scientific reports

OPEN

Integrated RNA sequencing analysis and machine learning identifies a metabolism-related prognostic signature in clear cell renal cell carcinoma

Yunxun Liu[1,2,4], Zhiwei Yan[1,2,4], Cheng Liu[3,4], Rui Yang[1,2], Qingyuan Zheng[1,2], Jun Jian[1,2], Minghui Wang[1,2], Lei Wang[1,2], Xiaodong Weng[1,2✉], Zhiyuan Chen[1,2✉] & Xiuheng Liu[1,2✉]

The connection between metabolic reprogramming and tumor progression has been demonstrated in an increasing number of researches. However, further research is required to identify how metabolic reprogramming affects interpatient heterogeneity and prognosis in clear cell renal cell carcinoma (ccRCC). In this work, single-cell RNA sequencing (scRNA-seq) based deconvolution was utilized to create a malignant cell hierarchy with metabolic differences and to investigate the relationship between metabolic biomarkers and prognosis. Simultaneously, we created a machine learning-based approach for creating metabolism-related prognostic signature (MRPS). Gamma-glutamyltransferase 6 (GGT6) was further explored for deep biological insights through in vitro experiments. Compared to 51 published signatures and conventional clinical features, MRPS showed substantially higher accuracy. Meanwhile, high MRPS-risk samples demonstrated an immunosuppressive phenotype with more infiltrations of regulatory T cell (Treg) and tumour-associated macrophage (TAM). Following the administration of immune checkpoint inhibitors (ICIs), MRPS showed consistent and strong performance and was an independent risk factor for overall survival. GGT6, an essential metabolic indicator and component of MRPS, has been proven to support proliferation and invasion in ccRCC. MRPS has the potential to be a highly effective tool in improving the clinical results of patients with ccRCC.

Keywords Clear cell renal cell carcinoma, Single-cell RNA sequencing, Cell metabolic reprogramming, Machine learning, Prognosis

Renal cell carcinoma (RCC) is a type of cancerous tumor originating from the renal epithelium. About half of all RCC cases in the United States are classified as clear cell RCC (ccRCC), which is the most frequently occurring specified histological subtype[1]. The recommended course of therapy for ccRCC is surgical resection. Unsatisfactory outcomes, nevertheless, will result in a poor prognosis due to delayed diagnosis and postoperative recurrence[2]. There is growing evidence that ICIs have therapeutic utility in ccRCC, a highly immunogenic tumor that is resistant to radiation treatment[3,4]. However, because of the heterogeneity of ccRCC and diverse tumor immune microenvironments (TIM), the application of ICIs is still restricted[5,6]. Therefore, in the era of customized therapy, it is vital to develop accurate biomarkers for the benefits of drug treatment and further improve the prognosis of ccRCC.

It has been proven that ccRCC cells alter physiological characteristics via a range of metabolic pathways, resulting in the induction of drug resistance and malignant behavior. Additionally, the research indicates that metabolic reprogramming, in which the activation of a certain pathway in tumor tissue may encourage tumor cell proliferation and invasion, may function as a predictive biomarker for ccRCC[7–9]. The underlying biological mechanisms are becoming more and more clear. For instance, ccRCC cells can drive the growth of tumors by

[1]Department of Urology, Renmin Hospital of Wuhan University, Wuhan 430060, China. [2]Institute of Urologic Disease, Renmin Hospital of Wuhan University, Wuhan 430060, China. [3]Department of Gynecology and Obstetrics, Renmin Hospital of Wuhan University, Wuhan 430060, China. [4]Yunxun Liu, Zhiwei Yan and Cheng Liu, contributed equally to this work. ✉email: wengxdx@163.com; drchenzy@whu.edu.cn; drliuxh@whu.edu.cn

starting a positive feedback loop between histone lactylation and PDGFRbeta signaling[10]. Additionally, aberrant lactate metabolism and mitochondrial dysfunction alter the performance of stromal cells and immune cells in the surrounding tissue, leading to inhibitory immunophenotype, which makes it difficult for patients with ccRCC to benefit from immunotherapy[11]. Moreover, medication resistance to targeted therapy is connected to the metabolic alterations in ccRCC cells[12].

However, earlier prognostic signatures for metabolism-related genes were based on bulk RNA-seq, which was insufficient to detect distinct cell types and impossible to analyze the complexity of intra tumor heterogeneity[13]. In contrast, scRNA-seq technology has grown in recent years and is now a potent tool for dissecting cells at the cellular level, allowing genome, transcriptome, and epigenome study at the individual cell level[14]. Unfortunately, due to the small number of patients from the scRNA-seq dataset, few studies have thoroughly examined the relationship between metabolic reprogramming as reflected in cell-level composition and inter-patient heterogeneity, which may account for the poor prognosis of ccRCC and the limited application of ICIs. To uncover new therapeutic targets, we integrated bulk RNA-seq and scRNA-seq of ccRCC to investigate the association between particular malignant cell subsets and patient prognosis.

After the deconvolution of gene expression, the prognostic effects of a subset of malignant cells known as META_active were examined at many levels in this investigation, which is characterized by metabolic reprogramming. We created MRPS based on genes in metabolically associated pathways and found notable variations in the TIM and clinical phenotype between various risk subgroups. A novel molecular marker, GGT6, was also discovered in ccRCC. These results highlight the importance of cell heterogeneity in ccRCC and set the groundwork for the development of therapeutical applications.

Methods

Collection of ccRCC datasets

Corresponding clinical data and mRNA expression data of GSE167573[15] were obtained from GEO (http://www.ncbi.nlm.nih.gov/geo/). Clinical and bulk RNA-seq data of ccRCC samples of TCGA-KIRC[16], E-MTAB-1980[17] and RECA-EU[18] were downloaded from UCSC Xena database (https://xenabrowser.net/datapages/), ArrayExpress database (https://www.ebi.ac.uk/biostudies/arrayexpress) and ICGC database (https://dcc.icgc.org/) respectively. The samples in the TCGA-KIRC database were divided into training and testing sets in the ratio of 1:1. E-MTAB-1980 and Meta cohort composed of GSE167573 database and RECA-EU database served as validation sets. The scRNA-seq expression profiles of 26 ccRCC samples were downloaded with GEO accessions: GSE152938[19], GSE207493[20], GSE224630[21].

Processing of scRNA-seq data

A standard quality control, transformation, and data integration based on "Seurat"[22] was performed: (a) cells with more than 5000 unique molecular identifier (UMI), genes detected in less than 10 cells, and cells with expressed genes fewer than 200 or more than 5000 were removed, while the proportion of mitochondria genes was limited to less than 15%. Additionally, cells with a gene to UMI ratio (in log10 space) greater than 0.8 were selected; (b) the gene expression matrixes were then normalized using the "SCTransform"[23] method to remove confounding sources of UMI, mitochondria genes percentage and cell cycle; (c) Principal component analysis (PCA) and "Harmony"[24] method were then used to integrate scRNA-seq datasets and perform batch calibration. The top 76 Harmony dimensions were identified where the percent change in variation between consecutive Harmony dimensions was less than 0.05%. Visualization employed "FindClusters" function at resolution 0.5 and the uniform manifold approximation and projection (UMAP) algorithm. The cell types were manually annotated based on highly expressed genes and canonical cell markers reported by the CellMarker2.0[25] database (Supplementary Table S1). Copy Number Variation (CNV) analysis based on "infercnv" (https://github.com/broadinstitute/inferCNV) was performed to estimate the chromosome copy number characteristics of malignant cells, in which NK cells were set as the control group.

Differential expression and functional analyses

The "FindAllMarkers" function in Seurat was used to detect the differential expressed genes (DEGs) between each type. The DEGs of malignant cells were used for functional analysis based on Kyoto Encyclopedia of Genes and Genomes (KEGG) database[26,27] and "clusterProfiler"[28] package. The significant pathways related to metabolism were selected and genes contained within metabolic pathways were defined as "META_process" gene set. To evaluate the availability of the META_process gene set in DEGs, Gene Set Enrichment Analysis (GSEA) was performed using the "clusterProfiler" package.

AUCell

To identify malignant cells with active META_process gene set, "AUCell" (https://github.com/aertslab/AUCell) was used to calculate each cell's "Area Under the Curve" (AUC) value. The distribution of AUC value across all the cells indicated the relative expression of the gene set. The ideal situation in the cell type of interest should be a bimodal distribution, and to ease the exploration of the distributions, the function "AUCell_exploreThresholds" was used to consider META_process active or inactive. Metabolism-related genes (MRGs) for downstream analysis were defined as entire genes in META_process gene set obtained from KEGG database.

ScRNA-seq deconvolution based on BayesPrism

Raw gene expression count matrix and reannotated cell type from scRNA-seq data and matrix of bulk RNA-seq expression of TCGA-KIRC were used as input for "BayesPrism"[29]. After removal of ribosomal protein genes, mitochondrial genes, genes in sex chromosomes and lowly transcribed genes, the shared protein coding genes were used for the construction of the Prism object, in which the "META_active" cells were specially treated as key

cells. The deconvolution profiles were then segmented by "surv_cutpoint" function to explore the relationship between cell proportion and survival.

Pseudotime and cell-cell communication analyses

The "Monocle2"[30] provided a toolkit for reconstructing the cell fate decisions and pseudo-time trajectories of malignant cells with different META_process. After filtering low-quality cells, we selected genes that were variable across cells with "dispersionTable" function. Using the DDRTree algorithm to reduce the scRNA-seq data dimensionality, malignant cells along the trajectory were ordered. "CytoTRACE"[31] method was used to predict the relative differentiation state of cells and determine the developmental sequence and origin point. Normalized count data was input into "CellChat"[32] analysis to explore the potential communication between core cells.

Construction and validation of machine learning-based signature

"Limma" package was used to identify DEGs between 534 ccRCC and 72 normal samples in TCGA-KIRC cohort with $|log2FC| > [mean(log2FC) + 2SD(log2FC)]$ and adjusted p-value < 0.05 (Benjamini Hochberg test). Afterwards, the DEGs and MRGs overlapped generating the differential expressed metabolism-related genes (DEMRGs) associated with ccRCC. After removing batch effects from several databases with the "sva"[33] package, co-detected DEMRGs were utilized in the subsequent analysis. To develop a consensus MRPS with high accuracy and stability performance, we integrated 10 machine learning algorithms including random survival forest (RSF), Ridge, elastic network (Enet), least absolute shrinkage and selection operator (Lasso), CoxBoost, stepwise Cox, partial least squares regression for Cox (plsRcox), supervised principal components (SuperPC), generalized boosted regression modeling (GBM), and survival support vector machine (survival-SVM). Algorithms such as RSF, stepwise Cox, CoxBoost, and Lasso have feature selection capabilities and could be combined to create a consensus model[34]. In total, 118 algorithm combinations based aforementioned ones were conducted to fit prediction models based on the leave-one-out cross-validation (LOOCV) framework in the TCGA training cohort. All models were detected in 3 validation cohorts (TCGA testing, E-MTAB-1980 and Meta). The best model was determined based on the highest average concordance index (C-index) across all cohorts. Patients were stratified into either high- or low-risk subgroups based on the median risk score derived from TCGA training cohort. The distribution of the best model was examined using the PCA method, and the prognostic value and prediction accuracy were assessed by the Kaplan-Meier (KM) curve and the receiver operating characteristic (ROC) curve.

Analysis of clinical characteristics and MRPS

Samples with multiple clinical characteristics were classified into the following subtypes, including age, gender, AJCC tumor stage, T stage, N stage, M stage, and Fuhrman grade. The distribution of clinical characteristics among different MRPS risk subgroups was assessed. Univariate and multivariate Cox regression analyses were performed to determine whether the MRPS independently served as an independent prognostic factor. In addition, using the "DynNom" package, a dynamic nomogram was developed to forecast the survival rates of patients with ccRCC. Calibration curves and C-index were used to verify the validity of nomogram.

Comparison of gene expression-based prognostic signatures in ccRCC

We gathered 51 reported ccRCC mRNA signatures from PubMed (https://pubmed.ncbi.nlm.nih.gov/) to compare their performance to that of MRPS. Due to restrictions in the validation cohort, the lncRNA and miRNA signatures were excluded. We collected the genes and corresponding coefficients constituted of these signatures, primarily created by the Lasso and Cox proportional-hazards models. In cases where particular coefficients were not supplied, we used the Cox model to rebuild them. Finally, the performance of MRPS compared to other signatures was then compared using the C-index.

TIM estimation and drug sensitivity

With the support of "IOBR"[35] package, the infiltration levels of immune cells in TIM were estimated by 7 algorithms ("TIMER", "CIBERSORT", "MCP counter", "EPIC", "xCell", "quanTIseq", "ssGSEA"). The distribution and correlation of immune cell types between MRPS were examined. Overall survival (OS) and RNA-seq data were retrieved from 3 clinical trials: CheckMate-009, CheckMate-010 and CheckMate-025[36]. To evaluate MRPS's prognostic prediction capacity, ccRCC patients who were treated with nivolumab (anti-PD-1) individually were gathered. Finally, the efficacy of anticancer drugs was assessed between the two subgroups based on the IC50 values acquired from the GDSC1, GDSC2, and CTRP2 databases using the "oncoPredict"[37] package.

Cell culture and transfection

CcRCC cell lines (769-P and 786-O) were purchased from the Chinese Academy of Sciences (Shanghai, China) and cultured in 1640 medium (Gibco, USA) containing 10% fetal bovine serum (Gibco, USA) and 1% penicillin-streptomycin. All cell lines were maintained in a 5% CO_2 and 37°C incubator. Cells in the logarithmic growth phase of 769-P and 786-O were seeded into six-well culture plates. When the cell density reached approximately 70%, 4μl of Lipo8000 (Beyotime Biotechnology, Shanghai, China), 100pmol of small interfering RNA (siRNA), and 125μl of serum-free 1640 medium were mixed to prepare the transfection reagent. The cells were then incubated at room temperature for 6 hours to allow for stabilization. After 48 hours of continued incubation, the knockdown efficiency of the siRNA targeting the gene could be assessed. The sequences of the siRNAs used were as follows: siRNA-1 (sense: 5'-CGGAUGGAAUCUCUAGCUAUUTT-3', antisense: 5'-AAUAGCUAGAGAUUC CAUCCGTT-3'), siRNA-2 (sense: 5'-CAGGGUAAUAACGAAAAAGCATT-3', antisense: 5'-UGCUUUUUCG UUAUUACCCUGTT – 3') (Sangon Biotech, shanghai, China).

Quantitative real-time PCR (qRT-PCR)

Total RNA was extracted from cells using TRIzol (Invitrogen, Carlsbad, CA), followed by reverse transcription into cDNA using the PrimeScript TM RT Kit (TaKaRa, Japan). qRT-PCR was performed with SYBR Green (Thermo Fisher Scientific, USA) in a CFX96 Touch™ Real-Time qPCR Detection System (Bio-Rad Laboratories, USA). GAPDH was utilized as experimental reference. The specific primers for the target gene as listed below: GAPDH (sense: 5'-GGAAGCTTGTCATCAATGGAAATC-3', antisense: 5'-TGATGACCCTTTTGGCTCCC-3'), GGT6 (sense: 5'-CTTTGTGTCTGGCGGTGGTA-3', antisense: 5'-AGTTCTGGACCTCAAAGCCC-3') (Sangon Biotech, shanghai, China).

Cell proliferation assay

The 5-ethynyl-2'-deoxyuridine (EdU) assay (RiboBio, Nanjing, China) was employed to assess the proliferative capacity of cells. Following transfection, cells were seeded into 24-well plates and treated with 50μM EdU working solution for 2 h. Subsequently, the cells were fixed, washed, and permeabilized. The cells were then incubated at room temperature in the dark for 30 min for EdU-positive cell staining. Then, Nuclear staining of the cells was performed. Finally, cells were observed and captured under an inverted fluorescence microscope, and the proliferation of cells was evaluated by calculating the ratio of EdU-positive cells (green cells) to the total number of DAPI-positive cells (blue cells).

Transwell migration and invasion assays

After transfection for 48 h, cells were seeded at a density of 5×10^4 cells/well in 8.0 Corning™ 24-well transwell plates (Corning, USA) for the assessment of cell migration and invasion capabilities. Before seeding, Matrigel (Corning, USA) was diluted in serum-free culture medium at a 1:9 ratio, and 50 μl of the dilution was added to each transwell chamber for invasion assays, while migration assays did not require this step. Meanwhile, 500 μL RPMI 1640 medium with 20% fetal bovine serum was added to the lower chamber. After incubation at 37 °C 5% CO_2 for 24 h, the cells on the bottom surface of the chamber were fixed with 4% paraformaldehyde for 20 min and with 0.1% crystal violet for 20 min. Under a high-magnification light microscope, cells in five randomly selected areas were counted.

Statistical analysis

Statistical analyses and graphical visualization were performed with R 4.2.0 and GraphPad Prism 9.0. Statistical significance was determined for comparisons between two groups with the Wilcoxon test. For comparisons of more than two groups, the ANOVA test was applied. The Benjamini-Hochberg test was used to control the false discovery rate and adjust P value. Spearman rank correlation analysis was used to determine the correlation between two variables. KM survival analysis was performed with the Logrank test. Two-sided Student t test was used to evaluate the C-index of MRPS compared to other signatures. The data are presented as mean ± SD and P value < 0. 05 was considered statistically significant. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$. ns, no significance.

Results

Identification of MRGs by scRNA-seq

As the first step to uncover the organization of cellular hierarchies in ccRCC, we collected the scRNA-seq profiles covering 26 ccRCC samples. After filtering, 81,335 qualified cells were used for subsequent analysis (Supplementary Figure S1A, B). PCA was performed, indicating main directions of cellular changes (Supplementary Figure S1C). With batch calibration from "Harmony" method, the last Harmony dimension where change of variation was more than 0.05% was identified (Supplementary Figure S1D). Based on marker genes (Supplementary Figure S1E, Supplementary Table S1), 33 cell clusters were identified using the UMAP algorithm and notably classified as malignant cells, endothelial cells, fibroblasts, macrophages, monocytes, dendritic cells, plasmacytoid dendritic cells (pDCs), NK cells, T cells, B cells or plasma cells, mast cells (Fig. 1A; Supplementary Figure S1F). The distribution of cell types in 26 ccRCC samples and clusters was shown in Supplementary Figure S1G. Considering the stability of CNV in immune cells, we chose NK cells as a control group. CNV analysis showed that there was an obvious copy number variation in malignant cells (Supplementary Figure S3A). Using the "FindAllMarkers" function and Wilcoxon test, significant DEGs of each cell type were identified (Fig. 1B). DEGs were particularly enriched in hypoxia-inducible factor 1 (HIF-1) signaling pathway and several metabolism-related pathways, according to functional analysis based KEGG database (Supplementary Table S2). These metabolism-related pathways included glycolysis/gluconeogenesis, biosynthesis of amino acids, glycine, serine and threonine metabolism, glutathione metabolism, carbon metabolism, arginine and proline metabolism and drug metabolism (Fig. 1C), which indicated that cell metabolism reprogramming should be considered as an important explanation for the malignancy of ccRCC. Consequently, we collected genes from KEGG metabolism-related pathways to generate a set of genes named META_process and integrated all genes in these pathways as MRGs. GSEA was used to examine if metabolism reprogramming was dynamic in malignant cells, and the results indicated that most genes were inclined to be active (adjust $P < 0.0001$) (Fig. 1D). AUCell analysis was used to determine the quantitative metabolic activity of each cell. This approach was successful in identifying cells that had the META_process gene set by assessing thresholds. Cells that had AUCell values higher above the cutoff point of 0.12 were classified as "META_active" cells; all other cells were classified as "META_inactive". (Fig. 1E). The distribution of AUCell values was significantly elevated in malignant clusters, whereas it was inferior in other clusters (Fig. 1F). Compared to other immune and stromal cells, ccRCC cells exhibited more metabolically active (Supplementary Figure S2, Supplementary Table S3).
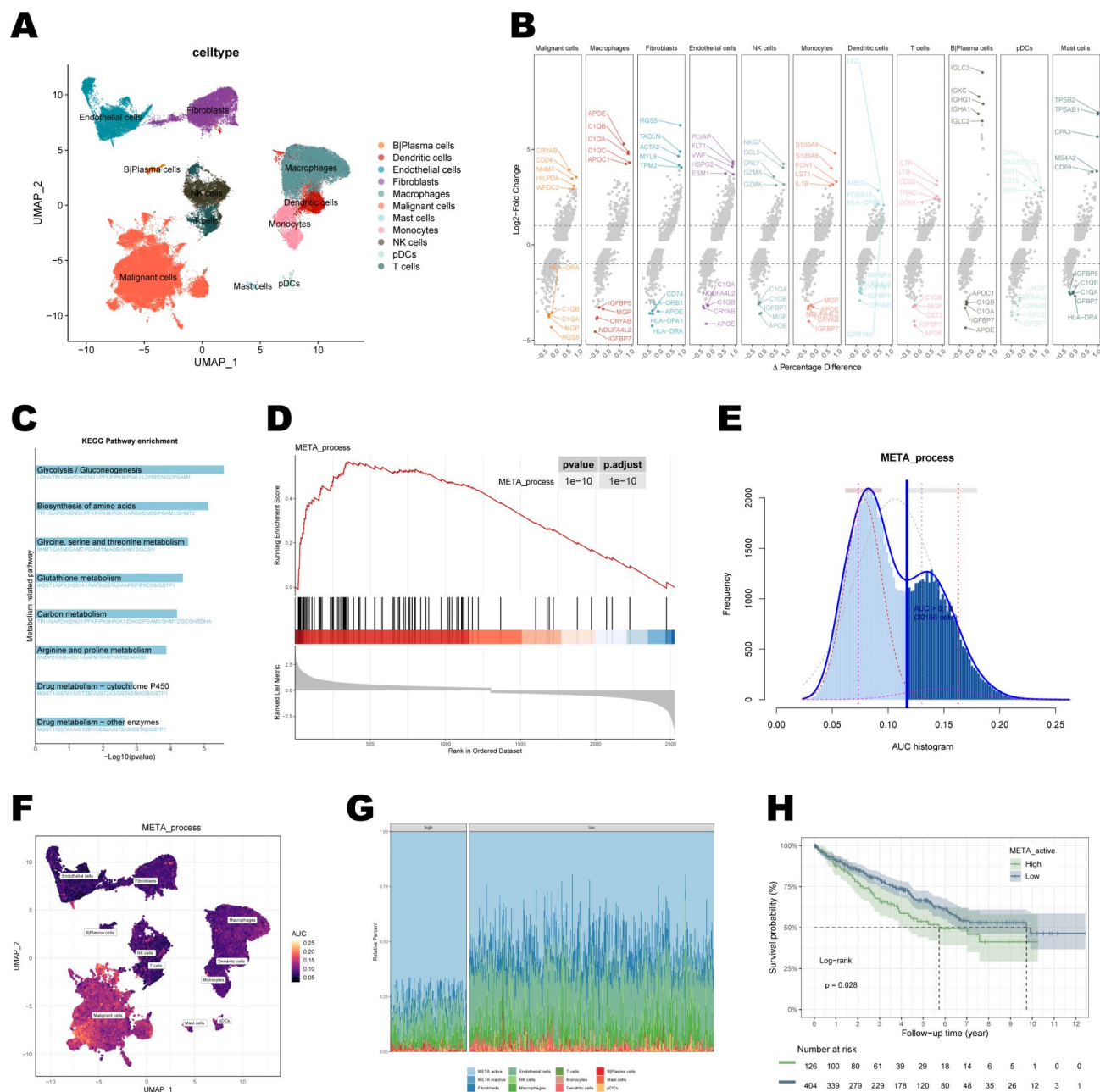
Fig. 1. Certain cell subsets of ccRCC correlated with metabolic passway and survival. (**A**) UMAP showing cells from all scRNA-seq samples. (**B**) Volcano plot of DEGs between different cell types. The top 5 genes with upregulated and downregulated were displayed. (**C**) KEGG pathway enrichment analysis of DEGs in malignant cells. (**D**) GSEA analysis of META_process gene set in DEGs. (**E**) AUCell values of all cells from ccRCC samples. Cells that had AUCell values of more than 0.12 were classified as META_active cells. (**F**) UMAP showing all cells based on the AUCell values. (**G**) Relative abundance of different cell proportions by gene expression deconvolution in each sample from high and low META_active cell composition. (**H**) KM analysis for survival probability of the two META_active cell proportion groups in TCGA-KIRC cohort.

ScRNA-seq deconvolution of META_active cells and its pseudotime and cell–cell communication analysis

The association between these redefining ccRCC cells proportion and their organizational structures with prognosis was then attempted to be clarified. After cleaning up the scRNA-seq reference and bulk mixture (Supplementary Figure S3B), the concordance of gene expression for different types of genes was checked, where protein coding genes were assigned as marker genes (Supplementary Figure S3C). Expression profile of META_active, META_inactive and other cells from scRNA-seq data enabled us to perform BayesPrism[29] to investigate the variability of metabolic characteristics among patient samples and their correlation with the molecular and clinical information of ccRCC in large-scale clinical datasets (Fig. 1G). Significant prognostic differences between

the high and low groups of META_active cell proportion were found based on the segmentation produced by the optimal cut-off analysis ($P < 0.05$, Fig. 1G, H; Supplementary Figure S3D); however, these differences were not maintained in median segmentation (Supplementary Figure S3E). Based on "Monocle2"[30] method, pseudotime analysis on malignant subtype cells was performed to explore their differentiation direction. The results demonstrated that malignant ccRCC cells eventually followed 2 different trajectories (Fig. 2A). Using "CytoTRACE" method, we confirmed that META_active cells (left) might originate from META_inactive ones (right) (Fig. 2B; Supplementary Figure S3F). The distribution of CytoTRACE score of different malignant subtype cells was examined by the Wilcoxon test (Fig. 2C) ($P < 0.0001$). Additionally, based on specific pathways and ligand receptors, we assumed cell-cell communication networks to predict intercellular communication. In particular, there is a higher frequency and intensity of interaction between the META_active subtype of cells and other immune and stromal cells (Fig. 2D). The heatmap shows that in the MIF and CD70 signaling pathways, META_active cells cellular communication occurred more frequently than META_inactive cells (Fig. 2E, Supplementary Figure S4A). Furthermore, in all the pathways as well as in the specific pathways, META_active cells were considerably more significant than META_inactive cells (Supplementary Figure S4B, C).

Construction and validation of MRPS based on machine learning

To investigate the stable biomarkers behind the observed association of the proportion of META_active cells and clinical outcomes of ccRCC,17 co-detected DEMRGs were evaluated after crossing 343 MRGs and DEGs between tumor and normal samples (Supplementary Figure S3G, H). Subsequently, we fitted 118 prediction models using 10 different machine learning algorithms, as described in method. The evaluation was conducted on the TCGA training cohort, with 3 cohorts (TCGA testing, E-MTAB-1980 and Meta) serving as external validation (Fig. 3A). It was obvious that the combination of RSF algorithm and GBM algorithm showed the ranked first mean (0.715) C-index and was selected as the optimal MRPS. In the RSF algorithm, when the number
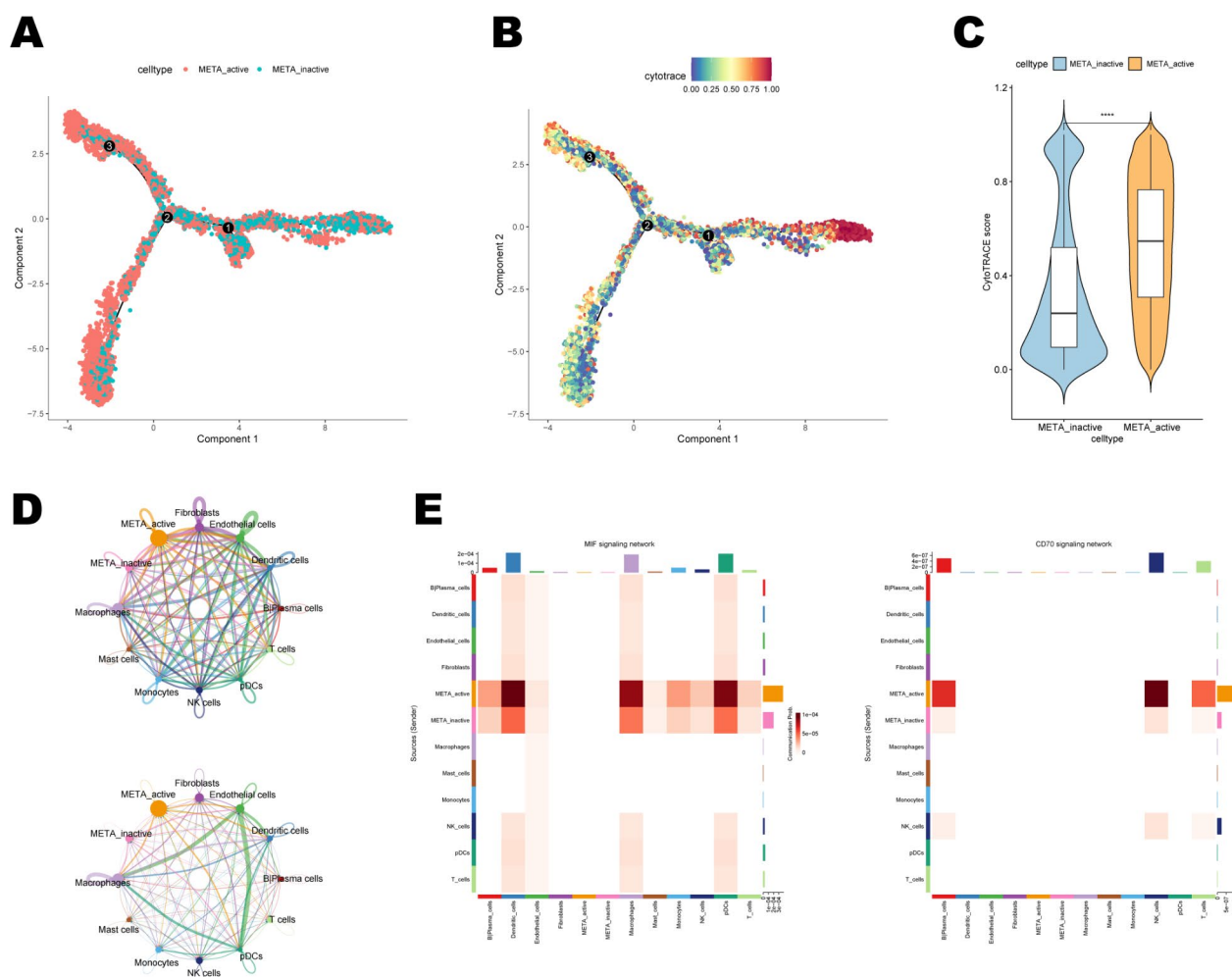

Fig. 2. Pseudotime and cell-cell communication analyses. (**A**, **B**) Combined application of CytoTRACE and Monocle2 to predict trajectories of malignant cells with different META_process states. (**C**) Volinplot showing the comparison of CytoTRACE score between different malignant subsets. (**D**) Number and strength of interactions between all cells types. (**E**) Heatmap showing the inferred intercellular communication network of MIF and CD70 signaling.
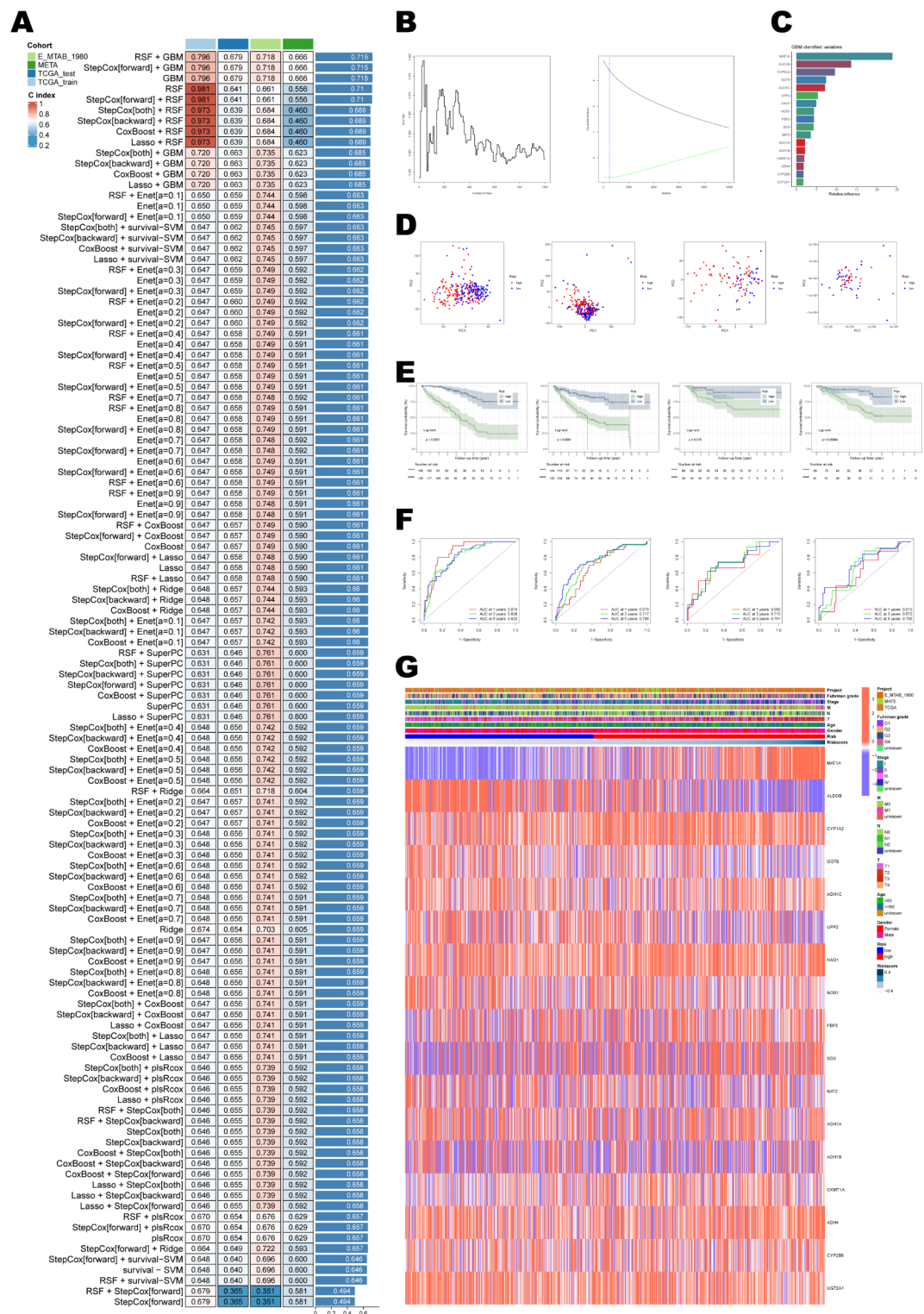
Fig. 3. Construction and validation of MRPS based on machine learning. (**A**) C-index of 118 prediction models using 10 machine learning algorithms across four cohorts. (**B**) Error rate and survival tree curve based on SVM (left). Error distribution in train (black) and validation (yellow) based on GBM (right). (**C**) Relative influence of 17 MRGs using the GBM algorithm. (**D**) PCA showing the distribution of low- and high-risk subgroups. (**E**) KM curves showing OS in the low- and high-risk subgroups. (**F**) ROC curves showing 1-, 3-, and 5-year OS rates in the low- and high-risk subgroups. (**G**) Heatmap of MRGs and clinical characteristics.

of surviving trees increased to 1000, the error ratio curve still did not tend to be completely stable (Fig. 3B), while in GBM regression, the optimal tree (iteration) was obtained when the Cox error reached its minimum value, which may be the reason for the overfitting phenomenon of RSF models. Next, we calculated risk score for each patient using GBM algorithm (Fig. 3C; Supplementary Table S4), and categorized them into either high- and low-risk groups according to the median risk score derived from TCGA training cohort (Supplementary Table S5-8).Two distribution tendencies were shown by PCA (Fig. 3D), and patients in the high-risk subgroup had substantially poorer outcome than those in the low-risk group TCGA training cohort and validation cohorts cohort (all $P < 0.05$), as shown in Fig. 3E. Consistently, the areas under the ROC curve (AUCs) for 1-, 3-, and 5-year OS rates were greater than 0.6 (Fig. 3F), in the TCGA training cohort (0.874, 0.808, and 0.823); TCGA testing cohort (0.690, 0.715, and 0.701); Meta cohort (0.613, 0.672, and 0.705). Heatmap represented clinical information, MRPS risk score, and DEMRGs expression level in all cohorts (Fig. 3G).

Relationship between MRPS and clinical characteristics

To investigate the relationship between the prognostic relevance of MRPS and clinical features in different cohorts, we evaluated MRPS risk scores across several categorized characteristics. In all the cohorts, high MRPS risk scores were shown to be more strongly correlated with unfavorable AJCC tumor stage, T stage, M stage, and Fuhrman grade, while no significant association was found with age (Fig. 4A; Supplementary Figure S5A, B). With univariate and multivariate Cox analyses, it was found that MRPS risk scores could serve as an independent prognostic factor for all samples in our collected cohort (Fig. 4B). Based on all the results above-mentioned, we meticulously developed a dynamic nomogram that combines the MRPS risk scores with clinical characteristics such as gender, T stage, N stage and M stage to predict 1-, 3-, and 5-year OS rates (Fig. 4C). The nomogram's accuracy in predicting 1-, 3-, and 5-year outcomes was validated by calibration curve (Fig. 4D). The higher predictive ability of the nomogram over the MRPS risk score and other clinical characteristics was further supported by the time-dependent C-index curve (Fig. 4E).

Comparison of MRPS with 51 previously published signatures in ccRCC

With the rapid development of high-throughput sequencing and big data technology, increasing numbers of machine learning-based signatures for predicting the survival probability of ccRCC samples have been established. To compare the performance of MRPS with other signatures, we collected and registered a total of 51 published signatures (Supplementary Table S9). Only mRNA signatures were selected for following analysis owing to the lack of sequencing information in GSE167573 datasets annotated by GPL20795. These signatures were associated with various biological processes, such as neutrophil extracellular trap, immune response, programmed cell death, mitophagy, hypoxia, methylation and glycolysis. Remarkably, MRPS exhibited the highest C-index among the remaining cohorts (TCGA training, Meta, and Combined all cohort), reduced dimensionally by two signatures in TCGA testing and E-MTAB-1980 (Fig. 5A). We observed that the majority of models showed strong performance in certain datasets but performed weakly in others, which might be because the models developed by overfitting and had poor generalization capabilities. These findings highlighted MRPS' excellent prediction ability and potential for extension.

MRPS related TIM infiltration landscape and potential therapeutic agents

Based on the 7 algorithms from "IOBR", we investigated the correlation between MRPS and TIM components at the bulk RNA-seq level. The heatmap shows the discrepancy in immune and stromal cell profiles calculated in high and low MRPS risk subgroups, through different algorithms (Fig. 6A). Further, in an evaluation of 22 immune cell landscapes through CIBERSORT, we found that there were no significant differences in most B and T cell subtypes, except for T follicular helper cells (Tfh), Treg and memory B cells (Fig. 6B). Moreover, a higher MRPS risk score was observed to be associated with a higher level of immune suppressive cells (Treg, TAM, etc.) (all $P < 0.05$, Fig. 6C), which indicated an immunosuppressive phenotype might exist in high-risk samples. Therefore, we gathered 181 advanced ccRCC samples that had received nivolumab (anti-PD-1) treatment to examine if there was a relationship between the signature and immunotherapy. KM survival analysis showed prolonged survival in low MRPS risk samples ($P < 0.05$, Fig. 6D). These collective results suggested that the high-risk subgroup identified by MRPS might not benefit from immunotherapy.

The IC50 values of chemotherapeutic and targeted drug of each sample were then compared based on the "oncoPredict" package to potential therapeutic agents for different MRPS risk subgroups. From the GDSC1, GDSC2, and CTRP2, we found that 7 drugs were negatively associated with risk scores and 2 were positively associated ($|R| > 0.4$ and $P < 0.05$, Fig. 7A). Specifically, patients in the high-risk subgroup exhibited higher sensitivity to BRD-K99584050, sepantronium bromide, SB216763, VER-155,008, niclosamide, pifithrin-alpha and NSC23766, while patients in the low-risk group showed greater sensitivity to myricetin and SGX-523 (Fig. 7B). In conclusion, these drugs might be potential treatments for ccRCC samples.

Knockdown of GGT6 promoted malignant proliferation and invasion of ccRCC cells

Many researches have proven the predictive role of MAT1A[38], ALDOB[39], and CYP1A2[40] in cancer, however, GGT6, another key gene in the signature, has yet to be explored. Downregulation expression levels of GGT6 in tumor tissues relative to paired normal tissues ($P < 0.0001$, Fig. 8A), and samples exhibiting high GGT6 expression levels displayed a significantly increased OS compared to those with low expression ($P < 0.05$, Fig. 8B). Initially, GGT6 knockdown cell lines were established to explore its impact on the proliferation and invasion of ccRCC cells. Two siRNA were used to silence GGT6, and the knockdown efficiency was validated by RT-qPCR ($P < 0.05$, Fig. 8C). Subsequent EdU assay results indicated that GGT6 knockdown significantly promoted the proliferation of renal cancer cells ($P < 0.05$, Fig. 8D). Furthermore, the results of the transwell assay also suggested that the migration and invasion capabilities of renal cancer cells were significantly enhanced after
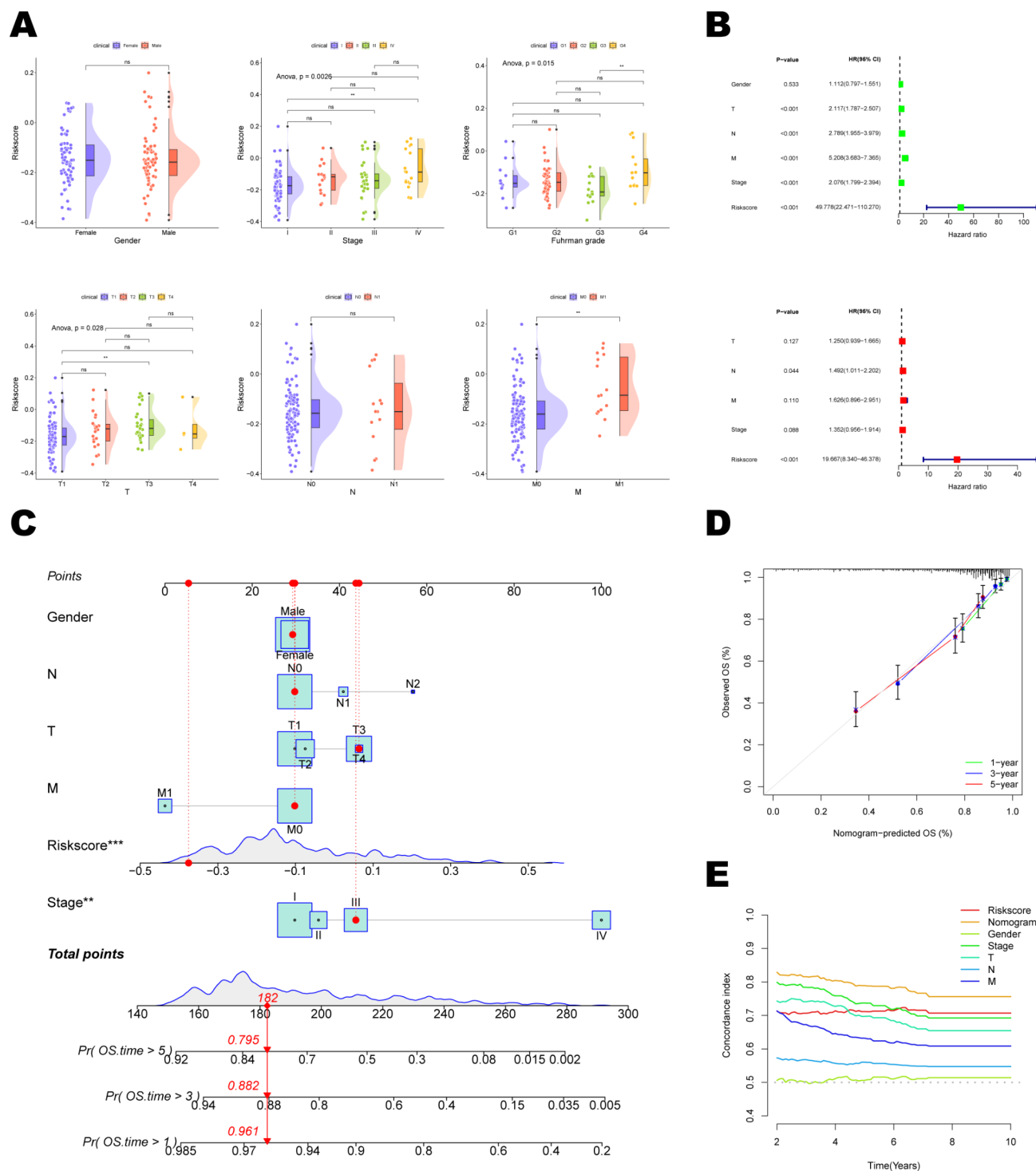
Fig. 4. Evaluation of MRPS with clinical characteristics and establishment of dynamic nomogram. (**A**) Relationship between gender, AJCC tumor stage, T stage, N stage, M stage, and Fuhrman grade with MRPS risk score in Meta cohort. (**B**) Univariate and multivariate Cox analysis of MRPS risk score and clinical characteristics. (**C**) Construction of the nomogram model. Dynamic nomogram is available at https://urology-liuyx.shinyapps.io/MRPS/. (**D**) The calibration curve of the nomogram. (**E**) The time-dependent C-index curves for the nomogram and clinical characteristics.
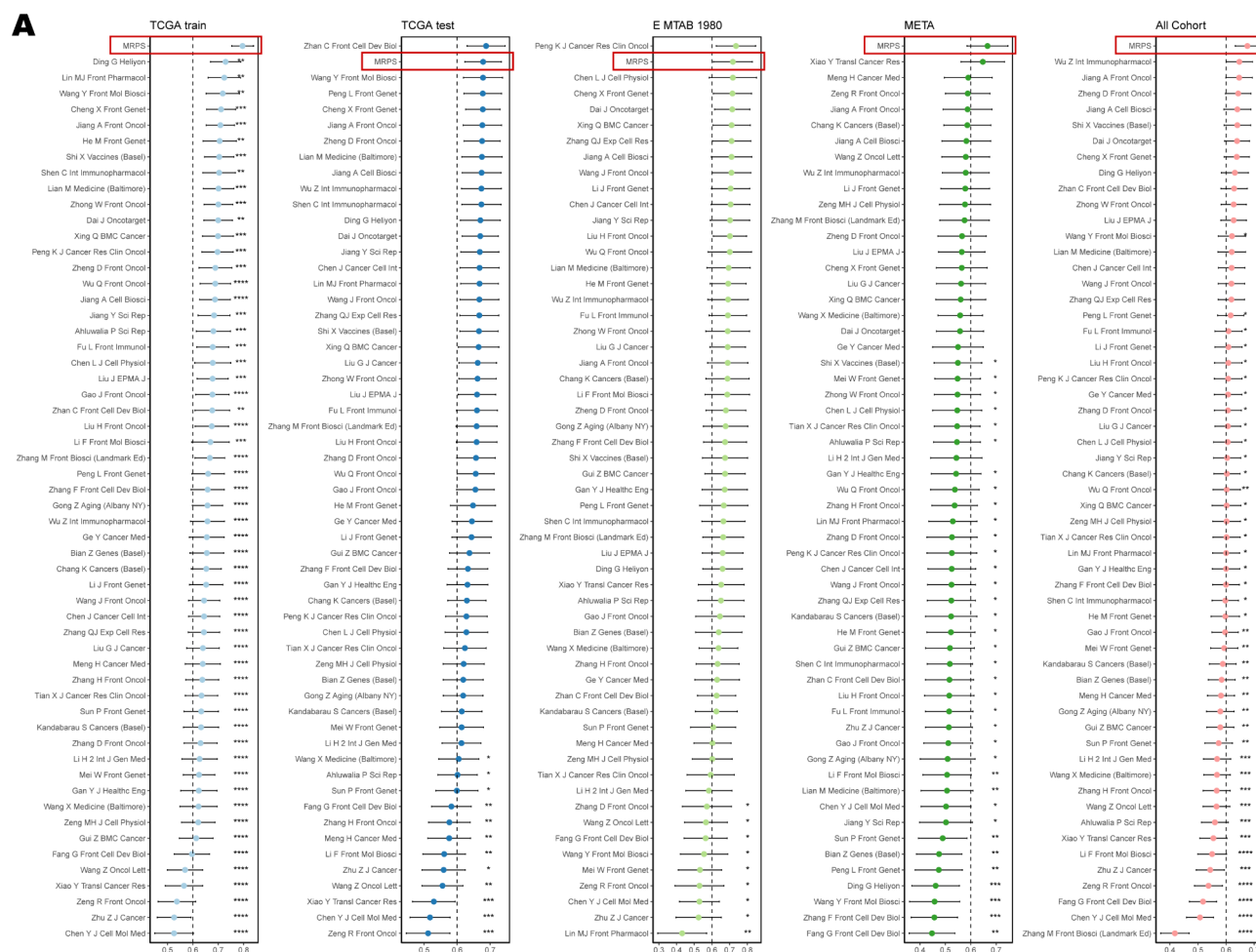
Fig. 5. Comparison of 51 polished gene expression-based prognostic signatures in ccRCC. (**A**) C-index analysis of MRPS and 51 published signatures in TCGA training (n=266), TCGA testing (n=264), E-MTAB-1980 (n=101), Meta (n=146) and all combined cohort (n=511).

GGT6 knockdown ($P < 0.05$, Fig. 8E, F). These findings collectively indicate that GGT6 might exert a suppressive effect on the occurrence and development of ccRCC.

Discussion

Cancer is characterized by unregulated cell development, which requires high energy levels to live in adverse environments such as low oxygen, dietary deficit, and chemotherapeutic drug intoxication[41–43]. Changes in metabolic pathways that control tumor energetics and biosynthesis, so-called metabolic reprogramming, is a prominent feature of ccRCC[44]. Inactivation of Von Hippel-Lindau (VHL) occurs in most sporadic ccRCC[44,45]. VHL increases proteasomal degradation of the transcription factor hypoxia-inducible factor (HIF) family, and its mutations activate pathways that reverse the effects of hypoxia in the normoxic environment, resulting in a condition of "pseudo-hypoxia"[44–46]. Such metabolic reprogramming, especially in Glycolysis and lipid metabolism, significantly influence ccRCC support, progression, and treatment choices through various biological processes[9,11,44,47]. This is further supported by our scRNA data, which showed that metabolism-related pathways were enriched in malignant cell populations along with HIF-1 pathway (Supplementary Table S2). The biosynthesis of glutathione (GSH) is equally important in ccRCC cells, as it is used to detoxify reactive oxygen species[48]. Consequently, as a center for several cellular signaling pathways, the adaptable and intricate metabolic network of ccRCC may offer novel targets for patient PET imaging, guided treatment, and enhanced prognosis[49,50].

Obesity might reflect people's metabolic status. However, the phenomenon of the "obesity paradox" existed among ccRCC patients[51,52]. It was thought that increased body mass should be responsible for the rise in ccRCC incidence[53–55], as obesity could trigger with metabolic abnormalities in the hypoxic microenvironment[56]. However, some studies suggested that higher BMI was associated with better survival[57,58].

Previous bulk RNA-seq based transcriptome analyses of ccRCC multi-patient samples focused mostly on prognostic gene screening and clinical outcome prediction, with less consideration given to intercellular heterogeneity (Supplementary Table S9). Meanwhile, scRNA-seq of ccRCC focuses on the research of cell origin[59], identification and function of new cell groups[60,61] and the heterogeneity of tumor cells[20]. However, it is
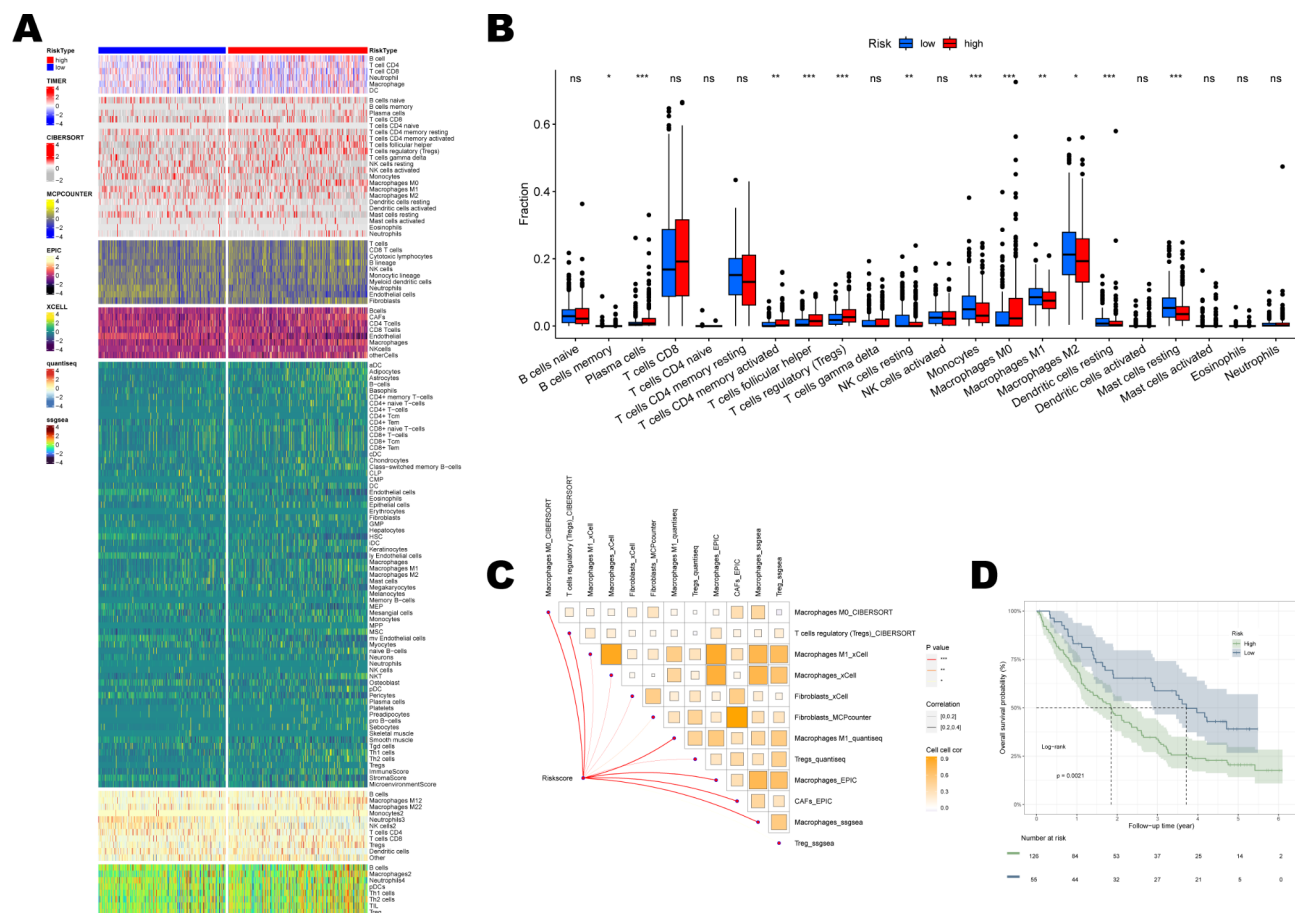
Fig. 6. Analysis of the tumor immune microenvironment in high- and low-risk subgroups. (**A**) 7 algorithms assess differences in immune cells infiltration between different risk subgroups. (**B**) Boxplot showing the relative fraction of infiltrating immune cells in patients between different risk subgroups. (**C**) Correlation analysis of MRPS risk score with immune suppressive cells estimated by 6 algorithms. (**D**) KM survival curve between high- and low-risk subgroups in CheckMate-009, CheckMate-010 and CheckMate-025 cohort.
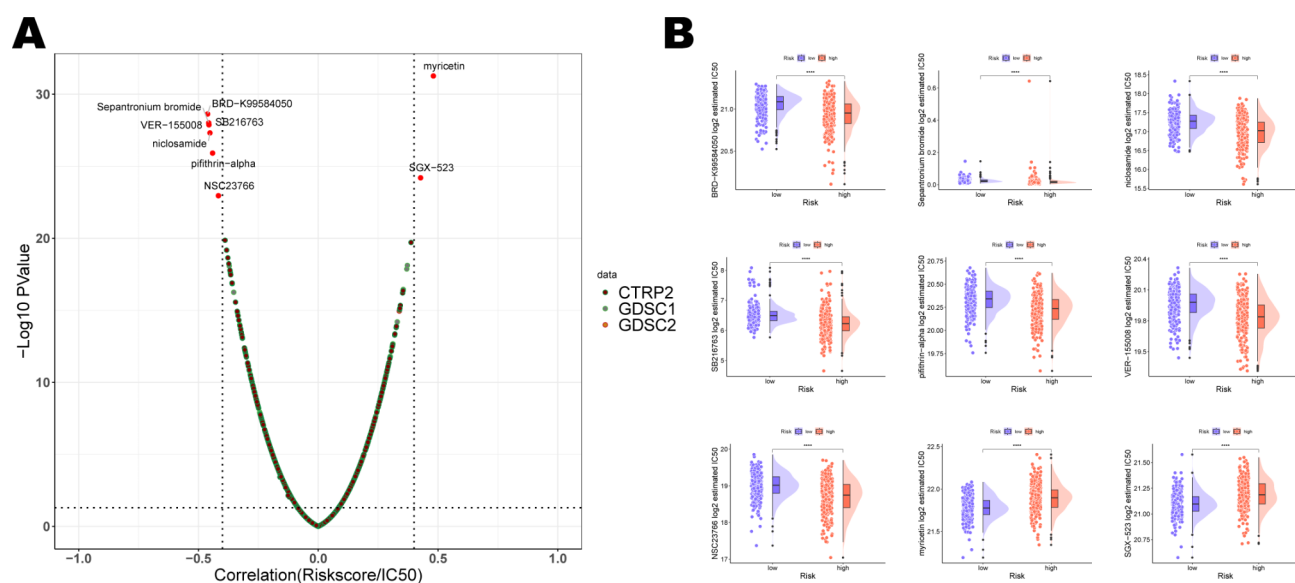
Fig. 7. Screening of therapeutic agents for ccRCC based on MRPS. (**A**) Correlation analysis of estimated IC50 values and MRPS risk score. (**B**) Boxplot showing the different distribution in estimated IC50 values of 9 candidate compounds between different risk subgroups.
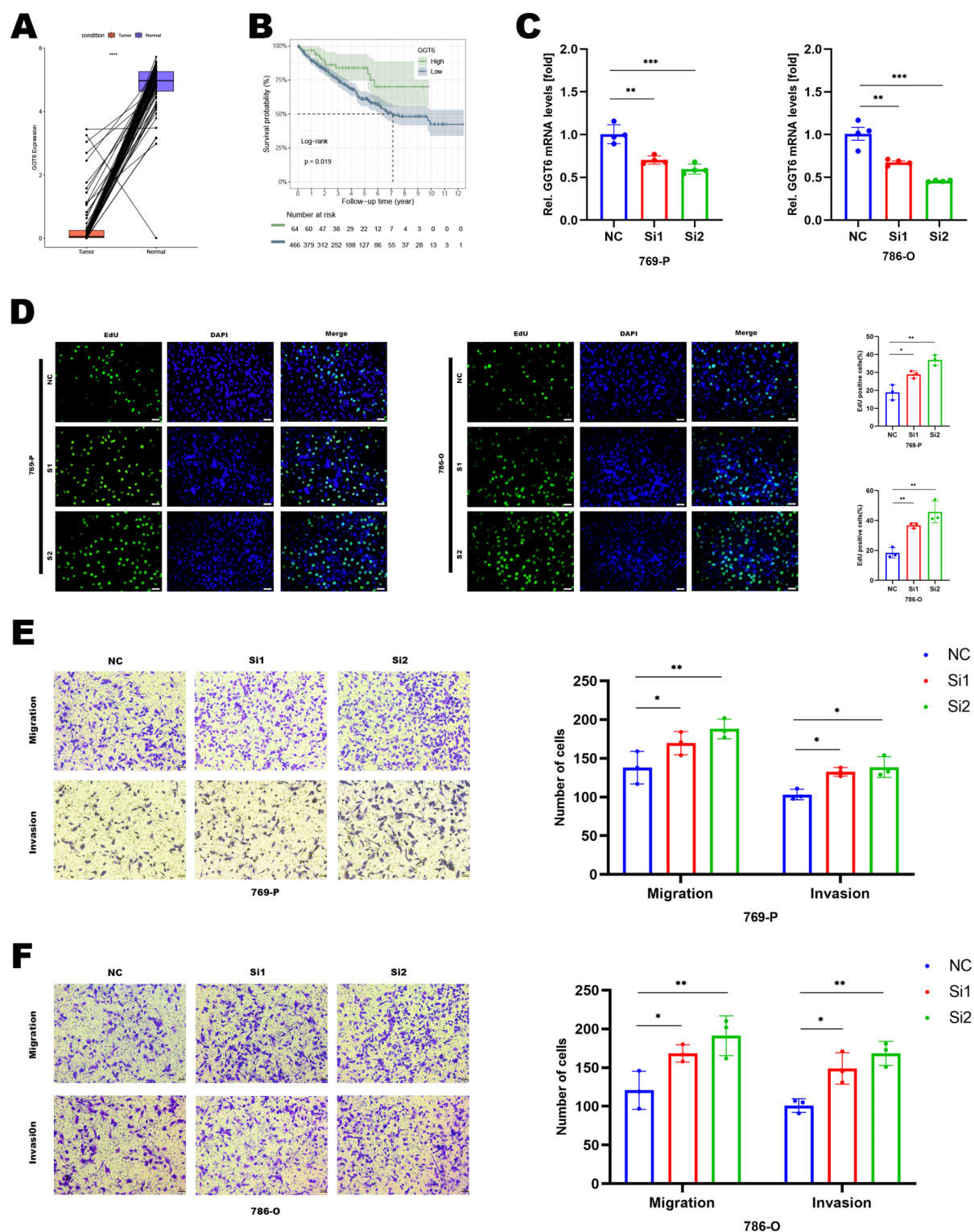
Fig. 8. Downregulation of GGT6 expression inhibits the proliferation, migration and invasion of ccRCC cells. (**A**) Boxplot showing the differential expression of GGT6 in paired tumor and normal tissues. (**B**) K-M survival curve analysis of high- and low- expression of GGT6 subgroup. (**C**) RT-qPCR results showed that siRNA silenced the mRNA expression of GGT6 in 769-P and 786-O. (**D**) Effect of GGT6 inhibition on ccRCC cell proliferation measured by EdU assay. (**E**, **F**) Effect of GGT6 inhibition on ccRCC cell migration and invasion ability was assessed by transwell assay.

challenging to correlate scRNA-seq analysis with prognosis because of the tiny sample size. Saout investigated the connection between certain ccRCC malignant cell subsets and patient prognosis via deconvolution approaches[21]. We employed a combination of scRNA-seq analysis and multi-patient sample sequencing to study the association between particular cell subsets and prognosis in ccRCC and categorized ccRCC patients based on unique cell metabolism reprogramming, further choosing therapeutic targets.

To learn more about the metabolic heterogeneity of malignant cells in ccRCC samples, we initially performed scRNA-seq analysis. By combining functional analysis with the AUCell method, we discovered a rise in the abundance of unique malignant cell subsets dubbed META_active cells, which demonstrate greater metabolic activity than other cells in the microenvironment (Fig. 1F; Supplementary Figure S2). Through the integration and analysis of transcriptome profiles of ccRCC samples from bulk RNA-seq and scRNA-seq, as well as patient-specific variations in cell composition and clinical outcomes from large cohorts, we have compiled the features of metabolism into a categorization system. According to the percentage of META_active malignant cells, samples were classified with high and low subgroups. The OS rate of the high META_active subgroup was worse in multiple cohorts, but the unstable performance suggested the need for more robust and simplified metabolism-related biomarkers to predict the prognosis of ccRCC patients(Fig. 1G, H; Supplementary Figure S3D, E). As opposed to META_inactive cells, META_active cells communicate with other cells in the microenvironment more often and exhibit specificity in the MIF and CD70 signaling pathways, which might be connected to the carcinogenesis and treatment resistance of ccRCC. Afterward, we combined MRGs with DEGs in tumor samples to create a stable and more clinically applicable predictive signature for evaluating the prognosis of ccRCC patients. We gathered 10 popular machine learning methods and created 118 models to overcome these constraints. MRPS, the best model for ccRCC prediction, may be regarded as the combination of RSF and GBM after thorough analysis. The 1, 3, and 5-year AUC values varied from 0.613 to 0.874 in the ROC curves, indicating that MRPS had a strong predictive effect on prognosis. It was also demonstrated that MRPS was an independent risk factor for the prognosis of ccRCC. The connection between MRPS and clinical characteristics was also examined. It appears that MRPS was not restricted to the predictive usefulness of OS, since the findings demonstrated a substantial correlation between risk scores and patients' distant metastases, tumor stage, and case grade. When it came to predicting the overall survival of patients with ccRCC, MRPS outperformed 51 published signatures and showed outstanding predictive ability across several cohorts. Thus, MRPS could be useful in determining the prognosis for ccRCC and categorizing patients into high or low-risk invasive subgroups.

The high-risk subgroup's increased cell abundance of invasive Treg and TAM resulted in a more inhibitory immunophenotype. Multiple algorithms were used in this study to infer immune cell infiltration in ccRCC, and all of them concluded that Treg and TAM were upregulated in high-risk subgroups. MIF interacted with CD74 + TAM, switching their activation status and forming an immunosuppressive/ngiogenic pro-tumor phenotype[62,63]. The interaction between CD70 and CD27 caused increased apoptosis and an exhausted state in T cells, thereby impeding anti-tumor immunity[64]. Therefore, we examined the predictive value of MRPS in immunotherapy in this study. Low-risk samples' OS was much longer than that of high-risk samples' ($P = 0.021$). These findings imply that MRPS could offer a new perspective on ICI responses against ccRCC.

Potential therapeutic drugs for ccRCC patients were identified via established MRPS. Among them, clinical trials have been conducted on cancer patients to confirm the effects of sepantronium bromide[65] and niclosamide[66]. Because of SGX-523's unanticipated renal toxicity[67], clinical development would no longer be conducted. Other medications' effectiveness and confounding have not yet been tested in clinical trials in cancer patients, and as technology advances, whether they can become the next generation of anti-tumor therapies in ccRCC patients must also be determined.

Finally, we looked at the possible physiological role that GGT6 members of the GGT family may have in the development and incidence of ccRCC. GGT6 knockdown enhanced the ability of ccRCC cells to proliferate and invade, indicating the viability of GGT6 as a ccRCC therapeutic target. We intend to further investigate its molecular mechanism in future studies.

Nevertheless, our study has several limitations. First, our findings are based on retrospective sequencing data and bioinformatics analysis; multicenter and prospective clinical studies are required to evaluate the robustness of MRPS and its relationship with obesity status. Second, more comprehensive experimental research is needed to understand the precise biological role of the genes in MRPS.

Conclusions

In conclusion, we created a stable and reliable prognostic signature of metabolic reprogramming including 17 genes based on a significant number of bioinformatics and machine learning programs. The signature's predictive power for clinical characteristics, TIM and immunotherapy was also assessed. This MRPS model is a potential tool to guide clinical judgment and personalized therapy of ccRCC patients.

Data availability

The datasets analysed during the current study are available in UCSC Xena (https://xenabrowser.net/datapages/), GEO (https://www.ncbi.nlm.nih/), ArrayExpress (https://www.ebi.ac.uk/biostudies/arrayexpress) and ICGC (https://dcc.icgc.org/), with accession numbers: TCGA-KIRC, GSE167573, GSE152938, GSE207493, GSE224630, E-MTAB-1980, and RECA-EU. No new algorithms were developed for this article. All code and data generated for analysis are available from the corresponding author upon reasonable request.

References

1. Patel, H. D. et al. Clinical stage migration and survival for renal cell carcinoma in the United States. *Eur. Urol. Oncol.* **2**(4), 343–348 (2019).
2. Noguchi, G. et al. Time-dependent change in relapse sites of renal cell carcinoma after curative surgery. *Clin. Exp. Metastasis* **35**(1–2), 69–75 (2018).
3. Rathmell, W. K. et al. Management of metastatic clear cell renal cell carcinoma: ASCO guideline. *J. Clin. Oncol.* **40**(25), 2957–2995 (2022).
4. Climent, C. et al. The role of immunotherapy in non-clear cell renal cell carcinoma. *Front. Oncol.* **13**, 941835 (2023).
5. Au, L. et al. Determinants of anti-PD-1 response and resistance in clear cell renal cell carcinoma. *Cancer Cell* **39**(11), 1497–1518 (2021). e11.
6. Lu, X. et al. Silencing of genes by promoter hypermethylation shapes tumor microenvironment and resistance to immunotherapy in clear-cell renal cell carcinomas. *Cell. Rep. Med.* **4**(11), 101287 (2023).
7. Li, M. et al. Monocarboxylate transporter 1 promotes proliferation and invasion of renal cancer cells by mediating acetate transport. *Cell. Biol. Int.* **45**(6), 1278–1287 (2021).
8. Ma, Y. et al. SIRT5-mediated SDHA desuccinylation promotes clear cell renal cell carcinoma tumorigenesis. *Free Radic Biol. Med.* **134**, 458–467 (2019).
9. Chen, C. et al. AUP1 regulates lipid metabolism and induces lipid accumulation to accelerate the progression of renal clear cell carcinoma. *Cancer Sci.* **113**(8), 2600–2615 (2022).
10. Yang, J. et al. A positive Feedback Loop between inactive VHL-Triggered histone lactylation and PDGFRbeta signaling drives clear cell renal cell carcinoma progression. *Int. J. Biol. Sci.* **18**(8), 3470–3483 (2022).
11. Siska, P. J. et al. Mitochondrial dysregulation and glycolytic insufficiency functionally impair CD8 T cells infiltrating human renal cell carcinoma. *JCI Insight* **2**(12) (2017).
12. Ye, J. et al. Upregulation of serine metabolism enzyme PSAT1 predicts poor prognosis and promotes proliferation, metastasis and drug resistance of clear cell renal cell carcinoma. *Exp. Cell. Res.* **437**(1), 113977 (2024).
13. Thind, A. S. et al. Demystifying emerging bulk RNA-Seq applications: the application and utility of bioinformatic methodology. *Brief. Bioinform.* **22**(6). (2021).
14. Papalexi, E. & Satija, R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat. Rev. Immunol.* **18**(1), 35–45 (2018).
15. Sun, G. et al. Integrated exome and RNA sequencing of TFE3-translocation renal cell carcinoma. *Nat. Commun.* **12**(1), 5262 (2021).
16. Cancer Genome Atlas Research N. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **499**(7456), 43–49 (2013).
17. Sato, Y. et al. Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat. Genet.* **45**(8), 860–867 (2013).
18. Zhang, J. et al. The international cancer genome consortium data portal. *Nat. Biotechnol.* **37**(4), 367–369 (2019).
19. Su, C. et al. Single-cell RNA sequencing in multiple pathologic types of renal cell carcinoma revealed novel potential tumor-specific markers. *Front. Oncol.* **11**, 719564 (2021).
20. Yu, Z. et al. Integrative single-cell analysis reveals transcriptional and epigenetic regulatory features of clear cell renal cell carcinoma. *Cancer Res.* **83**(5), 700–719 (2023).
21. Saout, J. R. et al. Single-cell deconvolution of a specific malignant cell population as a poor prognostic biomarker in low-risk clear cell renal cell carcinoma patients. *Eur. Urol.* **83**(5), 441–451 (2023).
22. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**(13), 3573–87e29 (2021).
23. Choudhary, S. & Satija, R. Comparison and evaluation of statistical error models for scRNA-seq. *Genome Biol.* **23**(1), 27 (2022).
24. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with harmony. *Nat. Methods.* **16**(12), 1289–1296 (2019).
25. Hu, C. et al. CellMarker 2.0: An updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. *Nucleic Acids Res.* **51**(D1), D870–D6 (2023).
26. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), 27–30 (2000).
27. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**(D1), D457–D462 (2016).
28. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**(5), 284–287 (2012).
29. Chu, T., Wang, Z., Pe'er, D. & Danko, C. G. Cell type and gene expression deconvolution with BayesPrism enables bayesian integrative analysis across bulk and single-cell RNA sequencing in oncology. *Nat. Cancer* **3**(4), 505–517 (2022).
30. Qiu, X. et al. Single-cell mRNA quantification and differential analysis with Census. *Nat. Methods* **14**(3), 309–315 (2017).
31. Gulati, G. S. et al. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* **367**(6476), 405–411 (2020).
32. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**(1), 1088 (2021).
33. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**(6), 882–883 (2012).
34. Liu, Z. et al. Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer. *Nat. Commun.* **13**(1), 816 (2022).
35. Zeng, D. et al. IOBR: Multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. *Front. Immunol.* **12**, 687975 (2021).
36. Braun, D. A. et al. Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nat. Med.* **26**(6), 909–918 (2020).
37. Maeser, D., Gruener, R. F. & Huang, R. S. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Brief. Bioinform.* **22**(6) (2021).
38. Frau, M., Feo, F. & Pascale, R. M. Pleiotropic effects of methionine adenosyltransferases deregulation as determinants of liver cancer progression and prognosis. *J. Hepatol.* **59**(4), 830–841 (2013).
39. Bu, P. et al. Aldolase B-mediated fructose metabolism drives metabolic reprogramming of colon cancer liver metastasis. *Cell. Metab.* **27**(6), 1249–1262 (2018).
40. Dai, Z. et al. CYP1A inhibitors: Recent progress, current challenges, and future perspectives. *Med. Res. Rev.* **44**(1), 169–234 (2024).
41. Schulze, A. & Harris, A. L. How cancer metabolism is tuned for proliferation and vulnerable to disruption. *Nature* **491**(7424), 364–373 (2012).
42. Pavlova, N. N. & Thompson, C. B. The emerging hallmarks of cancer metabolism. *Cell. Metab.* **23**(1), 27–47 (2016).
43. Thakur, C. & Chen, F. Connections between metabolism and epigenetics in cancers. *Semin Cancer Biol.* **57**, 52–58 (2019).
44. Wettersten, H. I., Aboud, O. A., Lara, P. N. Jr. & Weiss, R. H. Metabolic reprogramming in clear cell renal cell carcinoma. *Nat. Rev. Nephrol.* **13**(7), 410–419 (2017).
45. Gossage, L., Eisen, T. & Maher, E. R. VHL, the story of a tumour suppressor gene. *Nat. Rev. Cancer.* **15**(1), 55–64 (2015).
46. Zhu, H., Wang, X., Lu, S. & Ou, K. Metabolic reprogramming of clear cell renal cell carcinoma. *Front. Endocrinol. (Lausanne)* **14**, 1195500 (2023).
47. Bacigalupa, Z. A. & Rathmell, W. K. Beyond glycolysis: Hypoxia signaling as a master regulator of alternative metabolic pathways and the implications in clear cell renal cell carcinoma. *Cancer Lett.* **489**, 19–28 (2020).

48. Miess, H. et al. The glutathione redox system is essential to prevent ferroptosis caused by impaired lipid metabolism in clear cell renal cell carcinoma. *Oncogene* **37**(40), 5435–5450 (2018).
49. Abu Aboud, O. et al. Glutamine addiction in kidney Cancer suppresses oxidative stress and can be exploited for real-time imaging. *Cancer Res.* **77**(23), 6746–6758 (2017).
50. Stine, Z. E., Schug, Z. T., Salvino, J. M. & Dang, C. V. Targeting cancer metabolism in the era of precision oncology. *Nat. Rev. Drug Discov.* **21**(2), 141–162 (2022).
51. Tan, S. K. et al. Obesity-dependent adipokine chemerin suppresses fatty acid oxidation to confer ferroptosis resistance. *Cancer Discov.* **11**(8), 2072–2093 (2021).
52. Aurilio, G. et al. The role of obesity in renal cell carcinoma patients: Clinical-pathological implications. *Int. J. Mol. Sci.* **20**(22) (2019).
53. van de Pol, J. A. A., George, L., van den Brandt, P. A., Baldewijns, M. & Schouten, L. J. Etiologic heterogeneity of clear-cell and papillary renal cell carcinoma in the Netherlands cohort study. *Int. J. Cancer* **148**(1), 67–76 (2021).
54. Wang, H. K. et al. Visceral fat accumulation is associated with different pathological subtypes of renal cell carcinoma (RCC): A multicentre study in China. *BJU Int.* **114**(4), 496–502 (2014).
55. McGuire, B. B. & Fitzpatrick, J. M. BMI and the risk of renal cell carcinoma. *Curr. Opin. Urol.* **21**(5), 356–361 (2011).
56. Ringel, A. E. et al. Obesity shapes metabolism in the tumor microenvironment to suppress anti-tumor immunity. *Cell* **183**(7), 1848–1866 (2020). e26.
57. Sanchez, A. et al. Transcriptomic signatures related to the obesity paradox in patients with clear cell renal cell carcinoma: A cohort study. *Lancet Oncol.* **21**(2), 283–293 (2020).
58. Hakimi, A. A. et al. An epidemiologic and genomic investigation into the obesity paradox in renal cell carcinoma. *J. Natl. Cancer Inst.* **105**(24), 1862–1870 (2013).
59. Zhang, Y. et al. Single-cell analyses of renal cell cancers reveal insights into tumor microenvironment, cell of origin, and therapy response. *Proc. Natl. Acad. Sci. U.S.A.* m**118**(24) (2021).
60. Krishna, C. et al. 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 – 77 e6 (2021).
61. Obradovic, A. et al. Single-cell protein activity analysis identifies recurrence-associated renal tumor macrophages. *Cell* **184**(11), 2988–3005e16 (2021).
62. Tessaro, F. H. G. et al. Single-cell RNA-seq of a soft-tissue sarcoma model reveals the critical role of tumor-expressed MIF in shaping macrophage heterogeneity. *Cell. Rep.* **39**(12), 110977 (2022).
63. Noe, J. T. & Mitchell, R. A. MIF-dependent control of tumor immunity. *Front. Immunol.* **11**, 609948 (2020).
64. Benhamouda, N. et al. Plasma CD27, a surrogate of the intratumoral CD27-CD70 Interaction, correlates with immunotherapy resistance in renal cell carcinoma. *Clin. Cancer Res.* **28**(22), 4983–4994 (2022).
65. Kelly, R. J. et al. A phase I/II study of sepantronium bromide (YM155, survivin suppressor) with paclitaxel and carboplatin in patients with advanced non-small-cell lung cancer. *Ann. Oncol.* **24**(10), 2601–2606 (2013).
66. Parikh, M. et al. Phase ib trial of reformulated niclosamide with abiraterone/prednisone in men with castration-resistant prostate cancer. *Sci. Rep.* **11**(1), 6377 (2021).
67. Infante, J. R. et al. Unexpected renal toxicity associated with SGX523, a small molecule inhibitor of MET. *Invest. New. Drugs* **31**(2), 363–369 (2013).

Acknowledgements

Author contributions

Y.L. and Z.Y. designed the overall study. Y.L., Z.Y., C.L., R.Y., Q.Z., J.J and M.W. performed data collection, analysis, and experiments. Y.L., Z.Y., C.L., X.W. and X.L. were responsible for graph organization. Y.L., Z.Y., C.L., L.W., X.W., Z.C. and X.L. reviewed, edited, and revised the data and manuscript. All authors have read and approved the final manuscript.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-85618-7.

Correspondence and requests for materials should be addressed to X.W., Z.C. or X.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.