

Original Research

An Integrated Machine Learning Framework Identifies Prognostic Gene Pair Biomarkers Associated with Programmed Cell Death Modalities in Clear Cell Renal Cell Carcinoma

Bohong Chen¹, Mingguo Zhou¹, Li Guo¹, Haoxiang Huang¹, Xinyue Sun², Zihe Peng³,
Dapeng Wu^{1,*}, Wei Chen^{1,*}¹Department of Urology, The First Affiliated Hospital of Xi'an Jiaotong University, 710061 Xi'an, Shaanxi, China²Department of Neurology, The First Affiliated Hospital of Xi'an Jiaotong University, 710061 Xi'an, Shaanxi, China³Department of Urology, The Second Affiliated Hospital of Xi'an Jiaotong University, 710004 Xi'an, Shaanxi, China*Correspondence: wudapeng1974@163.com (Dapeng Wu); chenwei_urology@xjtu.edu.cn (Wei Chen)

Academic Editor: Sung Eun Kim

Submitted: 20 November 2023 Revised: 23 February 2024 Accepted: 27 February 2024 Published: 21 March 2024

Abstract

Background: Clear cell renal cell carcinoma (ccRCC) is a common and lethal urological malignancy for which there are no effective personalized therapeutic strategies. Programmed cell death (PCD) patterns have emerged as critical determinants of clinical prognosis and immunotherapy responses. However, the actual clinical relevance of PCD processes in ccRCC is still poorly understood. **Methods:** We screened for PCD-related gene pairs through single-sample gene set enrichment analysis (ssGSEA), consensus cluster analysis, and univariate Cox regression analysis. A novel machine learning framework incorporating 12 algorithms and 113 unique combinations were used to develop the cell death-related gene pair score (CDRGPS). Additionally, a radiomic score (Rad_Score) derived from computed tomography (CT) image features was used to classify the CDRGPS status as high or low. Finally, we conclusively verified the function of PRSS23 in ccRCC. **Results:** The CDRGPS was developed through an integrated machine learning approach that leveraged 113 algorithm combinations. CDRGPS represents an independent prognostic biomarker for overall survival and demonstrated consistent performance between training and external validation cohorts. Moreover, CDRGPS showed better prognostic accuracy compared to seven previously published cell death-related signatures. In addition, patients classified as high-risk by CDRGPS exhibited increased responsiveness to tyrosine kinase inhibitors (TKIs), mammalian Target of Rapamycin (mTOR) inhibitors, and immunotherapy. The Rad_Score demonstrated excellent discrimination for predicting high versus low CDRGPS status, with an area under the curve (AUC) value of 0.813 in the Cancer Imaging Archive (TCIA) database. PRSS23 was identified as a significant factor in the metastasis and immune response of ccRCC, thereby validating experimental *in vitro* results. **Conclusions:** CDRGPS is a robust and non-invasive tool that has the potential to improve clinical outcomes and enable personalized medicine in ccRCC patients.

Keywords: programmed cell death; clear cell renal cell carcinoma; machine learning; single-cell RNA-seq; radiomics; PRSS23

1. Introduction

Renal cell carcinoma (RCC) is among the top ten most commonly diagnosed cancers globally, with clear cell RCC (ccRCC) being the predominant histological subtype and accounting for approximately 70% of RCC cases [1].

The global incidence of ccRCC is increasing and this subtype has the worst prognosis of all RCC variants, with the highest rates of invasion, metastasis, and mortality [2]. In recent surveillance work, 73,820 new cases of ccRCC were diagnosed in the United States in 2019 and 14,770 deaths reported [1]. Tumor, Node, Metastasis (TNM) stage continues to be the most pertinent prognostic indicator for ccRCC in clinical practice. However, the prognostic accuracy of TNM is less than ideal, with substantial variation in survival between individuals with identical TNM stage [3]. ccRCC also exhibits lower responsiveness to radiotherapy and chemotherapy modalities compared to other RCC subtypes. Although immune checkpoint inhibitor (ICI) therapy

has emerged as the standard immunotherapeutic approach for advanced RCC, only a small proportion of patients show meaningful and long-lasting benefits. The expression of programmed death ligand 1 (PD-L1), tumor mutation burden (TMB), neoantigen load (NAL), and mismatch repair deficiency (dMMR, or microsatellite instability-high [MSI-H]) are potential biomarkers of sensitivity to ICI therapy. However, these techniques are limited by issues such as spatial and temporal heterogeneity, poor precision, and the presence of minor cell subpopulations [4–6]. To achieve tailored treatment plans for ccRCC, it is therefore important to develop effective models that can improve the accuracy of prognosis and the selection of patients for immunotherapy.

Programmed cell death (PCD) is intrinsically linked to both tumorigenesis and therapeutic efficacy. Tumors undergo diverse modes of cell death in response to varying environmental conditions [7]. The most recent guidelines put forth by the Nomenclature Committee on Cell



Death (NCCD) define PCD as encompassing a spectrum of processes including apoptosis, necroptosis, ferroptosis, pyroptosis, netotic cell death, entotic cell death, lysosome-dependent cell death, parthanatos, autophagy-dependent cell death, oxeiptosis, and alkaliptosis. Apoptosis is a regulated process that removes injured or surplus cells in an orderly manner without triggering inflammation. It involves condensation, nuclear cleavage, and macrophage-mediated vesicle engulfment [8]. Autophagy-dependent cell death involves stepwise lysosomal degradation and is central to metabolic homeostasis and nutrient recycling [9]. Both apoptosis and autophagy have been implicated in mediating chemoresistance in glioblastoma [10]. Alkaliptosis is a recently described modality of PCD that is governed by intracellular alkalinization [11]. Cuproptosis represents a novel form of cell death that relies on mitochondrial respiration and proceeds through direct binding of copper to fatty acylated constituents of the tricarboxylic acid (TCA) cycle [12]. Recent studies have identified disulfidptosis as a novel PCD mechanism characterized by disulfide stress [13]. Entotic cell death represents a process of active cellular invasion that is restricted to viable cells and pericellular regions. This non-apoptotic pathway proceeds independently of apoptotic executioner activation [14]. Ferroptosis is defined by the iron-dependent accumulation of cytotoxic lipid peroxides to lethal concentrations. Suppression of ferroptosis has been found to confer resistance to PD-1/PD-L1 immune checkpoint blockade [15,16]. Lysosome-dependent cell death proceeds through the activity of lysosomal hydrolases released into the cytosol subsequent to organelle membrane permeabilization [17]. The pivotal characteristic of necroptosis centers around the assembly of necrosomes and involves multifaceted progression [18]. Netotic cell death arises from the extrusion of neutrophil extracellular traps (NETs), which are reticulated chromatin structures deployed by granulocytes in response to infection or tissue damage [19]. Oxeiptosis harnesses the reactive oxygen species-sensing capability of KEAP1 to execute a distinct cell death program. It represents a novel cellular pathway that is likely to intersect with other modes of regulated cell death [20]. Parthanatos represents a specialized form of regulated cell death mediated by the overactivation of PARP-1 nuclease [21]. Pyroptosis is characterized by cell swelling, lysis, and the copious release of proinflammatory mediators [22]. It has been implicated in the translocation of programmed cell death ligand 1 (PD-L1) and the therapeutic response to PD-1 checkpoint blockade [23]. In summary, it has been established that PCD is a critical determinant of tumorigenesis and metastasis [18].

The generation of PCD-associated genetic signatures to categorize patients into discrete risk strata has been proposed to improve the accuracy of prognostic modeling [24]. However, prior efforts at generating quantitative models have been hampered by biological noise and technical biases. An alternative approach based on relative gene ex-

pression to generate gene pair (GP) signatures may be worthwhile since it eliminates the requirement for data standardization, thus allowing the development of reliable models [25]. However, this approach is constrained by necessary intrusive procedures and excessive cost to the patient. A potentially non-invasive alternative is the application of radiomics, which makes use of the quantitative characteristics of medical imaging [26,27].

The aim of this study was to identify PCD-associated gene pairs which can predict the prognosis and outcome of immunotherapy in ccRCC patients through a novel computational framework. Additionally, we investigated associations between radiomic features extracted from contrast-enhanced computed tomography (CT) and the cell death-related gene pair score (CDRGPS). Transcriptomic and clinical data for ccRCC patients were obtained from The Cancer Genome Atlas (TCGA-KIRC) database and ArrayExpress (E-MTAB-1980). These were used as discovery and validation cohorts, respectively. Cell death-related gene pairs were identified through single sample gene set enrichment analysis (ssGSEA), consensus clustering, and univariate Cox regression. A machine learning framework incorporating 12 algorithms and 113 combinations was used to develop the CDRGPS. The prognostic utility of CDRGPS was evaluated and incorporated into a nomogram. CDRGPS was applied in single-cell analysis to assess biological differences between high- and low-risk subgroups. We also examined the prediction of drug response and efficacy of immunotherapy. In addition, a radiomic score (Rad_Score) was derived from the CT imaging-quantified CDRGPS status. Finally, the role of PRSS23 in ccRCC was investigated (Fig. 1).

2. Materials and Methods

2.1 Data Collection

RNA sequencing (RNA-seq) data and corresponding clinical data for ccRCC were obtained from The Cancer Genome Atlas (TCGA) (<https://cancergenome.nih.gov>) cohort. After excluding samples with incomplete or absent essential clinicopathological annotations, the final TCGA-Kidney Renal Clear Cell Carcinoma (KIRC) cohort comprised 518 patients. An additional validation cohort of 101 ccRCC patients with gene expression microarray data and clinical details was sourced from the E-MTAB-1980 dataset in the ArrayExpress (<https://www.ebi.ac.uk/biostudies/arrayexpress>) repository [28].

2.2 PCD-Related Gene Sources

PCD-associated genes were selected from Kyoto Encyclopedia of Genes and Genomes (KEGG) databases, Gene Set Enrichment Analysis (GSEA) gene sets, and published reviews [29,30]. Specifically, a total of 580 apoptosis genes, 52 pyroptosis genes, 87 ferroptosis genes, 367 autophagy genes, 15 entotic cell death genes, 101 autophagy genes, 14 cuproptosis genes, 9 parthanatos genes, 8 netotic cell death genes, 7 alkaliptosis genes, 220 lysosome-

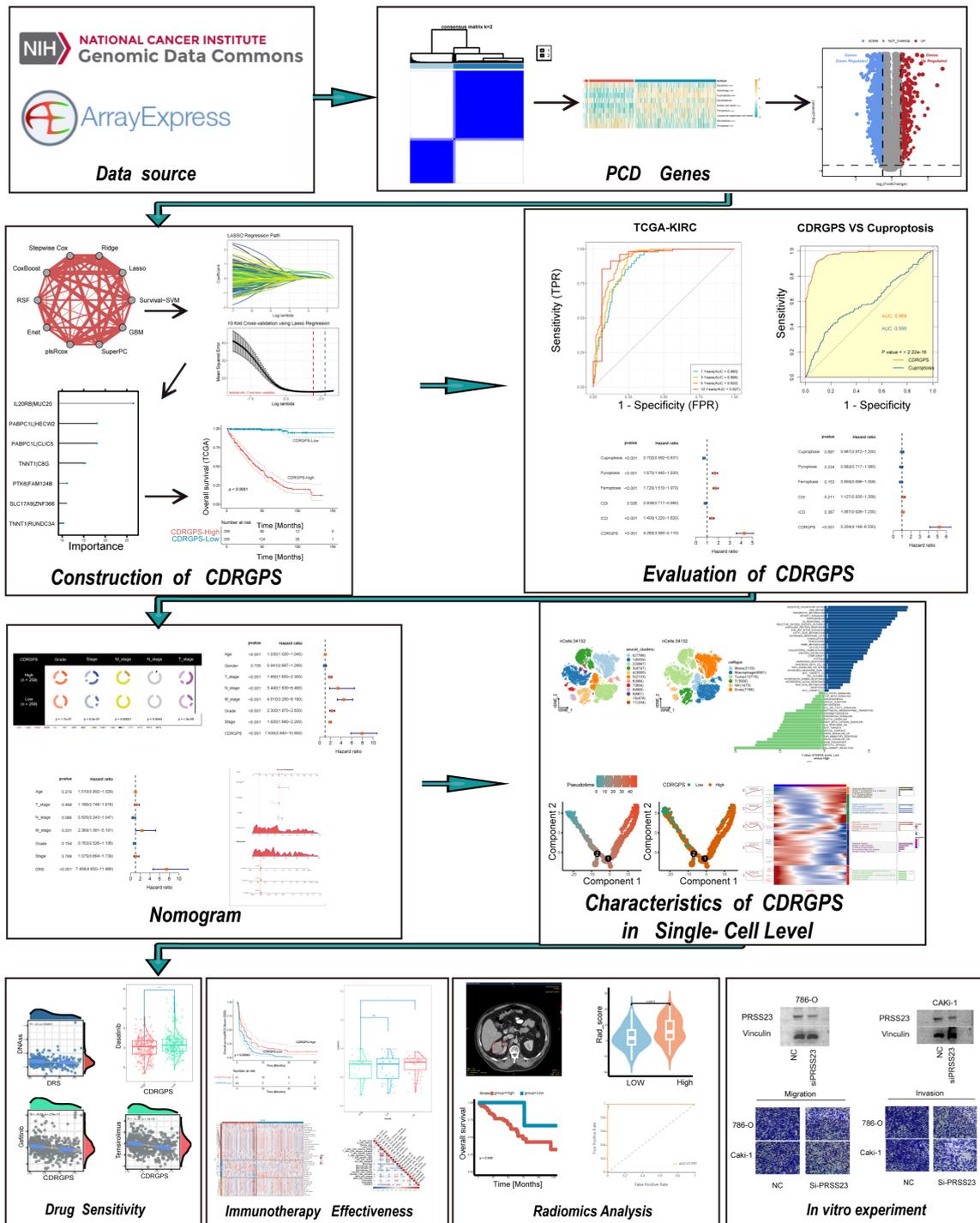


Fig. 1. Flowchart for this study. PCD, programmed cell death; CDRGPS, cell death-related gene pair score; TCGA-KIRC, The Cancer Genome Atlas - Kidney Renal Clear Cell Carcinoma.

dependent cell death genes, 15 disulfidptosis genes, and 5 oxceptosis genes were compiled (**Supplementary Table 1**).

2.3 Consensus Clustering Analysis

To quantify PCD pathway activity in each sample, ssGSEA was conducted using the GSVA R package 1.44.5 with gene expression profiles. The ssGSEA scores were then utilized to perform unsupervised consensus clustering

with the ConsensusClusterPlus R package 1.60.0 (1000 iterations, 80% resampling rate, Spearman correlation) to categorize ccRCC into distinct PCD subclusters.

2.4 Development of Cell Death Related Gene Pair Signature

Differentially expressed genes (DEGs) between the two PCD subclusters were identified using the DESeq2 R

package 1.42.0. Univariate Cox regression analysis was then performed to identify DEGs that were significantly correlated with overall survival (OS) in both the training and test cohorts ($p < 0.05$). The prognostic DEGs identified were utilized to generate gene pairs following the approach described previously by Hong *et al.* [31]. Briefly, genes were paired as A|B, with the pair designated as 1 if the expression of gene A exceeded that of gene B. The pair was otherwise designated as 0. Gene pairs where 0 or 1 exceeded a prevalence of 20% were considered effective. To create a CDRGPS with high precision and generalizability, 12 machine learning algorithms were integrated: Lasso, Ridge, StepAIC, Extreme Gradient Boosting (XGBoost), random forest (RF), elastic net (Enet), partial least squares regression for generalized linear models (plsRglm), generalized boosted regression modeling (GBM), Naive-Bayes, linear discriminant analysis (LDA), generalized linear model boosting (glmBoost), and support vector machine (SVM). A total of 113 permutations of these 12 algorithms were then evaluated in a 10-fold cross-validation framework using the TCGA-KIRC training dataset for variable selection and model building. We subsequently validated model performance using E-MATB-1980 as an external test cohort. The concordance index for each model was calculated across the training set and the external testing set. The prediction accuracy of each model was then ranked using the mean C-index. The set of algorithms that demonstrated both reliable performance and clinical relevance was chosen, leading to the development of a signature known as CDRGPS that was capable of predicting the OS of ccRCC patients.

2.5 Prognostic Value of CDRGPS and Potential Clinical Translation

Based on the model generated above, a score was determined for each sample in the training and test datasets. Using the median value as the cutoff point, patients were then divided into CDRGPS-high and -low groups. The “survminer” R package 0.4.9 and Kaplan-Meier (KM) analysis were used to compare the two groups in terms of OS, progression-free survival (PFS), disease-free survival (DFS), and disease-specific survival (DSS). Furthermore, the “timeROC” package 0.4 was employed to perform receiver operating characteristic (ROC) analysis for evaluating the sensitivity and specificity of CDRGPS in predicting the OS of ccRCC patients in the training and test cohorts. Additionally, the area under the curve (AUC) of CDRGPS was compared to that of other PCD-related signatures. Univariate and multivariate Cox regression analyses were also performed to determine whether CDRGPS was an independent prognostic factor for ccRCC patients. To improve the prognostic accuracy of CDRGPS, a nomogram was constructed using the “rms” package. This nomogram integrated CDRGPS and clinical characteristics, allowing quantification of the survival outcomes of ccRCC patients.

The timeROC curve, calibration curves, and decision curve analysis (DCA) were used to thoroughly assess the performance of the nomogram.

2.6 Single-Cell RNA-Seq Data Collection and Processing

Single-cell RNA sequencing (scRNA-seq) data was obtained from 7 ccRCC samples in the GSE156632 collection [32]. Cell clustering and dimension reduction were performed using the Seurat package [33]. Principal component analysis (PCA) was performed using “RunPCA”. A K-nearest neighbor analysis was then conducted using the “FindNeighbors” function. Complex expression profiles were visually represented by downscaling and facilitated by the “RunTSNE” function. This was followed by cell annotation based on marker genes associated with various cell types. Finally, pseudotime analysis was conducted using the monocle R package 2.26.0.

2.7 Tumor Stemness and Drug Sensitivity Analysis

Based on prior research, we derived six tumor stemness indices using messenger ribonucleic acid (mRNA) expression and methylation signatures: RNA expression-based stemness score (RNAss), Epigenetically regulated RNA expression-based stemness score (EREG.EXPss), DNA methylation-based stemness score (DNAss), Epigenetically regulated DNA methylation-based stemness score (EREG-METHss), Differentially methylated probes-based stemness score (DMPss), and Enhancer Elements/DNA methylation-based stemness score (ENHss) [34]. With regard to personalized treatment, the pRRophetic R package 0.5 [35] was used to predict the half-maximal inhibitory concentration (IC50) of chemotherapy drugs that are commonly administered to ccRCC patients. This was based on their distinct CDRGPS level.

2.8 Integrated Evaluation of CDRGPS-Based Immunotherapy Efficacy and Immune-Omics Molecular Characterisation

To quantify immune infiltrating cells and the ESTIMATE score, we utilized CIBERSORT [36], MCP-counter (Microenvironment Cell Populations-counter) [37], the QUANTISEQ algorithm [38], and the ESTIMATE (Estimation of Stromal and Immune cells in Malignant Tumors using Expression data) algorithm [39]. Additionally, we used data from ssGSEA to compute the cancer immunity cycle, which reflects the functionality of chemokines and immunomodulators [40]. Furthermore, to explore the predictive ability of CDRGPS for immunotherapy, we used data from the RCC-Braun_2020 cohort consisting of 181 ccRCC patients treated with nivolumab, of which 57 showed clinical benefit, 57 showed intermediate clinical benefit, and 67 showed no clinical benefit [41].

2.9 Development of the Radiomics Model

Preoperative CT scans for 267 samples and stored in The Cancer Imaging Archive (TCIA) were used to extract radiomics features for use in model building [42]. The volume of interest (VOI) for tumor segmentation was performed manually using the 3D Slicer program (<http://www.itksnap.org/pmwiki/pmwiki.php>). The entire tumor was manually segmented for each axial slice by two radiologists, both with >5 years of experience. The Pyradiomics 2.2.0 Python package (version 3.0, Python Software Foundation, Wilmington, DE, USA) was then used to extract 1688 radiomics characteristics from each VOI. Spearman's rank correlation analysis was used to determine which features should be chosen. Features with correlation coefficients >0.9 were classified as redundant and eliminated from further consideration. LASSO regression and 5-fold cross-validation were used to modify the penalty parameter in order to identify features with the highest predictive value for CDRGPS. This allowed creation of the radiomics score (Rad_Score), which involved linear combination of the selected features and their weighting according to the respective coefficients.

2.10 Cell Culture

Human renal cell carcinoma cell lines 786-O and CAKi-1 obtained from the American Type Culture Collection were cultured under standard conditions in a humidified incubator at 37 °C and 5% CO₂. The cell line identity was validated by STR profiling. And the cells were tested before and after the experiments for mycoplasma contamination detected by MycoStrip™ kit (InvivoGen, HongKong, China) based on isothermal polymerase chain reaction (PCR). Silencing of the target gene, *PRSS23*, was achieved using small interfering RNA (siRNA) purchased from Fenghuishengwu Associates. The *PRSS23* siRNA sequence was 5' GCGGCAGAUUTTATGGCTAUGA 3'. The 786-O and CAKi-1 cell lines were seeded at a density of 4 × 10⁵ cells per well in 6-well plates and allowed to grow until complete adherence was achieved. Both cell lines then underwent transfection using the lipo3000 compound.

2.11 Western Blotting

The 786-O and CAKi-1 cells were rinsed twice using chilled Phosphate-Buffered Saline (PBS) and lysed on ice by the addition of 250 μL lysis buffer. They were then centrifuged for 15 minutes at 14,000 rpm and 4 °C, and the debris removed. The combined protein was heated at 100 °C for 10 minutes. Equal amounts of protein samples were resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using a 10% polyacrylamide gel under a constant voltage of 120 V. The separated proteins were subsequently transferred to polyvinylidene difluoride (PVDF) membranes. These were blocked with 5% non-fat dry milk in Tris-buffered saline containing 0.1% Tween-20 (TBST) for 60 minutes at room temperature.

This was followed by overnight incubation at 4 °C with primary antibodies against PRSS23 (Abcam, Cambridge, UK, ab201182) and vinculin (Abcam, ab129002). After washing with TBST, the membranes were incubated with appropriate horseradish peroxidase-conjugated secondary antibodies for 60 minutes at room temperature. The immunoreactive bands were visualized using enhanced chemiluminescence reagents after further washing with TBST. Analysis of band densitometry was performed using Image Lab software and ImageJ to determine relative protein expression levels. These were normalized to the vinculin loading control.

2.12 Transwell Migration and Invasion Assays

Cell migration and invasion assays were performed using 24-well Transwell chambers containing 8 μm pore size polycarbonate membrane inserts (Thermo Fisher, Waltham, MA, USA). For invasion assays, the upper surface of the Transwell inserts was precoated with 100 μL of Matrigel basement membrane matrix (BD Biosciences, Franklin Lakes, NJ, USA) and incubated overnight at 4 °C to allow gelling. Cells were harvested and resuspended at a density of 3 × 10⁴ cells/mL in serum-free medium. A 100 μL cell suspension was added to the upper chamber of each insert, while the lower chambers were filled with 600 μL of culture medium containing 10% fetal bovine serum as a chemoattractant. Following 24-hour incubation at 37 °C to allow for cell migration or invasion through the membrane, the non-migrated cells on the upper surface were removed using a cotton swab. Migrated or invaded cells on the underside of the membrane were fixed with 4% paraformaldehyde solution for 20 minutes at room temperature and subsequently stained with 0.5% crystal violet dye for 5 minutes. The number of stained cells in five random fields per insert was counted under a light microscope to quantify the extent of migration and invasion.

2.13 Statistical Analysis

The statistical analysis software R 4.2.1 (The R Foundation for Statistical Computing, Vienna, Austria) was used for all analyses. The Wilcoxon test was used for non-parametric comparisons between two variables with non-normal distributions. Kaplan-Meier survival analysis and the log-rank test were used to compare OS, DFS, PFS, and DSS between subgroups. This was carried out using the survival R package 3.5-5. Univariate and multivariate Cox regression analyses were conducted to identify independent prognostic variables. Model performance was evaluated using ROC analysis, and the AUC was calculated using the R package timeROC. Differences in clinical traits between two CDRGPS subclusters were examined using the chi-squared test. $p < 0.05$ was used to indicate statistical significance, unless otherwise noted. Associations between CDRGPS, immune cell infiltration, and tumor stemness index were examined using Spearman's correlation analysis.

3. Results

3.1 ccRCC Classification Based on 9 Programmed Cell Death Patterns

The overall workflow for this study is outlined in Fig. 1. We utilized the RNA-seq data of 13 PCD-related genes from 518 ccRCC patients in the TCGA to perform ssGSEA and subsequently univariate Cox regression analysis. Nine of the 13 PCD patterns exhibited significant associations with the clinical outcome of ccRCC ($p < 0.1$; **Supplementary Table 2**). These included apoptosis, autophagy, cuproptosis, disulfidptosis, ferroptosis, pyroptosis, entotic cell death, lysosome-dependent cell death, and necroptosis. We next applied an unsupervised consensus clustering method based on the 9 prognosis-related PCD scores to categorize ccRCC samples into potential subgroups, ranging from $k = 2$ to $k = 5$. The cumulative distribution function (CDF) curves derived from the consensus clustering matrix revealed that $k = 2$ was the ideal number of subgroups (Fig. 2A–D). The two consensus clusters, C1 and C2, displayed distinct PCD patterns. C1 exhibited elevated levels of apoptosis, necroptosis and pyroptosis, while C2 was enriched for autophagy, cuproptosis, entotic cell death and ferroptosis (Fig. 2E). Patients belonging to C1 displayed reduced DSS, DFS, OS, and PFS compared to the C2 group ($p < 0.001$, log-rank test; Fig. 2F).

3.2 Construction of a Cell Death Related Gene Pairs Score Based on Integrative Machine Learning

To further examine the transcriptional heterogeneity of PCD subpopulations in ccRCC, differential analysis was conducted to contrast the two subclusters. A total of 920 genes were found to display substantial variation in expression level across the delineated groups (**Supplementary Fig. 1A**). Of these, 303 putative genes linked to PCD processes were identified as potential predictors of OS, as revealed by univariate Cox regression analysis ($p < 0.05$). These 303 prognosis-associated genes were then used to establish gene pairs, leading to the identification of 159 such gene pairs associated with prognostic outcomes. The ensemble of gene pairs, encompassing a subset of 51 genes, was derived through meticulous curation ($p < 0.05$; **Supplementary Fig. 1B**).

A compendium of 12 distinct machine learning algorithms were synergistically harnessed, comprising Lasso, Ridge, StepAIC, XGBoost, RF, Enet, partial least squares regression for generalized linear models(plsRglm), GBM, NaiveBayes, LDA, glmBoost and SVM. By employing a rigorous 10-fold cross-validation methodology, these algorithms were strategically amalgamated to identify the most resilient CDRGPS, as indicated by an elevated C-index performance metric. This iterative process was meticulously performed with both the training dataset and an independent external test dataset, as shown in Fig. 3A. Of the 113 models examined, the five prediction models with the highest average C-index were obtained exclusively using the RF

algorithm. These five predictive models showed good efficacy not only within the training dataset, but also with external validation datasets, exhibiting a C-index greater than the 0.75 threshold. Following an exhaustive screening process, the combined Lasso + RF configuration emerged as a particularly discerning predictive model, characterized by excellent accuracy with minimal variables. The final iteration of the CDRGPS was finally derived through synergistic integration of the Lasso and RF algorithms. Of note, the Lasso algorithm discerned a selection of seven pre-eminent cell death-related gene pairs (Fig. 3B), while the RF algorithm identified the most robust predictive model (Fig. 3C). A tailored risk score was calculated for each sample in the training and test cohorts, thus allowing patients to be categorized into CDRGPS-high or CDRGPS-low groups according to the median value. In both the TCGA-KIRC and E-MTAB-1980 cohorts, patients in the CDRGPS-high group had significantly worse OS than those in the CDRGPS-low group ($p < 0.001$, log-rank test; Fig. 3D,H). Similarly, the CDRGPS-high group had significantly worse DFS, PFS, and DSS ($p < 0.001$, log-rank test; Fig. 3E–G).

3.3 Evaluation of the CDRGPS Model

The discriminative performance of the CDRGPS was assessed by ROC analysis. The 1-, 3-, 5- and 10-year AUCs were 0.880, 0.896, 0.920 and 0.927, respectively, in the TCGA-KIRC, and 0.776, 0.763, 0.797 and 0.701 in the E-MTAB-1980 dataset (Fig. 4A,B). These findings highlight the powerful discriminative ability of the CDRGPS. Recent developments in big data analytics and sequencing technology have enabled considerable progress in the development of prognostic and predictive gene expression signatures for various diseases through the application of machine learning techniques. We next evaluated the prognostic accuracy of the CDRGPS compared to five alternative cell death-related signatures. These alternative signatures are based on cuproptosis, ferroptosis, immunogenic cell death, pyroptosis, and a composite of 12 combined PCD models, with a particular focus on ccRCC patient outcomes [7,30,43–47]. Significantly higher AUCs in both the training and external test cohorts demonstrated better performance by the CDRGPS in predicting patient survival compared to the other five risk scores (Fig. 4C–I, **Supplementary Fig. 2A–G**). Moreover, we performed univariate and multivariate Cox regression analyses on both the training and test cohorts for each of the relevant signatures. In both cohorts, the CDRGPS model was the only independent predictor of ccRCC patient outcome (Fig. 4J,K; **Supplementary Fig. 2H,I**). This finding underscores the robustness and reliability of the CDRGPS.

3.4 Construction of the Nomogram

Given that clinical characteristics are routinely used to assess the prognosis of ccRCC patients, we performed an in-depth investigation into the associations between

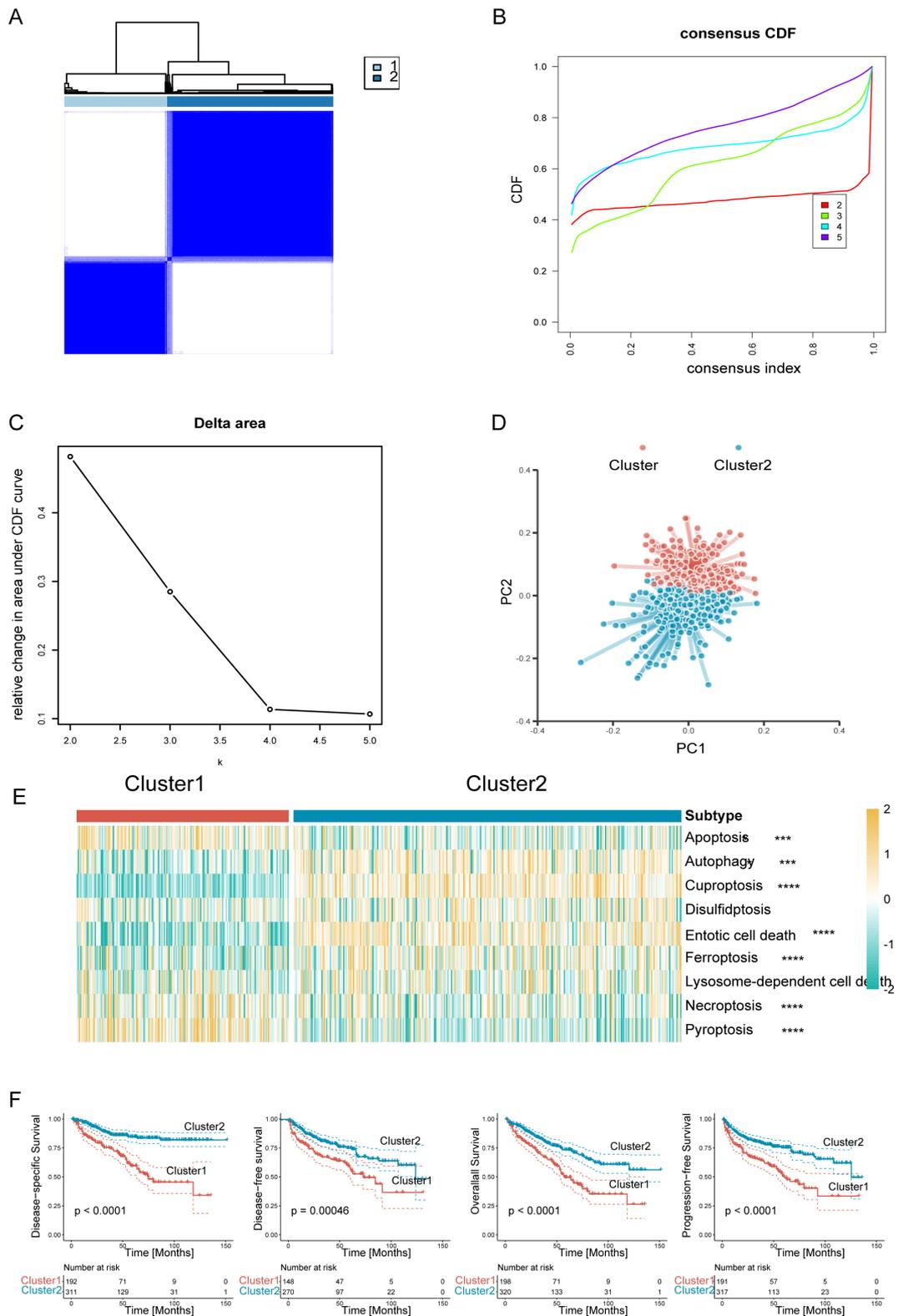


Fig. 2. Clear cell renal cell carcinoma (ccRCC) was classified into two programmed cell death (PCD) subclusters. (A) Consensus clustering matrix. (B) Consensus clustering cumulative distribution function (CDF) curves. (C) Delta area under CDF curves. (D) The principal component plot for the two clusters of ccRCC. (E) Heatmaps of two ccRCC subclusters associated with PCD based on single-sample gene set enrichment analysis (ssGSEA) scores. (F) Survival analysis of the two subtypes in the TCGA-KIRC dataset. Disease-specific survival (DSS), disease-free survival (DFS), overall survival (OS), and progression-free survival (PFS) are shown. **** $p < 0.0001$, *** $p < 0.001$.

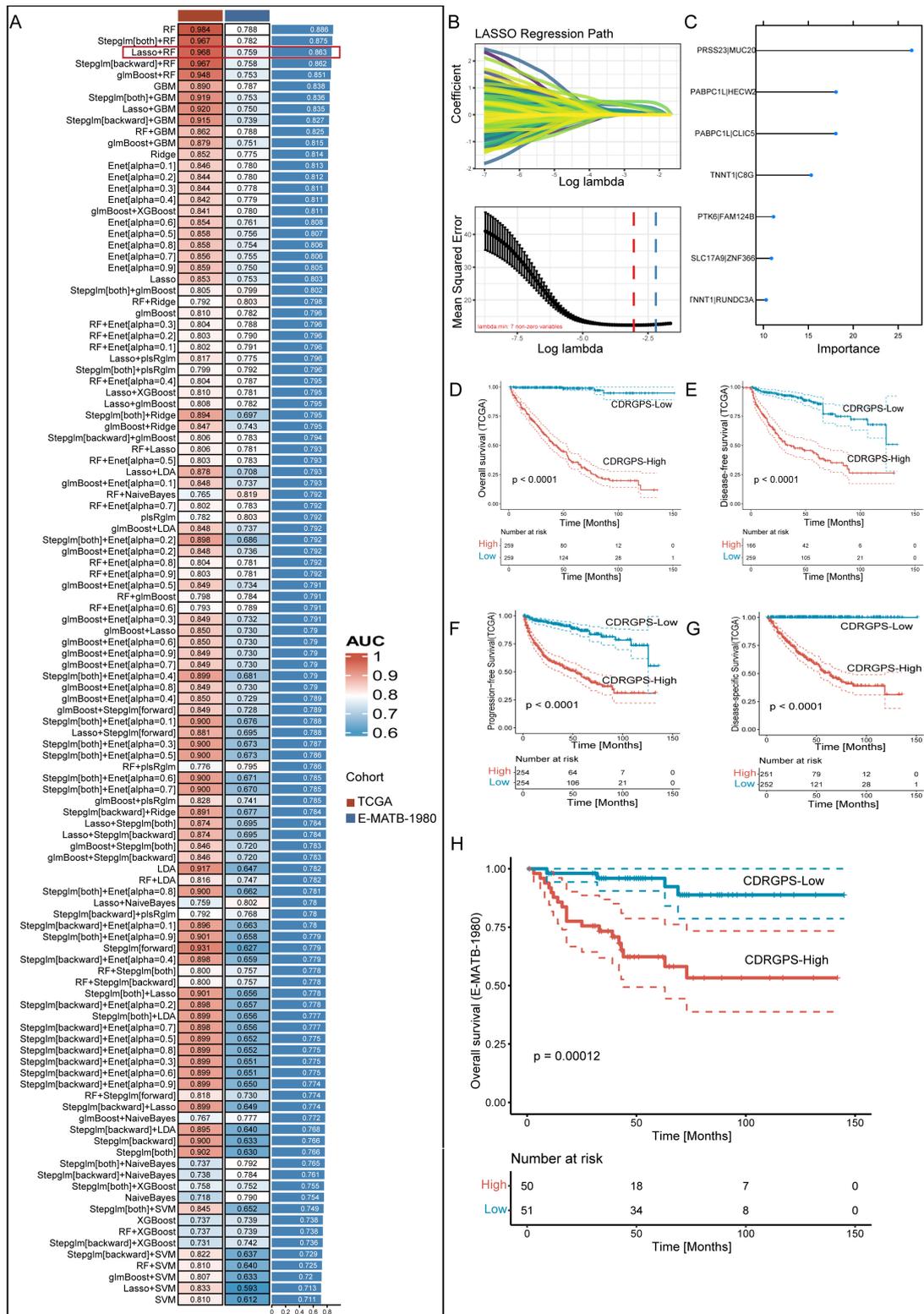


Fig. 3. A consensus CDRGPS was created and confirmed using a machine learning-based integrative process. (A) A total of 113 different predictive models were generated using a 10-fold cross-validation framework. The C-index for each model was then calculated across all datasets. **(B)** Visualization of Least Absolute Shrinkage and Selection Operator (LASSO) regression in the training cohort. The optimal λ was obtained when the partial likelihood deviance reached the minimum value. **(C)** Random forest analysis. **(D–G)** Survival analysis for OS, DFS, PFS and DSS in CDRGPS-high and -low ccRCC patients from the TCGA-KIRC cohort. **(H)** Survival analysis (OS) of CDRGPS-high and -low ccRCC patients in the E-MATB-1980 cohort.

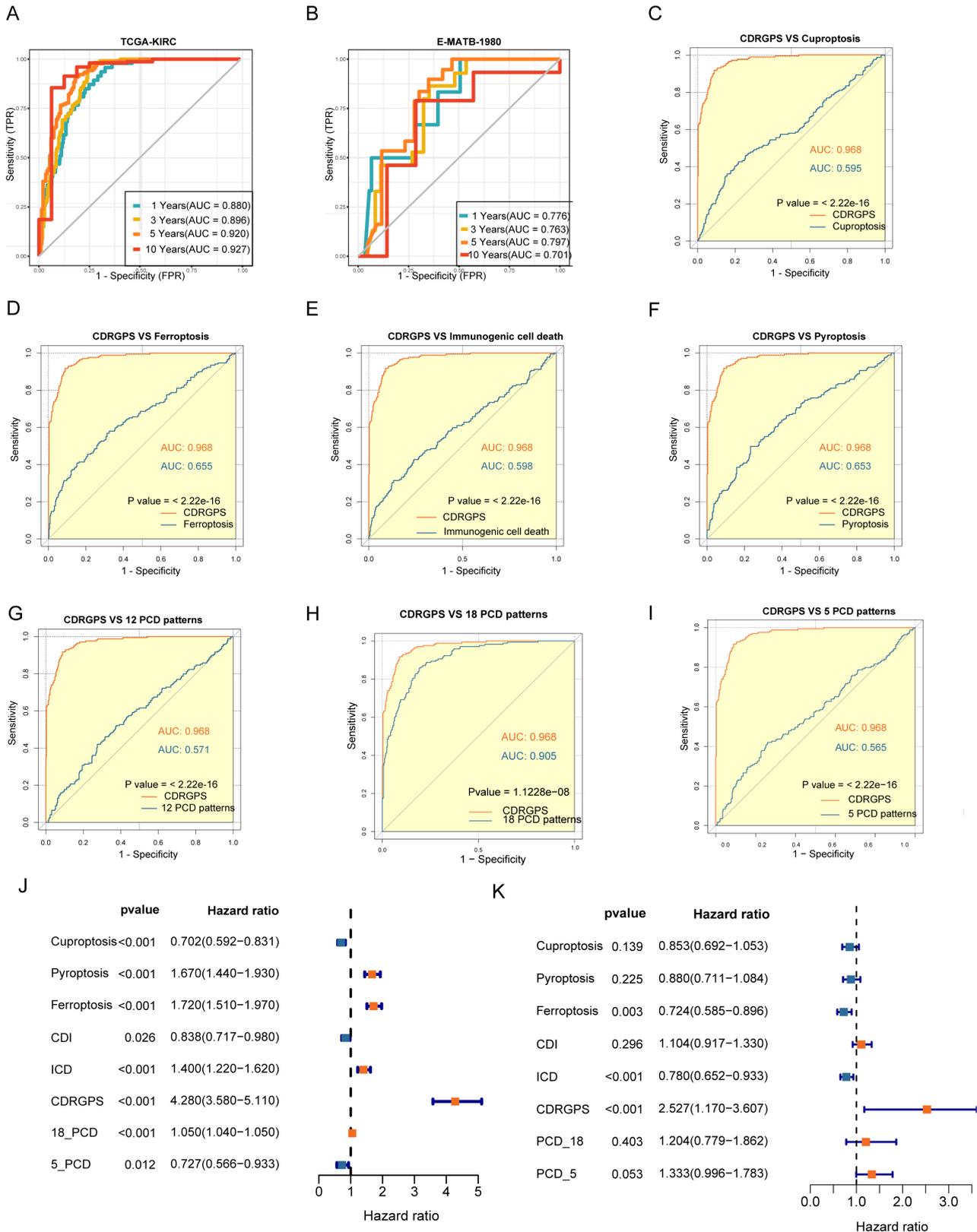


Fig. 4. Evaluation of the CDRGPS model. (A,B) Time-dependent receiver operating characteristic (ROC) analysis for the prediction of OS at 1-, 3-, and 10-years in the training and test cohorts. (C–I) Area under curve (AUC) analysis of the CDRGPS and other PCD-related models in the training cohort. (J,K) Results of univariate and multivariate analyses of OS using the CDRGPS and other PCD-related models.

CDRGPS and various clinical parameters. Significant differences in grade, stage and TNM status were found between the CDRGPS-high and -low groups in the training cohort ($p < 0.001$, chi-squared test, Fig. 5A). The CDRGPS model continued to exhibit statistical significance for OS following adjustment for potential confounding variables such as age, gender, TNM stage, American Joint Committee on Cancer (AJCC) stage, grade and M stage (Fig. 5B,C). To enhance the practical utility of CDRGPS, a nomogram was constructed integrating independent prognostic factors, including CDRGPS and M stage, along with key clinical variables such as T stage and N stage, developed through multivariate Cox regression analysis (Fig. 5D). The AUC values for the nomogram at 1-, 3-, 5- and 10-year intervals were 0.892, 0.923, and 0.948, respectively (Fig. 5E). Furthermore, the calibration curves showed high concordance between the predictions generated by the nomogram and the actual observations (Fig. 5F). Additionally, DCA unambiguously demonstrated that the nomogram offered a greater net therapeutic benefit compared to the other signatures (Fig. 5G).

3.5 Renal Cancer Cells with CDRGPS-High Display Biological Traits Related to the Immune System and Malignancy

Additional analyses were conducted using single-cell transcriptomic profiles from 7 ccRCC samples to determine if CDRGPS could distinguish discrete biological features at the single-cell level in ccRCC. After eliminating low-quality cells and carrying out normalization, integration and PCA, 11 clusters comprising 34,132 cells were identified (Fig. 6A). Six distinct cell types based on the marker genes were identified by subcluster annotation (**Supplementary Table 3**): B cells, endothelial cells, macrophages, monocytes, tumor cells, and T cells (Fig. 6B,C). Analysis of single-cell sequencing data using GSEA and GSVA revealed that tumor cells with high CDRGPS demonstrated a more robust regulation of tumor-related immunity, including pathways such as IMMUNE_RESPONSE, IL2_STAT5_SIGNALING, INTERFERON_GAMMA_RESPONSE, INTERFERON_ALPHA_RESPONSE, and the humoral immune response. Additionally, these cells displayed pronounced malignant biological characteristics, such as TGF_BETA_SIGNALING, PI3K_AKT_MTOR_SIGNALING, apoptosis, epithelial to mesenchymal transition, the tumor necrosis factor-mediated signaling pathway, and a greater response to hypoxia (Fig. 6D–H). We next used pseudo-time analysis to study correlations between CDRGPS and the developmental trajectory of malignant cells, given the diversity of cell developmental stages within tumor tissue. This investigation yielded a cell trace plot that effectively illustrates the fluctuations in CDRGPS across pseudo-time (Fig. 6I). Cells exhibiting an elevated CDRGPS were

mostly clustered close to the terminus node of the branching tree, whereas those with lower CDRGPS values were primarily found in the root of the branch tree. Differential gene expression analysis was performed relative to pseudo-time progression. Heatmaps were generated to visualize changes in gene expression patterns aligned with increasing pseudo-time, and the corresponding increase in CDRGPS. Interestingly, the up-regulated genes mostly had immune response-related functions (Fig. 6J).

3.6 The CDRGPS Shows Good Predictive Accuracy for the Effects of Immunotherapy

A rigorous, systematic investigation was performed to evaluate the effect of CDRGPS on immunotherapy in ccRCC. We initially studied the RCC-Braun_2020 cohort due to the complete prognostic and treatment-related data available for this patient cohort. The CDRGPS-high group exhibited more favorable prognostic outcomes, suggesting greater benefit from immunotherapy (Fig. 7A). In addition, the group showing clinical benefit (CB) had a significantly higher CDRGPS than the patient group showing no clinical benefit (NCB) ($p < 0.05$; Fig. 7B). We next analyzed the TME in order to investigate the underlying mechanisms responsible for the different responses to immunotherapy between the two groups. The ESTIMATE algorithm was used to calculate immune scores, stromal scores, ESTIMATE scores, and tumor purity scores for the CDRGPS subgroups. The CDRGPS-high group showed markedly elevated immune and ESTIMATE scores, together with a diminished tumor purity score (Fig. 7C). We next counted the number of immune cells in each sample in order to quantify differences in immune cell infiltration between the CDRGPS-high and -low groups. We employed Quantiseq, Timer and Mep_counter methodologies with RNA-sequencing data to assess the level of immune cell infiltration in ccRCC patients. This revealed that plasma cells, T cell CD8, B cells and cytotoxic lymphocytes were more prevalent in the CDRGPS-high group (Fig. 7D). Moreover, a substantial portion of cells exhibited a positive correlation between the CDRGPS and the level of immune infiltrate (Fig. 7E). We also investigated possible cellular mechanisms associated with CDRGPS by analyzing the cancer immunity cycle. In the TCGA-KIRC datasets, the CDRGPS-high subgroup showed increased activity in six of the seven phases of the cancer immunity cycle. These steps encompassed antigen release (Step 1), cancer antigen presentation (Step 2), priming and activation (Step 3), recruitment of tumor-infiltrating immune cells (Step 4), recognition of cancer cells by T cells (Step 6), and the killing of cancer cells (Step 7) (Fig. 7F). Prior investigations reported that increased expression of immune checkpoints is associated with a better reaction to ICI [48,49]. We therefore evaluated the level of immune checkpoint expression in different CDRGPS subgroups. As shown in Fig. 7G, the CDRGPS-high subgroup showed increased expression of all nine immune checkpoints. To

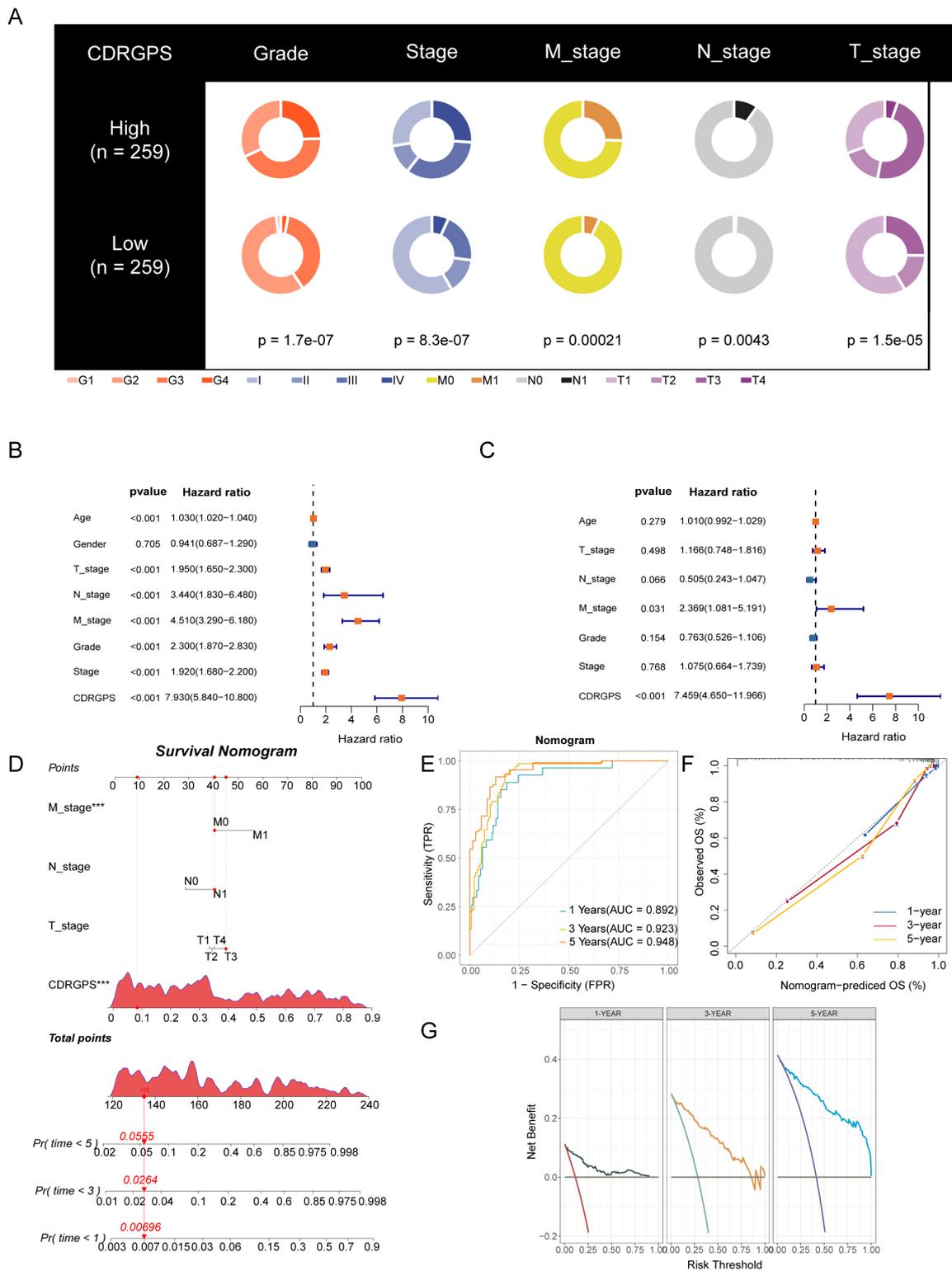


Fig. 5. Development and validation of the nomogram. (A) Clinical traits and their correlation with the CDRGPS-low and CDRGPS-high groups. (B,C) Results of univariate and multivariate analysis for OS in relation to clinical traits and CDRGPS in the TCGA-KIRC cohort. (D) The nomogram was constructed using the CDRGPS, T_stage, N_stage, and M stage. (E) ROC curves demonstrating the ability of the nomogram to predict outcomes at 1-, 3- and 5-years. (F) Nomogram calibration curves for 1-, 3- and 5-year OS. (G) Decision curve analysis (DCA) demonstrating the net benefit using the nomogram. *** $p < 0.001$.

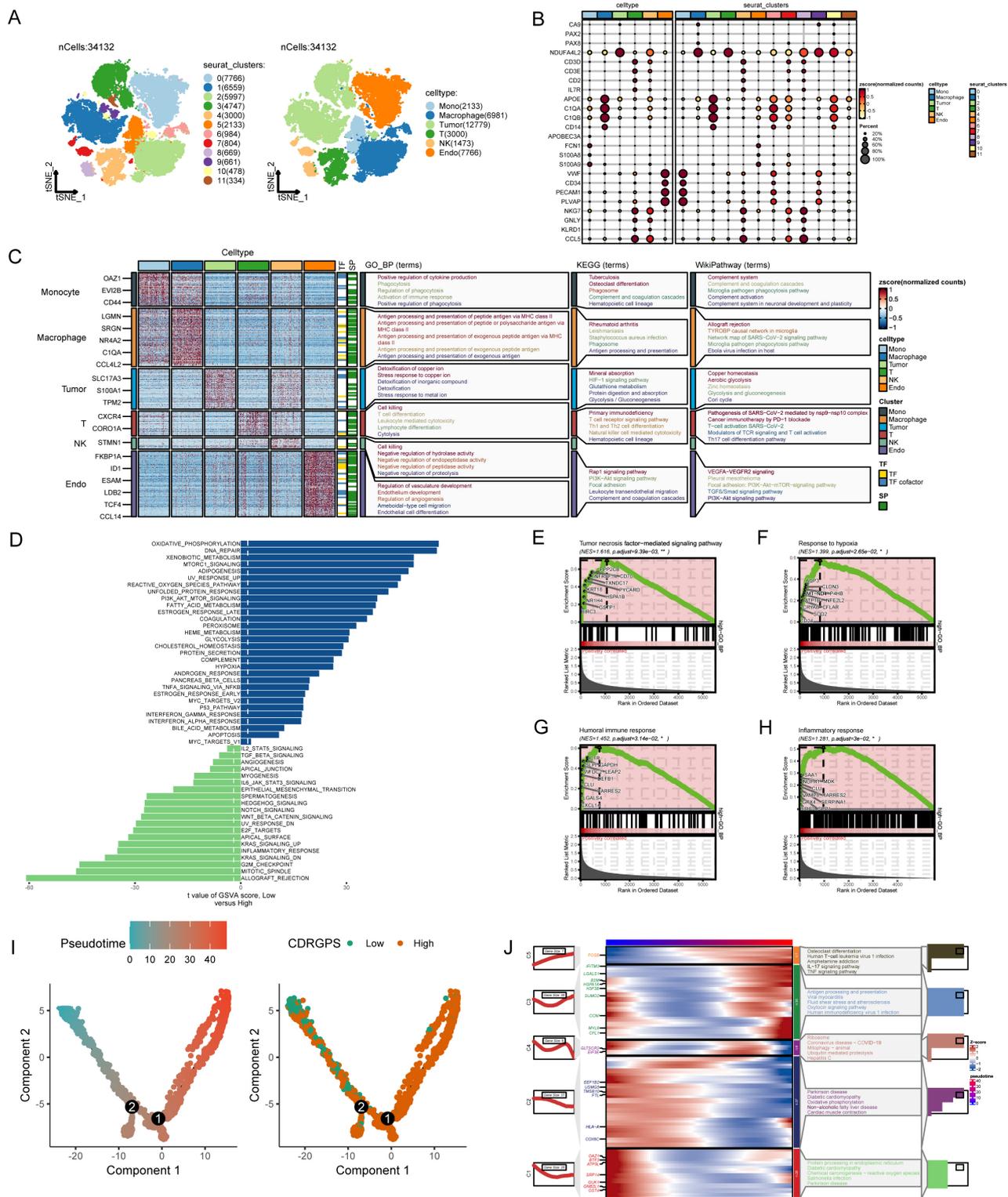


Fig. 6. Analysis of single-cell RNA sequencing data. (A) Plots of 11 cell clusters and 6 different cell types using t-Distributed Stochastic Neighbor Embedding (t-SNE). (B) The expression of marker genes in six cell clusters. (C) Marker gene heatmap of each cell subpopulation. (D) Gene set variation analysis of renal cancer cells with different levels of CDRGPS. (E–H) Gene set enrichment analysis of renal cancer cells with different CDRGPS levels. (I) Pseudo-time trajectory plot showing the association between CDRGPS and pseudo-time progression. (J) Heatmap displaying the top GO_BP terms and scaled expression of dynamic genes along a pseudo-time axis for renal carcinoma cells with varying CDRGPS levels. GO_BP, Gene Ontology Biological Process; KEGG, Kyoto Encyclopedia of Genes and Genomes.

further affirm the predictive ability of CDRGPS for patient response to immunotherapy, we examined TMB data acquired through the TCGA-KIRC cohort. Although not reaching statistical significance, the CDRGPS-high group showed a markedly higher TMB (Fig. 7H).

3.7 Associations between the CDRGPS and Drug Sensitivity

First-line therapy for advanced RCC typically involves tyrosine kinase inhibitors (TKIs) and mammalian Target of Rapamycin (mTOR) inhibitors. However, both intrinsic and acquired drug resistance pose persistent challenges, largely due to the subpopulation of tumor cells known as tumor-initiating cells (TICs) or cancer stem cells (CSCs). Six tumor stemness indices based on mRNA expression and DNA methylation signatures were obtained from earlier publications [50–52]. Correlations between the CDRGPS and these six indices were subsequently calculated. With the exception of DMPss, all indices showed statistically significant correlations with CDRGPS. Moreover, these correlations were consistently negative, with the exception of RNass (Fig. 8A). To further confirm these correlations, the drug response to selected TKIs (dasatinib, gefitinib, imatinib) and to the mTOR inhibitor temsirolimus was examined in patients stratified by CDRGPS. The CDRGPS-high group exhibited lower half-maximal inhibitory concentrations (IC50) for dasatinib, gefitinib, and temsirolimus (Fig. 8B). Furthermore, a negative association was observed between the CDRGPS and the IC50 values for these three drugs (Fig. 8C). These observations imply that ccRCC patients with CDRGPS-high may respond more favorably to therapy with TKIs and mTOR inhibitors.

3.8 Construction of a Radiomics Score (Rad_Score) Associated with CDRGPS

We selected 14 of 1688 features via LASSO regression to develop a Rad_Score that was predictive of high/low CDRGPS in TCGA-KIRC (Fig. 9A,B). The Rad_Score varied considerably between the CDRGPS-high and -low groups ($p < 0.001$, Fig. 9C). A total of 267 patients in the TCIA database were classified into High-Rad_Score and Low-Rad_Score groups using an optimized cutoff. The Rad_Score showed an AUC value of 0.813 for discriminating between CDRGPS-high and -low groups (Fig. 9D). Patients with a Low-Rad_Score had significantly better OS than those with a High-Rad_Score ($p < 0.05$; Fig. 9E).

3.9 PRSS23 as a Possible Treatment Target for RCC

Preliminary experiments suggest that serine protease 23 (PRSS23) may be a promising target for the suppression of metastasis and for immunotherapy in RCC patients. Analysis of gene expression data from TCGA and GTEx revealed upregulation of PRSS23 in ccRCC tissues compared to normal tissues (Fig. 10A). Immunohistochemical analysis confirmed that PRSS23 protein expression was

higher in RCC tissues compared to matching normal tissues (Fig. 10B). GSEA found that epithelial-mesenchymal transition and the IL2-Stat5 signaling pathway were positively associated with PRSS23 expression (Fig. 10C). Moreover, the TIMER database revealed associations between elevated PRSS23 expression and multiple immune cell categories, including CD8⁺ T cells, CD4⁺ T cells, macrophages, neutrophils and dendritic cells (Fig. 10D). siRNA-mediated knockdown of PRSS23 expression was confirmed by Western blotting (Fig. 10E,F, **Supplementary Fig. 3**) and was found to attenuate the migratory and invasive abilities of 786-O and CAKi-1 RCC cells (Fig. 10G), suggesting that PRSS23 may promote RCC metastasis.

4. Discussion

ccRCC is a significant global health concern due to its high incidence, metastatic potential, and poor prognosis. This type of RCC is quite common worldwide and exhibits substantial heterogeneity in terms of its molecular characteristics, clinical behavior, and natural disease progression [53]. The management of ccRCC is particularly challenging because of this diversity. The TNM staging system developed by the AJCC is a commonly used tool in clinical decision-making for ccRCC. However, the significant molecular and biological diversity of ccRCC mean that it is not adequately represented by anatomical TNM classification alone. This limitation can potentially lead to either overtreatment or undertreatment of patients based solely on TNM staging [54].

In recent years, cancer treatment has been revolutionized by immunotherapy targeting immune checkpoint pathways. Specifically, immune checkpoint inhibitors (ICIs) such as anti-PD-1/PD-L1 agents have become the standard of care for advanced RCC due to their ability to reverse T cell exhaustion and stimulate anti-tumor immune responses. However, despite improvements in RCC treatment by ICIs, long-lasting responses occur in only a minority of patients [55,56]. These challenges underscore the pressing need to develop more refined prognostic models for ccRCC that can assist clinical decision-making.

In the present study, ccRCC was classified into two distinct clusters. Cluster-1 showed elevated levels of apoptosis, necroptosis, and pyroptosis, while cluster-2 was enriched for autophagy, cuproptosis, entotic cell death, and ferroptosis. Importantly, cluster-1, was associated with worse prognosis than cluster-2. By performing differential expression analysis of the two clusters and univariate Cox regression analysis, we identified the critical genes contributing to the unique PCD patterns in ccRCC. Previous models often suffer from bias due to their reliance on absolute gene expression values. Moreover, the choice of algorithm used by researchers may be influenced by their personal preferences and biases. The current study devised a novel gene pair signature based on the relative expres-

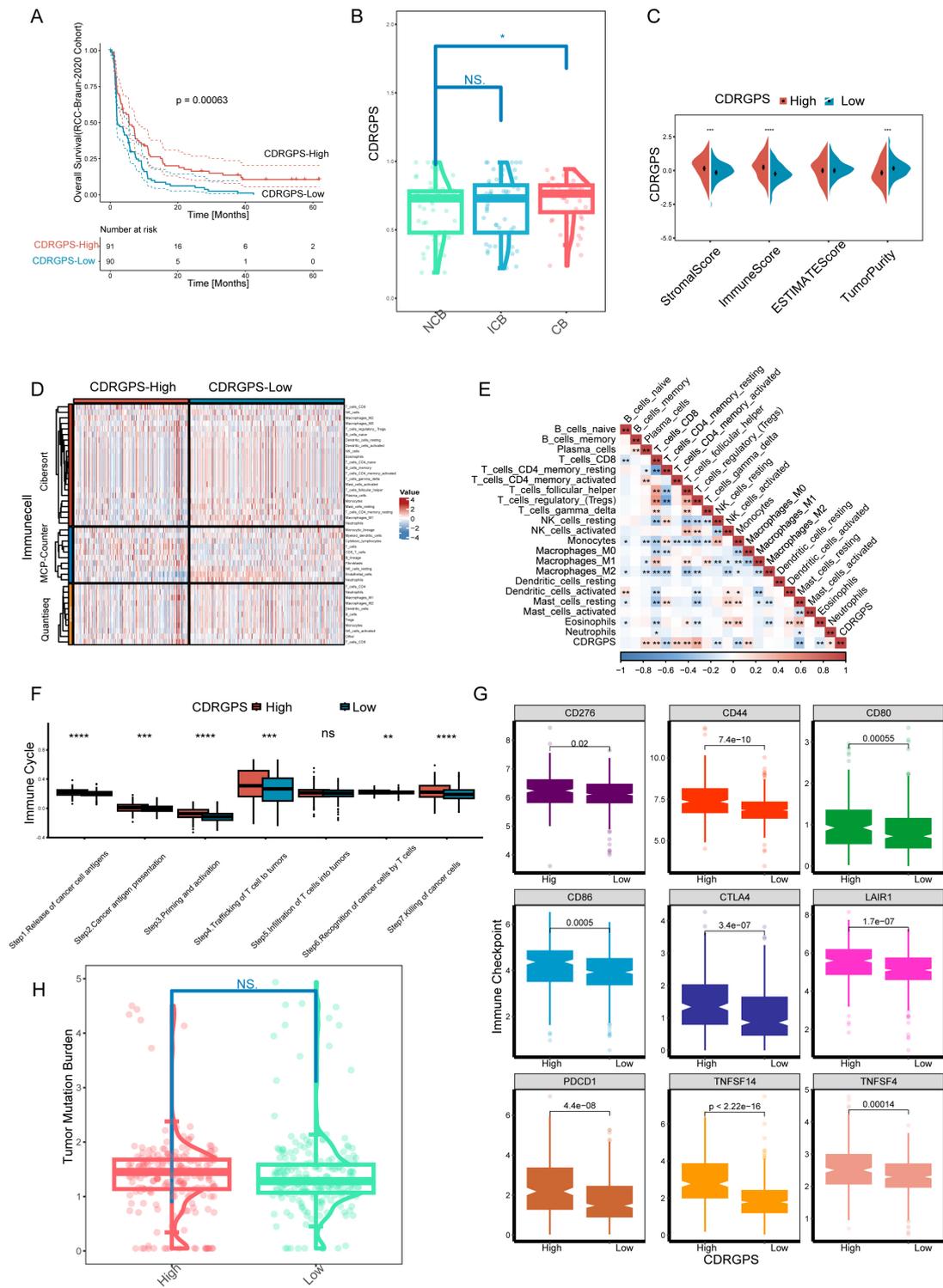


Fig. 7. Implications of the CDRGPS for immunotherapy. (A) Survival analysis of CDRGPS-high and -low patient groups in the immunotherapy cohort. (B) The distribution of CDRGPS in patient groups with different response to immunotherapy. (C) Comparison between CDRGPS-high and -low groups for the stromalscore, immunescore, ESTIMATE score, and tumor purity. (D) The abundance of infiltrating immune cells in the CDRGPS-high and -low groups was assessed using multiple algorithms. (E) Heatmap based on the Spearman r value between the CDRGPS and immune cell infiltration in the TCGA-KIRC cohort. (F) Comparison of the seven-step anticancer immunity cycle between CDRGPS-high and -low groups. (G) Comparison of immune checkpoint expression profiles between CDRGPS-high and -low groups. (H) Violin plots showing the distribution of tumor mutation burden (TMB) scores in CDRGPS-high and -low groups. **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns, non-significant.

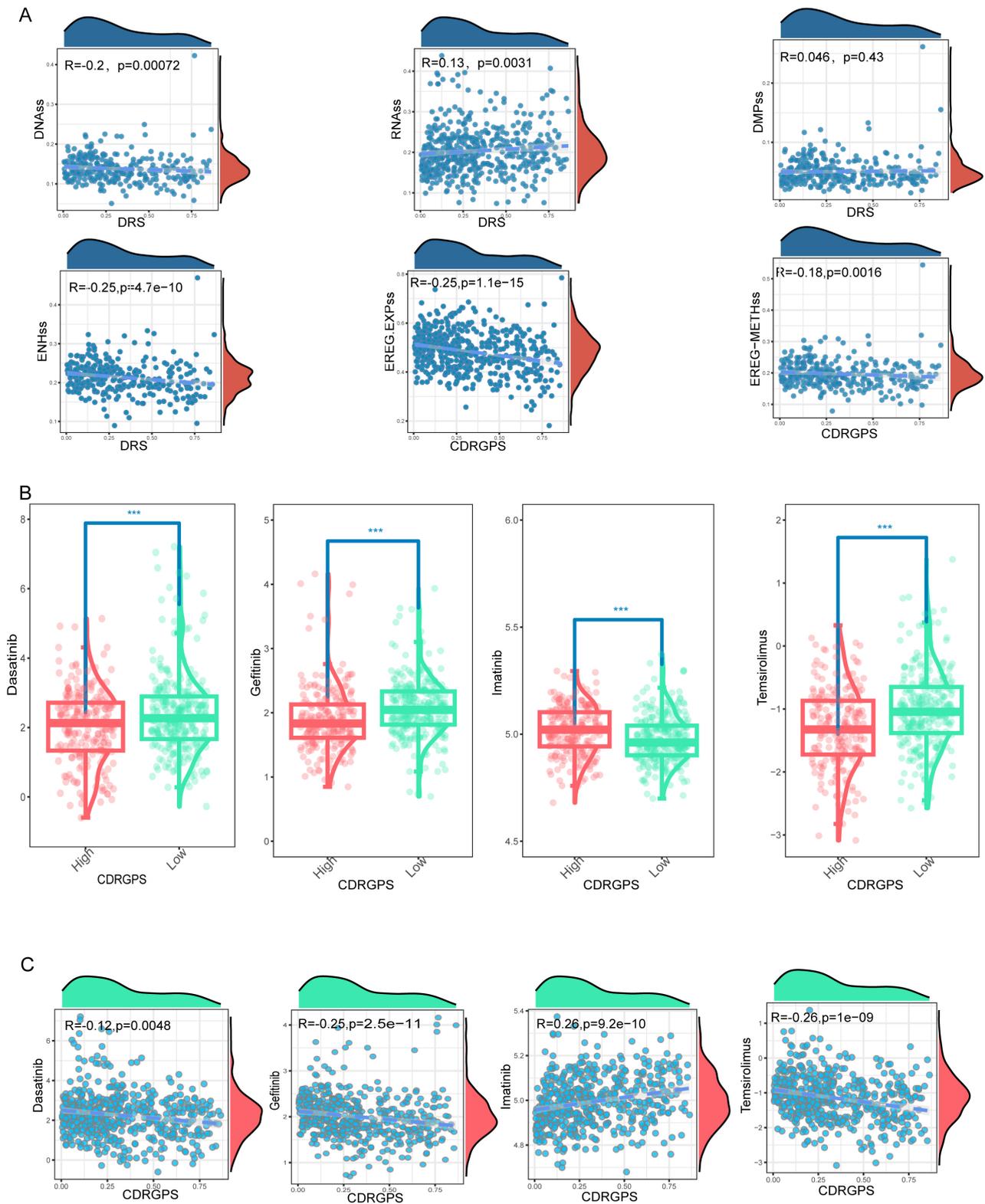


Fig. 8. Associations between CDRGPS and drug sensitivity. (A) Correlations between CDRGPS and tumor stemness indexes in ccRCC, including RNAss, EREG.EXPss, DNAss, EREG-METHss, DMPss, and ENHss. (B) Comparison of the susceptibility of CDRGPS-high and CDRGPS-low patients to TKIs such as dasatinib, gefitinib, and imatinib, and to the mTOR inhibitor temsirolimus. (C) The relationship between CDRGPS and the half-maximal inhibitory concentration (IC₅₀) of small molecule drugs including dasatinib, gefitinib, imatinib, and temsirolimus in ccRCC. $***p < 0.001$.

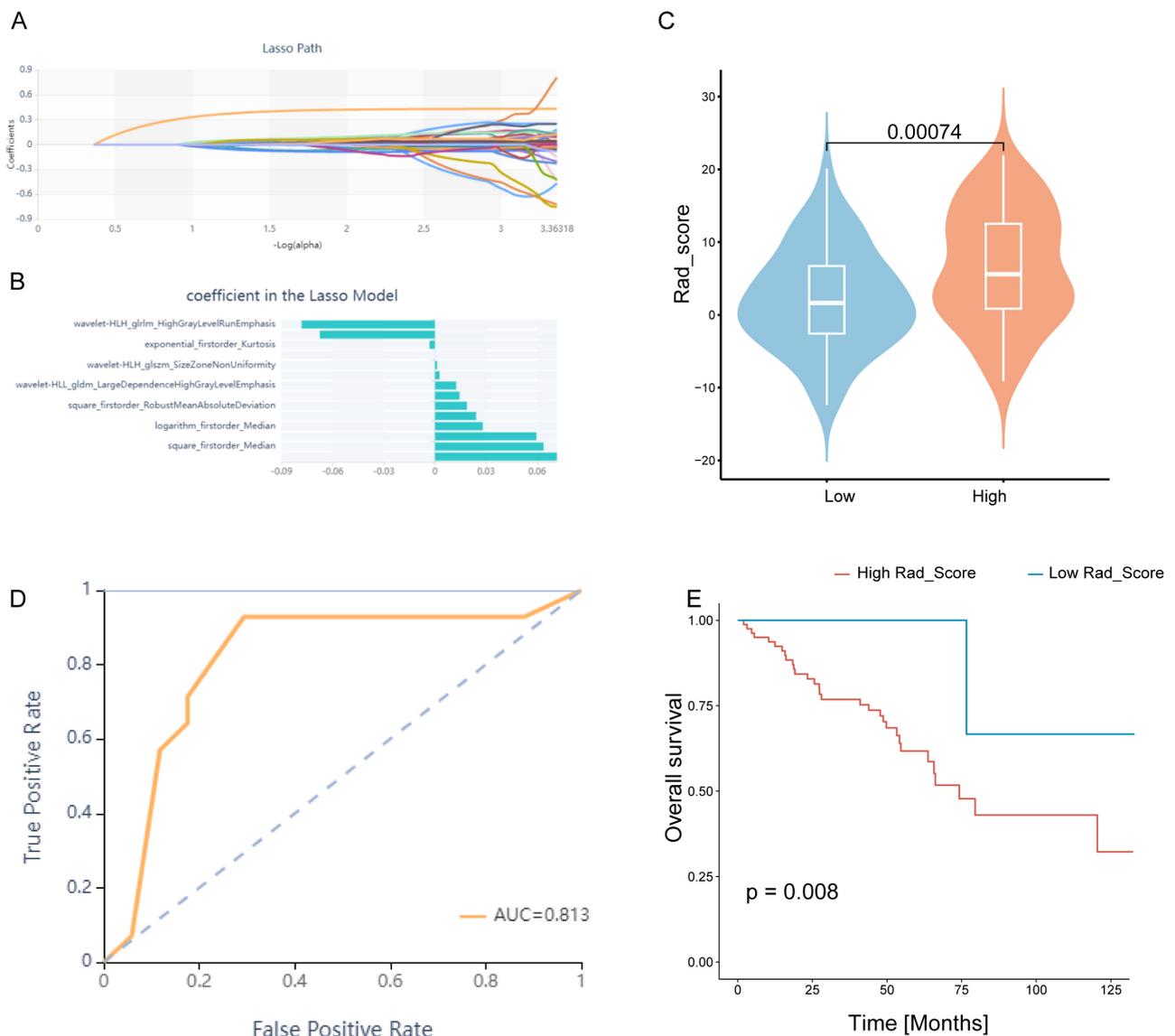


Fig. 9. Construction and evaluation of the Rad_score. (A) Lasso regression. (B) Plot of coefficient profiles for radiomics characteristics associated with the Rad_Score. (C) Differences in Rad_Score between the CDRGPS-high and -low groups. (D) The predictive value of Rad_Score for distinguishing the CDRGPS-high and -low groups. (E) Survival analysis (OS) of patients with high and low Rad_Score.

sion order within samples, rather than absolute expression levels. This eliminates the requirement for data normalization, thereby improving reliability and generalizability. Following the evaluation of 113 combinations of 12 machine learning algorithms, Lasso and RF were found to be the best model based on PCD-related gene pairs. This generated a more straightforward and transferable model by significantly reducing the dimensionality of the variables and exposing underlying patterns. Therefore, the CDRGPS may accurately predict the prognosis and efficacy of immunotherapy for ccRCC patients. Prognostic analysis of both the training and test cohorts revealed that CDRGPS was a predictor of OS. Furthermore, ROC analysis indicated

that CDRGPS consistently exhibited high accuracy and stable performance in both cohorts. These findings highlight the significant clinical potential of CDRGPS.

To further validate the robustness of our CDRGPS signature, we compared it to five previously published gene signatures that are based on various combinations of PCD-related genes. Few of these prognostic models have been used in clinical practice or undergone careful external validation. Univariate and multivariate Cox regression analysis revealed that only the CDRGPS model exhibited statistical significance as an independent prognostic factor in ccRCC. Comparative performance assessments further demonstrated the superior prognostication ability of

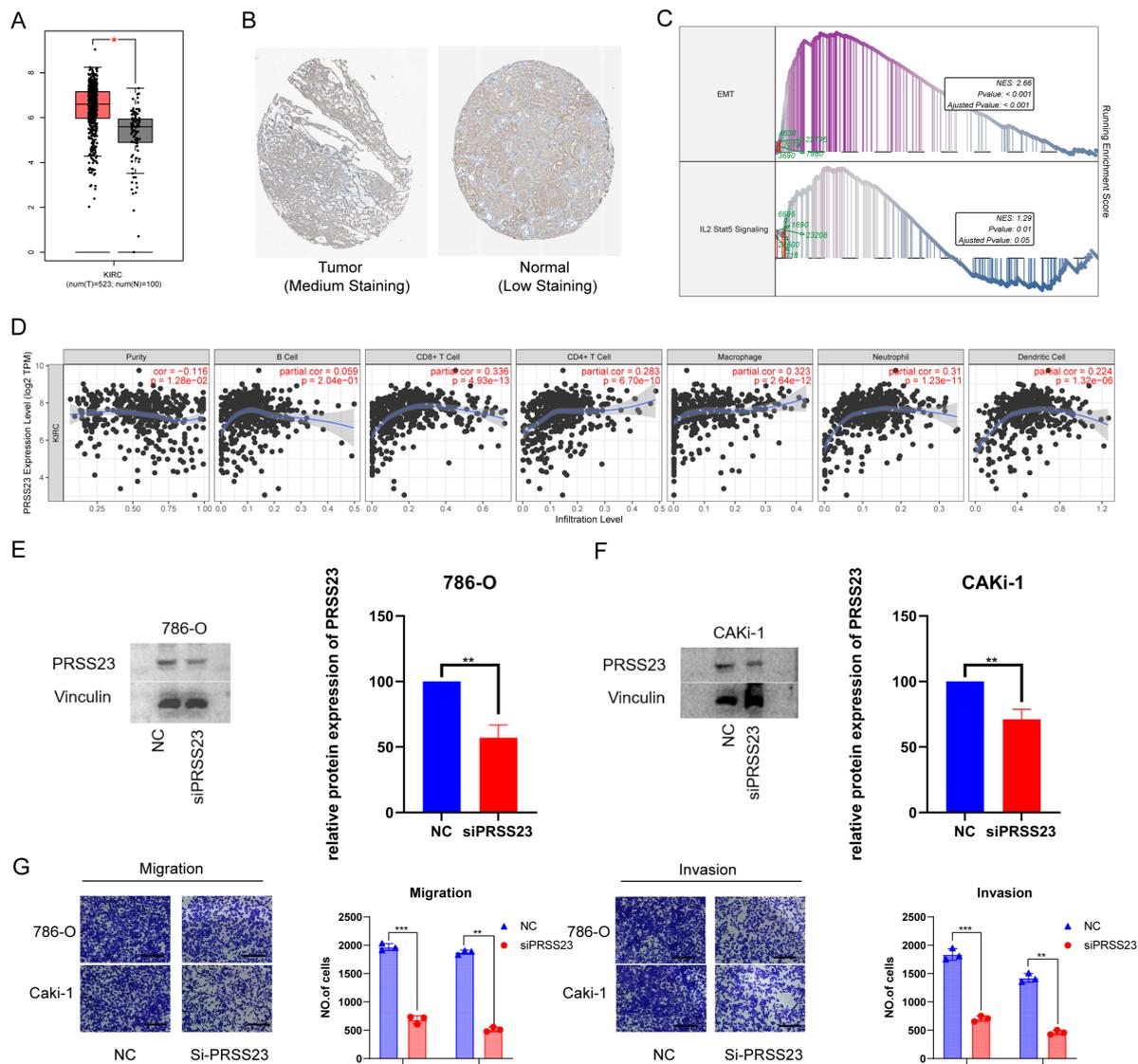


Fig. 10. The putative biological role of PRSS23 in RCC immunology and metastasis. (A) Comparison of PRSS23 expression level between tumor and normal tissues in RCC from the GEPIA database. (B) PRSS23 expression in RCC from The Human Protein Atlas database Version (HPA, image available from v23.proteinatlas.org). (C) The TCGA-KIRC cohort was split into two subgroups according to the median expression of PRSS23: PRSS23-high and PRSS23-low subgroups. GSEA found that the differentially expressed genes exhibited significant associations with immunity, epithelial-mesenchymal transition (EMT), and immune-related pathways. (D) Associations between PRSS23 and seven different immune cell types were revealed using TIMER analysis. (E) Knockdown of PRSS23 in 786-O and CAKi-1 cells. (F) Confirmation of PRSS23 knockdown by Western blot. (G) Transwell assay was used to assess the ability of RCC cells to migrate and invade following PRSS23 knockout (magnification: 400 \times , scale bars = 50 μ m). * p < 0.05, ** p < 0.01, *** p < 0.001. NES, Normalized Enrichment Score.

CDRGPS versus all other models across different datasets. Most models were proficient when applied to their own training dataset, but displayed weaker performance in independent external datasets, likely because of overfitting and poor generalizability. Subsequently, independent prognostic factors and crucial clinical variables were integrated to develop a concise and precise nomogram for prognostic

prediction in patients. Moreover, calibration and decision curve analyses upheld the clinical utility of the nomogram for providing valuable prognostication.

In the current study we leveraged the GSE156632 cohort, comprising scRNA-seq profiles of seven ccRCC samples, to characterize the heterogeneity of the TME. Quality control and annotation of the scRNA-seq data

using established marker genes identified six predominant cell types: endothelial cells, tumor cells, T cells, natural killer (NK) cells, macrophages and monocytes. Analysis of single-cell sequencing data using GSEA and GSVA revealed that tumor cells with high CDRGPS show enhanced regulation of tumor-related immunity pathways, including IMMUNE_RESPONSE, IL2_STAT5_SIGNALING, INTERFERON_GAMMA_RESPONSE, INTERFERON_ALPHA_RESPONSE, and humoral immune response. Additionally, these cells display enhanced malignant biological characteristics such as TGF_BETA_SIGNALING, PI3K_AKT_MTOR_SIGNALING, apoptosis, epithelial-to-mesenchymal transition, tumor necrosis factor-mediated signaling, and increased response to hypoxia. Furthermore, pseudotime trajectory analysis of tumor cell states showed that genes associated with high CDRGPS were linked to pathways related to carcinogenesis and immune function.

The conventional initial therapy used for advanced ccRCC is TKIs and mTOR inhibitors. However, the current prognostic markers utilized in clinical practice are unable to accurately identify individuals who may benefit from this treatment. To address this critical unmet need, we assessed the ability of CDRGPS to predict which patients would benefit most from TKIs and mTOR inhibitors. Mounting evidence has implicated PCD pathways in the response to these inhibitors [57,58]. Our findings revealed an inverse association between high CDRGPS and tumor stemness features. Additionally, a high CDRGPS was predictive of increased responsiveness to TKI and mTOR inhibitor therapy. Collectively, these results indicate the CDRGPS signature may serve as a powerful tool to inform treatment decisions and allow the selection of ccRCC patients for personalized therapy.

The advent of cancer immunotherapy has radically altered the treatment of solid tumors, including ccRCC. Nonetheless, ICIs are only effective in some individuals, and hence their over-utilization can result in substantial cost and severe adverse events. In this study, CDRGPS-high patients displayed abundant immune cell infiltration, indicative of an “immune-hot” tumor phenotype. Furthermore, CDRGPS-high tumors showed increased expression of numerous inhibitory checkpoint molecules, in line with the finding that integral immune checkpoint pathways are frequently co-opted by malignancies to attenuate anti-tumor immunity [59,60]. This suggests that adverse outcomes in CDRGPS-high ccRCC may be due to suppressed immune function caused by increased immune checkpoint signaling. Consequently, immune checkpoint blockade to restore anti-neoplastic immune responses may be an effective therapeutic strategy for CDRGPS-high patients. Furthermore, these patients also exhibited higher TMB and stronger anti-cancer immune cycle activity. Elevated TMB can increase the generation of immunogenic neoantigens derived from somatic mutations, subsequently driving the proliferation and acti-

vation of cytotoxic T lymphocytes against the tumor [61]. Taken together, our findings indicate that CDRGPS may be a useful biomarker to identify ccRCC patients most likely to derive clinical benefit from immunotherapy. This conclusion was validated in an independent cohort of patients undergoing immunotherapy.

Enzyme-Linked Immunosorbent Assay (ELISA), Immunohistochemistry (IHC) and RNA-sequence data can be used to evaluate CDRGPS. However, these methods are expensive and complex, thereby limiting their scalability. Enhanced CT examinations are commonly used in China for renal cancer diagnosis, even in resource-limited regions. The use of CT image-based radiological features to determine a patients’ CDRGPS therefore has significant clinical and economic potential. Earlier studies on ccRCC mainly used radiomics to predict post-nephrectomy clinical outcomes or nuclear grading, with few studies having explored the use of CT images to determine gene signatures [62,63]. In the present study, our radiomics model-based score (Rad_Score) showed good performance for determining the CDRGPS of patients in the TCIA database. Based on CT images, the Rad_Score can accurately reflect the CDRGPS and thus serve as a rapid prognostic indicator to help clinicians in treatment decision-making.

All of the genes in the gene pair model used in this study had significant involvement in cancer. PRSS23 is a serine protease that has been linked to tumor growth in a variety of malignancies and is markedly up-regulated in cancer stem cells [64,65]. Gene expression profiling studies have reported increased PRSS23 expression in several tumor types including breast cancer [66], prostate carcinoma [67], papillary thyroid carcinoma [68], and pancreatic cancer [69]. However, more research is required to determine the precise function of PRSS23 in RCC and its relationship to immunity. Our study showed that PRSS23 was positively associated with EMT and immune cell infiltration. Specifically, PRSS23 expression correlated positively with several immune cell types, including CD8⁺ T cells, CD4⁺ T cells, macrophages, neutrophils, and dendritic cells. Furthermore, our *in vitro* experiments provided evidence that PRSS23 may have a promoting effect on RCC metastasis. These findings indicate possible functions for PRSS23 in immune cell regulation and RCC metastasis, suggesting that it could also serve as a target for novel treatment approaches.

Despite the potential clinical utility of CDRGPS in ccRCC, certain limitations of this study warrant consideration. Firstly, although CDRGPS was rigorously evaluated and validated in the training and external datasets, large-scale prospective studies across multiple centers are needed to confirm these findings. Another limitation of the current investigation is the lack of direct *in vitro* or *in vivo* experimental confirmation of the results. Finally, the sample size in the TCGA database is relatively small.

5. Conclusions

By integrating diverse bioinformatics approaches and machine learning methods, we have built a robust and reliable model termed CDRGPS that has prognostic utility in ccRCC and the ability to predict treatment outcomes for targeted therapies and immunotherapy. A key accomplishment is development of the Rad_Score radiomic biomarker, which provides a non-invasive means of estimating patient CDRGPS status. Overall, CDRGPS holds significant promise for enabling precision oncology by guiding clinical decision-making and surveillance strategies. This should lead to individualized patient care based on minimally invasive and cost-effective approaches.

Consent for Publication

All authors have given their approval for this paper to be published.

Availability of Data and Materials

The corresponding author can be contacted for free to receive any data or R code used in this work. The datasets presented in this study can be found in online repositories.

Author Contributions

The research project was designed by DW and WC. BC and XS performed the data acquisition, and interpretation through a comprehensive review of the relevant literature. BC performed the data analysis and visualization. BC, LG, and MZ performed the *in vitro* experiments. BC, MZ, LG and WC wrote the manuscript. BC, ZP and HH contributed to preparing the figures and tables. DW and WC revised the manuscript. All authors contributed to editorial changes in the manuscript. All authors approved the final manuscript to be published. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

Not applicable.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at <https://doi.org/10.31083/j.fbl2903121>.

References

- [1] Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA: a Cancer Journal for Clinicians*. 2023; 73: 17–48.
- [2] Moch H, Amin MB, Berney DM, Compérat EM, Gill AJ, Hartmann A, *et al.* The 2022 World Health Organization Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. *European Urology*. 2022; 82: 458–468.
- [3] Martínez-Salamanca JI, Huang WC, Millán I, Bertini R, Bianco FJ, Carballido JA, *et al.* Prognostic impact of the 2009 UICC/AJCC TNM staging system for renal cell carcinoma with venous extension. *European Urology*. 2011; 59: 120–127.
- [4] Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *The Lancet. Oncology*. 2016; 17: e542–e551.
- [5] Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, *et al.* Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Annals of Oncology: Official Journal of the European Society for Medical Oncology*. 2019; 30: 44–56.
- [6] Cortes-Ciriano I, Lee S, Park WY, Kim TM, Park PJ. A molecular portrait of microsatellite instability across multiple cancers. *Nature Communications*. 2017; 8: 15180.
- [7] Su W, Hong T, Feng B, Yang Z, Lei G. A unique regulated cell death-related classification regarding prognosis and immune landscapes in non-small cell lung cancer. *Frontiers in Immunology*. 2023; 14: 1075848.
- [8] Zhao R, Kaakati R, Lee AK, Liu X, Li F, Li CY. Novel roles of apoptotic caspases in tumor repopulation, epigenetic reprogramming, carcinogenesis, and beyond. *Cancer Metastasis Reviews*. 2018; 37: 227–236.
- [9] Amaravadi RK, Kimmelman AC, Debnath J. Targeting Autophagy in Cancer: Recent Advances and Future Directions. *Cancer Discovery*. 2019; 9: 1167–1181.
- [10] Hombach-Klonisch S, Mehrpour M, Shojaei S, Harlos C, Pitz M, Hamai A, *et al.* Glioblastoma and chemoresistance to alkylating agents: Involvement of apoptosis, autophagy, and unfolded protein response. *Pharmacology & Therapeutics*. 2018; 184: 13–41.
- [11] Song X, Zhu S, Xie Y, Liu J, Sun L, Zeng D, *et al.* JTC801 Induces pH-dependent Death Specifically in Cancer Cells and Slows Growth of Tumors in Mice. *Gastroenterology*. 2018; 154: 1480–1493.
- [12] Tsvetkov P, Coy S, Petrova B, Dreishpoon M, Verma A, Abusamad M, *et al.* Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science (New York, N.Y.)*. 2022; 375: 1254–1261.
- [13] Liu X, Olszewski K, Zhang Y, Lim EW, Shi J, Zhang X, *et al.* Cystine transporter regulation of pentose phosphate pathway dependency and disulfide stress exposes a targetable metabolic vulnerability in cancer. *Nature Cell Biology*. 2020; 22: 476–486.
- [14] Wen S, Niu Y, Lee SO, Chang C. Androgen receptor (AR) positive vs negative roles in prostate cancer cell deaths including apoptosis, anoikis, entosis, necrosis and autophagic cell death. *Cancer Treatment Reviews*. 2014; 40: 31–40.
- [15] Stockwell BR, Friedmann Angeli JP, Bayir H, Bush AI, Conrad M, Dixon SJ, *et al.* Ferroptosis: A Regulated Cell Death Nexus Linking Metabolism, Redox Biology, and Disease. *Cell*. 2017; 171: 273–285.
- [16] Jiang Z, Lim SO, Yan M, Hsu JL, Yao J, Wei Y, *et al.* TYRO3 induces anti-PD-1/PD-L1 therapy resistance by limiting innate immunity and tumoral ferroptosis. *The Journal of Clinical Investigation*. 2021; 131: e139434.
- [17] Aits S, Jäättelä M. Lysosomal cell death at a glance. *Journal of Cell Science*. 2013; 126: 1905–1912.
- [18] Su Z, Yang Z, Xu Y, Chen Y, Yu Q. Apoptosis, autophagy,

- necroptosis, and cancer metastasis. *Molecular Cancer*. 2015; 14: 48.
- [19] Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, *et al.* Neutrophil extracellular traps kill bacteria. *Science (New York, N.Y.)*. 2004; 303: 1532–1535.
- [20] Scaturro P, Pichlmair A. Oxeiptosis: a discreet way to respond to radicals. *Current Opinion in Immunology*. 2019; 56: 37–43.
- [21] Fatokun AA, Dawson VL, Dawson TM. Parthanatos: mitochondrial-linked mechanisms and therapeutic opportunities. *British Journal of Pharmacology*. 2014; 171: 2000–2016.
- [22] Tang R, Xu J, Zhang B, Liu J, Liang C, Hua J, *et al.* Ferroptosis, necroptosis, and pyroptosis in anticancer immunity. *Journal of Hematology & Oncology*. 2020; 13: 110.
- [23] Hou J, Zhao R, Xia W, Chang CW, You Y, Hsu JM, *et al.* PD-L1-mediated gasdermin C expression switches apoptosis to pyroptosis in cancer cells and facilitates tumour necrosis. *Nature Cell Biology*. 2020; 22: 1264–1275.
- [24] Dai Z, Zhang N, Zhou R, Zhang H, Zhang L, Wang Z, *et al.* Identification of a single cell-based signature for predicting prognosis risk and immunotherapy response in patients with glioblastoma. *Clinical Immunology (Orlando, Fla.)*. 2023; 251: 109345.
- [25] Cirenajwis H, Lauss M, Planck M, Vallon-Christersson J, Staaf J. Performance of gene expression-based single sample predictors for assessment of clinicopathological subgroups and molecular subtypes in cancers: a case comparison study in non-small cell lung cancer. *Briefings in Bioinformatics*. 2020; 21: 729–740.
- [26] Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016; 278: 563–577.
- [27] Limkin EJ, Sun R, Derclé L, Zacharakí EI, Robert C, Reuzé S, *et al.* Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Annals of Oncology: Official Journal of the European Society for Medical Oncology*. 2017; 28: 1191–1206.
- [28] Sato Y, Yoshizato T, Shiraishi Y, Maekawa S, Okuno Y, Kamura T, *et al.* Integrated molecular analysis of clear-cell renal cell carcinoma. *Nature Genetics*. 2013; 45: 860–867.
- [29] Tang D, Kang R, Berghe TV, Vandenabeele P, Kroemer G. The molecular machinery of regulated cell death. *Cell Research*. 2019; 29: 347–364.
- [30] Zou Y, Xie J, Zheng S, Liu W, Tang Y, Tian W, *et al.* Leveraging diverse cell-death patterns to predict the prognosis and drug sensitivity of triple-negative breast cancer patients after surgery. *International Journal of Surgery (London, England)*. 2022; 107: 106936.
- [31] Hong W, Liang L, Gu Y, Qi Z, Qiu H, Yang X, *et al.* Immune-Related lncRNA to Construct Novel Signature and Predict the Immune Landscape of Human Hepatocellular Carcinoma. *Molecular Therapy. Nucleic Acids*. 2020; 22: 937–947.
- [32] Zhang M, Zhai W, Miao J, Cheng X, Luo W, Song W, *et al.* Single cell analysis reveals intra-tumour heterogeneity, microenvironment and potential diagnosis markers for clear cell renal cell carcinoma. *Clinical and Translational Medicine*. 2022; 12: e713.
- [33] Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*. 2018; 36: 411–420.
- [34] Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, *et al.* Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell*. 2018; 173: 338–354.e15.
- [35] Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One*. 2014; 9: e107468.
- [36] Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods in Molecular Biology (Clifton, N.J.)*. 2018; 1711: 243–259.
- [37] Shi J, Jiang D, Yang S, Zhang X, Wang J, Liu Y, *et al.* LPAR1, Correlated With Immune Infiltrates, Is a Potential Prognostic Biomarker in Prostate Cancer. *Frontiers in Oncology*. 2020; 10: 846.
- [38] Plattner C, Finotello F, Rieder D. Deconvoluting tumor-infiltrating immune cells from RNA-seq data using quanTIseq. *Methods in Enzymology*. 2020; 636: 261–285.
- [39] Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-García W, *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications*. 2013; 4: 2612.
- [40] Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity*. 2013; 39: 1–10.
- [41] Braun DA, Hou Y, Bakouny Z, Ficial M, Sant’ Angelo M, Forman J, *et al.* Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nature Medicine*. 2020; 26: 909–918.
- [42] Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, *et al.* The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *Journal of Digital Imaging*. 2013; 26: 1045–1057.
- [43] Bian Z, Fan R, Xie L. A Novel Cuproptosis-Related Prognostic Gene Signature and Validation of Differential Expression in Clear Cell Renal Cell Carcinoma. *Genes*. 2022; 13: 851.
- [44] Zhao GJ, Wu Z, Ge L, Yang F, Hong K, Zhang S, *et al.* Ferroptosis-Related Gene-Based Prognostic Model and Immune Infiltration in Clear Cell Renal Cell Carcinoma. *Frontiers in Genetics*. 2021; 12: 650416.
- [45] 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. *EPMA Journal*. 2023; 14: 275–305.
- [46] Gan Y, Zhang Z, Wang X, Li A, Fan Y, Zhang Q. Pyroptosis-Related Gene Signature Predicts the Prognosis of ccRCC Using TCGA and Single-Cell RNA Seq Database. *Journal of Healthcare Engineering*. 2022; 2022: 8224618.
- [47] Qin H, Abulaiti A, Maimaiti A, Abulaiti Z, Fan G, Aili Y, *et al.* Integrated machine learning survival framework develops a prognostic model based on inter-crosstalk definition of mitochondrial function and cell death patterns in a large multicenter cohort for lower-grade glioma. *Journal of Translational Medicine*. 2023; 21: 588.
- [48] Xu-Monette ZY, Zhou J, Young KH. PD-1 expression and clinical PD-1 blockade in B-cell lymphomas. *Blood*. 2018; 131: 68–83.
- [49] Qin S, Xu L, Yi M, Yu S, Wu K, Luo S. Novel immune checkpoint targets: moving beyond PD-1 and CTLA-4. *Molecular Cancer*. 2019; 18: 155.
- [50] Makhov P, Joshi S, Ghatalia P, Kutikov A, Uzzo RG, Kolenko VM. Resistance to Systemic Therapies in Clear Cell Renal Cell Carcinoma: Mechanisms and Management Strategies. *Molecular Cancer Therapeutics*. 2018; 17: 1355–1364.
- [51] Atkins MB, Tannir NM. Current and emerging therapies for first-line treatment of metastatic clear cell renal cell carcinoma. *Cancer Treatment Reviews*. 2018; 70: 127–137.
- [52] Lee TKW, Guan XY, Ma S. Cancer stem cells in hepatocellular carcinoma - from origin to clinical implications. *Nature Reviews. Gastroenterology & Hepatology*. 2022; 19: 26–44.
- [53] Hsieh JJ, Purdue MP, Signoretti S, Swanton C, Albiges L, Schmidinger M, *et al.* Renal cell carcinoma. *Nature Reviews. Disease Primers*. 2017; 3: 17009.
- [54] Salazar R, Taberero J. New approaches but the same flaws in the search for prognostic signatures. *Clinical Cancer Research*:

an Official Journal of the American Association for Cancer Research. 2014; 20: 2019–2022.

- [55] Raimondi A, Randon G, Sepe P, Claps M, Verzoni E, de Braud F, *et al.* The Evaluation of Response to Immunotherapy in Metastatic Renal Cell Carcinoma: Open Challenges in the Clinical Practice. *International Journal of Molecular Sciences*. 2019; 20: 4263.
- [56] Borcherdig N, Vishwakarma A, Voigt AP, Bellizzi A, Kaplan J, Nepple K, *et al.* Mapping the immune environment in clear cell renal carcinoma by single-cell genomics. *Communications Biology*. 2021; 4: 122.
- [57] Chen P, Huang HP, Wang Y, Jin J, Long WG, Chen K, *et al.* Curcumin overcome primary gefitinib resistance in non-small-cell lung cancer cells through inducing autophagy-related cell death. *Journal of Experimental & Clinical Cancer Research: CR*. 2019; 38: 254.
- [58] Long S, Wang Y, Chen Y, Fang T, Yao Y, Fu K. Pan-cancer analysis of cuproptosis regulation patterns and identification of mTOR-target responder in clear cell renal cell carcinoma. *Biology Direct*. 2022; 17: 28.
- [59] Toor SM, Sasidharan Nair V, Decock J, Elkord E. Immune checkpoints in the tumor microenvironment. *Seminars in Cancer Biology*. 2020; 65: 1–12.
- [60] Topalian SL, Taube JM, Anders RA, Pardoll DM. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nature Reviews. Cancer*. 2016; 16: 275–287.
- [61] McGranahan N, Furness AJS, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, *et al.* Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science (New York, N.Y.)*. 2016; 351: 1463–1469.
- [62] Deniffel D, McAlpine K, Harder FN, Jain R, Lawson KA, Healy GM, *et al.* Predicting the recurrence risk of renal cell carcinoma after nephrectomy: potential role of CT-radiomics for adjuvant treatment decisions. *European Radiology*. 2023; 33: 5840–5850.
- [63] Zhou H, Mao H, Dong D, Fang M, Gu D, Liu X, *et al.* Development and External Validation of Radiomics Approach for Nuclear Grading in Clear Cell Renal Cell Carcinoma. *Annals of Surgical Oncology*. 2020; 27: 4057–4065.
- [64] Han B, Yang Y, Chen J, He X, Lv N, Yan R. PRSS23 knockdown inhibits gastric tumorigenesis through EIF2 signaling. *Pharmacological Research*. 2019; 142: 50–57.
- [65] Seno A, Kasai T, Ikeda M, Vaidyanath A, Masuda J, Mizutani A, *et al.* Characterization of Gene Expression Patterns among Artificially Developed Cancer Stem Cells Using Spherical Self-Organizing Map. *Cancer Informatics*. 2016; 15: 163–178.
- [66] Zajchowski DA, Bartholdi MF, Gong Y, Webster L, Liu HL, Munishkin A, *et al.* Identification of gene expression profiles that predict the aggressive behavior of breast cancer cells. *Cancer Research*. 2001; 61: 5168–5178.
- [67] Ikon ET, Pang ALY, Johnson W, Cashman K, Zimmerman S, Muralidhar S, *et al.* U94 alters FN1 and ANGPTL4 gene expression and inhibits tumorigenesis of prostate cancer cell line PC3. *Cancer Cell International*. 2005; 5: 19.
- [68] Jarzab B, Wiench M, Fajarewicz K, Simek K, Jarzab M, Oczko-Wojciechowska M, *et al.* Gene expression profile of papillary thyroid cancer: sources of variability and diagnostic implications. *Cancer Research*. 2005; 65: 1587–1597.
- [69] Jones S, Zhang X, Parsons DW, Lin JCH, Leary RJ, Angenendt P, *et al.* Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science (New York, N.Y.)*. 2008; 321: 1801–1806.