Gao *et al. Biology Direct* (2026) 21:108
https://doi.org/10.1186/s13062-026-00780-w

Biology Direct

RESEARCH Open Access

Single-cell analysis reveals that NRG1/3–ERBB4 signaling affects metabolic reprogramming and immune escape in Wilms tumor

Zhiqiang Gao[3,4], Jie Lin[1,2], Huafei Tang[1,2], Ao Li[1,2], Rui Li[1,2], Ao Xu[1,2], Rui Hu[1,2], Shuai Xu[1,2], Maolin Zhang[1,2], Wenming Yang[1,2], Haijing Huang[1,2], Zheng Zhang[1,2] and Feng Liu[1,2,3,4*]

Abstract

Background Wilms tumor presents a heterogeneous tumor microenvironment. This study aimed to characterize the tumor microenvironment and identify prognostic genes to improve therapeutic strategies.

Methods We integrated single-cell RNA sequencing and transcriptomic data from paired tumor and normal tissues. Bioinformatics analyses included assessments of cellular heterogeneity, trajectories, and cell–cell communication. Prognostic genes were identified with differential expression, Cox regression, and machine-learning analyses. Furthermore, functional characterization, immune infiltration patterns, and therapeutic targets were systematically investigated. Potential therapeutic compounds were predicted using drug databases and a graph-based deep learning framework to predict compound–protein interactions.

Results Single-cell RNA sequencing revealed 17 cell clusters, with tumor-specific epithelial cells and renal progenitor cells. Pseudotime trajectory analysis revealed dynamic differentiation, highlighting NRG1/3–ERBB4 signaling. Intersection of the transcriptomic and single-cell data identified 405 key genes. A prognostic model incorporating eight prognostic genes (*PRLR, SLC16A7, SGIP1, PPARGC1A, CDHR5, GRB7, FKBP10*, and *UGT2B7*) stratified patients into high- and low-risk groups ($p < 0.0001$), with area under the curve values > 0.6 for 1-, 2-, and 3-year survival prediction. High-risk patients had elevated regulatory T cell infiltration and immune checkpoint genes (*TNFRSF9* and *KIR3DL3*). Chitosan was identified as a multitarget agent that interacts with the eight prognostic proteins.

Conclusions This study defined tumor cellular architecture and identified eight prognostic genes with potential clinical value.

Keywords Wilms tumor, Single-cell RNA sequencing, Prognostic gene, Tumor microenvironment, Molecular docking

*Correspondence:
Feng Liu
hx2liuf@163.com
[1]Department of Pediatric Urology, West China Second University Hospital, Sichuan University, Cheng Du, Si Chuan, China
[2]Key Laboratory of Birth Defects and Related Diseases of Women and Children, Ministry of Education, Sichuan University, Cheng Du, Si Chuan, China
[3]Department of Urology, Children's Hospital of Chongqing Medical University, National Clinical Research Center for Children and Adolescents' Health and Diseases, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing, China
[4]Chongqing Key Laboratory of Structural Birth Defect and Reconstruction, Chongqing, China

Introduction

Wilms tumor (WT), the most common pediatric renal malignancy, accounts for 6–7% of all childhood solid tumors. It primarily affects children aged 2–5 years [1], occurs rarely in neonates or adolescents, and shows no significant sex enrichment. WT is typically unilateral, although 5–9% of cases are bilateral, and approximately 7% are multifocal [2]. Children with WT often present without overt clinical symptoms; the most common finding is an abdominal mass [3]. Whereas WT commonly originates in one kidney, it frequently metastasizes to distant organs such as the lungs and liver. Tumors display marked heterogeneity [4], meaning that cell populations with varying degrees of differentiation respond differently to treatment. Surgical resection remains the cornerstone of clinical management. However, patients classified as high-risk (those with metastasis, recurrence, or drug resistance) face a poorer prognosis, negatively impacting their quality of life [5]. Therefore, exploring WT pathogenesis and developing reliable biomarkers and potential therapeutic targets are of paramount importance [6]. Despite significant improvements in patient survival rates achieved with current multidisciplinary, comprehensive treatments, the prognosis for advanced or recurrent cases remains suboptimal. Furthermore, challenges such as chemotherapy resistance remain critical issues that require urgent attention.

Traditional bulk RNA sequencing (RNA-seq), which enables analysis of the entire tumor tissue, provides only an "average" gene expression profile across all cells. This approach often masks signals from critical, low-abundance cell subpopulations, such as tumor stem cells or drug-resistant cells [7]. Consequently, a technology capable of dissecting tumors at the single-cell level is required.

Single-cell RNA sequencing (scRNA-seq) offers a high-resolution method for dissecting complex tissues [8]. By analyzing the transcriptomes of thousands of individual cells, scRNA-seq enables unbiased identification of distinct cell types and states, reconstruction of developmental trajectories, and mapping of cell communication networks. Detailed analysis of individual tumor cells can reveal the molecular basis of their heterogeneity and interactions with the tumor microenvironment (TME) [9, 10]. Recent studies have indicated that immune heterogeneity within the TME is closely associated with tumor progression, metabolic adaptation, and treatment response, highlighting the importance of high-resolution investigations into tumor–immune interactions [11, 12]. Furthermore, emerging evidence suggests that metabolic and immune regulatory programs in the TME could influence tumor behavior and therapeutic sensitivity jointly [13]. However, single-cell characterization of cellular architecture and the tumor–immune microenvironment remains insufficient for WT.

In this study, we used high-throughput scRNA-seq to construct a comprehensive WT cellular atlas. We systematically analyzed the cellular heterogeneity of malignant and non-malignant regions across a WT cohort to identify specific malignant cell populations and their associated gene expression signatures that correlate with adverse clinical outcomes, uncovering novel prognostic biomarkers. Concurrently, we characterized the WT immune landscape, including the functional states of key immune cell subsets, and investigated the ligand–receptor interactions that facilitate communication between tumor cells and their microenvironment. Our findings offer new insights into WT prognosis and identify potential therapeutic targets in both the tumor and its microenvironment, providing the basis for improved risk stratification and innovative therapeutic strategies.

Materials and methods

Data acquisition

Three paired tumor and adjacent normal tissue samples were collected at the Children's Hospital of Chongqing Medical University. This study received ethical approval from the Ethics Committee of the affiliated Children's Hospital of Chongqing Medical University and was conducted in accordance with the principles outlined in the Declaration of Helsinki. All subjects' parents signed written informed consent. All samples were obtained from patients before chemotherapy, and the pathological WT type was confirmed by the pathology department. scRNA-seq data were collected from the three paired tumor and adjacent normal tissue samples for gene screening and single-cell clustering analyses. Bulk mRNA transcriptome data were downloaded from the Gene Expression Omnibus database (accession: GSE197047) [14], comprising eight WT and eight adjacent normal tissue samples. Additionally, clinical and transcriptomic data from WT samples were accessed from The Cancer Genome Atlas TARGET-WT cohort [15], including 124 primary tumor samples and six solid tissue normal controls. Survival information was available for 50 WT patients.

scRNA-seq data processing

Transcriptomic data from scRNA-seq of WT and adjacent normal tissue samples were processed with the Seurat package (v.4.1.0) [16]. The PercentageFeatureSet function was used to calculate quality-control metrics, including total RNA molecule counts per cell (nCount_RNA), the number of unique genes detected per cell (nFeature_RNA), and the percentage of mitochondrial gene expression (percent.mt). Cells were filtered based on the following criteria: cells with <200 or >2,500 detected genes; cells with <400 or >5,000 total RNA counts; and cells with mitochondrial gene percentages > 5%.

Post-quality-controlled data were visualized using Seurat to display the distributions of nFeature_RNA, nCount_RNA, and percent.mt across samples. The filtered single-cell data were normalized using the NormalizeData function in Seurat, with parameters set to LogNormalize and scale.factor = 10,000. To detect highly variable genes, the FindVariableFeatures function was used to select the top 2,000 genes exhibiting the highest cell-to-cell variability.

Cell clustering and annotation

Following data normalization, single-cell transcriptomic data were scaled using the ScaleData function in the Seurat package. Principal component analysis (PCA) was carried out on the top 2,000 highly variable genes using the RunPCA function. Subsequently, to identify significant principal components (PCs), the JackStraw function was applied. The ElbowPlot function in Seurat was used to generate a scree plot. Uniform Manifold Approximation and Projection (UMAP) was performed with the RunUMAP function for dimensionality reduction of significant PCs. Cell clusters were identified using the FindClusters function with the resolution parameter set to 0.4, followed by UMAP visualization of cluster distributions. To determine cell types, the FindAllMarkers function was applied with thresholds min.pct = 0.2 and only.pos = TRUE to identify cluster-specific marker genes. Automated cell-type annotation was performed with the SingleR package (v.2.0.0) [17] with the following three reference datasets: HumanPrimaryCellAtlasData, BlueprintEncodeData, and ImmuneCellExpressionData. Manual curation was performed by comparing marker genes with the CellMarker database (http://bio-bigdata.hrbmu.edu.cn/CellMarker/) and the published literature [18]. Annotated clusters were visualized using UMAP, and marker gene expression patterns were displayed. Visualization of cell type distributions was achieved with the ggplot2 package (v.3.3.2) [19]. To assess potential patient-specific bias, PCA plots colored by patient identity were generated.

Cell cycle analysis

Cell cycle phase assignment was performed using Seurat, which classifies cells into G1, S, or G2/M phases. For each annotated cell cluster, the proportions of cells in the G1, S, and G2/M phases were calculated and visualized using ggplot2. The UMAP map was generated using ggplot2 to visualize the distribution of cells across different cycles. Additionally, a stacked bar plot was created with ggplot2 to display the percentage distribution of cell cycle phases across all clusters.

Tumor subpopulation clustering and pseudotime trajectory analysis

Tumor sample cells were reclustered using the FindClusters function in Seurat, with a resolution parameter (resolution = 0.6). Developmental trajectories were restructured using the Monocle package (v.2.30.0) [20]. Cells were ordered along pseudotime by specifying a root node corresponding to the least differentiated state. The root state was inferred based on the expression patterns of cluster-specific marker genes and their differentiation features. CytoTRACE analysis was performed to estimate the developmental potential of individual cells. Branched expression trajectories were analyzed using Branched Expression Analysis Modeling. Additionally, an alluvial plot generated with the R package ggalluvial (v.0.12.5) [21] was used to illustrate the correspondence between tumor subpopulations and patient origins. Stacked bar plots created with ggplot2 were used to quantify subpopulation proportions across individual WT samples. Subsequently, cluster-specific differentially expressed genes (cDEGs) were detected using the FindAllMarkers function in the Seurat package (adjusted $p < 0.05$, $|\log_2$ fold change (FC)| >1). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were carried out with the ClusterProfiler package (v.3.16.0) [22], retaining pathways meeting an adjusted $p < 0.05$ and FoldEnrichment > 1. The top 20 significant pathways were visualized using ggplot2. Cell–cell communication networks were resolved using the R package CellChat (v.1.6.1) [23] ($p < 0.05$, $\log_2$mean (Molecule 1, Molecule 2) ≥0.1). The R package CellPhoneDB (v.2.0) [24] was used to identify interactions of ligand–receptor pairs.

Subpopulation merging and identification

Merged cell clusters were visualized using UMAP. Key cell clusters exhibiting distinct separation between WT and adjacent normal tissues were selected. The R package ReactomeGSA (v.1.12.0) [25] was used to perform pathway enrichment analysis on key cell clusters with the analyze_sc_clusters function (adjusted $p < 0.05$).

Single-cell differential expression profiling

Differential expression analysis of key cell clusters between WT and normal groups was performed using the FindMarkers function in Seurat with stringent thresholds (adjusted $p < 0.05$, $|\log_2 FC| > 2$). Significant single-cell differentially expressed genes (sDEGs) were visualized using a Manhattan plot and heatmap, highlighting the top 10 upregulated and downregulated sDEGs (ranked by $|\log_2 FC|$). Expression of the top 50 sDEGs was quantified and visualized with ggplot2. Subsequently, a ranked gene list was generated by sorting all genes by their $\log_2 FC$ values. The "c2.cp.kegg.v7.5.1.symbols.gmt" gene set was retrieved from the Molecular Signature Database

(MSigDB; http://software.broadinstitute.org/gsea/msigdb/) as a reference dataset. Gene set enrichment analysis (GSEA) was conducted with the ClusterProfiler package (adjusted $p < 0.05$, |normalized enrichment score [NES]| >1).

Key gene screening

Differential expression analysis between the WT and normal groups was conducted with the DESeq2 package (v.1.38.0) [26] in the GSE197047 dataset. Genes with $|\log_2 FC| > 2$ and adjusted $p < 0.05$ were defined as differentially expressed genes (DEGs2). To visualize the differential expression patterns, a volcano plot was generated using ggplot2. Additionally, a heatmap was constructed using the ComplexHeatmap package (v.2.14.0) [27] to display expression profiles of the top 20 DEGs2 across samples. To detect key genes, single-cell-derived sDEGs were intersected with DEGs2 using the VennDiagram package (v.1.7.3) [28]. Overlapping genes were identified as key genes and were functionally annotated.

Functional enrichment analysis and protein–protein interaction network

To elucidate the biological processes associated with the identified key genes, Gene Ontology (GO) enrichment analysis was carried out with the ClusterProfiler package. The analysis encompassed the following three categories: biological process, cellular component, and molecular function (adjusted $p < 0.05$). Similarly, KEGG pathway enrichment analysis was performed using ClusterProfiler (adjusted $p < 0.05$). To investigate the protein-level interactions among key genes, the identified genes were uploaded to the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (http://www.string-db.org/), and a protein–protein interaction (PPI) network was constructed. Interaction confidence scores were filtered using a threshold of ≥0.4 to retain high-confidence interactions. The resulting interaction data from STRING were imported into Cytoscape (v.3.7.2) [29] for network visualization.

Prognostic gene screening and validation

Univariate Cox regression analysis was carried out with the survival package (v.3.5–3) [30] to evaluate the association between key genes and overall survival based on 50 WT patients with complete survival data from the TARGET-WT cohort. Genes meeting the threshold of $p < 0.01$ and a hazard ratio (HR) $\neq 1$ were defined as candidate prognostic genes. To eliminate false-positive candidates, proportional hazards (PH) assumption testing was conducted. Only genes satisfying the PH assumption ($p > 0.05$) were retained for subsequent analyses.

Prognostic model construction and validation

To refine candidate prognostic genes and construct a robust survival prediction model, least absolute shrinkage and selection operator (LASSO) regression analysis with five-fold cross-validation was performed using the glmnet package (v.4.1.4) [31]. The LASSO algorithm introduces a penalty term that reduces model complexity and mitigates potential overfitting. The LASSO algorithm was applied with the parameter family = cox to penalize overfitting and select genes strongly associated with survival outcomes. The optimal penalty parameter (lamba.min) was determined. A risk score coefficient was derived from the LASSO model as follows: Risk Score = Σ Coefficient (i) × Expression of Gene (i), where Coefficient (i) represents the weight assigned to each gene by the LASSO model (with positive coefficients indicating increased risk with higher expression and negative coefficients indicating reduced risk), and Expression of Gene (i) denotes the normalized expression level of the corresponding gene. Subsequently, the optimal cutoff value for risk stratification was calculated using the surv_cutpoint function. This cutoff divided the 50 TARGET-WT patients into high- and low-risk groups. To evaluate survival differences between the two groups, Kaplan–Meier (K–M) curves were generated with the survminer package (v.0.4.9) [32]. Time-dependent receiver operating characteristic (ROC) curves for 1-, 2-, and 3-year overall survival were plotted using the timeROC package (v.0.4) [33] to quantify the risk model's predictive accuracy. An area under the curve (AUC) >0.6 was considered indicative of acceptable model performance. Risk score and survival status distributions across groups were analyzed. Furthermore, to validate the biological relevance of prognostic genes, the expression levels of these genes in the two groups were compared. A heatmap was generated with the ComplexHeatmap package (v.2.14.0) to visualize differential expression patterns.

Prognostic gene distribution analysis across key cell clusters

To investigate the spatial distribution of prognostic genes within key cell clusters and assess their expression patterns across distinct subpopulations, they were mapped onto the UMAP coordinates. For quantitative analysis, violin plots were generated with ggplot2 to visualize the expression levels of each prognostic gene.

GSEA and correlation analysis for prognostic genes and *IGF2BP3*

To ascertain the biological pathways associated with prognostic genes and insulin-like growth factor 2 mRNA-binding protein 3 (IGF2BP3) in WT, GSEA was performed. IGF2BP3 is an RNA-binding protein that is significantly upregulated in WT and promotes cell

proliferation and survival by stabilizing *IGF2* mRNA and enhancing IGF2/IGF1R axis activity. Additionally, IGF2BP3 promotes the invasion and metastasis of WT cells by regulating the EMT(Epithelial-Mesenchymal Transition) process. For each prognostic gene, WT patients with survival information in the TARGET-WT cohort were divided into high- and low-expression groups, based on the median expression value of these genes. Differential expression analysis between the two groups was carried out with the DESeq2 package (v.1.38.0). Prognostic genes were ranked in $\log_2$FC descending order. GSEA was subsequently implemented using the ClusterProfiler package (v.3.16.0) to identify enriched pathways. The analysis employed the c2.cp.kegg.v7.5.1.symbols.gmt gene set from the MSigDB database (|NES| >1, $p < 0.05$, and false discovery rate (FDR) <0.25). The top 5 most significantly enriched pathways were selected for visualization. To explore the IGF2BP3-specific pathway, the same GSEA workflow was applied. To assess the associations between prognostic genes and *IGF2BP3*, Spearman correlation analysis was implemented with the R package psych (v.2.2.5) [34]. The analysis utilized transcriptomic data from WT samples with survival information in TARGET-WT. For each prognostic gene and *IGF2BP3*, pairwise correlation coefficients (cor) were calculated (|cor| >0.3, $p < 0.05$).

Immune cell infiltration analysis based on risk stratification

To evaluate heterogeneity in the immune microenvironment between the high- and low-risk groups, immune cell infiltration analysis was performed. A curated gene set comprising 28 immune cell type–specific signatures was obtained from published literature [35]. Single-sample GSEA was instigated with the GSVA package (v.1.49.4) [36] to calculate the relative infiltration levels of these immune cells across the 50 WT samples with survival information. Differences in immune cell infiltration between the two groups were assessed using the Wilcoxon test. Immune cell types showing statistically significant differences ($p < 0.05$) were defined as differentially infiltrated immune cells. To explore associations between risk scores, prognostic genes, and differentially infiltrated immune cells, Spearman correlation analysis was administered using psych (|cor|) >0.3, $p < 0.05$).

Immune checkpoint gene analysis for risk stratification

To investigate the correlation between immune checkpoint gene expression and risk scores in WT samples, immune checkpoint genes were analyzed based on the established high- and low-risk groups. A list of 79 immune checkpoint genes was obtained from the published literature [37] (Additional File 1). Differential expression of these genes between the two groups was compared using the Wilcoxon test. Immune checkpoint genes with statistically significant expression differences ($p < 0.05$) were identified as differentially expressed. Subsequently, cor values were calculated using the psych package (v.2.2.5) to evaluate the relationships between risk scores and differentially expressed immune checkpoint genes (|cor| >0.3, $p < 0.05$).

Drug prediction and compound–protein interaction prediction

Potential therapeutic compounds targeting WT were identified using two complementary approaches. First, drug prediction was performed using the Drug Signatures Database (DsigDB; http://dsigdb.tanlab.org/) to identify compounds that interact with proteins encoded by the prognostic genes and IGF2BP3. Second, compounds interacting with these proteins were retrieved from the Comparative Toxicogenomics Database (CTD; https://ctdbase.org/). After merging these two sources and removing duplicates, unique compounds were obtained. Compounds lacking Simplified Molecular Input Line Entry System (SMILES) representations and compounds with an excessively long SMILES string (deemed too structurally complex for reliable prediction) were excluded. Consequently, the remaining compounds were retained for subsequent analysis. To screen for compounds with potential binding affinity to the proteins encoded by these genes and evaluate their therapeutic relevance, a GraphBAN framework was used to perform compound–protein interaction (CPI) prediction. GraphBAN integrates graph neural networks with attention mechanisms to specifically predict interactions between drug compounds and target proteins. The model was trained on an integrated dataset comprising interactions from three established CPI databases: BindingDB (https://www.bindingdb.org), the KIBA dat aset (https://pubs.acs.org/doi/10.1021/ci400709d), and the BioSNAP dataset (http://snap.stanford.edu/biodata). SMILES strings for the retained compounds were programmatically acquired from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/). Molecular features were extracted using two complementary approaches: (1) molecular graph representations processed via Graph Convolutional Neural Networks, and (2) molecular embeddings generated using ChemBERTa. Amino acid sequences and Protein Data Bank structural information for the target proteins were obtained from the UniProt database (https://www.uniprot.org; accessed October 16, 2025). Protein embeddings were constructed using a combination of one-dimensional convolutional neural networks applied to the amino acid sequences and features derived from evolutionary scale modeling. The extracted compound molecular graph features and protein embedding features were input into the GraphBAN model. GraphBAN utilizes a bidirectional attention

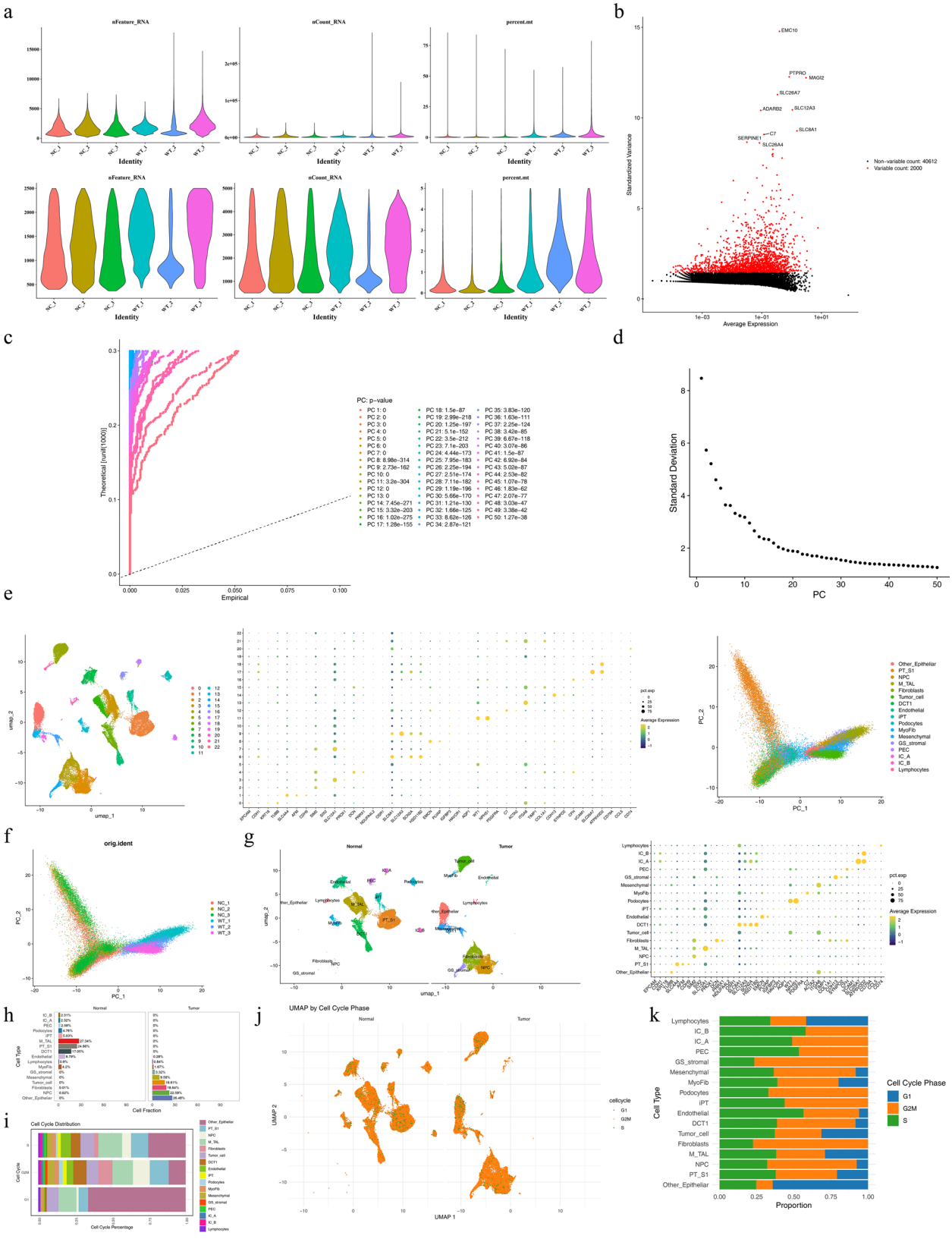


Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Single-cell transcriptomic atlas. (**a**) Quality-control metrics of single-cell RNA sequencing data. From left to right: number of genes detected per cell (nFeature_rna), total mRNA molecules per cell (nCount_rna), mitochondrial gene percentage (percent_mt). Top: pre-QC cells; bottom: post-QC cells. (**b**) Identification of highly variable genes. Red dots represent highly variable genes, and black dots represent nonvariable genes. (**c**) Principal component analysis (PCA) of the top 50 principal components (PCs). The X-axis shows the empirical distribution of the observed data, and the Y-axis shows the theoretical distribution under the null model. (**d**) Scree plot of PCA dimensionality reduction. The X-axis represents the PCs, and the Y-axis represents the SD, showing the variance contribution. (e-1) UMAP visualization of unsupervised cell clustering. (e-2) marker gene expression patterns across annotated cell types. The X-axis shows marker genes, and the Y-axis shows cell types. Yellow intensity reflects expression levels. (e-3) PCA of cell populations. (**f**) UMAP plot colored by patient identity, showing the distribution of cells across different samples. (g-1) subgroup clustering results after UMAP dimensionality reduction. (g-2) Comparative cell type composition between normal and tumor tissues. (**h**) Proportional distribution of major cell types across samples. (**i**) Cell cycle phase distribution across annotated cell types. The X-axis shows cell cycle phases, and the Y-axis shows proportional cell type distribution. (**j**) UMAP projection of cell cycle phases within subpopulations. Colors represent different cell cycle phases. (**k**) Cell cycle phase proportions within each cellular subpopulation. The X-axis shows the percentage of cells, and the Y-axis shows annotated cell types

mechanism to capture and weight local interaction patterns between compound and protein pairs. The model outputs the interaction probability (ranging from 0 to 1) for each compound–protein pair. A probability threshold of >0.5 was used to define potential binding interactions. Potential drug CPI pairs identified by GraphBAN were visualized using Cytoscape.

Statistical analysis

Bioinformatics analyses were performed in R (v.4.3.1). Comparative analysis of data from different groups was conducted using the Wilcoxon test ($p < 0.05$). For analyses involving multiple comparisons, p-values were adjusted using the Benjamini–Hochberg FDR method, and adjusted p-values were used to determine statistical significance.

Results

Single-cell landscape of WT reveals distinct cell type compositions and proliferative states

Initial quality-control filtering yielded 46,301 cells and 42,612 genes (Fig. 1a). Highly variable genes, such as *EMC10*, *PTPRO*, *MAGI2*, and *SLC26A7*, were identified (Fig. 1b). PCA demonstrated significant explanatory power for the first 50 PCs (all $p < 0.05$), with the variance contribution stabilizing beyond PC35 (Fig. 1c–d). Consequently, the first 35 PCs were selected for clustering, yielding 23 distinct cell clusters (Fig. 1e–1). Cell type annotation resolved 17 cell types, including other epithelial, proximal tubule S1 (PT_S1), renal progenitor cells (NPCs), medullary thick ascending limb (M_TAL), tumor cells, distal convoluted tubule segment 1, endothelial, injured proximal tubule cells (iPT), podocytes, myofibroblast (myoFib), mesenchymal, glomerular mesangial cells (GS_stromal), parietal epithelial cells, renal collecting duct epithelial cells A (IC_A), IC_B, fibroblasts, and lymphocytes (Figs. 1e–2–3). PCA visualization colored by patient identity revealed partially overlapping distributions across samples, rather than clear separation by patient origin, suggesting that clustering patterns were not solely driven by patient-specific effects (Fig. 1f). Comparative analysis revealed marked compositional differences between tumor and normal tissues. M_TAL and PT_S1 dominated the normal samples, whereas tumor specimens exhibited enrichment of other epithelial cells and NPCs (Fig. 1g–h). Cell cycle phase distribution analysis revealed cell type–specific proliferation patterns, with other epithelial cells showing elevated G1-phase proportions compared to the PT_S1 and M_TAL populations, which exhibited higher G2M-phase contributions (Fig. 1i). Notably, G2M-phase cells constituted the majority across all cell clusters (Fig. 1j). Further stratification revealed that G2M/S-phase dominance was conserved across most cell clusters, except for other epithelial cells and lymphocytes, where G1-phase cells predominated (Fig. 1k). These findings revealed distinct proliferative states between tumor-associated and normal renal epithelial clusters.

Dynamic cellular trajectories and intercellular communication in the TME

PCA revealed strong discriminatory power across the first 30 PCs, enabling tumor cells to be stratified into 17 distinct clusters (Figs. 2a–1–2). Pseudotemporal trajectory analysis delineated three developmental states (state1, early; state2/3, late), suggesting a potential differentiation continuum in which tumor cells transition from mature states (left trajectory) to recently differentiated phenotypes (right trajectory; Figs. 2b–1–2). Cluster-level trajectory mapping demonstrated an early differentiation characteristic of cluster 5 vs. late-stage features in clusters 0 and 2 (Figs. 2b–3). CytoTRACE analysis revealed distinct differentiation states across clusters. Specifically, clusters 14, 4, and 22 exhibited higher developmental potential, suggesting earlier differentiation states, whereas clusters 1, 3, and 7 showed more differentiated features (Figs. 2b–4). Sample-specific cluster distribution analysis demonstrated substantial TME heterogeneity across specimens. WT_1 was dominated by cluster 1, WT_2 by cluster 0, and WT_3 by cluster 2 (Fig. 2c). Differential expression analysis identified 4,162 cDEGs, with KEGG enrichment revealing 87 significantly enriched pathways ($p < 0.05$), such as the ribosome, Rap1 signaling, and the PI3K-Akt pathway (Fig. 2d and Additional File 2).

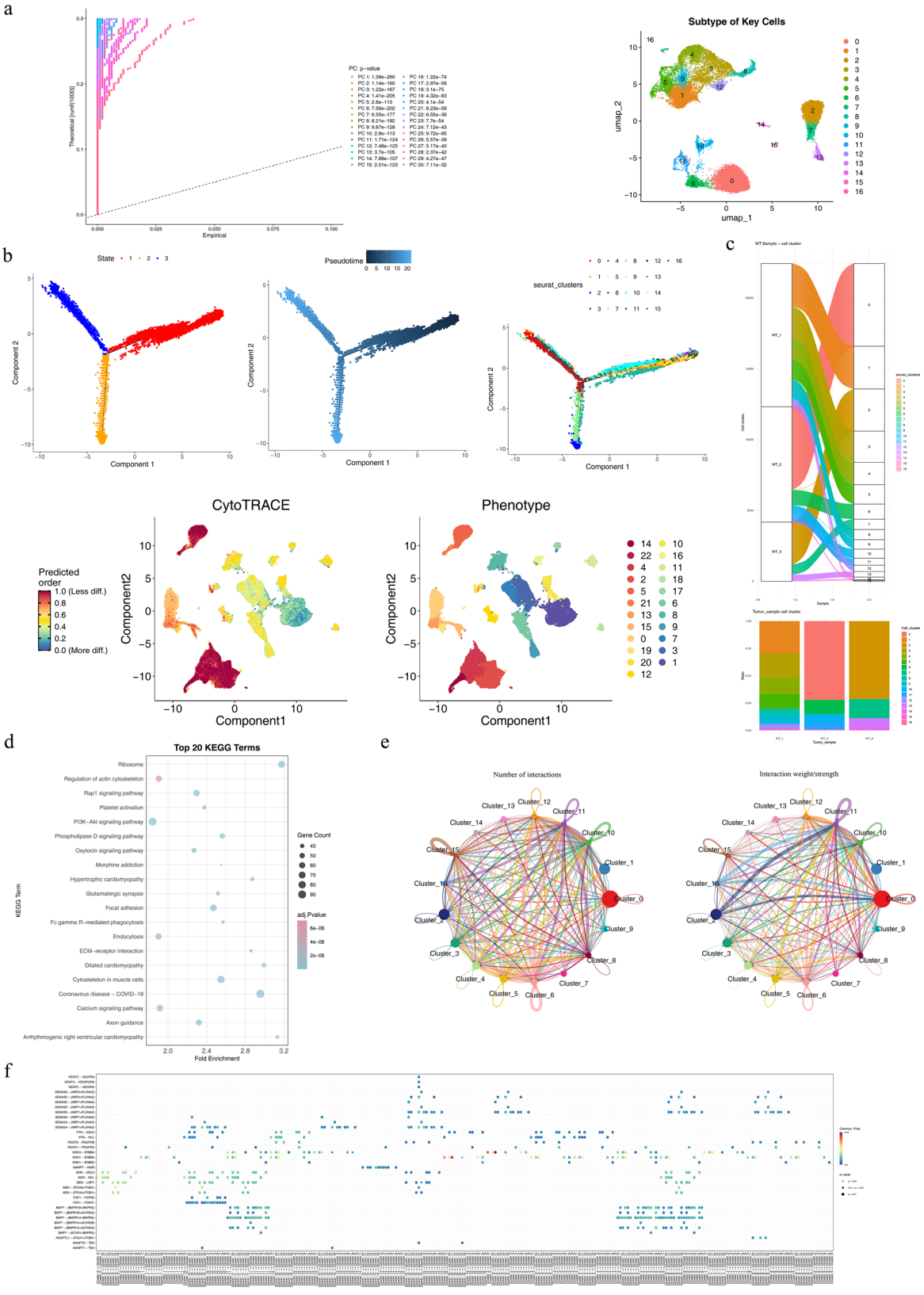


Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Single-cell characterization of Wilms tumor heterogeneity, developmental trajectories, and cell–cell communication. (**a-1**) PCA results of tumor cell samples. (**a-2**) Clustering of tumor cell samples. (**b-1**) Distribution of cell subpopulations. (**b-2**) Developmental state distribution of cells. (**b-3**) Classification of cell clusters based on differentiation status. (**b-4**) CytoTRACE analysis of tumor cell differentiation states. (**c-1**) Correspondence between cell subpopulations and sample origins. The left labels show sample names, and the right labels indicate cluster IDs. (**c-2**) Proportion of subpopulation cells across individual samples. Stacked bars show the relative proportions of each cluster per sample. (**d**) Functional enrichment of differentially expressed genes across subpopulations. Red indicates more significant pathways, and the dot size demonstrates gene count. (**e-1**) Interaction frequencies among all cells. The heatmap shows the number of interactions between cell types. (**e-2**) Interaction strength between cell types. Network edges represent communication intensity between populations. (**f**) Ligand–receptor interaction probability. Red bubbles indicate stronger predicted interactions between pairs

Cell–cell communication analysis uncovered an interaction network, with cluster 2 exhibiting higher communication frequency and intensity with clusters 10 and 16 (Fig. 2e). Furthermore, cell–cell communication analysis revealed multiple ligand–receptor interactions across tumor cell clusters. Among these interactions, NRG1–ERBB4 and NRG3–ERBB4 displayed the strongest signaling intensity and statistical significance ($p < 0.01$) in the interaction network. Specifically, NRG1–ERBB4 signaling connected cluster 16 with clusters 0, 10, and 11, whereas NRG3–ERBB4 signaling linked cluster 2 to clusters 0, 10, and 11 (Fig. 2f). These findings collectively characterize the dynamic evolution and heterogeneity within the TME and suggest potential signaling interactions associated with tumor progression.

Epithelial cells might be central regulators of the TME

We identified 10 distinct epithelial clusters, including PT_S1, NPC, and IC_A/B, which were collectively categorized as epithelial. Similarly, four tumor-associated fibroblast subtypes (myoFib, fibroblasts, mesenchymal, and GS_stromal) were collectively identified as cancer-associated fibroblasts. UMAP visualization demonstrated complete segregation between normal and tumor epithelial cells with minimal overlap (Fig. 3a–d). Thus, epithelial cells were designated as the key cell clusters for subsequent analyses. Functional enrichment of epithelial cells revealed marked activation of ion transport pathways, such as Na-permeable kainate receptors and proton-coupled neutral amino acid transporters, as well as abacavir metabolism (Fig. 3e). Moreover, comparative analysis identified 1,360 sDEGs between tumor and adjacent epithelial cells, with 669 upregulated genes (e.g., *COL2A1*, *SFRP2*, and *MDK*) and 691 downregulated genes (e.g., *UMOD*, *GP2*, and *AFM*) in tumor epithelial cells (Figs. 3f–1–2–3). Subsequently, GSEA identified 13 tumor-associated pathways, including the ribosome, adipocytokine signaling pathway, and aldosterone-regulated sodium reabsorption (Fig. 3g). The distinct molecular signature of tumor epithelial cells suggests their potential involvement in oncogenic niche formation. These findings collectively indicate that epithelial cells may contribute to TME remodeling.

Identification of key genes in the TME

Differential expression analysis of the GSE197047 dataset revealed 2,094 DEGs, including 1,253 upregulated and 841 downregulated genes, in tumor tissues compared to adjacent tissues (Fig. 4a–b). Intersection analysis with single-cell transcriptomic data identified 405 sDEGs and DEGs2, designated as key genes (Fig. 4c). Subsequently, GO analysis highlighted 337 terms, with top terms including apical plasma membrane, renal system development, and solute/sodium symporter activity (Fig. 4d and Additional File 3). KEGG pathway analysis also implicated four pathways, including one carbon pool by folate, protein digestion and absorption, parathyroid hormone synthesis, secretion, and action, and aldosterone-regulated sodium reabsorption (Fig. 4e). Additionally, PPI network analysis resolved 330 key genes into a cohesive network with 1,057 interactions. EZH2, EGF, and PPARG may play pivotal regulatory roles in TME modulation (Fig. 4f).

Prognostic model construction and validation

To establish a robust prognostic signature, single-variable Cox regression analysis revealed 11 candidate prognostic genes, including seven risk genes (*PRLR*, *SLC16A7*, *SGIP1*, *PPARGC1A*, *MAP3K15*, *MME*, and *UGT2B7*; HR > 1) and four protective genes (*CDHR5*, *GRB7*, *FKBP10*, and *HSD11B2*; HR < 1; Fig. 5a). PH assumption testing excluded *MAP3K15* ($p < 0.05$), retaining 10 PH-compliant genes for subsequent modeling (Fig. 5b). LASSO regression analysis optimized model performance at lambda.min = 0.1268364, selecting eight genes (*PRLR*, *SLC16A7*, *SGIP1*, *PPARGC1A*, *CDHR5*, *GRB7*, *FKBP10*, and *UGT2B7*) with non-zero coefficients to construct the final prognostic signature (Fig. 5c). The risk score, calculated as a weighted sum of gene expression levels (coefficients provided in Table 1), effectively stratified patients into high- and low-risk groups. K–M analysis revealed significantly shorter survival for patients in the high-risk group ($p < 0.0001$; Fig. 5d). 50 patients with WT having complete survival information were included; all experienced events during follow-up. The follow-up time ranged from 0.49 to 7.16 years, with a median follow-up time of 1.55 years. Time-dependent ROC curves indicated modest predictive ability, with AUC values exceeding 0.6 at 1-, 2-, and 3-year intervals (Fig. 5e). Risk score distribution analysis revealed a progressive decline

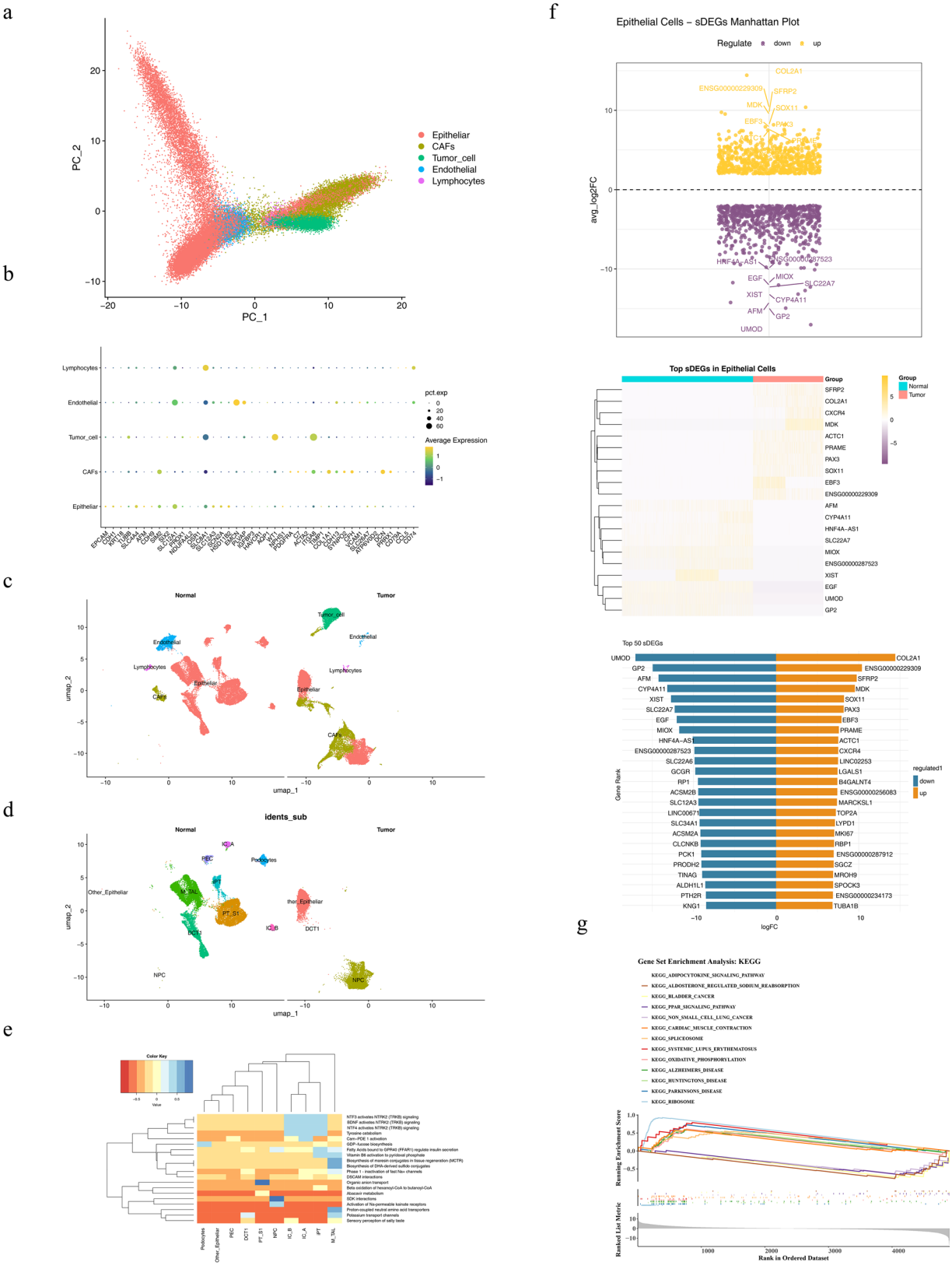


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Single-cell characterization of epithelial cell heterogeneity and functional remodeling in Wilms tumor. (**a**) PCA of all cell clusters. (**b**) Marker gene expression heatmap across the annotated cell clusters. (**c**) UMAP visualization of single-cell clusters after annotation, colored by tissue group. (**d**) UMAP projection of epithelial cell subsets stratified by tissue origin. (**e**) Functional enrichment analysis of key epithelial subpopulations. The intensity indicates enrichment significance. (f-1) Manhattan plot displaying sDegs in tumor vs. normal epithelial cells. (f-2) heatmap of epithelial sDEG expression patterns. (f-3) symmetric bar plot showing expression levels of the top 50 sDegs. (**g**) GSEA revealing 13 tumor-associated pathways in epithelial cells

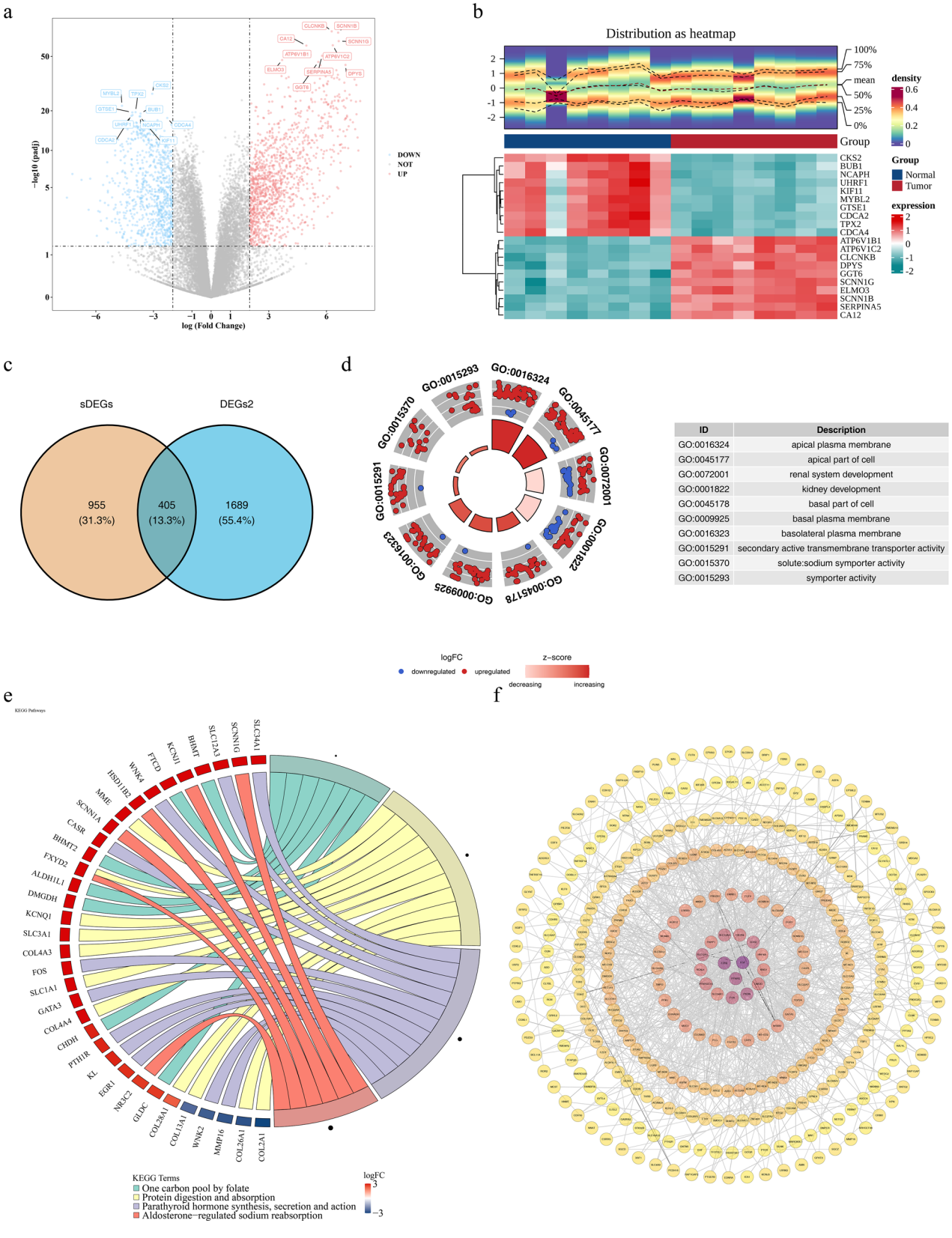
in survival time with increasing scores (Fig. 5f). Expression profiling of prognostic genes showed elevated levels of risk-associated genes (*PRLR*, *SLC16A7*, *SGIP1*, *PPARGC1A*, and *UGT2B7*) in the high-risk groups, whereas *CDHR5*, *GRB7*, and *FKBP10* exhibited low expression in the high-risk group (Fig. 5g). The model showed preliminary prognostic value for risk stratification in the TARGET-WT cohort, although further validation in larger independent cohorts is required.

Functional characterization of prognostic genes

The expression patterns of the eight signature genes (*PRLR*, *SLC16A7*, *SGIP1*, *PPARGC1A*, *CDHR5*, *GRB7*, *FKBP10*, and *UGT2B7*) were mapped across 10 epithelial clusters. UMAP visualization demonstrated broad expression of these genes across all epithelial subpopulations (Fig. 6a). Notably, *GRB7* and *UGT2B7* exhibited peak expression in iPT cells, whereas other genes showed distinct subcluster-specific dominance (Fig. 6b). Functional enrichment analysis revealed pathway divergence among prognostic genes. The protein products of *CDHR5*, *FKBP10*, *GRB7*, *IGF2BP3*, and *SGIP1* were significantly enriched in the ribosome, oxidative phosphorylation, and neurodegenerative pathways (Parkinson's disease and Alzheimer's disease). GRB7 was additionally associated with cardiac muscle contraction, whereas SGIP1 was linked to cytokine–receptor interactions (Figs. 6c–1–5). PPARGC1A coordinates extracellular matrix (ECM) remodeling (ECM–receptor interactions) and cancer-related pathways (focal adhesion; Figs. 6c–6). PRLR and SLC16A7 converged on cardiomyopathy pathways (hypertrophic cardiomyopathy); PRLR was also implicated in neuroactive ligand–receptor signaling and SLC16A7 in antigen presentation (Figs. 6c–7–8). UGT2B7 is uniquely associated with systemic lupus erythematosus alongside metabolic and neurodegenerative pathways (Fig. 6c-9). Correlation analysis uncovered coordinated regulatory networks. *UGT2B7* and *SGIsP1* showed the strongest positive association (cor = 0.44, $p < 0.001$). Furthermore, *IGF2BP3* was positively correlated with *PRLR* (cor = 0.35, $p < 0.01$) and *PPARGC1A* (cor = 0.32, $p < 0.05$). Conversely, *GRB7* and *SGIP1* showed the strongest negative relationship (cor = –0.42, $p < 0.05$; Fig. 6d). These results collectively delineate an organized functional architecture in which prognostic genes operate through distinct yet interconnected pathways. The observed correlation networks may highlight coordinated transcriptional programs governing epithelial plasticity, suggesting that these genes may be involved in TME remodeling.

Immune landscape and checkpoint profiling across risk groups

Immune infiltration analysis revealed distinct microenvironmental features, with neutrophils and B cells showing lower relative infiltration levels than other immune cells (Fig. 7a). Comparative immune infiltration analysis between the high- and low-risk groups revealed significant differences in four immune cell types. Monocyte infiltration was markedly reduced in high-risk patients, whereas gamma-delta T cells (γδ T cells), natural killer T cells (NKTs), and regulatory T cells (Tregs) showed elevated infiltration levels in this group (Fig. 7b). Additionally, risk scores demonstrated strong positive correlations with NKTs (cor = 0.50, $p < 0.001$) and Tregs (cor = 0.49, $p < 0.001$; Fig. 7c). Furthermore, γδ T cells were positively associated with *SGIP1* (cor = 0.515, $p < 0.001$) but inversely associated with *CDHR5* (cor = –0.466, $p < 0.001$) and *GRB7* (cor = –0.435, $p < 0.01$) expression. Monocytes exhibited a strong positive association with *CDHR5* (cor = 0.403, $p < 0.0001$), whereas NKTs showed dual associations: a positive correlation with *SGIP1* (cor = 0.528, $p < 0.001$) and *UGT2B7* (cor = 0.369, $p < 0.0001$), and negative correlations with *GRB7* (cor = –0.362, $p < 0.0001$) and *FKBP10* (cor = –0.382, $p < 0.0001$). Tregs demonstrated coordinated regulation with *SGIP1* (cor = 0.518, $p < 0.0001$), *PPARGC1A* (cor = 0.465, $p < 0.0001$), and *UGT2B7* (cor = 0.426, $p < 0.0001$), but antagonism with *GRB7* (cor = –0.374, $p < 0.0001$; Fig. 7d). Immune checkpoint analysis identified six differentially expressed genes (*BTN2A2*, *HLA-DRB5*, *ICOS*, *KIR3DL3*, *PDCD1*, and *TNFRSF9*) between risk groups (Fig. 7e). Risk scores exhibited a positive association with *TNFRSF9* (cor > 0.5, $p < 0.001$) and *KIR3DL3* (cor > 0.3, $p < 0.01$). Conversely, *BTN2A2* exhibited an inverse association (cor < –0.03, $p < 0.01$; Fig. 7f). These findings collectively delineate an immune microenvironment in high-risk patients characterized by enhanced infiltration of Treg subsets and coordinated expression of immune checkpoint genes. The inverse relationship between monocyte infiltration and risk scores, coupled with the strong associations between prognostic genes and immune cell dynamics, suggests potential interactions between epithelial transformation and immune regulatory alterations.



ID	Description
GO:0016324	apical plasma membrane
GO:0045177	apical part of cell
GO:0072001	renal system development
GO:0001822	kidney development
GO:0045178	basal part of cell
GO:0009925	basal plasma membrane
GO:0016323	basolateral plasma membrane
GO:0015291	secondary active transmembrane transporter activity
GO:0015370	solute:sodium symporter activity
GO:0015293	symporter activity

Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Identification and functional characterization of key genes in the Wilms tumor microenvironment (TME). (**a-1**) Volcano plot of DEGs in Wilms tumor vs. normal tissues. Red points indicate upregulated genes. Blue points indicate downregulated genes. Gray points represent nonsignificant genes. The top 10 most significantly upregulated and downregulated genes are labeled. (**a-2**) Heatmap of DEGs across Wilms tumor and normal samples. The upper panel displays expression density with quantiles and mean lines. The lower panel shows gene expression levels, with red indicating high expression and cyan indicating low expression. Sample groups are color-coded, with red for tumors and blue for normal tissues. (**b**) Venn diagram of key gene overlaps between sDegs and bulk RNA-seq DEGs. The intersection revealed 405 shared genes identified as key genes in the TME. (**c**) GO enrichment analysis of key genes in the Wilms TME. Bar height represents term significance, and color intensity reflects the z-score. The outer ring displays gene expression levels per term, with red indicating upregulation and blue indicating downregulation. KEGG pathway descriptions are provided for reference. (**d**) KEGG pathway enrichment of key genes. Genes are grouped by associated pathways, with red indicating upregulation and blue indicating downregulation in tumor versus normal tissues. (**e**) Protein–protein interaction network of key genes. Node darkness corresponds to degree centrality, and edge thickness represents interaction strength. EZH2, EGF, and PPARG emerged as central regulatory hubs

Identification of chitosan as a potential multitarget interaction candidate

To predict compounds that potentially interact with the nine identified target genes and to identify novel therapeutic candidates, we leveraged bioinformatics approaches. In total, 247 compounds were obtained from DsigDB, whereas 482 compounds with interaction relationships to the target genes were retrieved from CTD. Following compound merging and deduplication, 708 unique compounds were identified. Subsequent filtering excluded 204 compounds lacking SMILES representations and one compound with excessively long SMILES, resulting in a final set of 503 compounds for combination with the nine target genes. Using GraphBAN for screening, 25 compound–protein pairs were identified, encompassing eight proteins (excluding GRB7) and eight compounds. Notably, chitosan interacted with all eight proteins, whereas potassium dichromate interacted with PPARGC1A, UGT2B7, IGF2BP3, and SGIP1. Additional interactions included perfluorooctane sulfonic acid, ivermectin, perfluoro-n-nonanoic acid, perfluorodecanesulfonic acid, sirolimus, and perfluorooctanoic acid (Fig. 8). These findings suggest that the identified drug compounds, particularly chitosan, with its broad interaction profile, might interact with the target proteins and serve as candidates for experimental validation.

Discussion

WT, the most common pediatric kidney malignancy, presents with a heterogeneous TME, with unclear implications for clinical prognosis [38–40]. Our single-cell analysis unveiled significant cellular differences between WT and normal kidney tissue, highlighting a distinct epithelial cell landscape in these tumors. More importantly, we identified eight TME-related core prognostic genes and constructed a prognostic risk assessment model. This model demonstrated preliminary prognostic value in the TARGET-WT cohort. Furthermore, our preliminary mechanistic exploration of these prognostic genes provides novel insights into WT pathogenesis and suggests potential therapeutic targets.

Single-cell transcriptomic analysis revealed significant differences in the cellular composition between WT and adjacent normal kidney tissue. Normal renal tissue was predominantly composed of functionally mature PT_S1 and M_TAL cells. This composition aligns with their crucial physiological roles in maintaining renal homeostasis, including reabsorption of substances and urine concentration, reflecting the mature and stable state of normal tissue [41, 42]. Conversely, other epithelial cell subpopulations and renal progenitor cells predominated in WT tissue. Notably, these tumor epithelial cells were primarily in the G1 phase. This observation suggests that they may be in a relatively quiescent yet "primed" state with proliferative potential, or it could reflect aberrant cell cycle checkpoint regulation that delays their progression into the S phase [43]. Crucially, our dynamic trajectory analysis suggests a potential differentiation trend from a mature to an epithelial/undifferentiated phenotype. This "atavistic" shift toward an undifferentiated state is a key hallmark of malignant tumor evolution. Undifferentiated cells typically show enhanced proliferative capacity, invasiveness, and treatment resistance [44–46]. This dynamic process suggests that WT may continuously generate more aggressive and drug-resistant cell clones during its development, driving persistent tumor progression and metastasis. This provides cellular-level evidence for understanding WT mechanisms. However, RNA velocity analysis was not performed in this study; therefore, the inferred trajectory direction should be interpreted with caution.

A key translational outcome of this study is the construction and validation of a prognostic model consisting of eight genes: *PRLR*, *SLC16A7*, *SGIP1*, *PPARGC1A*, *CDHR5*, *GRB7*, *FKBP10*, and *UGT2B7*. This model effectively distinguished the survival outcomes of WT patients; those in the high-risk group had a significantly shorter median survival. Furthermore, ROC curve analysis demonstrated that the model had modest predictive ability at different time points. These findings suggest that the model may have value as a molecular tool for WT risk stratification. In clinical practice, this model could provide a preliminary basis for identifying WT patients with different risk profiles, although additional validation is required before clinical application. For example, it could direct high-risk patients toward more

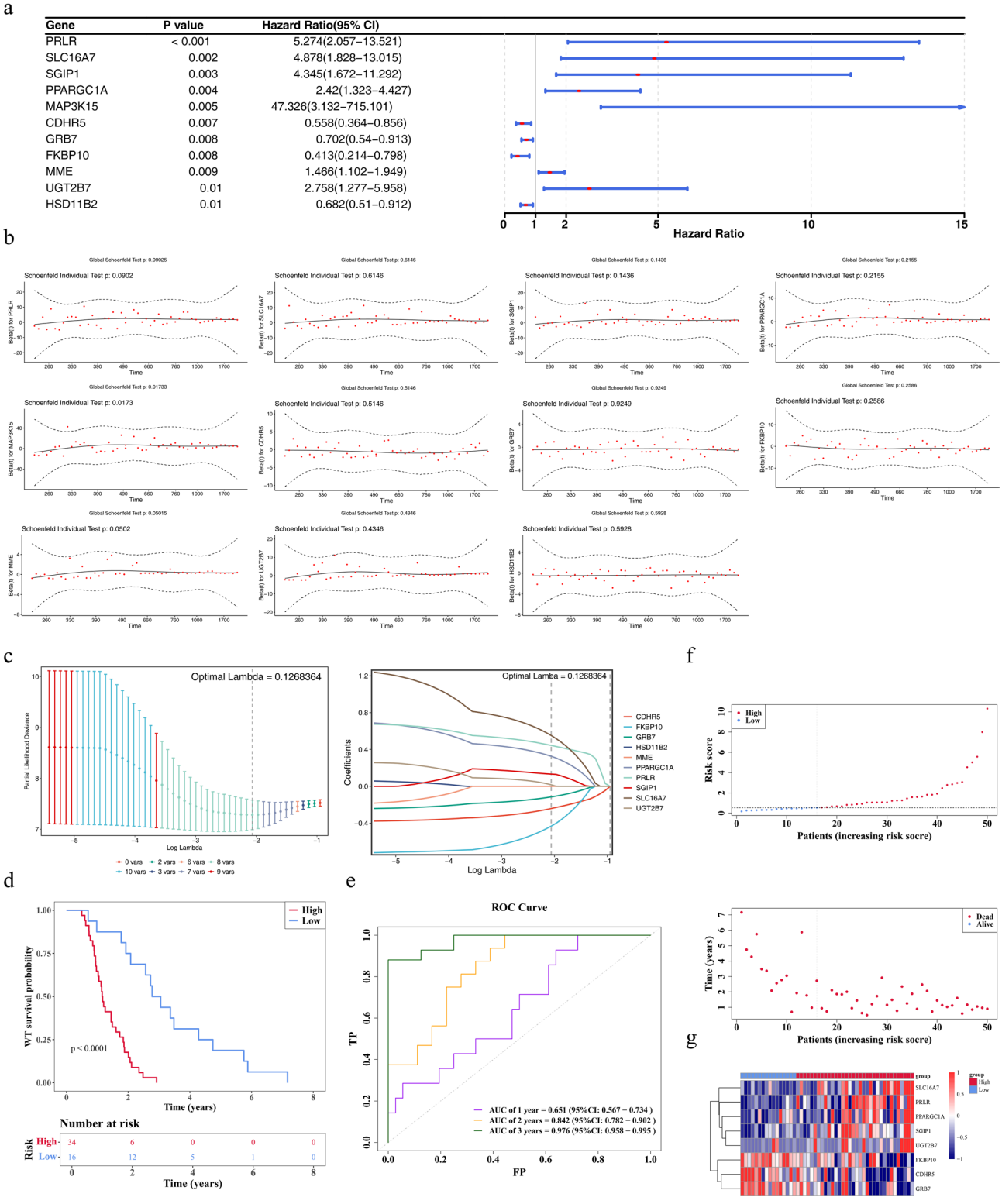


Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 Construction and validation of the 8-gene prognostic signature for Wilms tumor survival prediction. (**a**) Univariate Cox regression analysis of candidate prognostic genes. The red dots represent hazard ratios (HR), and the blue lines indicate confidence intervals. (**b**) Schoenfeld residual test for the proportional hazards (PH) assumption. The solid line is a smoothed spline fit, and the dashed lines represent ±2 SE ranges around the fit. Parallel lines relative to the X-axis suggest agreement with the PH assumption, and the coefficients β1, β2, and β3 are assumed to be time-invariant. (**c-1**) LASSO regression analysis for prognostic gene selection. Cross-validation curve showing retained genes across different regression models, with the lower X-axis as log(λ) and the upper X-axis as the corresponding gene count. (**c-2**) LASSO coefficient profiles of prognostic genes. Colored lines represent regression trajectories of individual genes, with the X-axis as log(λ). (**d**) Kaplan–Meier survival analysis of the training cohort. (**e**) Time-dependent ROC curves of the training cohort. The X-axis, the false-positive rate, reflects the proportion of negative samples incorrectly predicted to be positive, whereas the Y-axis, the true positive rate, indicates correctly predicted positive samples. AUC values measure model discrimination, with higher values signifying better performance. (**f-1**) risk score distribution in the training cohort. Red dots denote high-risk samples, and blue dots represent low-risk samples. (**f-2**) survival status between the risk groups in the training cohort. Red dots indicate deceased patients, blue dots represent survivors, and the dashed line shows the optimal risk score threshold. (**g**) Expression profiles of prognostic genes in the training cohort

intensive adjuvant chemotherapy regimens or enrollment in clinical trials of novel targeted drugs [47–49].

To further understand our prognostic model, we used GSEA to examine the potential functions of these core genes. The prolactin receptor (PRLR) is a transmembrane cytokine receptor that, upon binding to its primary ligand prolactin (PRL), initiates intracellular signaling cascades, primarily activating the JAK2–STAT5 pathway [50]. In addition to JAK–STAT, PRLR signaling exhibits cross-talk with other pathways, such as MAPK/ERK, PI3K/AKT, and Src [51]. In renal cancer, components of the prolactin signaling axis, including PRLR, may be dysregulated, contributing to renal cancer cell proliferation [52]. Our study connected PRLR to neuroactive ligand–receptor interaction pathways. Given the known neuromodulatory effects of PRL and the significant influence of the nervous system on tumor growth and metastasis through various neuropeptides and neurotransmitters [53, 54], PRLR in WT may not only regulate intrinsic tumor cell behavior but also mediate cross-talk between tumor cells and neural components within the TME, potentially influencing tumor progression.

Monocarboxylate transporter 2 (SLC16A7 or MCT2) is a key member of the solute carrier family 16 (SLC16), a group of proton-coupled monocarboxylate transporters. *SLC16A7* expression is dysregulated in various malignancies, affecting tumor initiation, progression, and immune responses [55]. In bladder cancer, decreased *SLC16A7* expression is associated with tumor progression, and restoration inhibits cancer cell proliferation while promoting $CD8^+$ T cell chemotaxis and tumor-killing capacity [56]. Conversely, studies in prostate cancer have reported increased *MCT2* expression in tumors [57], suggesting a context-dependent role. The precise balance of lactate import and export, influenced by various MCT isoforms, dictates tumor metabolic flexibility and invasiveness. Our GSEA analysis linked *SLC16A7* to antigen presentation pathways, suggesting a potential role in modulating the immune status of the WT microenvironment.

UDP-glucuronosyltransferase 2 family polypeptide B7 (UGT2B7) is an important member of the UGT superfamily, and its role is often associated with the metabolism of endogenous hormones and chemotherapeutic drugs, which influences both carcinogenesis and treatment response. In many cancers, altered UGT expression affects the bioavailability of procarcinogens or anticancer drugs. For instance, in breast cancer, UGT2B7 plays a role in the metabolism of estrogen and its derivatives; altered activity affects the estrogen level, which is a critical driver of hormone-sensitive breast cancers [58]. The kidney is a primary site of UGT expression, where these enzymes contribute to detoxification and substance elimination. Given that WT is a pediatric kidney cancer, the metabolic machinery of renal tissue, including UGT2B7, is intrinsic to its cellular environment. The metabolism of certain chemotherapeutic drugs used in WT treatment may be influenced by UGT2B7 activity, affecting drug pharmacokinetics and efficacy. Our study identified *UGT2B7* as a core prognostic gene in WT. A notable finding from our drug screening analysis was that estradiol (E2) may target proteins encoded by six genes in our prognostic model, including *UGT2B7* itself. This suggests a more complex relationship between hormone signaling and the identified prognostic gene network in WT than previously understood. Given the direct role of UGT2B7 in E2 metabolism, this finding suggests a potential feedback loop or regulatory axis in which E2 levels, influenced by *UGT2B7*, may affect the expression or activity of other prognostic genes, impacting tumor behavior.

Peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PPARGC1A or PGC-1α) is a transcriptional coactivator critical for regulating cellular energy metabolism. It coordinates the expression of genes involved in mitochondrial biogenesis, oxidative phosphorylation, fatty acid oxidation, and gluconeogenesis [59]. In some cancers, *PPARGC1A* is a tumor suppressor, and its downregulation promotes glycolysis and tumor growth. Conversely, in other cancer types, particularly highly metabolic tumors, elevated *PPARGC1A* expression fuels mitochondrial oxidative phosphorylation, supporting rapid proliferation and drug resistance [60]. Some studies have indicated that PPARGC1A influences tumor initiation, progression, and prognosis in

Table 1 LASSO-derived coefficients and functional annotations of the 8 prognostic genes in Wilms tumor

Gene	coef
PRLR	0.44423766
SLC16A7	0.55279684
SGIP1	0.13282253
PPARGC1A	0.32674898
CDHR5	−0.24155525
GRB7	−0.11540579
FKBP10	−0.43312477

renal cancer [61]. Our study identified *PPARGC1A* as a core prognostic gene in WT; its role as a key regulator of mitochondrial biogenesis and energy metabolism aligns with findings in other cancers. Aberrant *PPARGC1A* expression may remodel tumor cell metabolic patterns to adapt to rapid proliferative demands.

Cadherin-related family member 5 (CDHR5), a member of the cadherin superfamily, is believed to participate in cell–cell adhesion. It may influence cell signaling pathways that control cell behavior, including proliferation and differentiation [62]. Aberrant cadherin expression is frequently observed in renal cancer, in which it promotes tumor growth and metastatic potential [63]. Given that WT originates from undifferentiated renal blastema, aberrations in normal cell adhesion and developmental processes are intrinsically related to its pathogenesis. We identified *CDHR5* as a core prognostic gene whose downregulation may weaken cell adhesion. Reduced *CDHR5* expression may lead to the more aggressive, dedifferentiated phenotype observed in high-risk tumors.

Growth factor receptor-bound protein 7 (GRB7) is an adaptor protein belonging to a small family of SH2-domain-containing proteins that interact with numerous receptor tyrosine kinases. It participates in various cellular processes, including cell proliferation, migration, and differentiation [64–66]. *GRB7* is a well-recognized oncogene in breast cancer and is often coamplified with *HER2*/*ERBB2* on chromosome 17q. This coamplification enhances HER2 signaling, promoting aggressive tumor phenotypes, increased proliferation, and therapy resistance. In bladder cancer, a key mechanism by which GRB7 exerts oncogenic effects is by activating the classical PI3K/AKT pathway, amplifying signals that drive tumor cell growth, survival, and invasion [67]. We found that *GRB7* is a core prognostic gene in WT and that high expression may drive malignant progression. This aligns with its role in other cancers as an oncogenic adaptor protein activating key procancer signaling pathways like RAS/RAF/MEK/ERK [64–66]. Our results suggest that GRB7 acts as a critical mediator of oncogenic signaling in WT cells, contributing to their proliferative and invasive capabilities.

FKBP prolyl isomerase 10 (FKBP10) is a member of the FK506-binding immunophilin family. It is a large endoplasmic reticulum-resident immunophilin involved in protein folding and quality control. Beyond protein folding, FKBPs interact with various signaling molecules, influencing diverse cellular processes, including immune responses, cell growth, and stress responses [68, 69]. In renal cancer, *FKBP10* is significantly upregulated, promoting tumor cell proliferation, migration, and invasion. It also enhances the progression of clear cell renal cell carcinoma by regulating LDHA phosphorylation and modulating sensitivity to HIF2α blockade. Furthermore, FKBP10 promotes glioblastoma cell proliferation by activating the AKT–CREB–PCNA axis [70, 71]. Given the pro-oncogenic role of FKBP10 in other tumor types, our study suggests a similar mechanism may exist in WT, implying that FKBP10 could be a key mediator of WT cell invasion and metastatic potential. Direct functional studies in WT models are needed to elucidate the precise mechanisms by which FKBP10 contributes to WT pathogenesis.

SH3-domain GRB2-like endophilin B1 (SGIP1) is a less characterized protein primarily functioning as an endocytic adaptor. In addition to endocytosis, SGIP1 has been implicated in autophagy and energy homeostasis regulation [72, 73]. While it has not been as widely studied as prominent oncogenes or tumor suppressors, evidence suggests it may influence tumor cell growth, metabolism, and patient prognosis. Our study suggests that SGIP1 may be associated with immune and metabolic pathways in WT. Future functional studies are required to elucidate the precise mechanisms by which SGIP1 participates in WT pathogenesis, including its specific roles in metabolic reprogramming and immune modulation, which could provide novel targets for therapeutic intervention.

Furthermore, we found overlap in the pathway functions of the m^6 A reader protein, IGF2BP3, and several prognostic genes (e.g., *CDHR5*, *FKBP10*, *GRB7*, and *SGIP1*), particularly in core metabolic pathways like ribosome biogenesis and oxidative phosphorylation. IGF2BP3 maintains cell stemness by enhancing the stability of stemness gene mRNAs like *SOX2* [74] and promoting tumor invasion and metastasis [75, 76]. This suggests that our prognostic gene network may synergize with IGF2BP3-mediated m^6 A modification regulation, thereby reprogramming tumor cell energy metabolism to maintain tumor stem cell characteristics and malignant phenotypes.

This study suggests potential immune regulatory alterations within the TME of high-risk patients. We observed a reduction in antitumorigenic monocyte infiltration, coupled with an increase in immunosuppressive Tregs, γδ T cells, and NKTs. These immune alterations may be associated with tumor–immune interactions

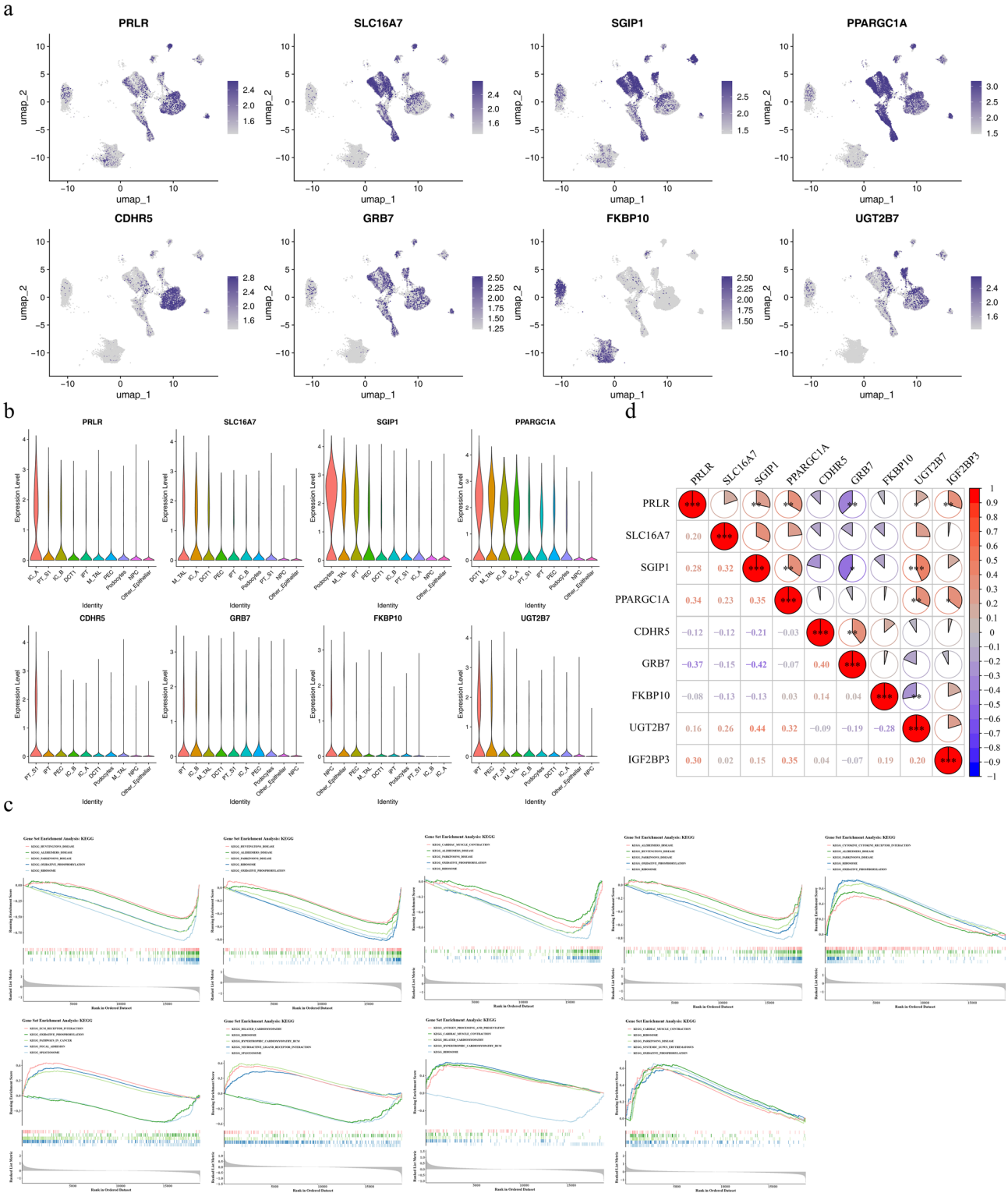


Fig. 6 Functional characterization and interaction networks of prognostic genes in Wilms tumor epithelial cells. (**a**) Spatial distribution of prognostic genes across epithelial cell clusters. The UMAP plot illustrates the expression patterns of the eight signature genes in distinct epithelial subpopulations. (**b**) Expression levels of prognostic genes in epithelial cell clusters. Violin plots depict the differential expression of each gene Across key cell types, with *GRB7* and *UGT2B7* showing peak expression in iPT cells. (**c**) Functional pathway enrichment of prognostic genes. GSEA results revealed distinct biological pathways associated with each gene, including ribosome biogenesis, oxidative phosphorylation, and neurodegenerative diseases. (**d**) Correlation network among prognostic genes and *IGF2BP3*. Spearman analysis highlighted significant co-expression relationships, with asterisks indicating statistical significance and numerical values representing correlation coefficients

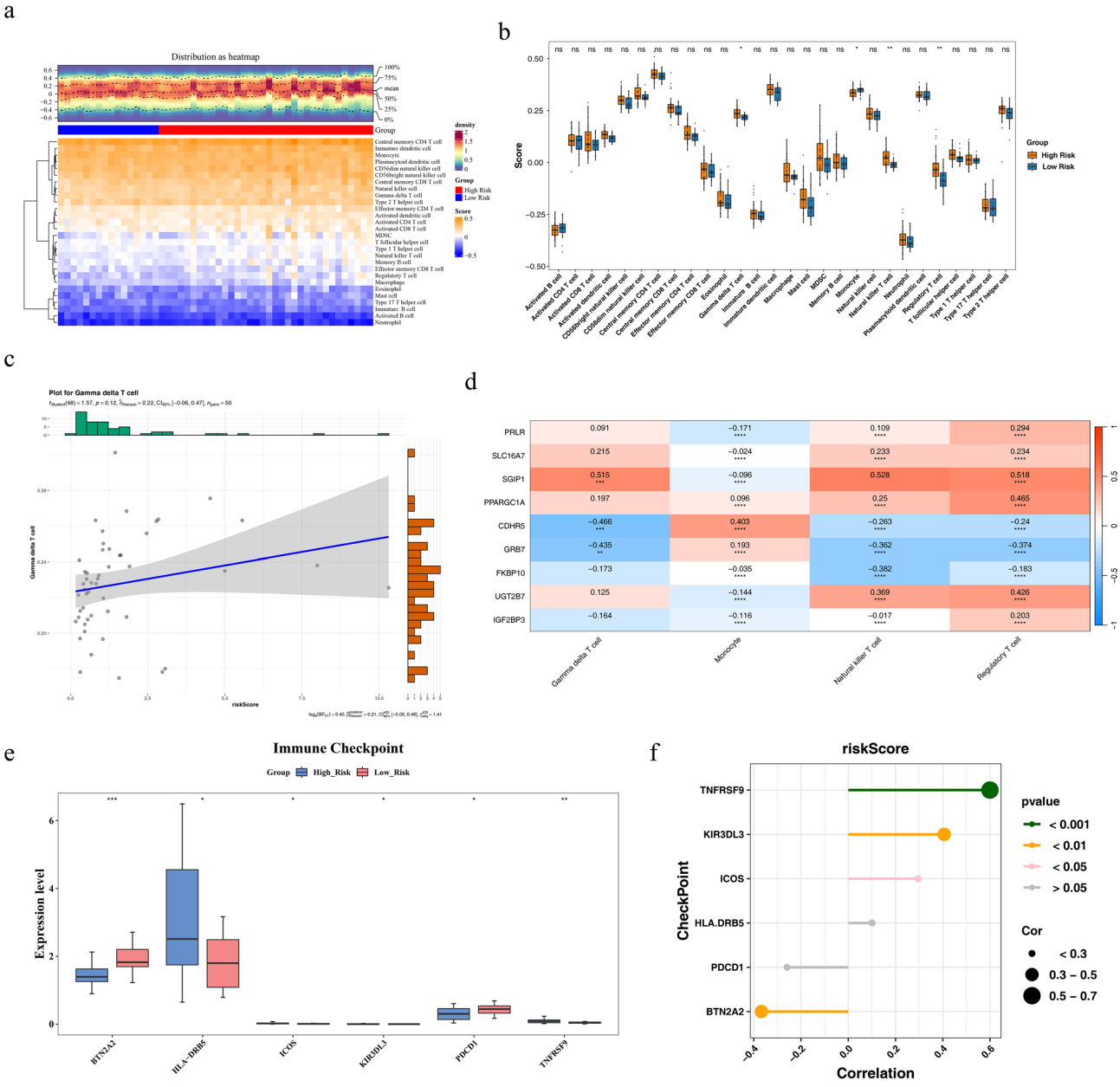


Fig. 7 Distinct immune landscape and checkpoint profiles stratify Wilms tumor risk groups. (**a-1**) Immune cell infiltration abundance in the high- and low-risk groups. Darker blue indicates lower infiltration abundance, and darker orange indicates higher infiltration abundance. (**a-2**) Differential immune cell infiltration analysis between the high- and low-risk groups. ***, **, and * denote *p*-values < 0.001, 0.01, and 0.05, respectively. (**b**) Correlation analysis between risk scores and differentially infiltrated immune cells. r-Spearman represents correlation values; the blue line indicates the trend, with a positive slope indicating a positive correlation. (**c**) Correlation analysis between prognostic genes and differentially infiltrated immune cells. ***, **, and * denote *p*-values < 0.001, 0.01, and 0.05, respectively. The numerical values represent correlation coefficients. Darker blue indicates a stronger negative correlation, and darker red indicates a stronger positive correlation. (**d**) Boxplot visualization of immune checkpoint expression between high- and low-risk groups. (**e**) Correlation analysis between risk scores and differentially expressed immune checkpoints. Colors represent different *p*-value ranges; larger dots indicate higher absolute correlation values |cor|

within the TME. We also identified a close association between the prognostic genes and immune cell infiltration. For example, monocyte infiltration was strongly positively correlated with *CDHR5* expression, suggesting a direct or indirect interaction that collectively impacts cell adhesion and tumor invasion. γδ T cells and NKT cells exhibited robust correlations with multiple prognostic genes, including *SGIP1*, *GRB7*, and *FKBP10*. This indicates that these immune cells undergo multifaceted signaling regulation within the TME and may exert either protumorigenic or antitumorigenic effects by influencing the expression of these genes. These findings

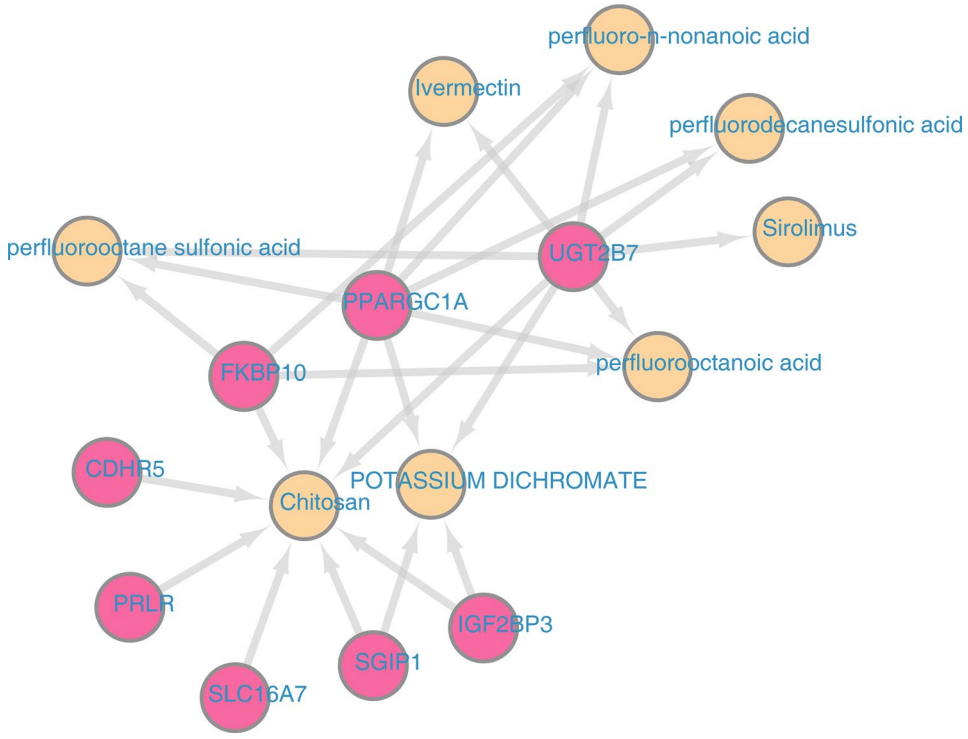


Fig. 8 Gene–drug interactions and molecular docking analysis of prognostic targets in Wilms tumor: gene–drug interaction network of prognostic genes and potential therapeutic agents. Blue nodes represent prognostic genes, and purple nodes indicate drug compounds

suggest a bidirectional cross-talk between epithelial transformation and immune evasion: tumor epithelial cells may influence immune microenvironment features by expressing specific prognostic genes. Conversely, the immune cells within this microenvironment and their secreted cytokines may provide feedback to tumor cells, promoting the maintenance and development of malignant phenotypes [77]. Intriguingly, our drug screening analysis identified E2 as a compound capable of targeting proteins encoded by six genes in our prognostic model, including *UGT2B7*, which plays a crucial role in hormone metabolism. This finding not only suggests a potentially overlooked role for hormone signaling pathways in WT but also opens new avenues for drug repurposing in WT treatment [78]. Although this requires experimental validation, it provides a preliminary basis for future studies exploring potential low-toxicity therapeutic strategies for WT. However, these drug predictions are primarily derived from computational analyses and require additional validation by experimental studies.

In summary, we comprehensively analyzed the complex cellular heterogeneity and immune microenvironment of WT by integrating single-cell and transcriptomic data. We not only unveiled the dynamic trajectory of tumor cell evolution toward an undifferentiated state but also identified an eight-gene prognostic model with potential predictive value. Crucially, this research illuminates the intricate interplay between epithelial cell plasticity and immune evasion, providing additional perspectives for understanding the malignant progression of WT. Despite our significant findings, this study has some limitations. First, the relatively limited number of WT samples with survival information ($n = 50$) may increase the risk of model overfitting. Second, this study lacked an independent external validation cohort, so the stability and generalizability of the prognostic model need to be validated in a larger dataset. ScRNA-seq analysis was only performed on three pairs of tumor–normal tissue samples, which may limit the generalizability of trajectory inference. Meanwhile, this analysis may also be affected by potential batch effects, and the results should be validated in larger single-cell cohorts. Single-cell RNA sequencing analysis used only three pairs of tumor-normal tissue samples, which may limit the generalability of trajectory inference. At the same time, this analysis may also be affected by potential batch effects. Therefore, the relevant results still need to be further validated in larger single-cell cohorts.Therefore, future research will focus on bridging this gap by conducting in-depth functional studies of the core genes in this eight-gene prognostic model. The aim is to elucidate their mechanisms of action at the molecular, cellular, and whole-organism levels, providing robust experimental support for the clinical translation of this prognostic model and the development of targeted therapies.

Conclusions

This study delineates the cellular and molecular landscape of WT using integrated single-cell and transcriptomic analyses. The enrichment and potential differentiation trajectories of tumor-specific epithelial and progenitor cells were revealed, highlighting the NRG1/3-Erbb4 signaling interaction. We established an 8-gene prognostic signature (comprised of *PRLR*, *SLC16A7*, *SGIP1*, *PPARGC1A*, *CDHR5*, *GRB7*, *FKBP10*, and *UGT2B7*), which stratifies patients into distinct risk groups with significant survival differences and was validated with time-dependent ROC analysis (AUC > 0.6). High-risk patients exhibited distinct immune microenvironment features characterized by elevated regulatory T cell infiltration and upregulated checkpoint genes (*TNFRSF9* and *KIR3DL3*). These genes play roles in metabolic reprogramming, immune evasion, and hormone response pathways. Meanwhile, the computational drug screening results suggest that chitosan may be a candidate compound worth further exploration in subsequent studies. Our findings may offer insights into WT pathogenesis and highlight a prognostic model with potential clinical relevance as well as candidate therapeutic targets. Although our results require validation, this work advances precision oncology approaches for WT management by linking cellular heterogeneity with clinical outcomes, identifying actionable molecular vulnerabilities within the tumor ecosystem.

Abbreviations

WT	Wilms tumor
TME	Tumor microenvironment
TLA	Three letter acronym
scRNA-seq	Single-cell RNA sequencing
AUC	Area under the curve
Bulk RNA-seq	Bulk RNA sequencing
GEO	Gene Expression Omnibus
TCGA	The Cancer Genome Atlas
UMAP	Uniform Manifold Approximation and Projection
BEAM	Branched Expression Analysis Modeling
cDEGs	Cluster-specific differentially expressed genes
KEGG	Kyoto Encyclopedia of Genes and Genomes
FC	Fold change
BP	Biological process
CC	Cellular component
MF	Molecular function
PH	Proportional hazards
LASSO	Least absolute shrinkage and selection operator
IGF2BP3	Insulin-like growth factor 2 mRNA-binding protein 3
DsigDB	Drug Signatures Database
PDB	Protein Data Bank
ECM	Extracellular matrix
γδ T cells	Gamma delta T cells
NKTs	Natural killer T cells
Tregs	Regulatory T cells
PT_S1	Proximal tubule S1 segment
M_TAL	Medullary thick ascending limb
PPARGC1A	Peroxisome proliferator-activated receptor gamma coactivator 1-alpha
PGC-1α	Peroxisome proliferator-activated receptor gamma coactivator 1-alpha
CDHR5	Cadherin-related family member 5
GRB7	Growth factor receptor bound protein 7
RTKs	Receptor tyrosine kinases
FKBP10	FKBP Prolyl Isomerase 10
UGT2B7	UDP-glucuronosyltransferase 2 family polypeptide B7
TME	Tumor microenvironment

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13062-026-00780-w.

Additional file 1

Additional file 2

Additional file 3

Acknowledgements

We would like to express our sincere gratitude to all individuals and organizations who supported and assisted us throughout this research. In conclusion, we extend our thanks to everyone who has supported and assisted us along the way.

Author contributions

Conceptualization,Zhiqiang Gao and Feng Liu; methodology,Jie Lin; software, Ao Li,Rui Li; validation, Ao Xu,Huafei Tang and Rui Hu; formal analysis, Shuai Xu; investigation, Maolin Zhang; resources, Wenming Yang; data curation, Haijing Huang; writing—original draft preparation,Zhiqiang Gao; writing—review and editing, Feng Liu; visualization, Zheng Zhang; supervision, Feng Liu; project administration, Feng Liu; funding acquisition, Feng Liu. All authors have read and agreed to the published version of the manuscript."

Funding

This research was funded by K Foundation of West China Second University Hospital, Sichuan University, grant number KZ317.

Data availability

The original data presented in the study are openly available in [the Gene Expression Omnibus (GEO) database] at [https://www.ncbi.nlm.nih.gov/geo/] or [GSE197047], and [the Cancer Genome Atlas (TCGA) database] at [https://portal.gdc.cancer.gov/] or [TARGET-WT].

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the affiliated Children's Hospital of Chongqing Medical University. Informed consent was obtained from all subjects involved in the study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 11 February 2026 / Accepted: 7 April 2026
Published online: 20 May 2026

References

1. Meadow W, Goldblatt AD, Lantos J. Current opinion in pediatrics: ethics and law 2001. Curr Opin Pediatrics. 2002;14(2):170–73. https://doi.org/10.1097/00008480-200204000-00005.
2. Spreafico F, Fernandez CV, Brok J, Nakata K, Vujanic G, Geller JI, et al. Wilms tumour. Nat Rev Dis Primers. 2021;7(1):75. https://doi.org/10.1038/s41572-021-00308-8.
3. Xiang B, Chen ML, Gao ZQ, Mi T, Shi Q-L, Dong J-J, et al. CCNB1 is a novel prognostic biomarker and promotes proliferation, migration and invasion in

Wilms tumor. BMC Med Genomics. 2023;16(1):189. https://doi.org/10.1186/s12920-023-01627-3.
4. Tian Xiaomao MW, Qinlin S, et al. Clinical analysis of Wilms tumors treated with CCCG-WT-2016 regimen in 43 cases. 2020;38(12):915–.
5. Ward E, Desantis C, Robbins A, Kohler B, Jemal A. Childhood and adolescent cancer statistics, 2014. CA Cancer J Clinicians. 2014;64(2):83–103. https://doi.org/10.3322/caac.21219.
6. Dome JS, Graf N, Geller JI, Fernandez CV, Mullen EA, Spreafico F, et al. Advances in Wilms tumor treatment and biology: progress through international collaboration. JCO. 2015;33(27):2999–3007. https://doi.org/10.1200/JCO.2015.62.1888.
7. Saeidian AH, Youssefian L, Vahidnezhad H, Uitto J. Research techniques made simple: whole-transcriptome sequencing by RNA-Seq for diagnosis of monogenic disorders. J Invest Dermatol. 2020;140(6):1117–26.e1. https://doi.org/10.1016/j.jid.2020.02.032.
8. Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y. Single-cell RNA sequencing technologies and applications: a brief overview. Clin Transl Med. 2022;12(3):e694. https://doi.org/10.1002/ctm2.694.
9. Binson VA, Thomas S, Subramoniam M, Arun J, Naveen S, Madhu S. A Review of machine learning algorithms for biomedical applications. Ann Biomed Eng. 2024;52(5):1159–83. https://doi.org/10.1007/s10439-024-03459-3.
10. Zhang Z, Wang ZX, Chen YX, Wu H-X, Yin L, Zhao Q, et al. Integrated analysis of single-cell and bulk RNA sequencing data reveals a pan-cancer stemness signature predicting immunotherapy response. Genome Med. 2022;14(1):45. https://doi.org/10.1186/s13073-022-01050-w.
11. Xia W, Liu J, Chen R, Junjie F, Lei W, Yize W, et al. Molecular subtypes and prognostic signature rooted in disulfidptosis highlight tumor microenvironment in lung adenocarcinoma. Chin J Cancer Res. 2025;37(5):796–820. https://doi.org/10.21147/j.issn.1000-9604.2025.05.11.
12. Gong Z, Du M, Li Y, Ye B, Huang Y, Gong H, et al. Machine learning identifies TIME subtypes linking EGFR mutations and immune states in lung adenocarcinoma. NPJ Digit Med. 2025;8(1):796. https://doi.org/10.1038/s41746-025-02172-2.
13. Zhang P, Liang X, Ye B, Xiaofei W, Yijie W, Zetian G, et al. Metabolic reprogramming signature predicts immunotherapy efficacy in lung adenocarcinoma: targeting SLC25A1 to overcome immune resistance. Chin J Cancer Res. 2025;37(6):1000–19. https://doi.org/10.21147/j.issn.1000-9604.2025.06.11.
14. Tian XM, Xiang B, Jin LM, Mi T, Wang J-K, Zhanghuang C, et al. Immune-related gene signature associates with immune landscape and predicts prognosis accurately in patients with Wilms tumour. Front. Immunol. 2022;13:920666. https://doi.org/10.3389/fimmu.2022.920666.
15. Gadd S, Huff V, Walz AL, Ooms AHAG, Armstrong AE, Gerhard DS, et al. A Children's Oncology group and TARGET initiative exploring the genetic landscape of Wilms tumor. Nat Genet. 2017;49(10):1487–94. https://doi.org/10.1038/ng.3940.
16. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with harmony. Nat Methods. 2019;16(12):1289–96. https://doi.org/10.1038/s41592-019-0619-0.
17. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol. 2019;20(2):163–72. https://doi.org/10.1038/s41590-018-0276-y.
18. Wang Y, Chen X, Tang N, Guo M, Ai D. Boosting clear cell renal carcinoma-specific drug discovery using a deep learning algorithm and single-cell analysis. IJMS. 2024;25(7):4134. https://doi.org/10.3390/ijms25074134.
19. Gustavsson EK, Zhang D, Reynolds RH, Garcia-Ruiz S, Ryten M. Ggtranscript: an R package for the visualization and interpretation of transcript isoforms using ggplot2. Bioinformatics. 2022;38(15):3844–46. https://doi.org/10.1093/bioinformatics/btac409.
20. Qiu X, Hill A, Packer J, Lin D, Ma Y-A, Trapnell C. Single-cell mRNA quantification and differential analysis with census. Nat Methods. 2017;14(3):309–15. https://doi.org/10.1038/nmeth.4150.
21. Brunson JC. Ggalluvial: layered grammar for alluvial plots. JOSS. 2020;5(49):2017. https://doi.org/10.21105/joss.02017.
22. Yu G, Wang L G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics: J Intgr Biol. 2012;16(5):284–87. https://doi.org/10.1089/omi.2011.0118.
23. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan C-H, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun. 2021;12(1):1088. https://doi.org/10.1038/s41467-021-21246-9.
24. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. Nat Protoc. 2020;15(4):1484–506. https://doi.org/10.1038/s41596-020-0292-x.
25. Griss J, Viteri G, Sidiropoulos K, Nguyen V, Fabregat A, Hermjakob H. ReactomeGSA - Efficient multi-Omics Comparative pathway analysis. Mol Cellular Proteomics. 2020;19(12):2115–25. https://doi.org/10.1074/mcp.TIR120.002155.
26. Love MI, Huber W, Anders S. Moderated estimation of Fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. https://doi.org/10.1186/s13059-014-0550-8.
27. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32(18):2847–49. https://doi.org/10.1093/bioinformatics/btw313.
28. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable venn and Euler diagrams in R. BMC Bioinf. 2011;12(1):35. https://doi.org/10.1186/1471-2105-12-35.
29. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504. https://doi.org/10.1101/gr.1239303.
30. Groeneveld CS, Chagas VS, Jones SJM, Robertson AG, Ponder BAJ, Meyer KB, et al. Rtnsurvival: an R/Bioconductor package for regulatory network survival analysis. Bioinformatics. 2019;35(21):4488–89. https://doi.org/10.1093/bioinformatics/btz229.
31. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization Paths for Cox's proportional hazards model via Coordinate Descent. J Stat Soft. 2011;39(5):1–13. https://doi.org/10.18637/jss.v039.i05.
32. Liu TT, Li R, Huo C, Li J-P, Yao J, Ji X-L, et al. Identification of CDK2-related immune forecast model and ceRNA in lung adenocarcinoma, a pan-cancer analysis. Front Cell Dev Biol. 2021;9:682002. https://doi.org/10.3389/fcell.2021.682002.
33. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat Med. 2013;32(30):5381–97. https://doi.org/10.1002/sim.5958.
34. Robles-Jimenez LE, Aranda-Aguirre E, Castelan-Ortega OA, Shettino-Bermudez BS, Ortiz-Salinas R, Miranda M, et al. Worldwide traceability of antibiotic residues from livestock in wastewater and soil: a systematic Review. Anim (Basel). 2021;12(1):60. https://doi.org/10.3390/ani12010060.
35. Xia MD, Yu RR, Chen DM. Identification of hub biomarkers and immune-related pathways participating in the progression of antineutrophil cytoplasmic antibody-associated glomerulonephritis. Front Immunol. 2022;12:809325. https://doi.org/10.3389/fimmu.2021.809325.
36. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinf. 2013;14(1):7. https://doi.org/10.1186/1471-2105-14-7.
37. Hu FF, Liu CJ, Liu LL, Zhang Q, Guo A-Y. Expression profile of immune checkpoint genes and their roles in predicting immunotherapy response. Briefings Bioinf. 2021;22(3). https://doi.org/10.1093/bib/bbaa176.
38. Saltzman AF, Cost NG, Romao RLP. Wilms tumor. Urol Clin N Am. 2023;50(3):455–64. https://doi.org/10.1016/j.ucl.2023.04.008.
39. Hont AB, Dumont B, Sutton KS, Anderson J, Kentsis A, Drost J, et al. The tumor microenvironment and immune targeting therapy in pediatric renal tumors. Pediatr Blood Cancer. 2023;70(S2):e30110. https://doi.org/10.1002/pbc.30110.
40. Wang Z, Jin L, Wang J, Tian X, Mi T, Li M, et al. Recruitment and polarization typing of tumor-associated macrophages is associated with tumor progression and poor prognosis in Wilms tumor patients. PLoS One. 2024;19(11):e0309910. https://doi.org/10.1371/journal.pone.0309910.
41. Hall AM, De Seigneux S. Metabolic mechanisms of acute proximal tubular injury. Pflugers Arch - Eur J Physiol. 2022;474(8):813–27. https://doi.org/10.1007/s00424-022-02701-y.
42. Dimke H, Schnermann J. Axial and cellular heterogeneity in electrolyte transport pathways along the thick ascending limb. Acta Physiologica. 2018;223(1):e13057. https://doi.org/10.1111/apha.13057.
43. Allen B, Savoy L, Ryabinin P, Bottomly D, Chen R, Goff B, et al. Upregulation of HOXA3 by isoform-specific Wilms tumour 1 drives chemotherapy resistance in acute myeloid leukaemia. Br J Haematol. 2024;205(1):207–19. https://doi.org/10.1111/bjh.19563.
44. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57–70. https://doi.org/10.1016/S0092-8674(00)81683-9.
45. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646–74. https://doi.org/10.1016/j.cell.2011.02.013.

46. Hanahan D. Hallmarks of cancer: new dimensions. Cancer Discov. 2022;12(1):31–46. https://doi.org/10.1158/2159-8290.CD-21-1059.
47. Huang J, Li Y, Pan X, Wei J, Xu Q, Zheng Y, et al. Construction of a Wilms tumor risk model based on machine learning and identification of cuproptosis-related clusters. BMC Med Inf Decis Mak. 2024;24(1):325. https://doi.org/10.1186/s12911-024-02716-8.
48. Li D, Ding C, Huang F, CHANG M, LU Z, LI G, et al. Bioinformatics analysis of the expression and prognostic value of PLAU gene in Wilms' tumor. Anticancer Res. 2024;44(9):3829–42. https://doi.org/10.21873/anticanres.17209.
49. Deng DX, Wen JJ, Cheng YF, Zhang X-H, Xu L-P, Wang Y, et al. Wilms' tumor gene 1 is an independent prognostic factor for pediatric acute myeloid leukemia following allogeneic hematopoietic stem cell transplantation. BMC Cancer. 2021;21(1):292. https://doi.org/10.1186/s12885-021-08022-0.
50. Bole-Feysot C, Goffin V, Edery M, Binart N, Kelly PA. Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. Endocr Rev. 1998;19(3):225–68. https://doi.org/10.1210/edrv.19.3.0334.
51. Goffin V, Binart N, Touraine P, Kelly PA. Prolactin: the new biology of an old hormone. Annu Rev Physiol. 2002;64(1):47–67. https://doi.org/10.1146/annurev.physiol.64.081501.131049.
52. Schmidt LS, Linehan W M. Genetic predisposition to kidney cancer. Semin Oncol. 2016;43(5):566–74. https://doi.org/10.1053/j.seminoncol.2016.09.001.
53. Di Chiaro P, Nacci L, Arco F, Brandini S, Polletti S, Palamidessi A, et al. Mapping functional to morphological variation reveals the basis of regional extracellular matrix subversion and nerve invasion in pancreatic cancer. Cancer Cell. 2024;42(4):662–81.e10. https://doi.org/10.1016/j.ccell.2024.02.017.
54. Frank DA. STAT3 as a central mediator of neoplastic cellular transformation. Cancer Lett. 2007;251(2):199–210. https://doi.org/10.1016/j.canlet.2006.10.017.
55. Halestrap AP. The SLC16 gene family – structure, role and regulation in health and disease. Mol Aspects Med. 2013;34(2–3):337–49. https://doi.org/10.1016/j.mam.2012.05.003.
56. Xu M, Zhou J, Lv J, Zhang Y. Tumor suppressing function of SLC16A7 in bladder cancer and its pan-cancer analysis. BMC Cancer. 2025;25(1):932. https://doi.org/10.1186/s12885-025-14345-z.
57. Pertega-Gomes N, Vizcaino J R, Felisbino S, Warren AY, Shaw G, Kay J, et al. Epigenetic and oncogenic regulation of SLC16A7 (MCT2) results in protein over-expression, impacting on signalling and cellular phenotypes in prostate cancer. Oncotarget. 2015;6(25):21675–84. https://doi.org/10.18632/oncotarget.4328.
58. Xu J, Zhou Y, He S, Wang Y, Ma J, Li C, et al. Activation of the YY1-UGT2B7 axis promotes mammary estrogen homeostasis dysregulation and exacerbates breast tumor metastasis. Drug Metab Dispos. 2024;52(5):408–21. https://doi.org/10.1124/dmd.124.001640.
59. Puigserver P, Spiegelman BM. Peroxisome proliferator-activated receptor-γ coactivator 1α (PGC-1α): transcriptional coactivator and metabolic regulator. Endocr Rev. 2003;24(1):78–90. https://doi.org/10.1210/er.2002-0012.
60. Qian X, Li X, Shi Z, Bai X, Xia Y, Zheng Y, et al. KDM3A senses oxygen availability to regulate PGC-1α-Mediated mitochondrial biogenesis. Mol Cell. 2019;76(6):885–95.e7. https://doi.org/10.1016/j.molcel.2019.09.019.
61. Ma T, Meng L, Wang X, Tian Z, Wang J, Liu X, et al. TNFSF13B and PPARGC1A expression is associated with tumor-infiltrating immune cell abundance and prognosis in clear cell renal cell carcinoma. Am J Transl Res. 2021;13(10):11048–64.
62. Ding X, Tian X, Liu W, Li Z. CDHR5 inhibits proliferation of hepatocellular carcinoma and predicts clinical prognosis. Ir J Med Sci. 2020;189(2):439–47. https://doi.org/10.1007/s11845-019-02092-7.
63. Bläsius FM, Meller S, Stephan C, Bläsius FM, Jung K, Ellinger J, et al. Loss of cadherin related family member 5 (CDHR5) expression in clear cell renal cell carcinoma is a prognostic marker of disease progression. Oncotarget. 2017;8(43):75076–86. https://doi.org/10.18632/oncotarget.20507.
64. Wang G, Wu Y, Su Y, Qu N, Chen B, Zhou D, et al. TCF12-regulated GRB7 facilitates the HER2+ breast cancer progression by activating Notch1 signaling pathway. J Transl Med. 2024;22(1):745. https://doi.org/10.1186/s12967-024-05536-6.
65. Chu PY, Tai YL, Shen TL. Grb7, a critical mediator of EGFR/ErbB signaling, in cancer development and as a potential therapeutic target. Cells. 2019;8(5):435. https://doi.org/10.3390/cells8050435.
66. Alcalde J, González-Muñoz M, Villalobo A. Grb7-derived calmodulin-binding peptides inhibit proliferation, migration and invasiveness of tumor cells while they enhance attachment to the substrate. Heliyon. 2020;6(5):e03922. https://doi.org/10.1016/j.heliyon.2020.e03922.
67. Zheng Y, Pei Y, Yang L, Zeng Z, Wang J, Xie G, et al. Upregulated GRB7 promotes proliferation and tumorigenesis of bladder cancer via phospho-AKT pathway. Int J Biol Sci. 2020;16(16):3221–30. https://doi.org/10.7150/ijbs.49410.
68. Ramadori G, Ioris RM, Villanyi Z, Firnkes R, Panasenko OO, Allen G, et al. FKBP10 regulates protein translation to sustain lung cancer growth. Cell Rep. 2020;30(11):3851–63.e6. https://doi.org/10.1016/j.celrep.2020.02.082.
69. Fu Y, Chen J, Ma X, Chang W, Zhang X, Liu Y, et al. Subcellular expression patterns of FKBP Prolyl Isomerase 10 (FKBP10) in colorectal cancer and its clinical significance. IJMS. 2023;24(14):11415. https://doi.org/10.3390/ijms241411415.
70. Liu R, Zou Z, Chen L, Feng Y, Ye J, Deng Y, et al. FKBP10 promotes clear cell renal cell carcinoma progression and regulates sensitivity to the HIF2α blockade by facilitating LDHA phosphorylation. Cell Death Dis. 2024;15(1):64. https://doi.org/10.1038/s41419-024-06450-x.
71. Cai HQ, Zhang MJ, Cheng ZJ, Yu J, Yuan Q, Zhang J, et al. FKBP10 promotes proliferation of glioma cells via activating AKT-CREB-PCNA axis. J Biomed Sci. 2021;28(1):13. https://doi.org/10.1186/s12929-020-00705-3.
72. Yan Z, Yan C. LINC01176 Hinders Thyroid cancer progression by Sponging miR-146b-5pto enhance SGIP1. EMIDDT. 2023;23(13):1637–48. https://doi.org/10.2174/1871530323666230417083447.
73. Aydin B, Arga KY. Co-expression network analysis elucidated a core module in association with prognosis of Non-functioning Non-invasive human pituitary adenoma. Front. Endocrinol. 2019;10:361. https://doi.org/10.3389/fendo.2019.00361.
74. Lu Y, Zhu J, Zhang Y, Li W, Xiong Y, Fan Y, et al. Lactylation-Driven IGF2BP3-Mediated serine metabolism reprogramming and RNA m6A—modification promotes lenvatinib resistance in HCC. Adv Sci. 2024;11(46):e2401399. https://doi.org/10.1002/advs.202401399.
75. Lin Z, Li J, Zhang J, Feng W, Lu J, Ma X, et al. Metabolic reprogramming Driven by IGF2BP3 promotes acquired resistance to EGFR inhibitors in Non–small cell lung cancer. Cancer Res. 2023;83(13):2187–207. https://doi.org/10.1158/0008-5472.CAN-22-3059.
76. Yang X, Bai Q, Chen W, Liang J, Wang F, Gu W, et al. M 6 A-Dependent modulation via IGF2BP3/MCM5/Notch axis promotes Partial EMT and LUAD metastasis. Adv Sci. 2023;10(20):e2206744. https://doi.org/10.1002/advs.202206744.
77. Sawada K, Chung H, Softic S, Moreno-Fernandez ME, Divanovic S. The bidirectional immune crosstalk in metabolic dysfunction-associated steatotic liver disease. Cell Metab. 2023;35(11):1852–71. https://doi.org/10.1016/j.cmet.2023.10.009.
78. Pushpakom S, Iorio F, Eyers PA, Escott KJ, Hopper S, Wells A, et al. Drug repurposing: progress, challenges and recommendations. Nat Rev Drug Discov. 2019;18(1):41–58. https://doi.org/10.1038/nrd.2018.168.

Publisher's Note