TITLE:

Deep visual multi-omics profiling reveals mechanisms that underly cancer cell differentiation and aggressiveness in clear cell renal cell carcinoma

Hella A. Bolck[1,2,*] Ede Migh[3], Andras Kriston[4], Natalia H. Zajac[5,6], Susanne Kreutzer[5], Tiberiu Totu[6,7,8], Peter Leary[5,6], Ferenc Kovacs[4], Dorothea Rutishauser[1], Sybille Pfammatter[5], Jonas Grossmann[5,6], Cassandra Litchfield[1], Marija Buljan[6,7], Niels J. Rupp[1], Peter Horvath[3,4,9*], Holger Moch[1*]

1) Department of Pathology and Molecular Pathology, University Hospital Zürich and University of Zürich, Zürich, Switzerland
2) Centre for Artificial Intelligence, Zürich University of Applied Sciences, Winterthur, Switzerland
3) Synthetic and Systems Biology Unit, Biological Research Centre, HUN-REN, Szeged, Hungary,
4) Single-Cell Technologies Ltd., Szeged, Hungary
5) Functional Genomics Center Zürich, Eidgenössische Technische Hochschule Zürich/University of Zürich, Zürich, Switzerland.
6) Swiss Institute of Bioinformatics, Lausanne, Switzerland
7) Nanomaterials in Health Laboratory, Swiss Federal Laboratories for Materials Science and Technology, St. Gallen, Switzerland
8) Department of Health Sciences and Technology, Eidgenössische Technische Hochschule Zürich, Zürich, Switzerland
9) Institute of AI for Health, Helmholtz Zentrum München, Neuherberg, Germany
*) Corresponding authors: Hella.Bolck@zhaw.ch, horvath.peter@brc.hu, holger.moch@usz.ch

WORKING TITLE:

Deep visual multi-omics analysis of clinically relevant ccRCC cells

Summary

Clear cell renal cell carcinoma (ccRCC) exhibits significant intra-tumoral heterogeneity (ITH) at both morphological and genetic levels, complicating treatment and contributing to disease progression[1,2]. Among these, rhabdoid ccRCCs stand out as highly aggressive tumors distinguished by cells with unique morphological features and poor clinical outcomes[3]. However, the correlation between distinct morphological phenotypes, specific molecular alterations, and their influence on tumor behavior remains poorly understood. Understanding the link between clinically relevant morphological features and their molecular underpinnings will be critical for developing more effective treatments targeting key cellular subsets responsible for tumor progression.

In this study, we integrated advanced AI-based image analysis with single-cell isolation and multi-omics profiling to dissect the link between clinically relevant morphological and molecular features of ccRCC

cells. Using a novel digital pathology workflow, we quantified low-grade, high-grade, and rhabdoid morphologies in ccRCC diagnostic images with unprecedented precision. Subsequently, isolation of two sets of 1,000 morphologically distinct cells for detailed mRNA and protein expression analyses, revealed significant increasing dysregulation associating with higher histopathological grades. Rhabdoid ccRCC cells (grade 4) demonstrated distinct molecular profiles, including upregulated FOXM1-driven proliferation, disrupted cell-matrix interactions, and enhanced immune evasion pathways. Despite high T-cell infiltration in rhabdoid areas, we identified a rhabdoid-specific immunosuppressive network driven by cytokines, IFN-beta, and integrin signaling, likely contributing to T-cell exhaustion. These findings provide a basis for novel therapeutic strategies targeting these pathways in combination with immunotherapy to improve outcomes for patients with aggressive rhabdoid ccRCC.

Key Points:

- ccRCC is characterized by well-established morphological heterogeneity but the correlation with the underlying molecular aberrations remained elusive.
- By integrating AI-based image analysis with single cell isolation and deep multi-omics profiling, we dissect the molecular intricacies of ccRCC, from targeted collection of 1,000 morphologically distinct cells.
- Our results demonstrate significant dysregulation of gene and protein expression correlating with higher histopathological grades in ccRCC.
- Aggressive ccRCC cells with rhabdoid differentiation (grade 4) display distinct molecular profiles, as they upregulate FOXM1-mediated proliferation, ECM remodeling and the immune evasion responses, suggesting new therapeutic avenues enhancing ICI efficacy in these patients.

Introduction

Renal cell carcinoma (RCC) comprises a spectrum of subtypes defined by a specific phenotype or molecular background[4]. Clear cell RCC (ccRCC) is the most common subtype, accounting for the majority of metastatic cases[5,6]. The hallmark of ccRCC is the loss of chromosome 3p concurrent with inactivation of the VHL tumor suppressor gene. VHL loss causes upregulation of hypoxic signaling and alters the functions of numerous pathways controlling cell survival, energy production, and angiogenesis. Thus, broad metabolic rewiring is a hallmark of ccRCC[7]. Additionally, these initiating genetic events are key drivers of clonal evolution, contributing to the profound genetic intra-tumor heterogeneity (ITH) of ccRCCs[1,8,9]. The metastatic potential of ccRCC is often indicated by morphological signs of tumor cell differentiation. The WHO/ISUP grading system has been developed to assess tumor aggressiveness and to guide treatment decisions[10,11]. This four-tiered system primarily assesses nucleolar prominence, with the highest grade (grade 4) being defined by distinct histological features such as sarcomatoid or rhabdoid differentiation[11]. Frequently, ccRCC present low- (grade 1/2) and high-grade (grade 3/4) phenotypes within the same tumor[12] challenging the precision of the WHO/ISUP grading system for individual patient prognosis[13]. Moreover, the link between this morphological ITH and the underlying molecular alterations remains unclear. Given the role of ITH in tumor evolution, adaptation, and drug resistance[14], integrating histological and molecular complexities is crucial for accurate prediction of prognosis, and effective treatment strategies.

Rhabdoid and sarcomatoid differentiation in ccRCC cells is a strong predictor of poor patient prognosis[11], as highlighted in the 2016 WHO/ISUP grading system[11]. Morphologically, the less-studied grade 4 rhabdoid ccRCC cells resembles rhabdomyoblasts and are observed in 3-7% of all RCCs[3,15,16] Recent clinical studies have shown that grade 4 RCC with sarcomatoid or rhaboid differentiation often exhibit significant responses to immune checkpoint inhibitor therapy (ICI)[17–20]. This is potentially linked to the with increased tumor-infiltrating lymphocytes (TILs) frequently seen in these high-grade tumors[5,21]. However, the distinct biology of grade 4 ccRCCs with rhabdoid differentiation and the mechanisms behind their ICI response remain largely unknown.

Advanced computational methods leveraging artificial intelligence (AI) and machine learning (ML) for recognizing histo-morphologically relevant phenotypes such as nuclear grades or rhabdoid differentiation are able to classify millions of cells from numerous tissue sections[22]. Their robustness

and efficiency offer a solution to expert inter-observer variability. Moreover they help to overcome the limitations of manual annotations at the single-cell level[22]. Single-cell -omics technologies, especially single-cell RNA sequencing (RNAseq), have been transformative in revealing molecular heterogeneity across tumor tissues[23,24]. Similarly, mass spectrometry (MS)-based proteomic techniques now allow for protein measurements at single cell resolution, further contributing to our understanding of cell-to-cell variation within tumors[25]. However, integrating single-cell histological classification with comprehensive molecular data remains challenging. Recent advancements have facilitated single-cell spatial characterization of tumor cells through AI-guided image analysis and ultra-sensitive proteomic workflows[26], but routine diagnostic hematoxylin and eosin (HE) staining has not been incorporated into these efforts. Additionally, previous work did not concurrently derive both protein and mRNA signatures to capture multi-omics profiles of deregulation in clinically relevant tumor subpopulations.

Here, we have integrated pathology-guided AI-based image analysis of diagnostic ccRCC samples with single-cell isolation and ultra-sensitive spatially resolved multi-omics data collection. This approach quantifies the expression levels of thousands of genes and proteins within histologically defined, clinically relevant cell subsets. We focused on delineating the molecular causes of rhabdoid differentiation, unravelling the relationships between low- and high-grade components coexisting within the same tumors. The insights derived from this comprehensive dataset will promote the identification of appropriate treatment strategies for ccRCC with rhabdoid differentiation and provide new perspectives particularly on the benefits of immune checkpoint inhibitors (ICI) for ccRCC patients with rhabdoid differentiation.

Results

Genome-wide genetic profiling of five ccRCCs with pronounced histo-morphological intra-tumor heterogeneity (ITH)

We selected five ccRCC cases exhibiting pronounced tumor-intrinsic micro-heterogeneity defined by different grade morphologies within a tumor (Fig. 1A and Suppl. Fig. 1A, Suppl. Table 1). Each sample underwent expert pathological assessment, confirming the co-occurrence of grade 1/2 (low grade), grade 3 (high grade), and rhabdoid (grade 4) ccRCC areas. Subsequently, we dissected gross tumor areas from all morphologies and adjacent normal tissue (Suppl. Fig. 1B) and subjected them to whole-genome DNA sequencing (WGS) (Suppl. Fig. 1C and 1D). Our analysis revealed that the tumor

mutational burden (TMB) remained consistent across samples aligning with previously reported values in ccRCC with and without rhabdoid features[8,20] (Fig. 1B). Arm-level chromosomal aberrations included the most prevalent structural alterations found in ccRCC, 3p loss and 5q gain (Fig. 1B and Suppl. Fig. 1E)[9]. Additionally, we consistently identified loss of chromosome 14q, a copy number variation (CNV) that has been associated with aggressive disease [27,28]. Our findings also include previously reported CNVs, such as gains in chromosomes 7q and 9p, as well as loss of chromosome 8q [28–30]. Notably, all ccRCC samples exhibited mutations in the *VHL* tumor suppressor gene (N=5) and harbored additional mutations in SETD2 (N=4), PBRM1 (N=4), BAP1 (N=2) as well as other frequently mutated ccRCC drivers (Fig. 1B, Suppl. Table 2 and 3)[8,28]. Thus, our genomic characterization largely corroborated findings from prior studies, providing confirmation of the typical molecular profiles of ccRCC samples.

A deep digital pathology workflow quantified morphological heterogeneity in diagnostic images

Next, we used artificial intelligence-based image analysis on the HE-stained diagnostic whole-slide images (WSIs) of these cases to quantify morphological heterogeneity automatically (Fig. 1C and 1D). This involved the BIAS software suite[26] for deep learning-based segmentation and phenotype classification. A pre-trained neural network segmented nuclei and subsequently features from both nuclei and surrounding areas were extracted[31]. An artificial neural network (ANN) classified regions as ccRCC or normal kidney tissue. A Mask R convolutional neural network (R-CNN) was trained by manual annotation of more than 1200 nucleoli to recognize grade 3 ccRCC cells (Suppl. Table 4). Similarly, we developed a model to segment rhabdoid differentiation in ccRCC cells (Grade 4) using a fine-tuned Mask R CNN that was trained by annotating more than 250 rhabdoid ccRCC cells (Suppl. Table 4). The results of these DL-based segmentation algorithms were integrated as additional features of putative tumor cells to train a classifier distinguishing grade 1/2 (low-grade), grade 3 (high-grade), and grade 4 (rhabdoid) ccRCC cells (Fig. 1D and 1E). Subsequently, we predicted these classes in two to three WSIs from each patient in our cohort and validated the model using a 10-fold self-validation test and on an unseen dataset in comparison to pathologist-annotated ground truth (Fig. 1F and 1G). Applying out workflow, we subsequently selected two sets of 1000 cells confidently classified as either grade 1/2 (low-grade), grade 3 (high-grade), grade 4 ccRCC cells with rhabdoid differentiation or normal kidney tubuli for each of the five ccRCC patient tissues and performed single-cell isolation (Fig. 1C).

mRNA and protein-level alterations in ccRCC cells with morphologic heterogeneity reflect core disease genes and pathways.

Following cell isolation, one replicate pool per morphological class was subjected to RNAseq analysis employing a protocol tailored for ultra-low sample inputs, which directly utilizes total RNA for library preparation input and relies on random priming. As a result, we obtained RNA expression data that mapped to 20,528 genes (Fig. 2A and Suppl. Table 5). The normal tubuli cells for patient 4 were found to display substantial tubular atrophy and interstitial fibrosis and were therefore excluded from further analysis. Principle component analysis (PCA) and hierarchical clustering analysis of the gene expression data revealed a clear distinction between the profiles of adjacent normal tubuli and ccRCC cells of various morphologies (Fig. 2A and Suppl. Fig. 2A). The PCA distances between ccRCC histologies were notably lower aligning with the expected similarity between these samples. However, a discernible trend of separation emerged between low- and higher-grade morphologies, particularly along PC1. Subsequently, we compared each of the ccRCC grade groups to the adjacent normal cells to determine differentially expressed genes (DEGs). The analysis revealed a significant increase of the number of both down- and upregulated mRNAs with higher grades (FDR <0.05 and |Log2FC| > 1) (Fig. 2B). Notably, the number of deregulated genes exhibited an over 4-fold increase between grade 1/2 ccRCC cells and the highest grade 4 rhabdoid cells, indicating a progressive dysregulation pattern.

Next, we investigated proteome alterations in ccRCC cells with heterogeneous histologies by subjecting the second pool of isolated cells to LC-MS adapting a highly sensitivity workflow compatible with ultra-low sample inputs[25]. Employing data-independent acquisition with parallel accumulation–serial fragmentation (diaPASEF) with an additional ion mobility dimension (TIMS), we successfully identified 4,382 proteins. Six samples were excluded from further analysis due to irregularities during laser microdissection (LMD) that resulted in lower protein counts and noticeable batch effects (Suppl. Table 5). PCA and hierarchical clustering analysis of the remaining proteomics data confirmed a distinct separation of the normal tubuli and all morphological classes of ccRCC cells (Fig. 2C and Suppl. Fig. 2B). Notably, the gradual transition among the ccRCC samples with increasing grades, as expressed by PCA distances, was more pronounced than observed in the transcriptomic data, particularly between cells with low (Grade 1/2) and higher (Grade 3/4) grades (Fig. 2C). Moreover, differential expression analysis between ccRCC grade morphologies and their adjacent normal tubuli cells revealed an

increasing number of differentially expressed proteins (DEPs) correlating with higher ccRCC grade, aligning with the trend observed in the transcriptomic data (Fig. 2D).

Overrepresentation analysis of the respective datasets revealed that downregulation of biological processes was more prominent than upregulation. Biological enrichment in DEG and DEP downregulated datasets uncovered terms mirroring known ccRCC-specific tumorigenic changes including oxidative phosphorylation and tricarboxylic acid cycle (TCA). Similarly, pathways related to normal kidney function (kidney development, glucuronidation, proton transmembrane transport) were increasingly downregulated with higher ccRCC grades (Fig. 2E and 2F, Suppl. Table 6 and 7). In contrast, biological pathways enriched in the upregulated datasets included response to hypoxia, DNA replication and cell-cell signaling. Notably, among the pathways upregulated specifically in grade 4/rhabdoid cells, we identified the immune response and related processes.

Taken together, our data established that AI-driven automated classification of ccRCC cells with distinct histological features followed by their individual LMD from formalin-fixed paraffin-embedded (FFPE) tissue from the classical diagnostic workup yielded suitable samples for deep molecular profiling.

Functional overlap between mRNA and protein dysregulation reveals progressive tumorigenic changes associate with increasing tumor grades

Even though regulation of transcription and protein expression can be distinct, we observed notable correlation between mRNA and protein differential expression in our data (Fig. 3A). We detected 40, 134 and 245 significant genes shared across the mRNA and protein data for grade 1/2, grade 3 and grade 4/rhabdoid cells, respectively (Suppl. Fig. 3A). Importantly, Vimentin (VIM) a key diagnostic marker indicating epithelial-to-mesenchymal transition (EMT) was found strongly elevated in both datasets, while other well-established ccRCC markers such as CA9 and CAV1, showed strong positivity in at least one of the -omics layers (Fig. 3A). Despite their overlap, the majority of regulated genes were distinct to one of the two data layers (Fig. 3A and Suppl. Fig. 3A) prompting us to investigate whether the regulated genes drive similar biological functions (Fig. 3B and Suppl. Fig. 3B). Enrichment analyses uncovered widespread metabolic reprogramming in glucose and lipid metabolism as well as the repression of pathways related to physiological kidney functions in both datasets. Interestingly, often these alterations were observed across all ccRCC grades (Suppl. Fig. 3B). However, the number of overlapping biological pathways increased concomitantly with tumor grade, reinforcing the link between

dysregulation and higher-grade morphologies[7]. The only upregulated biological pathway that was shared by mRNA and protein datasets and distinct to ccRCC cells with rhabdoid differentiation was the "immune response" suggesting a crucial role for the transformation and aggressiveness of ccRCC,

Consistent with the almost universal loss of VHL in ccRCC, upregulation of HIF-1 signaling was evident across all grade groups at both the mRNA and protein levels and HIF1α transcriptional targets, such as VEGFA and GLUT1 (SLC2A1), showed significant increases in both -omics layers. (Fig. 3C). Concomitantly, HIF-signaling-mediated metabolic adaptations were evident from the data indicated by elevated expression of glycolytic enzymes across all ccRCC morphologies (e.g., ENO2, PFKB, PKM) and repression of oxidative phosphorylation, signified by downregulation of electron transport chain enzymes (NDUFS1, NDUFV1 – Complex I; COX5A, COX6C – Complex IV; ATP6V1A, ATP6VB1 – Complex V) and TCA cycle components (ACO2, SDHA) (Fig. 3C). Moreover, we observed significant alterations in lipid metabolism, with fatty acid beta-oxidation enzymes (ACAT1, ECHS1, EHHADH) being among the most significantly downregulated processes in both datasets. We also found evidence for progressive dysregulation of the one-carbon metabolism, a crucial pathway for tumor progression that has been recently described in relation to ccRCC aggressiveness (Suppl. Fig. 3C)[32]. Reflecting the progressive loss of cellular identity during tumorigenesis and morphological transition into higher-grade components, we observed a gradual downregulation of proximal renal tubule genes and enzymes related to physiological kidney functions, such as glucuronidation and transmembrane transport, which correlated with higher tumor grades (Suppl. Fig. 3C). Furthermore, grade 3 and grade 4/rhabdoid cells exhibited significant upregulation of essential stem cell markers such as CD44 and DCLK1[33,34].

Integrated network analysis unveils cell-matrix interactions and enhanced proliferation are key alterations characterizing rhabdoid ccRCC cells

As the direct overlap between mRNA and protein expression was limited (Fig. 3B), we employed integrative network-based analysis of transcriptomic and proteomic profiles to assess whether dysregulation on both layers would point to similar deregulated processes, offering deeper insights into the mechanisms underlying ccRCC tumor progression and histological heterogeneity. Specifically, we combined differentially expressed elements for both data layers and constructed grade-specific networks based on knowledge-based high confidence protein-protein interactions (PPI) networks. After constructing the networks, we used the current flow betweenness centrality metric in order to identify

central nodes that connect elements in each network[35]. The top 30% most central elements of each network were further examined (Fig. 4A and Suppl. Table 10). Specifically, in the grade 4/rhabdoid-specific network, downregulated central nodes included factors implicated in renal functions (e.g. ACE2, NHERF1, ALDH3A2) as well as cell adhesion and EMT (e.g. CDH1). Conversely, in the network of upregulated factors, central elements reflected cellular adaptations related to processes such as EMT, cell migration, adhesion (e.g., VIM, ANXA2), and the Warburg effect (e.g., SLC2A1). Thus, the central network nodes that integrated contributions of mRNA and protein dysregulations corroborated the reprogramming of cellular processes towards a more aggressive and metastatic phenotype in grade 4/rhabdoid cells. We next applied MONET, a network modularization tool, which was constructed based on top performing algorithms for identifying modules (i.e. communities or subnetworks) of disease-related genes in the networks. We subsequently uncovered functionally related network modules associating with specific biological processes that could describe distinct changes in ccRCC cells with different grade morphologies (Fig. 4B). In Grade 4/rhabdoid ccRCC cells we determined a number of unique modules with biological functions not shared by the other morphologies (Fig. 4B and Suppl. Fig. 4A). Downregulated modules contained largely metabolic processes previously associated with ccRCC tumorigenesis underscoring that progressive transcriptional and translational dysregulation correlated with higher tumor grades.

Pathways associated with upregulated elements provided important and novel insights into the biology of rhabdoid differentiation. For instance, analysis of cellular regulators that were identified with a high centrality in the network constructed for cells with rhabdoid differentiation highlighted Forkhead box transcription factor 1 (FOXM1) as a prominent hub in Module IV (Fig. 5A). As FOXM1 plays a critical role in regulating the cell cycle, it is implicated in the control of genes essential for cell proliferation and progression (TRRUST v2)[36]. Of note, this node was of much lower significance in low- and high-grade ccRCC networks (Fig. 5A and Suppl. Fig. 4A) indicating an elevated potential for proliferation in cells with the rhabdoid phenotype. Importantly, high FOXM1 expression was associated with poor prognosis in the TCGA-KIRC cohort (Suppl. Fig. 4B) underscoring the relevance of its overexpression for patient prognosis. Additionally, the cell cycle subnetwork in rhabdoid cells encompassed several molecules that function downstream of the FOXM1 transcription factor such as DNA replication factors (CLSPN, CDC6), elements implicated in the G1/S- (CDC25A, CCNA2, CDK3) and G2/M-checkpoints (PLK1,

CDC25A/C, BUB1) (Fig. 5A, 5B and Suppl. Fig. 4C). Their dysregulation likely contributes to the increased proliferative potential of rhabdoid ccRCC cells. Additionally, in this subnetwork, we observed rhabdoid ccRCC cell-specific upregulation of BIRC5 (Fig. 5B), a well-established inhibitor of apoptosis protein (IAP).

Module V in the G4/rhabdoid-specific network encompassed elements involved in focal adhesion, integrin cell surface interactions, homeostasis and proteoglycans implicated in cancer (Fig. 5C, Suppl. Table 10) thus emphasizing the significance of tumor-cell driven re-organization of the extracellular matrix (ECM) for ccRCC progression and metastatic dissemination. This module contained several genes that were specific for the grade 4/rhabdoid network and that had high centrality scores. Among these, were integrins (ITGA4, ITGA10, ITGB2, ITGB7) with known roles in influencing tumor growth, metastasis, and the interactions within the tumor microenvironment (TME)[37]. Importantly, among these factors ITGA4 was associated with poor prognosis in the TCGA-KIRC cohort (Suppl. Fig. 5A). Further nodes in this module included other upregulated elements with known functions in modulating the TME, such as FN1, PLAUR, VEGFC and CD38. These factors have been linked to EMT thereby promoting cancer cell migration and invasion. Conversely, we observed downregulation of THBS1, CDH1 and EPCAM specifically in the rhabdoid network (Fig. 5D and Suppl. Table 10) underscoring the overall dysregulation of cell-matrix interactions and enhanced activation of EMT. Immunohistochemistry (IHC) analysis confirmed an almost complete absence of E-cadherin protein expression in the rhabdoid component of morphologically heterogeneous ccRCC cases, suggesting its potential utility as a companion biomarker to aid accurate diagnosis of grade 4 rhabdoid differentiation (Suppl. Table 8 and 9, Fig. 5E). The relevance of this finding was further underscored by the association of low CDH1-mRNA expression with poor survival in the TCGA KIRC cohort (Suppl. Fig. 5B).

These results not only shed light on the specific mechanisms contributing to the malignant potential of rhabdoid ccRCC cells, but they also highlight the molecular alterations underlying their structural hallmarks, such as discohesive growth patterns and high mitotic activity, which are likely driven by altered cell adhesion and enhanced proliferation.

Immune evasion and dysregulation are hallmarks of rhabdoid ccRCC cells

Our analysis of transcriptome and proteome data suggested that immune response dysregulation was specific to rhabdoid ccRCC cells across both data layers (Fig. 2E and 2F, Fig. 3B). To investigate this

further, we analyzed CD8+ T-cells infiltration in a larger cohort of ccRCCs with grade 4/rhabdoid areas. We detected CD8+ T-cells in almost all grade 4/rhabdoid ccRCC cases of our extended cohort (Suppl. Table 8 and 9) indicating inflamed phenotypes of varying degrees (Fig. 6A). In our grade4/rhabdoid-specific mRNA expression data we also observed significant upregulation of the immune checkpoint molecule PD-L1 (CD274). Of note, this was not the case for other well-studied checkpoint factors such as CTL-4 or LAG3 (Suppl. Fig. 5C). Integrated network analysis highlighted an immunosuppressive network largely specific to grade 4/rhabdoid ccRCC cells, which was distinct from lower-grade cells (Modules III, XI and XXIII) (Fig. 6B). In these modules, we found high expression of chemokines including CCL4, CCL20, CXCL1 and CXCL13 (Fig. 6C), molecules that can interact with the TME and immune cells to foster an immunosuppressive milieu[38–40]. In addition, elevated type I interferon (IFN) signaling, particularly IFN-alpha and IFN-beta, emerged as a prominent feature in module III. High expressions of STAT1/2, the main signaling molecules of the IFN pathway, was associated with grade 3 and grade 4/rhabdoid cells (Fig. 6B and 6C). Upregulation of STAT1/2 transcriptional targets were predominantly enriched for interferon- beta signaling molecules (OAS1, MX1 and IFI44) (Fig. 6B and 6C, Suppl. Fig. 4C)[41]. Highly interestingly, we also identified upregulation of CD38 in grade 4/rhabdoid ccRCC cells, a known target of IFN-beta signaling (Fig. 5D). CD38 associated with the cell adhesion network module in our integrated network analysis (Module V , Fig. 5C) but was more recently described to suppress cytotoxic T cells activity when overexpressed on tumor cells[42] suggesting CD38 overexpression may have a potential role in immune evasion in rhabdoid ccRCC. Additionally, we observed CD38 upregulation in RNAseq data from cases with rhabdoid differentiation in the CheckMate 025 cohort, irrespective of the treatment arm (Fig. 6D). However, due to the limited number of samples with rhabdoid differentiation and accompanying RNAseq data in the nivolumab arm (N = 8), it was not feasible to draw definitive conclusions regarding the clinical benefit of immune checkpoint inhibitors (ICI) in this subgroup. Besides CD38, also ITGB2, another factor with roles at the intersection of ECM remodeling and immune modulation, showed prominent upregulation in grade 4/rhabdoid ccRCC cells (Fig. 5C and 5D) likely contributing to a state of immune suppression in rhabdoid ccRCC[43].

Take together, our analysis revealed a substantial immunosuppressive network driven by signaling pathways activated specifically in rhabdoid ccRCC cells. Upregulated factors like CD38 and ITGB2 appear to exert immunomodulatory functions that inhibit the activity of abundant CD8+ T-cells within the rhabdoid ccRCC TME.

Discussion

Morphological intra-tumor heterogeneity (ITH) is a well-documented phenomenon, where for example different grade morphologies can co-occur within the same tumor section[44,45]. Modern microscopy can generate high-resolution maps of these diagnostic tissue sections, enabling the computer-assisted assessment of histo-morphological diversity. AI- and ML-based methods have been successfully applied in biomedical imaging for segmentation and classification tasks across various cancer types[46–50]. Our AI-driven automated classification system can reliably grade and quantify ccRCC cells in HE-stained digital WSI. This enhances the accuracy of histopathological assessments and helps to overcome the need for manual expert annotations.

Previous studies have shown the profound genetic ITH in ccRCC tissue samples[1,2] but a correlation with morphological features was not possible with their approaches. Combining AI-based image analysis with single-cell isolation and deep multi-omics profiling of selected cells provided us with unprecedented insights into the relationship between molecular and morphological heterogeneity in ccRCC. This revealed a trajectory of progressive dysregulation in ccRCC cells increasing grades on both, mRNA and protein levels, indicating that non-genetic alterations likely contribute to malignancy in ccRCC. Importantly, mRNA and protein measurements provided complementary information and consistently with previous data, we observed that overall downregulation of pathways was more prominent than upregulation during ccRCC tumor cell differentiation[32,51]. As tumor grade increased, ccRCC mRNA and protein data showed overlapping functional alterations, highlighting progressive dysregulation of metabolic pathways such as one-carbon metabolism and decreased expression of proximal renal tubule genes essential for kidney function[32]. Notably, grade 4/rhabdoid ccRCC cells appeared to diverge most from highly differentiated kidney tubular cells, showing increased expression of stem cell markers like CD44 and DCLK1[33,34], implying a potential role of stem cell-like features for the aggressiveness in high-grade ccRCC. As the acquired expression profiles reflected many known ccRCC tumorigenic changes[8,51], we were confident in the accuracy of mRNA and protein dysregulation portraits obtained from as little as 1000 individually collected ccRCC cells from FFPE sections.

While recent studies have explored the molecular landscape and therapeutic vulnerabilities of RCC with sarcomatoid differentiation in considerable detail[19,20], the biology of rhabdoid ccRCC remained understudied[17]. Bakouny and colleagues profiled the molecular features in grade 4 ccRCC cells but did

not identify molecular characteristics that distinguished rhabdoid differentiation in RCCs[20]. However, the majority of cases of their cohort comprised sarcomatoid RCCs, raising the question whether rhabdoid ccRCC is truly similar to or biologically distinct from sarcomatoid cases, as their morphologies suggest. Previous molecular investigations have associated the rhabdoid component with BAP1 mutations[20,45]. However, our genome-wide genetic profiling of five ccRCC tissue samples with rhabdoid differentiation did not corroborate that BAP1 mutations were a feature in most cases. We also did not observe the loss of chromosome 11p, which was previously reported as a distinct feature in ccRCC with rhabdoid differentiation[52]. These findings suggest that genetic alterations specific to rhabdoid differentiation in ccRCC remain uncertain. Instead, our study highlighted the specific transcriptome and proteome alterations in a homogeneous cohort of ccRCCs with rhabdoid differentiation. We identified upregulation of FOXM1-driven proliferative signaling and ECM remodeling as key molecular changes associated with high-grade and rhabdoid phenotypes, underscoring their role in high metastatic potential. Interestingly, FOXM1 has been associated with tumor progression and poor prognosis in patients with ccRCC before[53]. Upregulations of integrins on the other hand, was shown to enhance tumor cell motility and interactions within the ECM, thus promoting tumor growth, cell migration, and metastatic dissemination, potentially by modulating matrix metalloproteinases (MMPs)[37]. Of note, integrins have been associated with worse prognosis in ccRCC and particularly ITGB2 was shown to be overexpressed in aggressive cancers[54–56]. On the other hand, enhanced proliferation and ECM remodeling likely explain the distinct morphological features observed in rhabdoid cells in ccRCC as these processes disrupt normal cell adhesion and structural integrity within the tumor, allowing cells to assume a more invasive, migratory phenotype[16]. Here we highlight distinct molecular alterations in rhabdoid ccRCC cells and suggest that targeting these networks could be crucial in preventing metastasis and improving patient survival.

Moreover, our study unveils a distinct immunosuppressive signaling network driven by rhabdoid ccRCC cells, involving well-known immune checkpoints like PD-L1 and novel immunomodulatory factors such as CD38 and ITGB2[43,57,58]. Interestingly, ITGB2, which we found overexpressed in rhabdoid ccRCC cells, has recently been identified to regulate the immune response within the TME besides its more established roles in the integrin-ECM signaling network that promote EMT and metastatic spread[43,59]. Additionally, elevated interferon signaling, chemokine, and integrin expression further characterized this immunosuppressive network. Consistently, we observed high STAT1/2 expression and elevated IFN-I

signaling in rhabdoid ccRCC cells which likely drive EMT, immune evasion, and an aggressive phenotype in inflammatory tumors[41,60]. STAT1/2 in particular are well recognized markers of aggressiveness in ccRCC[32]. These findings suggest that rhabdoid ccRCC cells play key roles in reshaping the TME, creating an immunosuppressive milieu that likely compromises CD8+ T-cell function and may mediate resistance to ICI[57]. These findings were further corroborated by two recent studies that highlight the significant role of the TME for ccRCC progression and advances disease showing that these exhibit enrichment of terminally exhausted CD8+ T cells[61,62]

Although clinical inter-patient heterogeneity and the limited number of samples pose challenges, our study is the first to integrate AI-driven image analysis, single-cell isolation, and deep multi-omics profiling to characterize the molecular aberrations in distinct clinically relevant and morphologically defined aggressive cell populations. Analysis of larger cohorts, such as the CheckMate 025 trial[63], revealed comparably low numbers of rhabdoid ccRCC samples with accompanying molecular data. Despite this limitation, our analysis of RNAseq data from the trial demonstrated elevated CD38 expression in rhabdoid ccRCCs, highlighting its potential relevance. However, the small sample size precluded any conclusions regarding the clinical benefit of ICIs in this subgroup. Taken together, our findings offer valuable insights into the molecular pathways underpinning grade 4/rhabdoid ccRCC, including FOXM1-driven cell cycle regulation, ECM remodeling, and immune evasion. These data provide a strong rationale for exploring combination therapies targeting these pathways to enhance the efficacy of ICIs[58] in this challenging patient population.

Methods

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Hella A. Bolck.

Materials availability

This study did not generate new unique reagents.

Data and code availability

Data
Any information required to reanalyze the data reported in this paper is available upon request.

Code
Any information required to reanalyze the data reported in this paper is available upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Patient cohort

The Department of Pathology and Molecular Pathology at University Hospital Zürich made available archival FFPE tissue samples for this study. It was approved by the local Ethics Committee (BASEC# 2019-01959) and in agreement with Swiss law (Swiss Human Research Act). All patients gave written consent. The tumor samples used in this study were diagnosed as ccRCC and the overall grade was noted from the diagnostic histopathologic patient reports according to the WHO/ISUP[11]. Two board-certified pathologists re-assessed all FFPE tissue samples in the context of this study to confirm the presence of malignant and normal areas as well as different grade morphologies. Associated clinical information is shown in Suppl. Table 1, 4 and 8.

METHOD DETAILS, QUANTIFICATION AND STATISTICAL ANALYSIS

Genomic characterization by whole-genome sequencing (WGS)

To extract tumor and normal DNA from our core cohort, we sampled all areas that were included in the image analysis and single-cell isolation workflow to cover all mutations that may be present at these sites. To facilitate this a 2 µm section was stained with hematoxylin and eosin (H&E) to evaluate the morphology and an experienced pathologist marked the areas for dissection (Suppl Fig. 1). This slide

was used as a reference image for the AVENIO Millisect System (Roche) and subsequently the marked areas were dissected from a consecutive unstained 10 µm section prepared on a glass slide. Tissue areas from at least two 10 µm sections were collected in a tube containing lysis buffer (Promega, A826E from Maxwell® 16 FFPE Tissue LEV DNA Purification Kit) and subjected to genomic DNA extraction on a Maxwell® 16 machine (Promega) following the manufacturer's instructions.

Whole-genome sequencing libraries were constructed according to Illumina TruSeq DNA PCR Free Library Prep protocol HT (Illumina Inc., San Diego, CA, USA). This included: (1) fragmentation of 400 ng genomic DNA to 350 bp inserts by Covaris LE220-plus, (2) cleanup of fragmented DNA, (3) end repair, (4) removal of large and small DNA fragments, (5) 3′-end adenylation and (6) adapter ligation. The resulting libraries were quantified and quality-assessed with the iSeq100 (Illumina). The samples were sequenced with the NovaSeq 6000 platform (Illumina) using S4 flow cells with 300 cycles (2 × 150 reads) and measured 2 × to reach an average coverage of 30X for normal kidney tissue and 120X for the tumor fraction per patient.

The raw sequencing files (bcf files) were converted to fastq format and demultiplexed in a single step using Illumina's bcl2fastq program via Illumina DRAGEN Bio-IT Platform (v3.8.4). First steps in secondary analysis, including read trimming, alignment to GRCh38, sorting, duplicate marking and somatic variant calling were done using Illumina DRAGEN Bio-IT Platform (version 3.9.5) [64] or Mutect2 [65] and Strelka2 [66]. To identify and annotate somatic variants, we annotated and converted the VCF files into MAF format using VEP [67]. The MAF files were analyzed using Maftools in R [68]. In addition, Maftools was used for comprehensive analysis and visualization, including tumor mutational burden (TMB) assessment, mutational signature analysis, and identification of significantly mutated genes. Genome-wide copy number aberrations were visualized with gistic (v2.0.23) using default parameters and log-transformed segmentation files from DRAGEN as input.

Sample preparation for morphological analysis

For each ccRCC case from our core cohort, we selected one to four representative FFPE blocks in order to include areas displaying low-grade, high-grade and rhabdoid areas as well as normal kidney histology. Subsequently, 10 µm sections were prepared using PEN MembraneSlides 1.0 (Carl Zeiss) and stained with H&E. Slides were digitized scanning z-stacks with seven z-slices in 0.5 µm increments using a Hamamatsu C9600 scanner at ×40 optical magnification. The scanner resolution was 0.228 µm per pixel. The optimal focus plane was selected manually before loading the images into the image analysis software (BIAS) to save processing time and storage. However, since BIAS features several pre-processing options, this step can be automatized by the software.

Image analysis

Using the BIAS software suite[26], we developed a pipeline to assign every cell to one of the following classes (i) grade 1/2 (low-grade) ccRCC, (ii) grade 3 (high-grade) ccRCC, (iii) grade 4/rhabdoid ccRCC, (iv) normal kidney, (v) other/junk. This workflow was composed of several steps: segmentation of nuclei, specification of cytoplasmatic areas, pre-classification into tumor and normal kidney cells, segmentation of nucleoli, segmentation of rhabdoid cells, classification into the distinct classes (i-v) mentioned above.

We used Artificial Neural Networks (ANN) for both segmentation and classification. The classification models were multilayer perceptron networks (MLPs) trained on features extracted from the segmented cells. To segment nuclei, we integrated nucleAIzer models into BIAS[31]. Dilating the nuclear contours by 10 µm created the cytoplasmatic masks that were used to extract features from the cytoplasm. These were aggregated with features of the 10 neighboring cells and an artificial neural network pre-classified the tissue into healthy and tumor.

Instance segmentation algorithms based on a Mask-RCNN implementation with a Resnet101 backbone achieved custom segmentation of prominent nucleoli and rhabdoid cells. Additional cases were annotated as training data (Suppl. Table 4). Subsequently, cRCC cells were classified into grade 1/2, grade 3, grade4/rhabdoid, normal kidney or other/junk by MLP classification models trained on features extracted from the segmented cells. Phenotype predictions were ranked by confidence and the 2000 best cells were manually curated. These were used as an extended training set to improve the ML classification further.

K-fold stratified cross-validation (where K=10) was performed using the core patient cohort (Suppl. Table 1). Moreover, we evaluated our approach using six independent grade 4 ccRCC cases with rhabdoid morphology and grade ITH (Supp. Table 8).

Sample preparation and single cell isolation

H&E-stained FFPE tissue sections on PEN MembraneSlides were subjected to coverslip removal by immersing them in xylene for 1-2 hours (until the coverslip loosened) followed by xylene removal and hydration using descending ethanol concentrations and air-drying. Subsequently, cells were isolated by LMD based on the morphologies assigned by the machine learning classification module of BIAS in the same sections. Relevant cell contours were transferred to a Leica LMD6 microscope for automated single-cell isolation. The Leica LMD6 system, equipped with an HC PL FLUOTAR L ×63/0.70 objective, was used for cell excision. Cells were excised based on nuclei contours plus a 6 µm offset. For each morphological class in each sample, two sets of 1,000 cells were microdissected and collected into PCR caps (VWR, 732-0548) or 384-well plates (Eppendorf® Twin tec PCR, 384 wells EP0030128508) To cells isolated for RNAseq analysis, 5 µl of PKD buffer (Qiagen) and 25 µl of RNA later (Qiagen) were added, followed by centrifugation at 10,000 g for 5 minutes. For proteome analysis we added 10µl of ammonium bicarbonate buffer to the isolated cells.

RNA sequencing and data analysis

RNA was extracted from microdissected single cells using the RNeasy FFPE kit (Quagen, Hilden, Germany). Entire RNA extracts were transferred into NGS Library preparation without QC. Due to the low concentration of total RNA the SMARTer® Stranded Total RNA-Seq Kit v2 -Pico Input Mammalian (A Takara Bio Company, California, USA) was used to prepare sequencing libraries.

Briefly, total RNA samples were reverse-transcribed using random priming into double-stranded cDNA in the presence of a template switch oligo (TSO). When the reverse transcriptase reached the 5’ end of the RNA fragment, the enzyme’s terminal transferase activity added non-templated nucleotides to the

3' end of the cDNA. The TSO paired with the added non-templated nucleotide, enabling the reverse transcriptase to continue replicating to the end of the oligonucleotide. This resulted in a cDNA fragment that contained sequences derived from the random priming oligo and the TSO. PCR amplification using primers binding to these sequences could then be performed. The PCR added full-length Illumina adapters, including dual barcodes for multiplexing. Ribosomal cDNA was cleaved by ZapR in the presence of the mammalian-specific R-Probes. Remaining fragments were enriched with a second round of PCR amplification using primers designed to match Illumina adapters (16 cycles). The quality and quantity of the enriched libraries were validated using the Tapestation (Agilent, Waldbronn, Germany). The product was a smear with an average fragment size of approximately 300 - 360 bp. The libraries were normalized to 5nM in Tris-Cl 10 mM, pH8.5 with 0.1% Tween 20.

A shallow sequencing approach on the iSeq (Illumina, Inc, California, USA) was used to check the read distribution among samples, rRNA content and reference genome mapping. The Novaseq 6000 (Illumina, Inc, California, USA) was used for deep sequencing according to the standard manufacturer's protocol. Sequencing configuration was single-end 100 bp. Demultiplexing was performed with bcl2fastq (Illumina, Inc, California, USA) allowing for 0 mismatch on Index 1 and 1 mismatch on Index 2. Between 70 million and 146 million reads were generated per sample.

RNA sequencing data analysis was performed using the SUSHI framework[69], which encompassed the following steps: Read quality was inspected using FastQC, and sequencing adaptors removed using fastp; Alignment of the RNASeq reads using the STAR aligner[70] and with the GENCODE human genome build GRCh38 (Release 42) as the reference[71]; the counting of gene-level expression values using the 'featureCounts' function of the R package Rsubread[72]; differential expression using the generalized linear model as implemented by the DESeq2 Bioconductor R package[73], and; Gene Ontology (GO) term pathway analysis using the hypergeometric over-representation and GSEA tests *via* the 'enrichGO' and `gseGO` functions respectively of the clusterProfiler Bioconductor R package[74]. For differential gene expression analysis (DEG) we contrasted gene expression in each ccRCC grade to that in the normal renal tubuli samples to obtain three ccRCC-specific comparisons termed (i) grade 1/2 (low-grade), (ii) grade 3 (intermediate/high-grade, (iii) grade 4/rhabdoid. Additional visualizations were generated in the exploreDE Interactive Shiny app[75]. All R functions were executed on R version 4.2.0 (R Core Team, 2020) and Bioconductor version 3.15.

Proteomics data acquisition and analysis

For sample lysis, the collected cells were spun down and 20 µl of 50 mM Triethylammonium bicarbonate (TEAB) were added to each tube or well. The 384-well plate was closed with sealing tape (Corning, CLS6569-100EA) and centrifuged at 2000 g for 2 min. All samples were incubated for 1 hour at 95 °C. For the plates a PCR thermo cycler with 384-well reaction module (Biorad S1000) was used, for the PCR tubes, the 96-well module (Eppendorf, Mastercycle X50a) was installed. The lid temperature for all incubation step was always 110 °C. Prior to incubation for 60 min at 75°C, 5 µl of 60 % acetonitrile in 50mM TEAB was added. For in-solution protein digestion, samples were cooled to room temperature and subjected to pre-digestion using 1 µl LysC (4 ng/µl) for 4 h at 37 °C. Subsequently, 1 µl Trypsin (4

ng/µl in TEAB) was added for overnight digestion at 37 °C. In the morning, another 4 ng of Trypsin was added and the samples were incubated for 5 hours at 37 °C[76]. The digestion was stopped by acidification, adding 10 µl 5 % aqueous trifluoroacetic acid and the sample were stored at -20 °C until LC-MS analysis.

For proteomics data acquisition, **h**alf of the digested samples were loaded on Evotips (Evosep Biosystems) following the provided instructions and analysed on an Evosep One LC coupled to a TIMS TOF Pro mass spectrometer (Bruker). The mass spectrometry proteomics data were handled using the local laboratory information management system b-fabric (LIMS) [77], and all relevant data have been deposited to the ProteomeXchange Consortium via the PRIDE (http://www.ebi.ac.uk/pride) partner repository with the data set identifier [tbd].

The acquired data-independent-acquisition (DIA) spectra were processed with DIA-NN (Version 1.8.1.8)[78], using a library free approach with the ensembl homo_sapiens data base (release-108) and oxidation at methionine as variable modification. The R package prolfqua[79] was used to process, handle and analyze the differential expression and to determine group differences, confidence intervals, and false discovery rates for all quantifiable proteins. Starting with the report.tsv files generated by DIA-NN, which included the precursor ion abundances for each raw file, we determined gene centric abundances by first aggregating the precursor abundances to peptidoform abundances after filtering for gValues < 0.01. Then, we employed the Tukeys-Median Polish to estimate gene abundances. This was possible with a mapping table where for each ensemble protein the gene name was listed (if there was no gene name listed, the ensemble protein identifier was used). Using the gene names instead of the protein identifier allowed us to roll-up the peptidoform abundances to gene names. Furthermore, before fitting the linear models, we transformed the gene abundances using the log2 transformation followed by robust scaling normalization.

For differential protein expression analysis (DEP) we contrasted protein expression in each ccRCC grade to that in the normal renal tubuli samples to obtain three ccRCC-specific comparisons termed (i) grade 1/2 (low-grade), (ii) grade 3 (intermediate/high-grade, (iii) grade 4/rhabdoid. Gene Ontology (GO) term pathway analysis was performed in the same way as in the RNA sequencing data analysis pipeline using the hypergeometric over-representation and GSEA tests *via* the 'enrichGO' and `gseGO` functions respectively from the clusterProfiler Bioconductor R package[74].

Immunohistochemistry (IHC)

Sections (2 µm) were prepared from FFPE specimens and were stained with haematoxylin and eosin (H&E) for histological evaluation. For IHC, sections (2 µm) were transferred to glass slides followed by antigen retrieval. IHC was performed using the Ventana Benchmark automated system (Roche, Ventana Medical Systems, Oro Valley, AZ) and Ventana reagents. The Optiview DAB IHC detection kit (Roche, Ventana Medical Systems) was used to stain with the antibody against CD8 (Clone C8/144B, cat. number M7103, DAKO, used 1:100). The staining procedure for CDH1 (E-cadherin clone 36, cat. number 610181, BD Biosciences, used 1:500) was carried out with the automated Leica BOND system using the Bond Polymer Refine Detection Kit (Leica Biosystems, Wetzlar, Germany).

Survival analysis

Survival curves were generated using the Gene Expression Profiling Interactive Analysis (GEPIA) web tool (http://gepia.cancer-pku.cn) to assess the prognostic significance of gene expression levels for CDH1, FOXM1 and ITGA4 in the TCGA Kidney Renal Clear Cell Carcinoma (TCGA-KIRC) patients. Patients were divided into high and low expression groups based on the median expression level. Kaplan-Meier survival curves were generated, and the log-rank test was used for statistical comparison. Hazard ratios (HR), 95% confidence intervals (CI), and p-values were calculated and displayed. Survival differences were considered significant if p-values were <0.05. HRs >1 indicated higher risk, while HRs <1 indicated a protective effect.

Statistical testing

We performed differential expression analysis (DEG/DEP) comparing ccRCC grade groups (grade 1/2, grade 3, grade 4/rhabdoid) to adjacent normal kidney tubule cells. To assess significant changes, we applied a false discovery rate (FDR) threshold of < 0.05 and considered log2 fold change (log2FC) greater than |1|. The significance levels for FDR are denoted as follows: **** <= 0.0001, *** <= 0.001, **<= 0.01, * <= 0.05. Additionally, we color-coded the log2FC values for visual representation.
For enrichment analysis in the transcriptomic and proteomic datasets, we conducted over-representation analysis using the hypergeometric test. Specifically, we examined Gene Ontology (GO) biological pathways. Benjamini-Hochberg adjusted p-values were calculated and Significance levels are denoted as follows: **** for p <= 0.0001, *** for p <= 0.001, ** for p <= 0.01, * for p <= 0.05.

Patient and upper-quartile (UQ) normalized transcripts-per-million (TPM) RNAseq data from CheckMate 025 (CM-025; NCT01668784) prospective clinical trial of the anti-PD-1 antibody nivolumab in ccRCC was obtained from the accompanying paper[63]. CM-025 was a randomized phase III trial which demonstrated an improved overall survival (OS) with nivolumab over the mTOR inhibitor everolimus. Patients were included based on the availability of RNAseq data and information on rhabdoid differentiation. After excluding cases with incomplete annotations, CD38 expression levels were plotted and analyzed in R.

Data integration by PPI network analysis

For comparison of proteomic and transcriptomic datasets we used the results from the differential expression and GO enrichment analyses performed for each data type independently between the normal and the three grades of the ccRCC. Clustering plots were generated using all genes with standard deviation in log2 fold change between samples of at least 0.5 and a maximum of 2000 genes. To highlight the synergies between the measured proteomics and transcriptomics profiles, we performed integrated network analysis using the NOODAI software platform[35]. Differentially expressed elements of the proteomics and transcriptomics analysis were used as input and searched against the default publicly available protein-protein interaction (PPI) databases containing only high confidence human interactors from STRING[80], BioGRID[81] and IntACT[82] databases were collected. For STRING

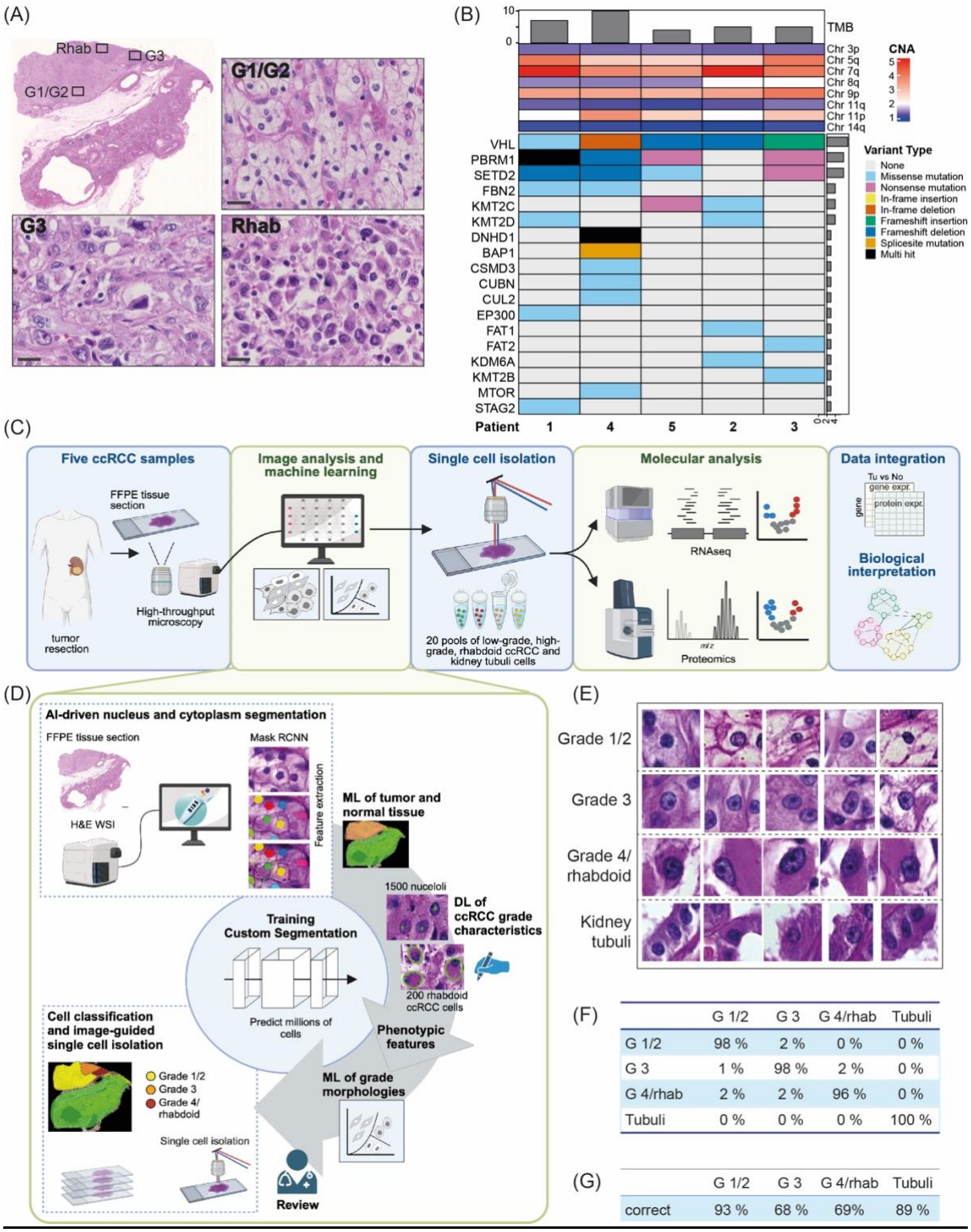
(version 11.5) complete interactome data from all sources was filtered to include only entries with a combined score above 0.7. For BioGRID (version 4.4.218), a mitab data file that contained a dataset of interactors with physical interactions supported by independent validations was used. It was filtered to exclude entries without an associated confidence value. For the IntACT database, the psimitab from 13/07/2022 was used and filtered to keep only interactions with a confidence score above 0.7. Interaction pairs obtained from the different databases were overlapped and merged into a joint dataset. This was used in NOODAI to construct a network, in which only interaction pairs with both members significantly differentially expressed in at least one of the data layers (transcriptomics or proteomics) were retained. Only the network with the highest number of connected members was further analyzed (without allowing edge repetition). In order to identify highly connective central nodes in the networks, the default current flow betweenness centrality metric was used[83].

Functional network modules[84] were subsequently identified through NOODAI by using the MONET software[85]. Modularity optimization with undirected edges was selected, setting the target average node degree within the identified modules to 10. The modules were default sorted based on the number of nodes they contained. To assess the robustness of the network decomposition and to identify the main signaling axes present in our PPI networks, NOODAI performed pathways overrepresentation analysis for individual modules by searching against the pre-loaded Reactome and KEGG databases. By default, background datasets were composed of all genes and proteins that were used to construct the full-scale network of the respective conditions. Results were filtered to retain significant pathways with at least five significant hits and an FDR below 0.05 (Benjamini-Hochberg[86] correction of Fisher's exact test p-values)., Non-specific pathways with more than 300 members were omitted. Pathways that did not have more than two different overrepresented members were considered redundant and were investigated in the context of their explanatory power. For the visualization of results, NOODAI generated circular charts that highlighted the top 30% of most central nodes (based on the current-flow betweenness centrality). For selected modules and full-size networks, the graphical representation was done in Cytoscape 3.10.1[87] using the webtool-generated edge files. For edge routing, the *organic routing* from yFiles plugin was used[88].

Additional R packages used for data handling and visualization of the NOODAI results included RColorBrewer[89], readr[90], stringr[91], gtools[92], gridBase[93], ComplexHeatmap[94], tidygraph[95], biomaRt[96] and readxl[97].

Figures

Figure 1



(F)	G 1/2	G 3	G 4/rhab	Tubuli
G 1/2	98 %	2 %	0 %	0 %
G 3	1 %	98 %	2 %	0 %
G 4/rhab	2 %	2 %	96 %	0 %
Tubuli	0 %	0 %	0 %	100 %

(G)	G 1/2	G 3	G 4/rhab	Tubuli
correct	93 %	68 %	69%	89 %

Figure 1: AI-driven, image-guided single cell isolation and multi-omics analysis of ccRCC with pronounced morphological intra-tumor heterogeneity including rhabdoid de-differentiation.

(A) Representative H&E images (Patient 4) indicating ITH of grade features: Grade 1/2 (low-grade, G1/2), grade 3 (high-grade, G3) and grade 4 ccRCC with rhabdoid differentiation (Rhab). Scale bars: WSI – 2 mm, high magnification panels - 100 µm.

(B) Overview of somatic alterations identified by WGS analysis in five samples used in this study, including TMB, selected copy number events and driver mutations.

(C) Overview of workflow combining AI-guided single-cell histological classification and isolation with transcriptomic and proteomic analysis.

(D) AI-driven segmentation workflow for grade 1/2 (low), grade 3 (high) and grade 4/rhabdoid ccRCC as well as normal-appearing kidney tubuli cells in diagnostic tissue image using BIAS.

(E) Example images of cells from the classes stated above identified by ML.

(F) Cross-validation accuracies of our image analysis workflow using a randomized 10-fold validation test.

(G) Validation using independent unseen data.

Figure 2

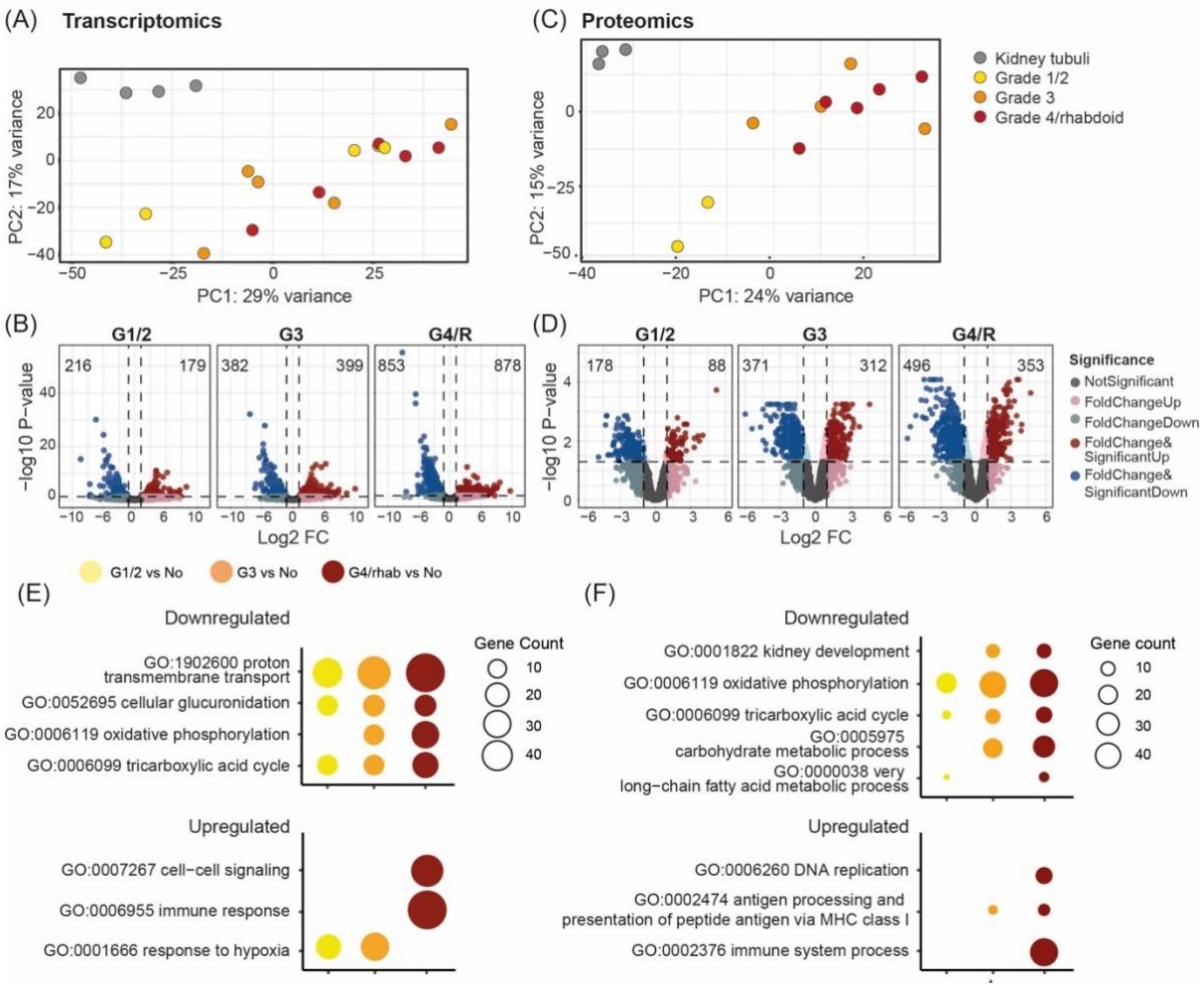


Figure 2: Differential Expression Analysis of mRNAs and proteins from individually isolated cells with distinct grade morphologies.

(A) PCA of RNAseq data for 20,528 genes from five patients.

(B) Volcano plot showing DEGs (FDR < 0.05, logFC > |1|) in ccRCC grade groups compared to adjacent normal tissues. Genes that were significantly down- or upregulated are presented with red/blue filled scatters and their number is indicated in the upper left and right corners, respectively.

(C) PCA of LC-MS data for 4,382 proteins from five patients.

(D) Volcano plot showing DEPs (FDR < 0.05, logFC > |1|) in ccRCC grade groups compared to adjacent normal tissues. Proteins that were significantly down- or upregulated are presented with red/blue filled scatters and their number is indicated in the upper left and right corners, respectively.

(E) Enrichment analysis of associated GO biological pathways generated from over-representation analysis (hypergeometric test) for the transcriptomic datasets indicating adjusted p-values.

(F) Same as (E) for proteomic dataset.

Figure 3

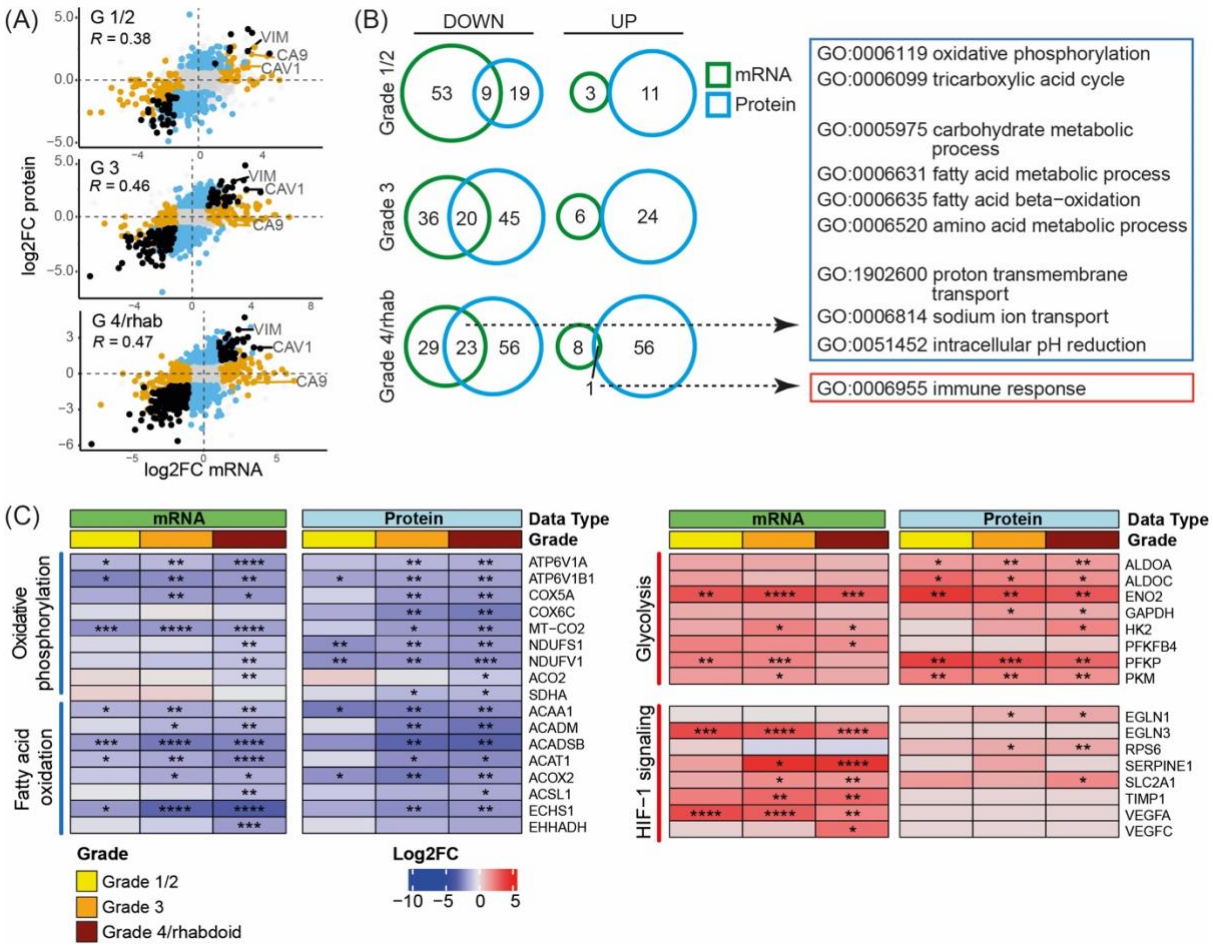


Figure 3: Functional overlap between mRNA and protein data.

(A) Scatterplot of differential mRNA and protein expression between ccRCC grades and adjacent normal cells (Log2FC) indicating the correlation between datasets. Each dot represents one gene. Negative values indicate a decrease and positive values an increase in expression between tumor and NAT

(B) Venn diagram showing numbers of enriched GO biological pathways in DEGs and DEPs for each grade group and the overlap between both data layers. Selected shared GO terms for grade 4/rhabdoid class of ccRCC cells are specified (for a complete overview, refer to Suppl. Fig. 4B).

(C) Heatmaps depicting metabolic rewiring and upregulation of HIF1-signaling during ccRCC de-differentiation evident on both data layers. Alterations of representative genes and proteins are depicted as log2FC with respect to adjacent normal cells. Diagram on the left side depicts downregulated factors (blue), while on the right side upregulated factors are shown (red). Genes were selected from Glycolysis (Reactome R-HSA-70171), Fatty acid metabolism (Reactome R-HSA-8978868), Oxidative phosphorylation (KEGG path:hsa00190) and HIF-1 signaling (KEGG path:hsa04066). FDR is indicated by"****" <= 0.0001, "***" <= 0.001, "**"<= 0.01, "*" <= 0.05.

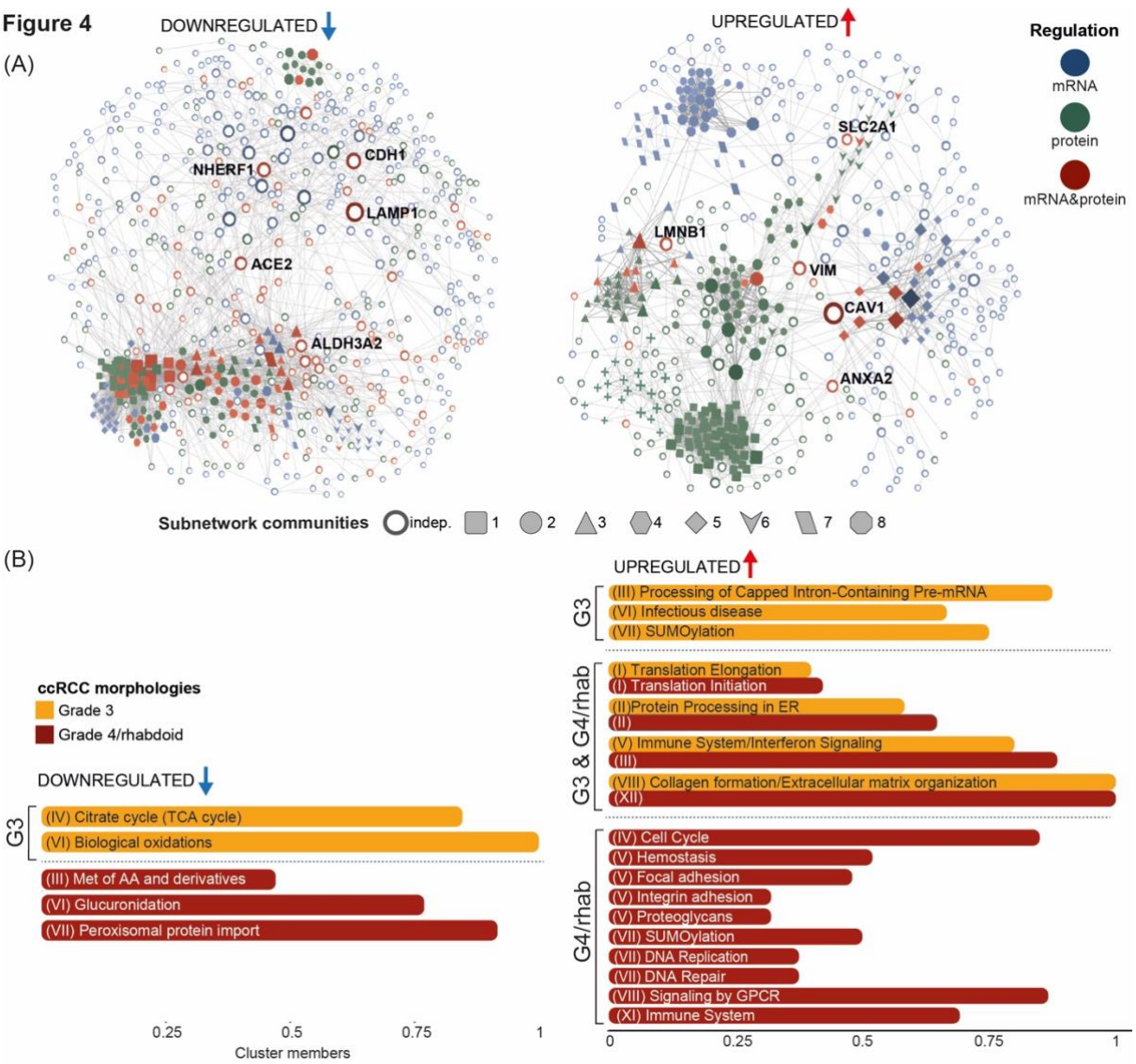


Figure 4: Multi-omics data integration uncovers distinct de-regulated processes in aggressive high-grade and rhabdoid ccRCC cells.

(A) PPI network distinct to G4/rhabdoid ccRCC cells built by integrating the significant elements from transcriptomic and proteomic datasets that were down- (N=706) or upregulated (N=501). Node sizes and darker colors correspond to higher centrality scores. The five top highly connected hub genes that associated with no specific subnetwork are highlighted. Elements within network communities with defined biological functions, identified after the network decomposition, are specified.

(B) Biological processes enriched in modules distinct to G3 and G4/rhabdoid cells with at least 10 members following decomposition of the PPI networks in (A) based on KEGG and Reactome annotations. Enriched terms were determined by an FDR < 0.05 (Fisher's exact test with Benjamini-Hochberg correction). Only non-redundant terms are presented, except where additional terms provided complementary information. The fraction of proteins in each module with the corresponding annotation is provided.

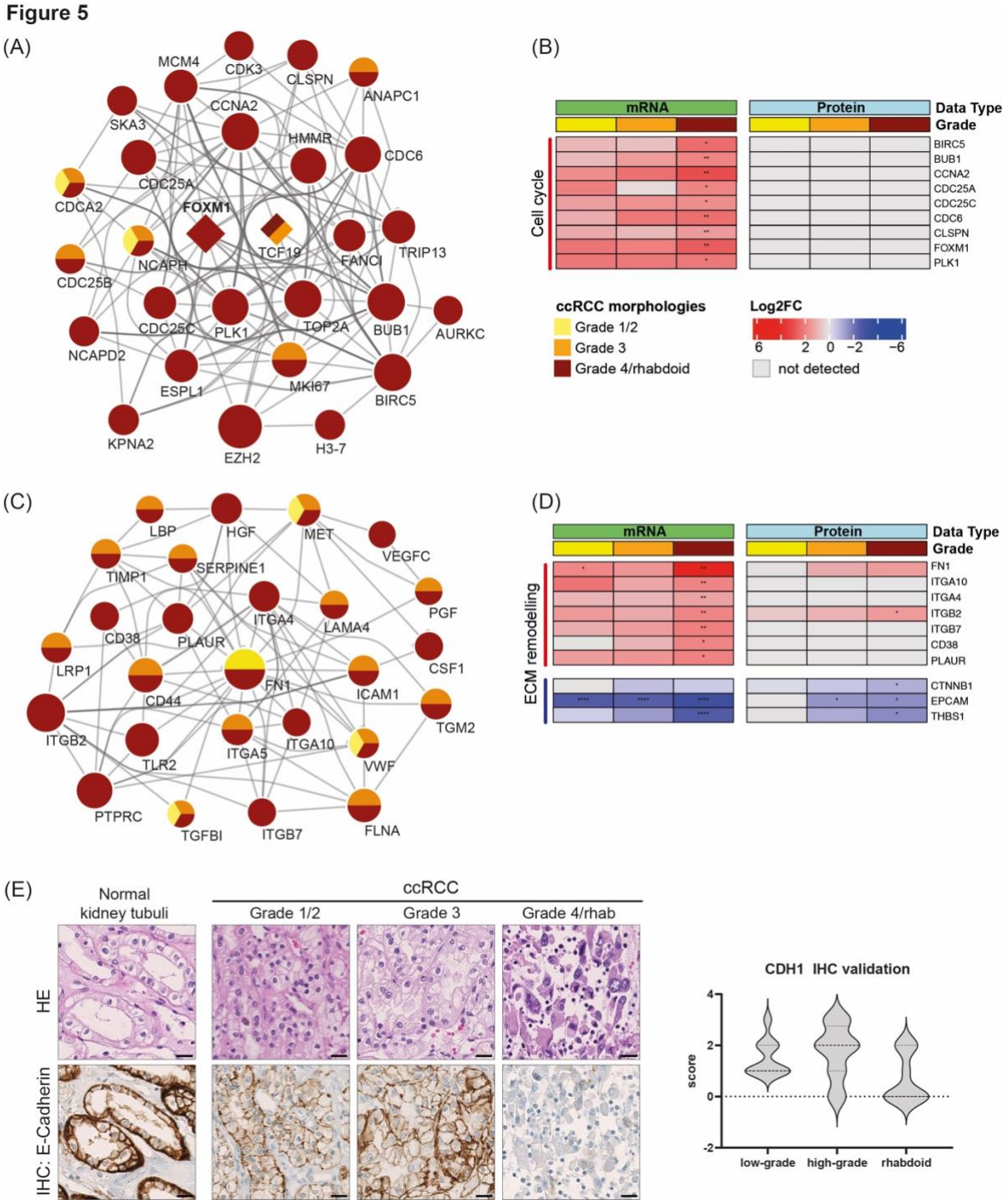


Figure 5: Grade 4/rhabdoid ccRCC cells are characterized by enhanced expression of cell cycle factors and dysregulation of cell-matrix interactions.

(A) Network decomposition highlighting a subnetwork associated with upregulation of cell cycle factors (Reactome R-HSA-1640170) in the grade 4/rhabdoid ccRCC cells (Module IV). Factors that are significantly upregulated also in lower grade ccRCC cells are indicated. Bigger nodes correspond to higher centrality values in the Grade4/rhabdoid ccRCC PPI network. Grades are color-coded as in (B).

(B) Heat maps highlighting selected elements from module IV and their respective dysregulation during ccRCC de-differentiation on both data layers. Alterations of representative genes and proteins are depicted as log2FC with respect to adjacent normal cells. FDR is indicated by “****” <= 0.0001, “***” <= 0.001, “**”<= 0.01, “*” <= 0.05.

(C) Network decomposition highlighting a subnetwork containing upregulated elements involved in focal adhesion (KEGG path:hsa04510), integrin cell surface interactions (Reactome R-HSA-216083), homeostasis (Reactome R-HSA-109582) and proteoglycans implicated in cancer (KEGG path:hsa05205) in the grade 4/rhabdoid ccRCC cells (Module V). Factors that are significantly upregulated in lower grade ccRCC cells are indicated. Bigger nodes correspond to centrality values in the Grade4/rhabdoid ccRCC PPI network. Color codes as in (B).

(D) Heat maps highlighting selected elements implicated in ECM remodelling and cell adhesion and their respective dysregulation during ccRCC de-differentiation on both data layers. Alterations of representative genes and proteins are depicted as log2FC with respect to adjacent normal cells. Color codes are depicted in (B). FDR is indicated by"****" <= 0.0001, "***" <= 0.001, "**"<= 0.01, "*" <= 0.05.

(E) IHC analysis of E-cadherin expression ccRCC with distinct grade morphologies. Representative images are shown for each grade group, the scale bar represents 20 μm. On the right: Expression levels were scored from 0 (no expression) to 3 (high expression) and the quantification of N=11 cases is shown.

Figure 6

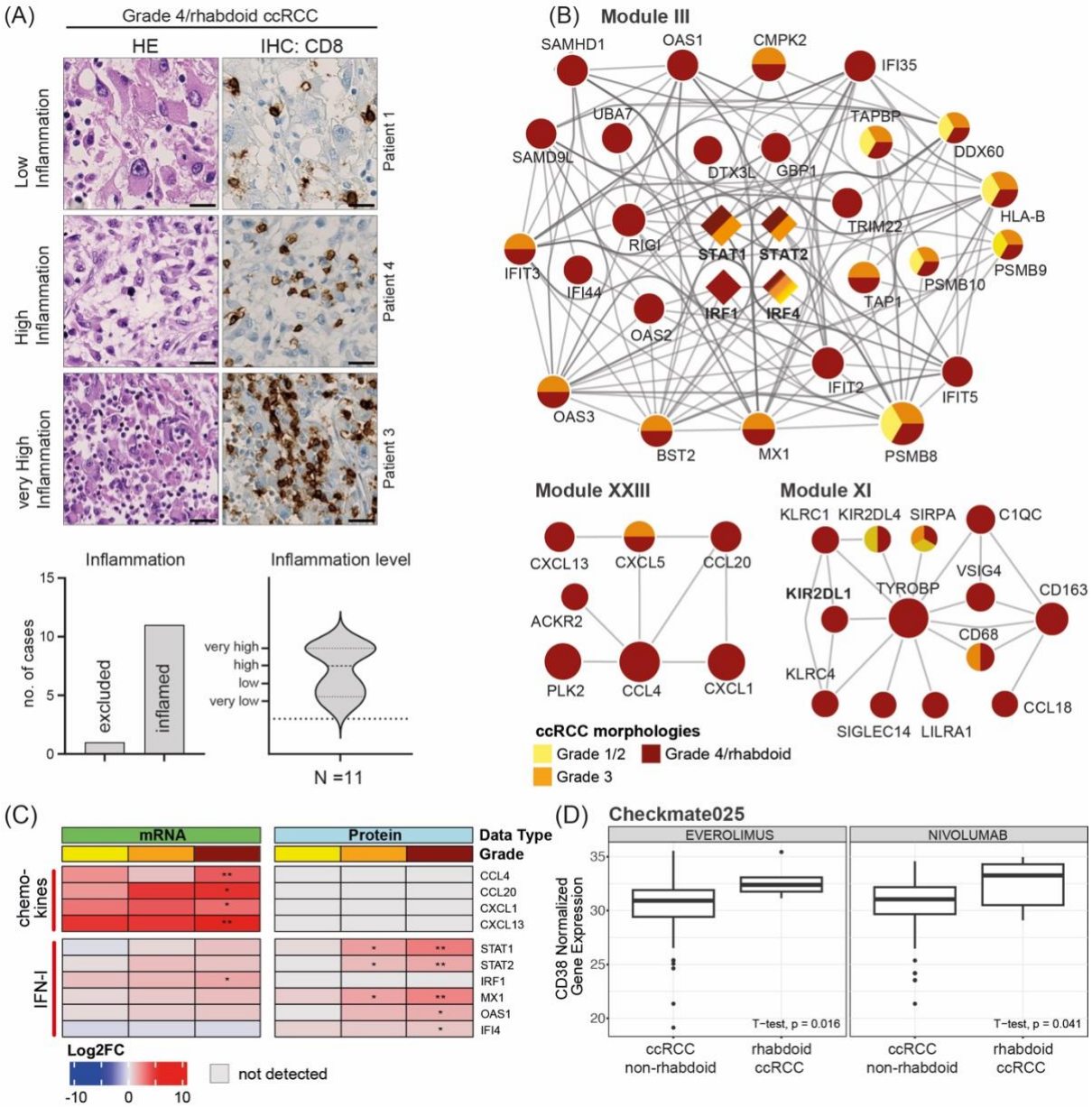


Figure 6: Aggressive ccRCC cells show upregulation of the immune response.

(A) IHC analysis using antibodies against CD8 depicting infiltration of CD8+ T cells into the rhabodoid component of the ccRCC cases in this study. CD8+T cells were scored and assigned to the depicted immune phenotypes (inflamed or excluded) by pathologist review. Inflamed tumors were further categorized into low, high or very high inflammation levels and representative images for each phenotype are shown. The scale bar represents 20 μm. We analyzed 12 cases with rhabdoid de-differentiation and quantified the frequency of inflammation (CD8+ T cells) as well as the level (low, high, very high) focussing on the rhabdoid component of the ccRCCs.

(B) Integrated network analysis showed a significant enrichment of upregulation of immune response elements in Modules III (Interferone Signaling, Rectome R-HSA-913531 and Immune System, Reactome R-HSA-168256), XI (Immune System, Reactome R-HSA-168256) and XXIII (Chemokine receptors bind chemokines, R-HSA-380108) of the grade4/rhabdoid ccRCC network. Factors that are significantly upregulated also in lower grade ccRCC cells are indicated. Bigger nodes correspond to higher centrality values in the grade4/rhabdoid ccRCC PPI network.

(C) Heat maps highlighting selected elements associated with the immune system-related pathways and their respective dysregulation during ccRCC de-differentiation on both data layers. Alterations of representative genes and proteins are depicted as log2FC with respect to adjacent normal cells. FDR is indicated by"****" <= 0.0001, "***" <= 0.001, "**"<= 0.01, "*" <= 0.05.

(D) Boxplots of normalized CD38 gene expression in non-rhabdoid vs. rhabdoid ccRCC patients from the CM-025 clinical trial cohort for the everolimus (Non-rhabdoid ccRCC N=108, Rhabdoid ccRCC N= 6) as well as nivolumab (Non-rhabdoid ccRCC N=139, Rhabdoid ccRCC N= 8) arms. Significance was tested using a Welch t-test.

Supplementary Figure Legends

Supplementary Figure 1: Image-guided single cell isolation and multi-omics analysis of ccRCC with pronounced morphological heterogeneity and rhabdoid de-differentiation.

(A) Representative H&E images indicating ITH of grade features: Grade 1/2 (low-grade, G1/2), grade 3 (high-grade, G3) and grade 4 ccRCC with rhabdoid differentiation (Rhab). Scale bars: WSI – 2 mm, high magnification panels - 100 µm.

(B) Overview images of WSI indicating the areas that were dissected to perform WGS.

(C) Mean coverage of ccRCC and adjacent normal kidney in WGS data for all five cases.

(D) Whole-genome mutation spectra indicating SNV classes, and their frequency, in our data highlighting the presence of C>T transitions caused by cytosine deamination from formalin fixation of the tissue samples.

(E) CNV profiles determined from WGS data. CNV loss/gain are indicated by color coding, blue indicated loss and red indicated gain.

Supplementary Figure 2: Analysis of mRNA and protein expression from individually selected cells with different grade morphologies.

(A) Correlation analysis for top100 variance features of the transcriptomic dataset obtained from interrogating cells isolated from normal kidney tubuli tissue or ccRCC with distinct grade morphologies (N=19). Correlation coefficients are indicated by the color code.

(B) Correlation analysis for proteomics samples isolated from normal kidney tubuli tissue or ccRCC with distinct grade morphologies (N=14). Correlation coefficients are indicated by the color code.

Supplementary Figure 3: Functional overlap between mRNA and protein data.

(A) Upset plots indicating the number of significants for mRNA and proteomics datasets for each grade group and their overlap.

(B) Bubble plot depicting all 24 shared GO terms for mRNA and protein differential expression in the grade 4/rhabdoid class of ccRCC cells and their overlap with samples from other grade morphologies.

(C) Heatmaps depicting selected alterations that mark kidney cell identity (proximal tubule cells markers and factors implicated in physiological kidney functions) as well as common ccRCC markers, cancer stem cell markers and factors implicated in one-carbon metabolism. Representative genes and proteins are depicted as log2FC with respect to adjacent normal cells. FDR is indicated by "****" <= 0.0001, "***" <= 0.001, "**"<= 0.01, "*" <= 0.05.

Supplementary Figure 4: Multi-omics data integration uncovers distinct de-regulated processes in aggressive cells and their transcriptional regulation.

(A) Biological processes enriched in modules shared by Grade 1/2 ccRCC grade morphology with at least 10 members following decomposition of the PPI networks in (A) based on KEGG and Reactome annotations. Enriched terms were determined by an FDR < 0.05 (Fisher's exact test

with Benjamini-Hochberg correction). Only non-redundant terms are presented, except where additional terms provided complementary information. The fraction of proteins in each module with the corresponding annotation is provided.

(B) Kaplan-Meier survival analysis of FOXM1 in TCGA-KIRC dataset (n = 516) showing overall survival. HR (high) = 0.66 p(HR) = 0.0079.

(C) Circular diagram highlighting the observed interaction between upregulated TFs (bottom) and their targets (top) within the top 30% nodes of the PPI networks based on the current-flow betweenness centrality for the three cell morphologies analyzed individually. Tracks mark the grade groups in which they were found to be significantly dysregulated (corresponds to color-code in (A)) and the bands color-code for the associated signaling pathways.

Supplementary Figure 5: Grade 4/rhabdoid ccRCC cells show distinct expression profiles.

(A) Kaplan-Meier survival analysis of ITGA4 with HR (high) = 1.8 p(HR) = 0.0022.

(B) As in (A) for CDH1 with HR (high) = 0.48 p(HR) = 5.9e-06.

(C) Heat maps highlighting selected elements associated with the immune system-related pathways and their respective dysregulation during ccRCC de-differentiation on both data layers. Alterations of representative genes and proteins are depicted as log2FC with respect to adjacent normal cells. FDR is indicated by"****" <= 0.0001, "***" <= 0.001, "**"<= 0.01, "*" <= 0.05.

References

1. Turajlic, S., Xu, H., Litchfield, K., Rowan, A., Horswell, S., Chambers, T., O'Brien, T., Lopez, J.I., Watkins, T.B.K., Nicol, D., et al. (2018). Deterministic Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal. Cell *173*, 595-610 e11. https://doi.org/10.1016/j.cell.2018.03.043.
2. Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., et al. (2012). Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. New England Journal of Medicine *366*, 883–892. https://doi.org/doi:10.1056/NEJMoa1113205.
3. Kara, O., Maurice, M.J., Zargar, H., Malkoc, E., Akca, O., Andrade, H.S., Ramirez, D., Caputo, P.A., Nelson, R.J., Rini, B., et al. (2016). Prognostic implications of sarcomatoid and rhabdoid differentiation in patients with grade 4 renal cell carcinoma. Int Urol Nephrol *48*, 1253–1260. https://doi.org/10.1007/s11255-016-1314-z.
4. Moch, H., Amin, M.B., Berney, D.M., Compérat, E.M., Gill, A.J., Hartmann, A., Menon, S., Raspollini, M.R., Rubin, M.A., Srigley, J.R., et al. (2022). The 2022 World Health Organization Classification of Tumours of the Urinary System and Male Genital Organs—Part A: Renal, Penile, and Testicular Tumours. European urology. https://doi.org/10.1016/j.eururo.2022.06.016.
5. Hsieh, J., Purdue, M., Signoretti, S., Swanton, C., Albiges, L., Schmidinger, M., Heng, D., Larkin, J., and Ficarra, V. (2017). Renal cell carcinoma. Nature Reviews Disease Primers *3*, 17009. https://doi.org/10.1038/nrdp.2017.9.
6. Monda, S.M., Lui, H.T., Pratsinis, M.A., Chandrasekar, T., Evans, C.P., and Dall'Era, M.A. (2023). The Metastatic Risk of Renal Cell Carcinoma by Primary Tumor Size and Subtype. Eur Urol Open Sci *52*, 137–144. https://doi.org/10.1016/j.euros.2023.04.015.
7. Hakimi, A.A., Reznik, E., Lee, C.-H., Creighton, C.J., Brannon, A.R., Luna, A., Aksoy, B.A., Liu, E.M., Shen, R., Lee, W., et al. (2016). An Integrated Metabolic Atlas of Clear Cell Renal Cell Carcinoma. Cancer Cell *29*, 104–116. https://doi.org/10.1016/j.ccell.2015.12.004.
8. The Cancer Genome Atlas Research Network (2013). Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature *499*, 43–49. https://doi.org/10.1038/nature12222.
9. Mitchell, T.J., Turajlic, S., Rowan, A., Nicol, D., Farmery, J.H.R., O'Brien, T., Martincorena, I., Tarpey, P., Angelopoulos, N., Yates, L.R., et al. (2018). Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. Cell *173*, 611-623 e17. https://doi.org/10.1016/j.cell.2018.02.020.
10. Delahunt, B., Eble, J.N., Egevad, L., and Samaratunga, H. (2019). Grading of renal cell carcinoma. *74*, 4–17. https://doi.org/10.1111/his.13735.
11. Moch, H., Cubilla, A.L., Humphrey, P.A., Reuter, V.E., and Ulbright, T.M. (2016). The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs—Part A: Renal, Penile, and Testicular Tumours. European urology *70*, 93–105. http://dx.doi.org/10.1016/j.eururo.2016.02.029.
12. Lopez, J.I., Guarch, R., Larrinaga, G., Corominas-Cishek, A., and Orozco, R. (2013). Cell heterogeneity in clear cell renal cell carcinoma. APMIS : acta pathologica, microbiologica, et immunologica Scandinavica *121*, 1187–1191. https://doi.org/10.1111/apm.12073.
13. Gulati, S., Martinez, P., Joshi, T., Birkbak, N.J., Santos, C.R., Rowan, A.J., Pickering, L., Gore, M., Larkin, J., Szallasi, Z., et al. (2014). Systematic evaluation of the prognostic impact and intratumour heterogeneity of clear cell renal cell carcinoma biomarkers. European urology *66*, 936–948. https://doi.org/10.1016/j.eururo.2014.06.053.
14. Burrell, R.A., McGranahan, N., Bartek, J., and Swanton, C. (2013). The causes and consequences of genetic heterogeneity in cancer evolution. Nature *501*, 338–345. https://doi.org/10.1038/nature12625.
15. Przybycin, C.G., McKenney, J.K., Reynolds, J.P., Campbell, S., Zhou, M., Karafa, M.T., and Magi-Galluzzi, C. (2014). Rhabdoid Differentiation Is Associated With Aggressive Behavior in Renal Cell Carcinoma: A Clinicopathologic Analysis of 76 Cases With Clinical Follow-up. The American Journal of Surgical Pathology *38*, 1260. https://doi.org/10.1097/PAS.0000000000000251.

16. Gökden, N., Nappi, O., Swanson, P.E., Pfeifer, J.D., Vollmer, R.T., Wick, M.R., and Humphrey, P.A. (2000). Renal Cell Carcinoma With Rhabdoid Features. The American Journal of Surgical Pathology *24*, 1329.
17. Hahn, A.W., Lebenthal, J., Genovese, G., Sircar, K., Tannir, N.M., and Msaouel, P. (2022). The significance of sarcomatoid and rhabdoid dedifferentiation in renal cell carcinoma. Cancer Treatment and Research Communications *33*, 100640. https://doi.org/10.1016/j.ctarc.2022.100640.
18. Wang, Z., Kim, T.B., Peng, B., Karam, J., Creighton, C., Joon, A., Kawakami, F., Trevisan, P., Jonasch, E., Chow, C.-W., et al. (2017). Sarcomatoid Renal Cell Carcinoma Has a Distinct Molecular Pathogenesis, Driver Mutation Profile, and Transcriptional Landscape. Clinical Cancer Research *23*, 6686–6696. https://doi.org/10.1158/1078-0432.Ccr-17-1057.
19. Malouf, G.G., Flippot, R., Dong, Y., Dinatale, R.G., Chen, Y.-B., Su, X., Compérat, E., Rouprêt, M., Mano, R., Blum, K.A., et al. (2020). Molecular characterization of sarcomatoid clear cell renal cell carcinoma unveils new candidate oncogenic drivers. Scientific Reports *10*, 701. https://doi.org/10.1038/s41598-020-57534-5.
20. Bakouny, Z., Braun, D.A., Shukla, S.A., Pan, W., Gao, X., Hou, Y., Flaifel, A., Tang, S., Bosma-Moody, A., He, M.X., et al. (2021). Integrative molecular characterization of sarcomatoid and rhabdoid renal cell carcinoma. Nat Commun *12*, 808. https://doi.org/10.1038/s41467-021-21068-9.
21. Şenbabaoğlu, Y., Gejman, R.S., Winer, A.G., Liu, M., Van Allen, E.M., de Velasco, G., Miao, D., Ostrovnaya, I., Drill, E., Luna, A., et al. (2016). Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. Genome Biol *17*, 231. https://doi.org/10.1186/s13059-016-1092-z.
22. Baxi, V., Edwards, R., Montalto, M., and Saha, S. (2022). Digital pathology and artificial intelligence in translational medicine and clinical practice. Modern Pathology *35*, 23–32. https://doi.org/10.1038/s41379-021-00919-2.
23. Kashima, Y., Sakamoto, Y., Kaneko, K., Seki, M., Suzuki, Y., and Suzuki, A. (2020). Single-cell sequencing techniques from individual to multiomics analyses. Experimental & Molecular Medicine *52*, 1419–1427. https://doi.org/10.1038/s12276-020-00499-2.
24. Marx, V. (2021). Method of the Year: spatially resolved transcriptomics. Nature methods *18*, 9–14. https://doi.org/10.1038/s41592-020-01033-y.
25. Brunner, A.-D., Thielert, M., Vasilopoulou, C., Ammar, C., Coscia, F., Mund, A., Hoerning, O.B., Bache, N., Apalategui, A., Lubeck, M., et al. (2022). Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. Molecular Systems Biology *18*, e10798. https://doi.org/10.15252/msb.202110798.
26. Mund, A., Brunner, A.-D., and Mann, M. (2022). Unbiased spatial proteomics with single-cell resolution in tissues. Molecular cell *82*, 2335–2349. https://doi.org/10.1016/j.molcel.2022.05.022.
27. Shen, C., Beroukhim, R., Schumacher, S.E., Zhou, J., Chang, M., Signoretti, S., and Kaelin, W.G. (2011). Genetic and Functional Studies Implicate HIF1α as a 14q Kidney Cancer Suppressor Gene. Cancer Discovery *1*, 222–235. https://doi.org/10.1158/2159-8290.Cd-11-0098.
28. Turajlic, S., Xu, H., Litchfield, K., Rowan, A., Chambers, T., Lopez, J.I., Nicol, D., O'Brien, T., Larkin, J., Horswell, S., et al. (2018). Tracking Cancer Evolution Reveals Constrained Routes to Metastases: TRACERx Renal. Cell *173*, 581-594 e12. https://doi.org/10.1016/j.cell.2018.03.057.
29. JOSEPH C. PRESTI, MONICA WILHELM, VICTOR REUTER, PAUL RUSSO, ROBERT MOTZER, and FREDERIC WALDMAN (2002). Allelic Loss On Chromosome 8 And 9 Correlates With Clinical Outcome In Locally Advances Clear Cell Carcinoma Of The Kidney. Journal of Urology *167*, 1464–1468. https://doi.org/doi:10.1016/S0022-5347(05)65346-3.
30. Gupta, S., Cheville, J.C., Jungbluth, A.A., Zhang, Y., Zhang, L., Chen, Y.-B., Tickoo, S.K., Fine, S.W., Gopalan, A., Al-Ahmadie, H.A., et al. (2019). JAK2/PD-L1/PD-L2 (9p24.1) amplifications in renal cell carcinomas with sarcomatoid transformation: implications for clinical management. Modern Pathology *32*, 1344–1358. https://doi.org/10.1038/s41379-019-0269-x.
31. Hollandi, R., Szkalisity, A., Toth, T., Tasnadi, E., Molnar, C., Mathe, B., Grexa, I., Molnar, J., Balind, A., Gorbe, M., et al. (2020). nucleAIzer: A Parameter-free Deep Learning Framework for Nucleus Segmentation Using Image Style Transfer. Cell Syst *10*, 453-458.e6. https://doi.org/10.1016/j.cels.2020.04.003.
32. Qu, Y., Feng, J., Wu, X., Bai, L., Xu, W., Zhu, L., Liu, Y., Xu, F., Zhang, X., Yang, G., et al. (2022). A proteogenomic analysis of clear cell renal cell carcinoma in a Chinese population. Nature Communications *13*, 2052. https://doi.org/10.1038/s41467-022-29577-x.

33. Weygant, N., Qu, D., May, R., Tierney, R.M., Berry, W.L., Zhao, L., Agarwal, S., Chandrakesan, P., Chinthalapally, H.R., Murphy, N.T., et al. (2015). DCLK1 is a broadly dysregulated target against epithelial-mesenchymal transition, focal adhesion, and stemness in clear cell renal carcinoma. Oncotarget *6*, 2193–2205. https://doi.org/10.18632/oncotarget.3059.
34. Fendler, A., Bauer, D., Busch, J., Jung, K., Wulf-Goldenberg, A., Kunz, S., Song, K., Myszczyszyn, A., Elezkurtaj, S., Erguen, B., et al. (2020). Inhibiting WNT and NOTCH in renal cancer stem cells and the implications for human patients. Nature Communications *11*, 929. https://doi.org/10.1038/s41467-020-14700-7.
35. Totu, T., Tomasoni, M., Bolck, H.A., and Buljan, M. (2024). NOODAI: A webserver for network-oriented multi-omics data analysis and integration pipeline. Preprint at bioRxiv, https://doi.org/10.1101/2024.11.08.622488 https://doi.org/10.1101/2024.11.08.622488.
36. Han, H., Cho, J.-W., Lee, S., Yun, A., Kim, H., Bae, D., Yang, S., Kim, C.Y., Lee, M., Kim, E., et al. (2018). TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. Nucleic Acids Research *46*, D380–D386. https://doi.org/10.1093/nar/gkx1013.
37. Arun, A.S., Tepper, C.G., and Lam, K.S. (2018). Identification of integrin drug targets for 17 solid tumor types. Oncotarget *9*, 30146–30162. https://doi.org/10.18632/oncotarget.25731.
38. Hu, J., Zhao, Q., Kong, L.-Y., Wang, J., Yan, J., Xia, X., Jia, Z., Heimberger, A.B., and Li, S. (2021). Regulation of tumor immune suppression and cancer cell survival by CXCL1/2 elevation in glioblastoma multiforme. Sci Adv *7*, eabc2511. https://doi.org/10.1126/sciadv.abc2511.
39. Zhang, L., Zhang, M., Wang, L., Li, J., Yang, T., Shao, Q., Liang, X., Ma, M., Zhang, N., Jing, M., et al. (2021). Identification of CCL4 as an Immune-Related Prognostic Biomarker Associated With Tumor Proliferation and the Tumor Microenvironment in Clear Cell Renal Cell Carcinoma. Front. Oncol. *11*. https://doi.org/10.3389/fonc.2021.694664.
40. Wang, D., Yang, L., Yu, W., Wu, Q., Lian, J., Li, F., Liu, S., Li, A., He, Z., Liu, J., et al. (2019). Colorectal cancer cell-derived CCL20 recruits regulatory T cells to promote chemoresistance via FOXO1/CEBPB/NF-κB signaling. J Immunother Cancer *7*, 215. https://doi.org/10.1186/s40425-019-0701-2.
41. Langbein, L.E., El Hajjar, R., He, S., Sementino, E., Zhong, Z., Jiang, W., Leiby, B.E., Li, L., Uzzo, R.G., Testa, J.R., et al. (2022). BAP1 maintains HIF-dependent interferon beta induction to suppress tumor growth in clear cell renal cell carcinoma. Cancer Letters *547*, 215885. https://doi.org/10.1016/j.canlet.2022.215885.
42. Piedra-Quintero, Z.L., Wilson, Z., Nava, P., and Guerau-de-Arellano, M. (2020). CD38: An Immunomodulatory Molecule in Inflammation and Autoimmunity. Front. Immunol. *11*. https://doi.org/10.3389/fimmu.2020.597959.
43. Xu, H., Zhang, A., Han, X., Li, Y., Zhang, Z., Song, L., Wang, W., and Lou, M. (2022). ITGB2 as a prognostic indicator and a predictive marker for immunotherapy in gliomas. Cancer Immunol Immunother *71*, 645–660. https://doi.org/10.1007/s00262-021-03022-2.
44. Lopez, J.I. (2016). Intratumor heterogeneity in clear cell renal cell carcinoma: a review for the practicing pathologist. APMIS : acta pathologica, microbiologica, et immunologica Scandinavica *124*, 153–159. https://doi.org/10.1111/apm.12500.
45. Singh, R.R., Murugan, P., Patel, L.R., Voicu, H., Yoo, S.Y., Majewski, T., Mehrotra, M., Wani, K., Tannir, N., Karam, J.A., et al. (2015). Intratumoral morphologic and molecular heterogeneity of rhabdoid renal cell carcinoma: challenges for personalized therapy. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc *28*, 1225–1235. https://doi.org/10.1038/modpathol.2015.68.
46. Arvaniti, E., Fricker, K.S., Moret, M., Rupp, N., Hermanns, T., Fankhauser, C., Wey, N., Wild, P.J., Rüschoff, J.H., and Claassen, M. (2018). Automated Gleason grading of prostate cancer tissue microarrays via deep learning. Scientific Reports *8*, 12054. https://doi.org/10.1038/s41598-018-30535-1.
47. L. Duran-Lopez, J. P. Dominguez-Morales, A. F. Conde-Martin, S. Vicente-Diaz, and A. Linares-Barranco (2020). PROMETEO: A CNN-Based Computer-Aided Diagnosis System for WSI Prostate Cancer Detection. IEEE Access *8*, 128613–128628. https://doi.org/10.1109/ACCESS.2020.3008868.
48. Diao, S., Hou, J., Yu, H., Zhao, X., Sun, Y., Lambo, R.L., Xie, Y., Liu, L., Qin, W., and Luo, W. (2020). Computer-Aided Pathologic Diagnosis of Nasopharyngeal Carcinoma Based on Deep

Learning. The American Journal of Pathology *190*, 1691–1700. https://doi.org/10.1016/j.ajpath.2020.04.008.
49. Lagree, A., Shiner, A., Alera, M.A., Fleshner, L., Law, E., Law, B., Lu, F.I., Dodington, D., Gandhi, S., Slodkowska, E.A., et al. (2021). Assessment of Digital Pathology Imaging Biomarkers Associated with Breast Cancer Histologic Grade. Curr Oncol *28*, 4298–4316. https://doi.org/10.3390/curroncol28060366.
50. Wang, X., Janowczyk, A., Zhou, Y., Thawani, R., Fu, P., Schalper, K., Velcheti, V., and Madabhushi, A. (2017). Prediction of recurrence in early stage non-small cell lung cancer using computer extracted nuclear features from digital H&E images. Sci Rep *7*, 13543. https://doi.org/10.1038/s41598-017-13773-7.
51. Wettersten, H.I., Aboud, O.A., Lara, P.N., and Weiss, R.H. (2017). Metabolic reprogramming in clear cell renal cell carcinoma. Nat Rev Nephrol *13*, 410–419. https://doi.org/10.1038/nrneph.2017.59.
52. Perrino, C.M., Hucthagowder, V., Evenson, M., Kulkarni, S., and Humphrey, P.A. (2015). Genetic alterations in renal cell carcinoma with rhabdoid differentiation. Human pathology *46*, 9–16. https://doi.org/10.1016/j.humpath.2014.09.001.
53. Xue, Y.-J., Xiao, R.-H., Long, D.-Z., Zou, X.-F., Wang, X.-N., Zhang, G.-X., Yuan, Y.-H., Wu, G.-Q., Yang, J., Wu, Y.-T., et al. (2012). Overexpression of FoxM1 is associated with tumor progression in patients with clear cell renal cell carcinoma. J Transl Med *10*, 200. https://doi.org/10.1186/1479-5876-10-200.
54. Chen, Y., Wang, Y., Cai, Z., Fan, X., and Zhang, Y. (2020). Integrin α7 is overexpressed and correlates with higher pathological grade, increased T stage, advanced TNM stage as well as worse survival in clear cell renal cell carcinoma patients: A retrospective study. J Clin Lab Anal *34*, e23034. https://doi.org/10.1002/jcla.23034.
55. Morozevich, G., Kozlova, N., Cheglakov, I., Ushakova, N., and Berman, A. (2009). Integrin alpha5beta1 controls invasion of human breast carcinoma cells by direct and indirect modulation of MMP-2 collagenase activity. Cell Cycle *8*, 2219–2225. https://doi.org/10.4161/cc.8.14.8980.
56. Li, C., Deng, T., Cao, J., Zhou, Y., Luo, X., Feng, Y., Huang, H., and Liu, J. (2023). Identifying ITGB2 as a Potential Prognostic Biomarker in Ovarian Cancer. Diagnostics *13*, 1169. https://doi.org/10.3390/diagnostics13061169.
57. Chen, L., Diao, L., Yang, Y., Yi, X., Rodriguez, B.L., Li, Y., Villalobos, P.A., Cascone, T., Liu, X., Tan, L., et al. (2018). CD38-Mediated Immunosuppression as a Mechanism of Tumor Cell Escape from PD-1/PD-L1 Blockade. Cancer Discovery *8*, 1156–1175. https://doi.org/10.1158/2159-8290.CD-17-1033.
58. Choueiri, T.K., and Motzer, R.J. (2017). Systemic Therapy for Metastatic Renal-Cell Carcinoma. The New England journal of medicine *376*, 354–366. https://doi.org/10.1056/NEJMra1601333.
59. Hamidi, H., and Ivaska, J. (2018). Every step of the way: integrins in cancer progression and metastasis. Nat Rev Cancer *18*, 533–548. https://doi.org/10.1038/s41568-018-0038-z.
60. Musella, M., Manic, G., De Maria, R., Vitale, I., and Sistigu, A. (2017). Type-I-interferons in infection and cancer: Unanticipated dynamics with therapeutic implications. Oncoimmunology *6*, e1314424. https://doi.org/10.1080/2162402X.2017.1314424.
61. Braun, D.A., Street, K., Burke, K.P., Cookmeyer, D.L., Denize, T., Pedersen, C.B., Gohil, S.H., Schindler, N., Pomerance, L., Hirsch, L., et al. (2021). Progressive immune dysfunction with advancing disease stage in renal cell carcinoma. Cancer Cell *39*, 632-648.e8. https://doi.org/10.1016/j.ccell.2021.02.013.
62. Bi, K., He, M.X., Bakouny, Z., Kanodia, A., Napolitano, S., Wu, J., Grimaldi, G., Braun, D.A., Cuoco, M.S., Mayorga, A., et al. (2021). Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma. Cancer cell *39*, 649-661.e5. https://doi.org/10.1016/j.ccell.2021.02.015.
63. Braun, D.A., Hou, Y., Bakouny, Z., Ficial, M., Angelo, M.S., Forman, J., Ross-Macdonald, P., Berger, A.C., Jegede, O.A., Elagina, L., et al. (2020). Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. Nat Med *26*, 909–918. https://doi.org/10.1038/s41591-020-0839-y.
64. Wright, M.N., Gola, D., and Ziegler, A. (2017). Preprocessing and Quality Control for Whole-Genome Sequences from the Illumina HiSeq X Platform. In Statistical Human Genetics: Methods and

Protocols, R. C. Elston, ed. (Springer New York), pp. 629–647. https://doi.org/10.1007/978-1-4939-7274-6_30.
65. Cibulskis, K., Lawrence, M.S., Carter, S.L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E.S., and Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol *31*, 213–219. https://doi.org/10.1038/nbt.2514.
66. Kim, S., Scheffler, K., Halpern, A.L., Bekritsky, M.A., Noh, E., Källberg, M., Chen, X., Kim, Y., Beyter, D., Krusche, P., et al. (2018). Strelka2: fast and accurate calling of germline and somatic variants. Nat Methods *15*, 591–594. https://doi.org/10.1038/s41592-018-0051-x.
67. McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., and Cunningham, F. (2016). The Ensembl Variant Effect Predictor. Genome Biology *17*, 122. https://doi.org/10.1186/s13059-016-0974-4.
68. Mayakonda, A., Lin, D.-C., Assenov, Y., Plass, C., and Koeffler, H.P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. *28*, 1747–1756. https://doi.org/10.1101/gr.239244.118.
69. Hatakeyama, M., Opitz, L., Russo, G., Qi, W., Schlapbach, R., and Rehrauer, H. (2016). SUSHI: an exquisite recipe for fully documented, reproducible and reusable NGS data analysis. BMC Bioinformatics *17*, 228. https://doi.org/10.1186/s12859-016-1104-8.
70. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.
71. Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J.E., Mudge, J.M., Sisu, C., Wright, J.C., Armstrong, J., Barnes, I., et al. (2021). GENCODE 2021. Nucleic Acids Research *49*, D916–D923. https://doi.org/10.1093/nar/gkaa1087.
72. Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res *41*, e108. https://doi.org/10.1093/nar/gkt214.
73. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.
74. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287. https://doi.org/10.1089/omi.2011.0118.
75. Leary, P., and Rehrauer, H. (2024). exploreDE Interactive Shiny App. (Zenodo). https://doi.org/10.5281/zenodo.10571252 https://doi.org/10.5281/zenodo.10571252.
76. Mund, A., Coscia, F., Kriston, A., Hollandi, R., Kovács, F., Brunner, A.-D., Migh, E., Schweizer, L., Santos, A., Bzorek, M., et al. (2022). Deep Visual Proteomics defines single-cell identity and heterogeneity. Nature Biotechnology. https://doi.org/10.1038/s41587-022-01302-5.
77. Panse, C., Trachsel, C., and Türker, C. (2022). Bridging data management platforms and visualization tools to enable ad-hoc and smart analytics in life sciences. Journal of Integrative Bioinformatics *19*. https://doi.org/10.1515/jib-2022-0031.
78. Demichev, V., Messner, C.B., Vernardis, S.I., Lilley, K.S., and Ralser, M. (2020). DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. Nat Methods *17*, 41–44. https://doi.org/10.1038/s41592-019-0638-x.
79. prolfqua: A Comprehensive R-Package for Proteomics Differential Expression Analysis | Journal of Proteome Research https://pubs.acs.org/doi/10.1021/acs.jproteome.2c00441.
80. Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., and Bork, P. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic acids research *47*, D607–D613.
81. Oughtred, R., Rust, J., Chang, C., Breitkreutz, B.-J., Stark, C., Willems, A., Boucher, L., Leung, G., Kolas, N., and Zhang, F. (2021). The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. Protein Science *30*, 187–200.
82. Kerrien, S., Aranda, B., Breuza, L., Bridge, A., Broackes-Carter, F., Chen, C., Duesbury, M., Dumousseau, M., Feuermann, M., and Hinz, U. (2012). The IntAct molecular interaction database in 2012. Nucleic acids research *40*, D841–D846.

83. Newman, M.E.J. (2005). A measure of betweenness centrality based on random walks. Social networks *27*, 39–54.
84. Choobdar, S., Ahsen, M.E., Crawford, J., Tomasoni, M., Fang, T., Lamparter, D., Lin, J., Hescott, B., Hu, X., and Mercer, J. (2019). Assessment of network module identification across complex diseases. Nature methods *16*, 843–852.
85. Tomasoni, M., Gómez, S., Crawford, J., Zhang, W., Choobdar, S., Marbach, D., and Bergmann, S. (2020). MONET: a toolbox integrating top-performing methods for network modularization. Bioinformatics *36*, 3920–3921.
86. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological) *57*, 289–300.
87. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research *13*, 2498–2504.
88. yWorks Gmb, H. (2024). yFiles.
89. Neuwirth, E., and Brewer, R.C. (2014). ColorBrewer palettes. R package version *1*.
90. Wickham, H., Hester, J., Francois, R., Bryan, J., Bearrows, S., and Posit, P.B.C. (2023). Package 'readr.' Read Rectangular Text Data. Available online: https://cran. r-project. org/web/packages/readr/readr. pdf (accessed on 28 June 2022).
91. Wickham, H., and Wickham, M.H. (2019). Package 'stringr.' Website: http://stringr. tidyverse. org, https://github. com/tidyverse/stringr.
92. Bolker, B., and Warnes, G. (2022). Lumley, T. gtools: Various R programming tools. R package version 3.9. 4.
93. Murrell, P. (2014). gridBase: Integration of base and grid graphics. R package version 0.4-7, 24.
94. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849.
95. Pedersen, T.L. (2023). tidygraph: A Tidy API for Graph Manipulation [R Package].
96. Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature Protocols *4*, 1184–1191.
97. Wickham, H., Bryan, J., Posit, P.B.C., Kalicinski, M., Valery, K., Leitienne, C., Colbert, B., Hoerl, D., and Miller, E. (2023). Package 'readxl.'