting in a more comprehensive and reliable model [77, 78]. Furthermore, as CCT5 expression varied across cancer types, a multi-gene signature could promote model stability and generalizability, making CCT5.Sig a robust pan-cancer signature for ICB response prediction [59, 75, 79].

Despite the insights, our study has several limitations that warrant further investigation. Firstly, while CCT5.

Sig demonstrated strong predictive accuracy for ICB response, the model primarily relied on transcriptomic data and machine learning algorithms. And experimental validation, such as CCT5 knockdown or overexpression in cancer models, needs to be carried out for validation. Future studies should incorporate functional assays to confirm CCT5's role in tumor proliferation, stemness, and immune regulation. Also, while our study offers a comprehensive analysis of CCT5's roles across pan-cancer, we acknowledge that integrating co-expression network analysis, such as hdWGCNA, could provide deeper insights into the underlying gene regulatory networks. By uncovering co-expression modules and identifying key regulatory factors, hdWGCNA could further elucidate how CCT5 interacts with other genes in the context of tumor progression, immune modulation, and cancer stemness. This is especially true in pan-cancer analyses focusing on transcription factors, combining hdWGCNA and transcriptional regulatory network (TRN) inference algorithms could facilitate the identification of shared transcriptional signatures and regulatory networks across multiple cancer types, thus enabling the discovery of universal biomarkers and therapeutic targets. In addition, our findings suggested that CCT5 functions as a pan-cancer biomarker, but tumor heterogeneity was not fully addressed. While multiple cancer types were analyzed, we did not perform a detailed evaluation of subtype-specific differences (e.g., HR-positive vs. triple-negative breast cancer) or the impact of clinical factors such as prior treatments, and/or metabolic conditions. These variables may influence CCT5s' prognostic significance, and future studies should incorporate stratified analyses to determine whether CCT5's effects are consistent across different patient populations. Furthermore, even though we applied ComBat to correct for batch effects, but dataset variability due to differences in sample collection, sequencing depth, and tumor microenvironment composition may still influence the results. Finally, while our study discusses CCT5 as a potential therapeutic target, we did not explore whether pharmacological modulation of CCT5 could enhance ICB response. Further research should investigate small-molecule inhibitors or RNA-based silencing strategies targeting CCT5 and assess their effects on tumor growth, immune infiltration,

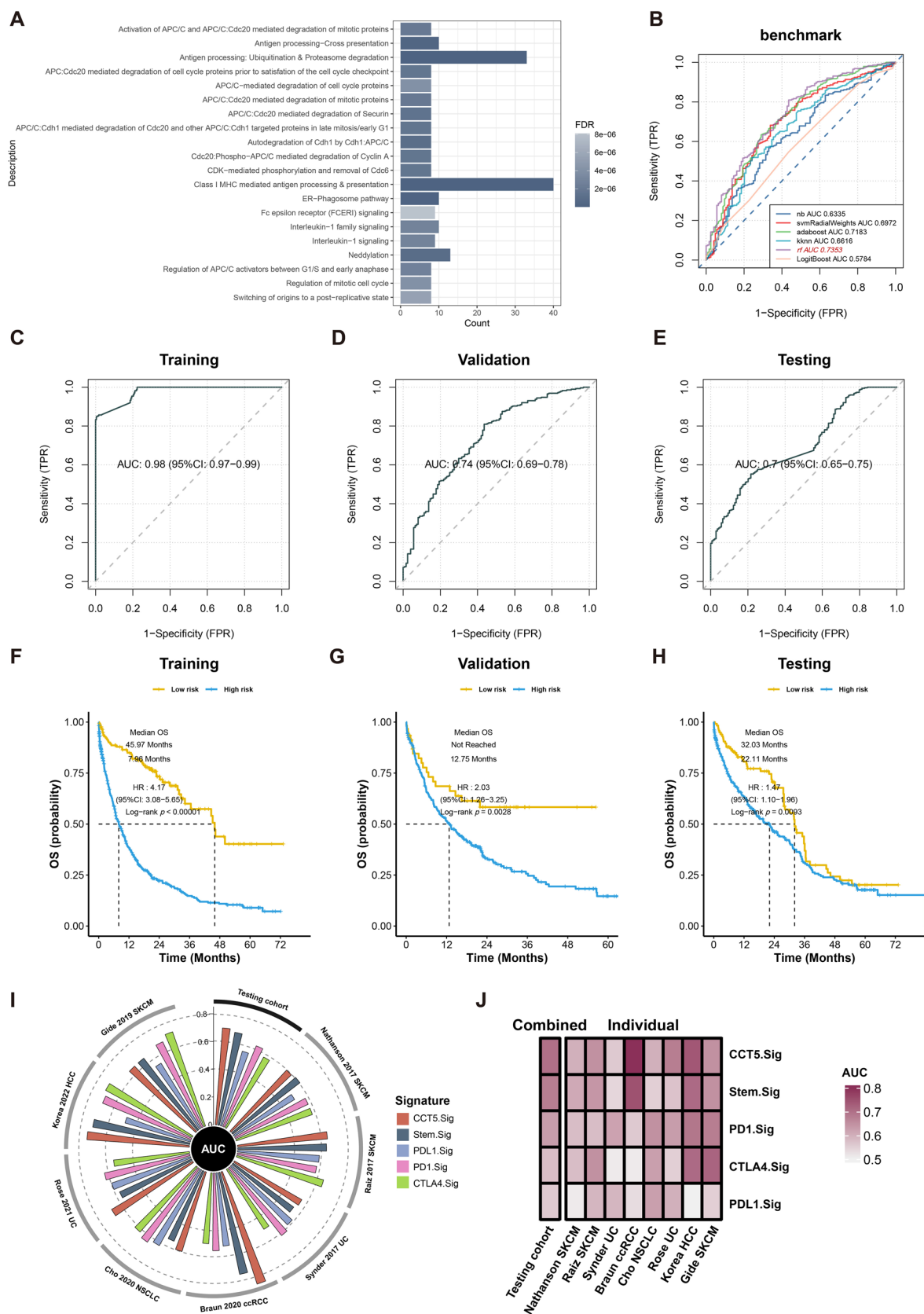


Fig. 8 Functional enrichment, predictive performance, and survival analysis of CCT5.Sig. **(A)** Bar plot displaying enriched biological pathways associated with CCT5 expression. **(B)** ROC curves comparing different predictive models. **(C–E)** ROC curves for CCT5.Sig performance in training **(C)**, validation **(D)**, and test **(E)** cohorts. **(F–H)** Kaplan–Meier survival curves comparing overall survival (OS) between high-risk and low-risk groups in training **(F)**, validation **(G)**, and test **(H)** cohorts. **(I–J)** Circus plots **(I)** and heatmap **(J)** show the comparison between the performance of the CCT5.Sig model and previously published pan-cancer models for response to immunotherapy on different testing sets

and therapy sensitivity [80]. If CCT5 inhibition proves effective, it could serve as a new therapeutic strategy to enhance immunotherapy efficacy.

Conclusion

We performed a comprehensive multi-dimensional pan-cancer analysis of CCT5, uncovering its critical roles in cell cycle regulation, DNA damage response, cancer stemness maintenance, and tumor microenvironment modulation. To translate these findings into clinical application, we developed CCT5.Sig, a machine learning-based model trained on pan-cancer ICB cohorts, enabling accurate immunotherapy response prediction and advancing personalized cancer treatment strategies.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00262-025-04071-7>.

Acknowledgements The authors extend their sincere gratitude to BioRender for providing some of the materials used in the figures. The authors also acknowledge the contributions of the TCGA data portal, UCSC Xena data portal, and all researchers who generously shared their valuable sequencing data, which made this study possible. Jiajun Li would like to thank Zhen Zhang and Ruihua Xu for their outstanding data visualization codes made publicly available on GitHub, which greatly facilitated this work.

Author contributions JL was involved in conceptualization, investigation, funding acquisition, writing—original draft, and writing—review and editing. NX was involved in writing—original draft and writing—review and editing. LH was involved in investigation, writing—original draft, and validation. JX was involved in visualization. YH was involved in visualization. DW was involved in validation. FC was involved in writing—review and editing. YW was involved in writing—review and editing. JJ was involved in writing—review and editing. YH was involved in visualization, investigation, and writing—review and editing. HY was involved in conceptualization, investigation, writing—original draft, and writing—review and editing. The authors read and approved the final manuscript.

Funding This work was supported by National College Students Innovation and Entrepreneurship Training Program, China (No. 202410343023).

Data availability No datasets were generated or analyzed during the current study.

Declarations

Competing interests The authors declare no competing interests.

Ethics approval and consent to participate The studies involving humans were approved by the ethics committee of The First Affiliated Hospital of Wenzhou Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Consent for publication This manuscript does not include details, images, or videos relating to an individual person; therefore, consent for publication is not needed beyond the informed consent provided by all study participants as described above.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. GBD 2021 Global Stillbirths Collaborators (2024) Global, regional, and national stillbirths at 20 weeks' gestation or longer in 204 countries and territories, 1990–2021: findings from the Global Burden of Disease Study 2021. *Lancet* 404(10466):1955–1988
2. Morad G, Helmink BA, Sharma P, Wargo JA (2021) Hallmarks of response, resistance, and toxicity to immune checkpoint blockade. *Cell* 184(21):5309–5337
3. Topalian SL, Forde PM, Emens LA, Yarchoan M, Smith KN, Pardoll DM (2023) Neoadjuvant immune checkpoint blockade: a window of opportunity to advance cancer immunotherapy. *Cancer Cell* 41(9):1551–1566
4. Roh SH, Kasembeli M, Bakthavatsalam D, Chiu W, Tweardy DJ (2015) Contribution of the type II chaperonin, TRiC/CCT, to oncogenesis. *Int J Mol Sci* 16(11):26706–26720
5. Zheng L, Chen X, Zhang L, Qin N, An J, Zhu J et al (2023) A potential tumor marker: chaperonin containing TCP-1 controls the development of malignant tumors. *Int J Oncol* 63(3):1–10
6. Xu J, Zhang Y, Liu C, Yan P, Yang Z (2021) Roles of the miR-139-5p/CCT5 axis in hepatocellular carcinoma: a bioinformatic analysis. *Int J Med Sci* 18(15):3556–3564
7. Guan B, Xu M, Zheng R, Guan G, Xu B (2023) Novel biomarkers to predict treatment response and prognosis in locally advanced rectal cancer undergoing neoadjuvant chemoradiotherapy. *BMC Cancer* 23(1):1099
8. Li Y, Liu C, Zhang X, Huang X, Liang S, Xing F et al (2022) CCT5 induces epithelial-mesenchymal transition to promote gastric cancer lymph node metastasis by activating the Wnt/beta-catenin signalling pathway. *Br J Cancer* 126(12):1684–1694
9. Meng Y, Yang L, Wei X, Luo H, Hu Y, Tao X et al (2021) CCT5 interacts with cyclin D1 promoting lung adenocarcinoma

- cell migration and invasion. *Biochem Biophys Res Commun* 567:222–229
10. Ooe A, Kato K, Noguchi S (2007) Possible involvement of CCT5, RGS3, and YKT6 genes up-regulated in p53-mutated tumors in resistance to docetaxel in human breast cancers. *Breast Cancer Res Treat* 101(3):305–315
 11. Wang Z, Yang L, Huang Z, Li X, Xiao J, Qu Y et al (2022) Identification of prognosis biomarkers for high-grade serous ovarian cancer based on stemness. *Front Genet* 13:861954
 12. Zhang Z, Wang ZX, Chen YX, Wu HX, Yin L, Zhao Q et al (2022) Integrated analysis of single-cell and bulk RNA sequencing data reveals a pan-cancer stemness signature predicting immunotherapy response. *Genome Med* 14(1):45
 13. Han Y, Wang Y, Dong X, Sun D, Liu Z, Yue J et al (2023) TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment. *Nucleic Acids Res* 51(D1):D1425–D1431
 14. Zhang M, Yang H, Wan L, Wang Z, Wang H, Ge C et al (2020) Single-cell transcriptomic architecture and intercellular cross-talk of human intrahepatic cholangiocarcinoma. *J Hepatol* 73(5):1118–1130
 15. Shiao SL, Gouin KH III, Ing N, Ho A, Basho R, Shah A et al (2024) Single-cell and spatial profiling identify three response trajectories to pembrolizumab and radiation therapy in triple negative breast cancer. *Cancer Cell* 42(1):70–84
 16. Lu Y, Yang A, Quan C, Pan Y, Zhang H, Li Y et al (2022) A single-cell atlas of the multicellular ecosystem of primary and metastatic hepatocellular carcinoma. *Nat Commun* 13(1):4594
 17. Li J, Wu C, Hu H, Qin G, Wu X, Bai F et al (2023) Remodeling of the immune and stromal cell compartment by PD-1 blockade in mismatch repair-deficient colorectal cancer. *Cancer Cell* 41(6):1152–1169
 18. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A et al (2021) A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet* 53(9):1334–1347
 19. Ma L, Hernandez MO, Zhao Y, Mehta M, Tran B, Kelly M et al (2019) Tumor cell biodiversity drives microenvironmental reprogramming in liver cancer. *Cancer Cell* 36(4):418–430
 20. Guo W, Zhang C, Wang X, Dou D, Chen D, Li J (2022) Resolving the difference between left-sided and right-sided colorectal cancer by single-cell sequencing. *JCI Insight* 7(1):e152616
 21. Ren Y, Huang Z, Zhou L, Xiao P, Song J, He P et al (2023) Spatial transcriptomics reveals niche-specific enrichment and vulnerabilities of radial glial stem-like cells in malignant gliomas. *Nat Commun* 14(1):1028
 22. Barkley D, Moncada R, Pour M, Liberman DA, Dryg I, Werba G et al (2022) Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. *Nat Genet* 54(8):1192–1201
 23. Andersson A, Larsson L, Stenbeck L, Salmen F, Ehinger A, Wu SZ et al (2021) Spatial deconvolution of HER2-positive breast cancer delineates tumor-associated cell type interactions. *Nat Commun* 12(1):6012
 24. Wu R, Guo W, Qiu X, Wang S, Sui C, Lian Q et al (2021) Comprehensive analysis of spatial architecture in primary liver cancer. *Sci Adv* 7(51):eabg3750
 25. Jerby-Arnon L, Shah P, Cuoco MS, Rodman C, Su MJ, Melms JC et al (2018) A cancer cell program promotes T cell exclusion and resistance to checkpoint blockade. *Cell* 175(4):984–997
 26. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L et al (2015) Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 350(6257):207–211
 27. Auslander N, Zhang G, Lee JS, Frederick DT, Miao B, Moll T et al (2018) Robust prediction of response to immune checkpoint blockade therapy in metastatic melanoma. *Nat Med* 24(10):1545–1549
 28. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S et al (2016) Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* 165(1):35–44
 29. Ascierto ML, Makohon-Moore A, Lipson EJ, Taube JM, McMiller TL, Berger AE et al (2017) Transcriptional mechanisms of resistance to anti-PD-1 therapy. *Clin Cancer Res* 23(12):3168–3180
 30. Nathanson T, Ahuja A, Rubinsteyn A, Aksoy BA, Hellmann MD, Miao D et al (2017) Somatic mutations and neoepitope homology in melanomas treated with CTLA-4 blockade. *Cancer Immunol Res* 5(1):84–91
 31. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS et al (2017) Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell* 171(4):934–949
 32. Liu D, Schilling B, Liu D, Sucker A, Livingstone E, Jerby-Arnon L et al (2019) Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat Med* 25(12):1916–1927
 33. Gide TN, Quek C, Menzies AM, Tasker AT, Shang P, Holst J et al (2019) Distinct immune cell populations define response to anti-PD-1 monotherapy and anti-PD-1/anti-CTLA-4 combined therapy. *Cancer Cell* 35(2):238–255
 34. Snyder A, Nathanson T, Funt SA, Ahuja A, Buros Novik J, Hellmann MD et al (2017) Contribution of systemic and somatic factors to clinical response and resistance to PD-L1 blockade in urothelial cancer: an exploratory multi-omic analysis. *PLoS Med* 14(5):e1002309
 35. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y et al (2018) TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 554(7693):544–548
 36. Rose TL, Weir WH, Mayhew GM, Shibata Y, Eulitt P, Uronis JM et al (2021) Fibroblast growth factor receptor 3 alterations and response to immune checkpoint inhibition in metastatic urothelial cancer: a real world experience. *Br J Cancer* 125(9):1251–1260
 37. Ascierto ML, McMiller TL, Berger AE, Danilova L, Anders RA, Netto GJ et al (2016) The intratumoral balance between metabolic and immunologic gene expression is associated with anti-PD-1 response in patients with renal cell carcinoma. *Cancer Immunol Res* 4(9):726–733
 38. Braun DA, Street K, Burke KP, Cookmeyer DL, Denize T, Pedersen CB et al (2021) Progressive immune dysfunction with advancing disease stage in renal cell carcinoma. *Cancer Cell* 39(5):632–648
 39. Miao D, Margolis CA, Gao W, Voss MH, Li W, Martini DJ et al (2018) Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science* 359(6377):801–806
 40. Zhu AX, Abbas AR, de Galarreta MR, Guan Y, Lu S, Koepfen H et al (2022) Molecular correlates of clinical response and resistance to atezolizumab in combination with bevacizumab in advanced hepatocellular carcinoma. *Nat Med* 28(8):1599–1611
 41. Jung H, Kim HS, Kim JY, Sun JM, Ahn JS, Ahn MJ et al (2019) DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nat Commun* 10(1):4278
 42. Cho JW, Hong MH, Ha SJ, Kim YJ, Cho BC, Lee I et al (2020) Genome-wide identification of differentially methylated promoters and enhancers associated with response to anti-PD-1 therapy in non-small cell lung cancer. *Exp Mol Med* 52(9):1550–1563
 43. Kim ST, Cristescu R, Bass AJ, Kim KM, Odegaard JJ, Kim K et al (2018) Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer. *Nat Med* 24(9):1449–1458
 44. van den Ende T, de Clercq NC, van Berge Henegouwen MI, Gisbertz SS, Geijsen ED, Verhoeven RHA et al (2021) Neoadjuvant chemoradiotherapy combined with atezolizumab for resectable

- esophageal adenocarcinoma: a single-arm phase II feasibility trial (PERFECT). *Clin Cancer Res* 27(12):3351–3359
45. Zhao J, Chen AX, Gartrell RD, Silverman AM, Aparicio L, Chu T et al (2019) Immune and genomic correlates of response to anti-PD-1 immunotherapy in glioblastoma. *Nat Med* 25(3):462–469
 46. Mangiola S, Doyle MA, Papenfuss AT (2021) Interfacing Seurat with the R tidy universe. *Bioinformatics* 37(22):4100–4107
 47. Satija R, Farrell JA, Gennert D, Schier AF, Regev A (2015) Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* 33(5):495–502
 48. Hu C, Li T, Xu Y, Zhang X, Li F, Bai J et al (2023) Cell Marker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. *Nucleic Acids Res* 51(D1):D870–D876
 49. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ et al (2016) Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 352(6282):189–196
 50. Evan GI, Vousden KH (2001) Proliferation, cell cycle and apoptosis in cancer. *Nature* 411(6835):342–348
 51. Gulati GS, Sikandar SS, Wesche DJ, Manjunath A, Bharadwaj A, Berger MJ et al (2020) Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* 367(6476):405–411
 52. Rao A, Barkley D, Franca GS, Yanai I (2021) Exploring tissue architecture using spatial transcriptomics. *Nature* 596(7871):211–220
 53. Ru B, Huang J, Zhang Y, Aldape K, Jiang P (2023) Estimation of cell lineages in tumors from spatial transcriptomics data. *Nat Commun* 14(1):568
 54. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 102(43):15545–15550
 55. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 1(6):417–425
 56. Chen D, Liu J, Zang L, Xiao T, Zhang X, Li Z et al (2022) Integrated machine learning and bioinformatic analyses constructed a novel stemness-related classifier to predict prognosis and immunotherapy responses for hepatocellular carcinoma patients. *Int J Biol Sci* 18(1):360–373
 57. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M et al (2019) Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. *Cancer Immunol Res* 7(5):737–750
 58. Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V et al (2018) Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep* 23(1):181–193
 59. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X et al (2018) Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med* 24(10):1550–1558
 60. Palermo B, Franzese O, Frisullo G, D'Ambrosio L, Panetta M, Campo G et al (2023) CD28/PD1 co-expression: dual impact on CD8⁺ T cells in peripheral blood and tumor tissue, and its significance in NSCLC patients' survival and ICB response. *J Exp Clin Cancer Res* 42(1):287
 61. Skoulidis F, Araujo HA, Do MT, Qian Y, Sun X, Cobo AG et al (2024) CTLA4 blockade abrogates KEAP1/STK11-related resistance to PD-(L)1 inhibitors. *Nature* 635(8038):462–471
 62. Bell HN, Zou W (2024) Beyond the barrier: unraveling the mechanisms of immunotherapy resistance. *Annu Rev Immunol* 42(1):521–550
 63. Krishna C, DiNatale RG, Kuo F, Srivastava RM, Vuong L, Chowell D et al (2021) Single-cell sequencing links multiregional immune landscapes and tissue-resident T cells in ccRCC to tumor topology and therapy efficacy. *Cancer Cell* 39(5):662–677
 64. Westcott PMK, Muiyas F, Hauck H, Smith OC, Sacks NJ, Ely ZA et al (2023) Mismatch repair deficiency is not sufficient to elicit tumor immunogenicity. *Nat Genet* 55(10):1686–1695
 65. Reed ER, Jankowski SA, Spinella AJ, Noonan V, Haddad R, Nomoto K et al (2023) beta-catenin/CBP activation of mTORC1 signaling promotes partial epithelial-mesenchymal states in head and neck cancer. *Transl Res* 260:46–60
 66. Ju M, Bi J, Wei Q, Jiang L, Guan Q, Zhang M et al (2021) Pan-cancer analysis of NLRP3 inflammasome with potential implications in prognosis and immunotherapy in human cancer. *Brief Bioinform* 22(4):bbaa345
 67. Wu Z, Uhl B, Gires O, Reichel CA (2023) A transcriptomic pan-cancer signature for survival prognostication and prediction of immunotherapy response based on endothelial senescence. *J Biomed Sci* 30(1):21
 68. Cheng X, Cao Y, Liu X, Li Y, Li Q, Gao D et al (2024) Single-cell and spatial omics unravel the spatiotemporal biology of tumour border invasion and haematogenous metastasis. *Clin Transl Med* 14(10):e70036
 69. Icard P, Fournel L, Wu Z, Alifano M, Lincet H (2019) Interconnection between metabolism and cell cycle in cancer. *Trends Biochem Sci* 44(6):490–501
 70. Phan TG, Croucher PI (2020) The dormant cancer cell life cycle. *Nat Rev Cancer* 20(7):398–411
 71. Weng J, Li S, Zhu Z, Liu Q, Zhang R, Yang Y et al (2022) Exploring immunotherapy in colorectal cancer. *J Hematol Oncol* 15(1):95
 72. Yang QC, Wang S, Liu YT, Song A, Wu ZZ, Wan SC et al (2023) Targeting PCSK9 reduces cancer cell stemness and enhances anti-tumor immunity in head and neck cancer. *iScience* 26(6):106916
 73. Kucukkose E, Heesters BA, Villaudy J, Verheem A, Cercel M, van Hal S et al (2022) Modeling resistance of colorectal peritoneal metastases to immune checkpoint blockade in humanized mice. *J Immunother Cancer* 10(12):e005345
 74. Chen E, Zou Z, Wang R, Liu J, Peng Z, Gan Z et al (2024) Predictive value of a stemness-based classifier for prognosis and immunotherapy response of hepatocellular carcinoma based on bioinformatics and machine-learning strategies. *Front Immunol* 15:1244392
 75. Cao R, Yang F, Ma SC, Liu L, Zhao Y, Li Y et al (2020) Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in Colorectal Cancer. *Theranostics* 10(24):11080–11091
 76. Wu B, Zhang B, Li B, Wu H, Jiang M (2024) Cold and hot tumors: from molecular mechanisms to targeted therapy. *Signal Transduct Target Ther* 9(1):274
 77. Chowell D, Yoo SK, Valero C, Pastore A, Krishna C, Lee M et al (2022) Improved prediction of immune checkpoint blockade efficacy across multiple cancer types. *Nat Biotechnol* 40(4):499–506
 78. Bareche Y, Kelly D, Abbas-Aghababazadeh F, Nakano M, Esfahani PN, Tkachuk D et al (2022) Leveraging big data of immune checkpoint blockade response identifies novel potential targets. *Ann Oncol* 33(12):1304–1317
 79. Wang XQ, Danenberg E, Huang CS, Egle D, Callari M, Bermejo B et al (2023) Spatial predictors of immunotherapy response in triple-negative breast cancer. *Nature* 621(7980):868–876
 80. Yang C, Geng H, Yang X, Ji S, Liu Z, Feng H et al (2024) Targeting the immune privilege of tumor-initiating cells to enhance cancer immunotherapy. *Cancer Cell* 42(12):2064–2081