Renal cell carcinoma histologic subtypes exhibit distinct transcriptional profiles

Pedro Barata, …, Chadi Nabhan, Rana R. McKay

J Clin Invest. 2024. https://doi.org/10.1172/JCI178915.

Graphical abstract

Find the latest version:
https://jci.me/178915/pdf
Renal Cell Carcinoma Histologic Subtypes Exhibit Distinct Transcriptional Profiles

Authors
Pedro Barata¹,²*, Shuchi Gulati³*, Andrew Elliott⁴, Hans J. Hammers⁵, Earle Burgess⁶, Benjamin A. Gartrell⁷, Sourat Darabi⁸, Mehmet A. Bilen⁹, Arnab Basu¹⁰, Daniel M. Geynisman¹¹, Nancy A. Dawson¹², Matthew R. Zibelman¹¹, Tian Zhang⁶, Shuanzeng Wei¹¹, Charles J. Ryan¹³, Elisabeth I. Heath¹⁴, Kelsey A. Poorman⁴, Chadi Nabhan⁴, Rana R. McKay¹⁵

*Both authors are first authors

Affiliation
1 Tulane Medical School, New Orleans, Louisiana; 2 University Hospitals Seidman Cancer Center, Cleveland, Ohio; 3 UC Davis Health System, Sacramento, California; 4 CARIS Life Sciences, Irving, Texas; 5 UT Southwestern Kidney Cancer Program, Dallas, Texas; 6 Levine Cancer Institute Atrium Health, Charlotte, North Carolina; 7 Montefiore Medical Center/Albert Einstein College of Medicine, Bronx, New York; 8 Hoag Memorial Presbyterian Hospital, Newport Beach, California; 9 Winship Cancer Institute, Emory University, Atlanta, Georgia; 10 University of Alabama, Birmingham, Alabama; 11 Fox Chase Cancer Center, Philadelphia, Pennsylvania; 12 Georgetown University Lombardi Comprehensive Cancer Center, Washington, DC; 13 University of Minnesota, Minneapolis, Minnesota; 14 Barbara Ann Karmanos Cancer Institute, Wayne State University, Detroit, Michigan; 15 Moores Cancer Center, UC San Diego, California

Word Count: 3748

Corresponding Author:
Pedro C. Barata, MD, MSc, FACP
Co-Leader Genitourinary (GU) Disease Team
Director of GU Medical Oncology Research Program
University Hospitals Seidman Cancer Center
Associate Professor of Medicine
Case Western Reserve University
Case Comprehensive Cancer Center
11100 Euclid Avenue, Lakeside Suite 1200, R 1215
Cleveland, OH 44106
Phone: 216-262-1214
Pedro.Barata@UHhospitals.org

Rana R. McKay, MD
Genitourinary Oncology Lead
Moores Cancer Center,
Conflicts of Interest
AE, KP, CN are employees of Caris Life Sciences. Other authors declare no conflicts of interest relevant to this work.

Funding:
No funding

Patient Summary
Renal cell carcinoma histologic subtypes have distinct expression of gene sets representing key molecular pathways with potential to personalize treatments for patients.

Abstract
Molecular profiling of clear cell RCC (ccRCC) tumors of clinical trial patients has identified distinct transcriptomic signatures with predictive value, yet data in non-clear cell variants (nccRCC) are lacking. We examined the transcriptional profiles of RCC tumors representing key molecular pathways, from a multi-institutional, real-world patient cohort, including ccRCC (N=508) and centrally-reviewed nccRCC (N=149) samples. ccRCC had increased angiogenesis signature scores compared to the heterogeneous group of nccRCC tumors (mean z-score 0.37 vs -0.99, P<0.001), while cell cycle, fatty acid oxidation (FAO)/AMPK signaling, fatty acid synthesis (FAS)/pentose phosphate signature scores were increased in one or more nccRCC subtypes. Among both ccRCC and nccRCC tumors, T-effector scores statistically correlated with increased immune cell infiltration and were more commonly associated with immunotherapy-related markers (PD-L1+/TMB-High/MSI-High). In conclusion, this study provides evidence of differential gene transcriptional profiles among ccRCC vs nccRCC tumors, providing new insights for optimizing personalized and histology-specific therapeutic strategies for patients with advanced RCC.

Keywords: renal cell carcinoma, non-clear cell, gene expression signatures, molecular subgroups, T-effector, sarcomatoid, angiogenic

Introduction
Renal cell carcinoma (RCC) is a common cancer among men and women in the United States, with an estimated 81,800 new cases and 14,890 deaths expected in 2023. (1) Clear cell
RCC (ccRCC) is the most common subtype, representing 70-80% of all RCCs. (2, 3) Other variant histologies, which have been historically lumped together as non-clear cell RCC, have distinct clinical features and pathogenesis including papillary, chromophobe, medullary, collecting-duct, MiT family translocation RCC, succinate dehydrogenase-deficient RCCs, hereditary leiomyomatosis and syndrome-associated RCC and unclassified RCC. (3) Across all RCC histologies, 15-20% harbor sarcomatoid dedifferentiation, which portends poor prognosis, increased likelihood of presenting with advanced stage, and worse survival across all stages. (5)

Over the past decade, the medical management of advanced RCC has significantly changed with the emergence of the immune checkpoint inhibitors and next generation tyrosine kinase inhibitors (TKIs). Currently, front line treatment options include combined immune-oncology (IO)-IO or IO-TKI based treatment. (6-9) Vascular endothelial growth factor (VEGF) TKIs continue to be relevant and efficacious either as monotherapy or in combination with immunotherapy. (10) Tumors with rhabdoid/sarcomatoid dedifferentiation are associated with improvement in clinical outcomes including overall response rate (ORR) and progression-free survival (PFS) with IO-based approaches. (11-14)

While we have made great strides in improving survival for RCC patients in the modern era, outcomes to therapy are heterogeneous, with a subset of patients demonstrating long-term durability while others demonstrate intrinsic resistance to treatment. (6, 8, 9, 15) Most importantly, to date, there are no clinically applicable predictive biomarkers to help optimize therapy selection in the clinic. Common markers of response to immune checkpoint inhibitors, such as programmed cell death ligand 1 (PDL1) expression and tumor mutation burden (TMB) are at times associated with higher responses, yet they have not been applied clinically given the presence of observed responses in the absence of these markers. (16-18)

Important work has been done to identify transcriptomic signatures in both localized and metastatic ccRCC. Particularly in metastatic ccRCC, gene expression signatures have been described based on markers of angiogenesis and those of immune activation. The phase 2 IMmotion 150 trial evaluated the clinical relevance of T effector/IFNγ (Teff) and Angiogenesis high/low gene expression signatures identified by RNA sequencing. (16) Herein, the high Teff high signature was associated with longer PFS in the atezolizumab + bevacizumab group versus sunitinib group. By contrast, a high angiogenic signature was associated with improved PFS in the sunitinib group. Subsequently, the randomized, global phase 3 IMmotion 151
integrated multi-omic analyses leading to identification of robust molecular clusters derived from analyses of 823 tumors from patients with advanced RCC, including 134 tumors with sarcomatoid features. A total of seven gene clusters were identified by non-negative matrix factorization including inflammatory and angiogenic signatures. Cluster 1 and 2 were characterized by angiogenic genes (enriched for vascular and VEGF pathway-related genes), clusters 4, 5 and 7 showed increased expression of inflammatory pathways, and cluster 3 and 6 were characterized by high myeloid and low T-effector gene expression patterns. Differential outcomes to therapy were observed in each of the clusters, beginning to shed light on the potential clinical applicability of a biomarker selection strategy utilizing the cluster classification.

Other phase 3 trials such as Javelin Renal 101 and CheckMate 214 also investigated the predictive value of transcriptomic signatures. Using a different methodology (“Javelin Renal 101 Immuno signature”), a novel 26-gene expression signature derived from 720 tumors from patients enrolled on the Javelin Renal 101 trial was associated with PFS to treatment with axitinib + avelumab versus sunitinib. In the exploratory analysis of CheckMate 214 including 213 samples (20% of total study cohort) the immune-based signatures, whose scores were derived from three IMmotion150 signatures, the JAVELIN Renal 101 signature and tumor inflammation signature (TIS), were associated with PFS in patients treated with immune checkpoint inhibitors, but failed to show an association with overall survival (Checkmate 214), and the association between angiogenic gene expression and anti-VEGF therapies was also not statistically significant.

Data on gene expression signatures and other molecular characterization in different RCC histologies beyond ccRCC are lacking. Here, we present data from an international, multi-institutional, real-world cohort of RCC patients who have undergone comprehensive molecular evaluation. We aim to describe the gene expression signatures, mutational profiles and protein expression patterns across the different RCC histologies, including tumors with sarcomatoid/rhabdoid features and non-clear cell pathologies.

Results

1 - Study cohort and patient characteristics

The study cohort comprised of a total of 657 patient samples, including both clear cell (ccRCC, N=508) and non-clear cell RCC (nccRCC, N=149) histologic subtypes (Table 1, Figure 1).
Sarcomatoid and/or rhabdoid features were present in 9.4% of the overall cohort, with a significantly higher frequency in patients with nccRCC (14.1% vs 8.1% ccRCC, P=0.03), and specifically in chromophobe (20.0% vs 8.1%, P=0.03) and mixed subtypes (23.5% vs 8.1%, p<0.01). Papillary RCC tumors were associated with an increased median age at the time of biopsy compared to ccRCC, while medullary RCC was associated with a younger median age. MiT translocation RCC was more frequent among women (87.5% vs 30.1% ccRCC, P<0.01).

Distributions of age, gender, and tissue specimen source (N=337 collected from primary site, and N=322 from metastatic site) were similar between ccRCC and nccRCC subtypes.

2 - Transcriptional characterization and stratification of RCC patient samples into molecular subgroups

Prior studies of RCC have described molecular subgroups with gene expression signatures that reflect activation of key molecular pathways, including T-effector (comprised of CD274, CD8A, EOMES, IFNG and PRF1) and angiogenic (comprised of ANGPTL4, CD34, ESM1, KDR, KDR, PECAM1 and VEGFA) gene sets, and these subgroups were further associated with differential outcomes to therapy.(17)(20) We performed gene expression profiling of ten signatures in a cohort of real-world RCC tumor samples and characterized signature scores by histologic subtype (Figure 2).

Angiogenesis signature scores were significantly higher in ccRCC compared to all nccRCC subtypes (mean z-score 0.37 vs -0.99, P<0.001), along with highest median expression of complement cascade (mean z-score 0.13 vs -0.44, P<0.001) and T-effector signature scores (mean z-score 0.08 vs -0.27, P<0.001) observed in ccRCC (Figure 2A-B). Chromophobe RCC had increased fatty acid oxidation (FAO)/AMPK signaling scores (mean z-score 0.38 vs -0.02 in ccRCC, P<0.05). Stromal scores were increased in medullary RCC (mean z-score 0.74 vs 0.11 in ccRCC, P<0.05), with decreased scores observed for multiple signatures in both subtypes. MiT translocation RCC had increased angiogenesis with decreased complement cascade (mean z-score -0.60 vs 0.13 in ccRCC, P<0.05) and stromal scores (mean z-score -0.51 vs 0.11 in ccRCC, P<0.05). Cell cycle (mean z-score 0.78 vs -0.03 in ccRCC, P<0.05) and fatty acid synthesis (FAS)/pentose phosphate scores (mean z-score 0.97 vs -0.14 in ccRCC, P<0.001) were significantly increased in collecting duct carcinoma. Papillary and mixed tumors had
increased fatty acid synthesis (FAS)/pentose phosphate scores (mean z-score 0.72 and 0.48, respectively, P<0.001 each).

We next examined gene expression signatures for associations with patient demographic features (Figure 2C). Compared to younger patients, older patients were associated with decreased myeloid inflammation (mean z-score -0.15 vs 0.01, P<0.05) and stromal expression scores (mean z-score -0.13 vs 0.21, P<0.001). RCC samples from female patients had increased angiogenesis (mean z-score 0.24 vs 0.05, P<0.001), FAO/AMPK signaling (mean z-score 0.23 vs -0.02, P<0.001), and FAS/pentose phosphate scores (mean z-score 0.15 vs -0.01, P<0.05), while complement cascade (mean z-score -0.09 vs 0.03, P<0.05) and Ω-oxidation scores (mean z-score -0.13 vs -0.05, P<0.05) were decreased compared to male patients. Additionally, metastatic samples had higher cell cycle (mean z-score 0.19 vs -0.20, P<0.001), FAS/pentose phosphate (mean z-score 0.15 vs -0.07, P<0.01), stroma (mean z-score 0.37 vs -0.24, P<0.001), myeloid inflammation (mean z-score 0.03 vs -0.20, P<0.001), and complement cascade scores (mean z-score 0.05 vs -0.07, P<0.001) compared to specimens collected from the kidney.

3 – Genomic alterations are differentially associated with molecular subgroups across RCC histologies

The most common alteration among ccRCC was VHL (78%, N=396), which was associated with lower FAS/pentose phosphate signature scores (mean z-score difference -0.15 compared to VHL-wildtype tumors, P<0.05) (Figure 3A). Other commonly mutated genes among ccRCC included PBRM1 (47.7%, N=240) that associated with high angiogenesis scores (mean z-score difference 0.20, P<0.01) and low FAS/pentose phosphate scores (mean z-score difference -0.19, P<0.05), SETD2 (23.6%, N=116) that associated with cell cycle (mean z-score difference 0.41, P<0.001), FAS/pentose phosphate (mean z-score difference 0.26, P<0.05), and myeloid inflammation scores (mean z-score difference 0.24, P<0.01), and KDM5C (16.7%, N=64) that associated with increased complement cascade (mean z-score difference 0.31, P<0.001) and Ω-oxidation signature scores (mean z-score difference 0.30, P<0.001). In chromophobe RCC, mutations in TP53 (mean z-score difference 1.09, P<0.05), PTEN (mean z-score difference 1.28, P<0.05), and RB1 were most prevalent and each associated with
increased cell cycle scores (mean z-score difference 1.42, P<0.05), along with increased stromal scores in tumors with TP53 (mean z-score difference 1.48, P<0.05) and PTEN mutations (mean z-score difference 1.73, P<0.05) (Figure 3B). Alterations in SETD2, NF2, ARID1 and MLH1 were identified in samples from collecting duct carcinoma, although none were significantly associated with gene signatures (Figure 3C). In papillary RCC, mutations in ARID1A (9.5%, N=6) associated with decreased angiogenesis (mean z-score difference -0.68, P<0.01), cell cycle (mean z-score difference -0.89, P<0.05), FAO/AMPK signaling (mean z-score difference -0.70, P<0.05), FAS/pentose phosphate (mean z-score difference -1.14, P<0.05), and stromal scores (mean z-score difference -0.75, P<0.05), while SETD2 (11.5%, N=7) associated with increased snoRNA (mean z-score difference 0.63, P<0.05) and decreased T-effector scores (mean z-score difference -0.38, P<0.05) (Figure 2D). In mixed tumors, mutations in VHL were associated with increased angiogenesis scores (mean z-score difference 0.68, P<0.05), while BAP1 associated with increase angiogenesis (mean z-score difference 1.08, P<0.05) and decreased FAS/pentose phosphate scores (mean z-score difference -1.10, P<0.05) (Figure 3E).

4 - Molecular subgroups are associated with distinct tumor microenvironments

The presence of tumor-infiltrating lymphocytes predicts response to checkpoint inhibitor therapy, and we hypothesized that the gene expression profiles of molecular subgroups would be associated with differences in tumor microenvironment composition. Using the Microenvironment Cell Population-counter method(21), the relative abundance of immune and stromal populations in the tumor microenvironment was estimated from cell type-specific transcripts levels. In both ccRCC and nccRCC, the T-effector signature positively correlated with the presence of cytotoxic lymphocytes (Spearman ρ = 0.9, P<0.001), T cells/CD8+ T cells (ρ = 0.9, P<0.001), NK cells (ρ = 0.7, P<0.001), monocytic lineage (ρ = 0.6, P<0.001) and myeloid dendritic cell abundance (ρ = 0.6, P<0.001), as well as with a ‘T cell-inflamed’ signature that has been associated with response to immunotherapy (ρ = 0.9, P<0.001), and the expression of multiple immune checkpoint genes (ρ = 0.05 to 0.8, P<0.001) (Figure 4A). Endothelial cell and fibroblast abundance had the strongest association with angiogenesis (ρ = 0.9, P<0.001) and stromal cell scores (ρ = 0.9, P<0.001), respectively, in both ccRCC and nccRCC subtypes. Median abundance of cytotoxic lymphocytes, CD8+ T cells, NK
cells, myeloid dendritic cells, and endothelial cells was highest in ccRCC, while B lineage, fibroblasts, neutrophils, and monocytic lineage abundance was highest in collecting duct, medullary, papillary, and mixed RCC subtypes, respectively (Figure 4B).

Sarcomatoid/rhabdoid features were present in 9.4% of the overall cohort and, compared to ccRCC (8.1%, N=41), were significantly more frequent in chromophobe (20.0%, N=6, P<0.05) and mixed (23.5%, N=8, P<0.01) RCC subtypes (Figure 4C). Overall, 15.0% (N=97) of RCC samples were PDL1+ (staining of ≥2+ intensity and ≥5% tumor cells using SP142 antibody), with significantly higher frequency of PDL1+ tumors in medullary (37.5%, N=3, P<0.05), MiT translocation (42.9%, N=3, P<0.05), papillary (24.2%, N=14, P<0.05), and mixed (26.5%, N=9, P<0.05) RCC compared to ccRCC (12.0%, N=60, P<0.05). The overall median TMB was 4 mutations/megabase, and TMB-high (≥10 mutations/megabase) was observed in 1.9% (N=12) of all RCC samples, most frequently among collecting duct carcinoma (33.3%, N=2, vs ccRCC 1.8%, N=9, P<0.01), and often concurrent with dMMR/MSI-H status.

5 – Sarcomatoid/rhabdoid features are associated with unique molecular and immune profiles

The presence of sarcomatoid/rhabdoid features in both clear cell and nccRCC subtypes was associated with increased T-effector, cell cycle, myeloid inflammation, and stromal signature scores, as well as decreased FAO/AMPK signaling scores (Figure 5A). Interestingly, several associations between gene alteration and signature score varied by histological subtype and the presence of sarcomatoid/rhabdoid features (Figure 5B). For example, SETD2 mutations were associated with lower stromal scores in ccRCC with sarcomatoid/rhabdoid features (mean z-score difference -0.87, P<0.05) but higher stromal scores in ccRCC without sarcomatoid/rhabdoid features (mean z-score difference -0.87, P<0.05). However, TP53 mutations were similarly associated with decreased complement cascade scores in nccRCC, regardless of sarcomatoid/rhabdoid features (mean z-score difference -0.84 in sarcomatoid/rhabdoid +, -0.99 in sarcomatoid/rhabdoid-, P<0.01), in addition to increased stromal in sarcomatoid/rhabdoid+ (mean z-score difference 1.47, P<0.05) and increased angiogenesis scores in sarcomatoid/rhabdoid- (mean z-score difference 0.43, P<0.01).
**Discussion**

Our analysis of a large cohort of real-world patient samples is concordant with recent trial reports on gene expression signatures in ccRCC.\(^{(14, 17, 19)}\) As data on nccRCC are sparse, our findings among a subpopulation of centrally confirmed cases of nccRCC subtypes provide valuable insights into the specific molecular pathways and immune microenvironment of each RCC subtype and their associations with other clinical markers of interest. A better understanding of the molecular underpinnings and gene expression patterns across RCC subtypes will be critical for informing therapeutic strategies for patients with variant histology RCC, a group that has historically been underrepresented in clinical trials and continues to represent an unmet need. Our comparative analyses of ccRCC and nccRCC subtypes revealed histology-specific and biomarker-associated expression of key molecular pathways to provide new insights for these rare patient populations.

Clear cell samples were predominant in this study cohort, with a similar proportion of cases (77%) to real-world prevalence rates.\(^{(22, 23)}\) Concordant with other large ccRCC cohorts such as the Cancer Genome Atlas Research Network\(^{(24)}\), DNA sequencing data revealed frequent alterations in genes controlling cellular oxygen sensing (eg, \(VHL\)), as well as chromatin remodeling genes such as \(PBRM1\), \(SETD2\), and \(BAP1\). Both angiogenic and myeloid inflammation scores were higher in ccRCC compared to nccRCC tumors. The most abundant immune cell types in ccRCC samples were CD8+ T-cells, macrophages and CD4+ T-cells, consistent with previous reports.\(^{(25)}\) However, it has been shown that clear cell tumors are clustered into distinct molecular subgroups with different distribution of immune cells; in our analysis, the differential association of cell population with molecular subgroups seem to support such findings.\(^{(25)}\) Single-cell transcriptomic profiling of immune cells have detected a higher proportion of exhausted CD8+ T cell in advanced disease compared to earlier stages\(^{(26)}\) and higher levels of co-inhibitory receptors and effector molecules in cytotoxic T cells among responders to immunotherapy.\(^{(27)}\) At the somatic level, \(PBRM1\) mutations have been associated with a less immunogenic tumor microenvironment, upregulated angiogenesis, and suggested more limited benefit from immunotherapy.\(^{(28-30)}\) The lack of clinical annotation and integration of single-cell sequencing prevented us from confirming these findings and require further validation in future real-world datasets.
Papillary RCC was the most represented nccRCC subtype in our analysis, as expected from epidemiology studies. Papillary is no longer subclassified into type 1 and type 2, yet we found molecular alterations reported historically present in type 1 subtype such as MET alterations and type 2 subtype including chromatin modification (eg, ARID1A, SETD2), NRF2 pathway (eg, FH, NF2L2) and the Hippo pathway (eg, NF2). The lower angiogenic scores relative to ccRCC is concordant with the observed lower activity of anti-VEGF inhibitors in these tumors. Further, the presence of inflammatory gene scores, immune-related markers, and immune cell populations in these tumors might help explain the clinical efficacy that immune checkpoint inhibitors have shown in these tumors, either as monotherapy or combined with anti-VEGF TKIs.

To a lesser extent, our cohort included patients with papillary and other nccRCC subtypes and we identified differential gene expression scores: chromophobe RCC had increased fatty acid oxidation (FAO)/AMPK signaling scores while stromal scores were increased in medullary RCC. Cell cycle and fatty acid synthesis (FAS)/pentose phosphate scores were significantly increased in collecting duct carcinoma. Chromophobe RCC is known to be associated with multiple losses of chromosomes 1, 2, 6, 10, 13, 17 and 21, and TP53 and PTEN are the two most frequently mutated genes. Genomic structural arrangements involving the TERT promoter region, as well as diffusely increased mitochondrial function and mitochondrial DNA alterations, are more common in chromophobe RCC, which was identified in our cohort as well. Sarcomatoid/rhabdoid features were frequently found (20%) in these tumors as previously reported, yet immunotherapies continue to show limited activity in these tumors. Of note, non-sarcomatoid chromophobe tumors had similar mutation frequencies of TP53 (61%), RB1 (15%), and PTEN (13%) as the overall analysis, along with similar expression of the ten gene sets representing key molecular pathways, with exception of the “stroma” gene set that enriched in chromophobe tumors with sarcomatoid features present.

Collecting duct samples, which are characterized by frequent genomic alterations involving NF2, SETD2, ARID1A, and SMARCB1, had the highest median myeloid inflammation expression score while among the lowest angiogenesis score. These findings may help to explain the clinical reports of relative success of mTOR inhibitors in the NF2-
mutated cases, as well as disease control rates with immune checkpoint inhibitors, while anti-
angiogenic therapies and chemotherapy are of limited value.(41) Owing to the rarity of MiT Translocation (tRCC), our cohort included only a limited
number of molecularly confirmed cases, which had a clear female predominance and younger
age at presentation, as expected.(42) Angiogenesis, complement cascade, and stroma
expression scores were decreased compared to ccRCC, but the lack of recurrent co-alterations
precluded further analysis of biomarker associations.

Finally, there was a strong association between sarcomatoid/rhabdoid+ tumors and
high myeloid inflammation scores and low angiogenic scores. While this association has been
observed in some trial reports (eg, IMmotion151) but not others (eg, CheckMate 214),
variations in methodologies of analysis and availability of tissue samples across these studies
limit cross trial comparisons of this correlative data.(19, 20) While we highlight results from a large dataset of genomically profiled distinct RCC tumors,
there are several limitations to this work. Limited clinical data available in the database
prevented us from investigating the presence of the gene expression scores by IMDC
prognostic risk groups. Similarly, the predictive value of the transcriptomic scores could not
be assessed. Rarer forms of RCC, such as collecting duct, medullary and translocation RCC,
were under-represented in this cohort and require molecular profiling of additional samples
in future studies to verify results. While we presume that most samples were submitted for
molecular profiling at the time of advanced disease based on clinical guidelines for molecular
testing, precise staging information was not available. The impact of systemic therapies on
the molecular characterization of tumors is largely unknown and tumor clonal heterogeneity
and evolution could not be assessed. Future studies in both clear cell such as OPTIC trial
(NCT05361720) and in variant RCC subtypes that incorporate gene expression scores, are
required to validate their predictive value, and help with patient selection.

In conclusion, despite these limitations, we were able to identify distinct
transcriptional profiles across multiple RCC histologies from a large cohort of real-world RCC
patient samples. The findings of our work are concordant with prior trial data, suggesting
potential clinical significance and therapeutic implications. Future directions include
independent prospective validation of these findings in the context of different systemic
therapies that are currently available or under development.
Figures and Tables:

Figure 1 – Study flow diagram

Figure 2 – RCC subtypes exhibit distinct gene expression profiles. Differential expression of 10 gene sets representing key molecular pathways by RCC subtype (A). Radial plots of the median gene signature expression level by RCC subtype (B) and patient demographics (C).

Figure 3 – Genomic alterations associated with gene signatures across RCC histologies. Oncoprint of the most commonly altered genes, with heatmap indicating the difference in gene signature score differences between biomarker-positive (i.e. mutated) and -negative tumors, in clear cell (A), chromophobe (B), collecting duct (C), papillary (D), and mixed tumors (E). Note: Genes with < 2 altered samples were excluded. *P<0.05.

Figure 4 – Gene signatures are associated with unique tumor microenvironments. (A) Heatmap of immunotherapy (IO)-related biomarkers, relative abundance of immune and stromal cell population estimated from RNA expression, and expression of key immune checkpoint genes across all RCC samples, with adjacent heatmap indicating the correlation strength with gene signatures. (B) Radial plot of the median relative abundance of cell types by RCC subtype. (C) Prevalence of IO-related biomarkers by RCC subtype.

Figure 5 – Sarcomatoid/rhabdoid features are associated with a distinct expression profiles. (A) Radial plot of the median gene signature expression level by RCC subtype. (B) Heatmap of gene signature score differences between biomarker-positive (i.e. mutated) and -negative tumors. Note: Genes with < 2 altered samples were excluded.

Supplemental Figure S1 - Radial plots of the median gene signature expression level by patient demographics.

Supplemental Figure S2 - Heatmap of gene signature score differences between biomarker-positive (i.e. mutated) and -negative tumors. Note: Genes with < 2 altered samples were excluded. *P<0.05.

Supplemental Figure S3 - Radial plots of the median gene signature expression level for each RCC subtype, including clear cell (A), chromophobe (B), collecting duct (C), medullary (D), MiT family translocation (E), mixed (F), and papillary (G). Black dotted line represents the overall study cohort median expression level.

Table 1 – Study cohort characteristics by RCC histological subtype

Table 2 – Study cohort characteristics by the presence of sarcomatoid/rhabdoid features

Methods

Sex as a biological variant

Samples from both males and females were involved in this research as the findings do apply to both groups.

Study cohort

Clinical physicians submitted formalin-fixed paraffin-embedded (FFPE) samples from patients with kidney cancer (N=657) to a commercial CLIA-certified laboratory for molecular
profiling (Caris Life Sciences, Phoenix, AZ) (Figure 1). All tumor samples categorized as variant histologies underwent central pathology review at Caris. Tumors classified as mixed subtypes included samples with histologic features of more than one subtype, most commonly papillary with clear cell changes, or unspecific features. The MiT family translocation subtype was confirmed by tumor genomic sequencing.

Clinical characteristics
Limited baseline clinical factors such as age and sex as a biological variable (male, female) were available and included in this study.

DNA Next-Generation Sequencing (NGS)
NGS was performed on isolated genomic DNA using the NextSeq platform (Illumina, Inc., San Diego, CA) for 592 cancer-relevant genes (N=375 samples) or the Illumina NovaSeq 6000 platform (Illumina, Inc., San Diego, CA) for whole exome sequencing (WES) (N=282 samples).

Prior to molecular testing, tumor enrichment was achieved by harvesting targeted tissue using manual microdissection techniques. A custom-designed SureSelect XT assay was used to enrich exonic regions of 592 whole-gene targets (Agilent Technologies, Santa Clara, CA).

All variants were detected with > 99% confidence based on allele frequency and amplicon coverage, with an average sequencing depth of coverage of > 500 and an analytic sensitivity threshold of 5% established for variant calling. For WES, a hybrid pull-down panel of baits designed to enrich for more than 700 clinically relevant genes at high coverage and high read-depth was used, along with another panel designed to enrich for an additional >20,000 genes at lower depth, and a 500Mb SNP backbone panel (Agilent Technologies) was added to assist with gene amplification/deletion measurements and other analyses. Genomic variants were classified by board-certified molecular geneticists according to criteria established by the American College of Medical Genetics and Genomics (ACMG). When assessing mutation frequencies of individual genes, ‘pathogenic,’ and ‘likely pathogenic’ were counted as mutations while ‘benign’, ‘likely benign’ variants and ‘variants of unknown significance’ were excluded.

RNA Whole Transcriptome Sequencing (WTS) and fusion detection
WTS uses a hybrid-capture method to pull down the full transcriptome from FFPE tumor samples using the Agilent SureSelect Human All Exon V7 bait panel (Agilent Technologies, Santa Clara, CA) and the Illumina NovaSeq platform (Illumina, Inc., San Diego, CA). FFPE
specimens underwent pathology review to discern the percent tumor content and tumor size; a minimum of 10% tumor content in the area for microdissection was required to enable enrichment and extraction of tumor-specific RNA. Qiagen RNA FFPE tissue extraction kit was used for extraction, and the RNA quality and quantity were determined using the Agilent TapeStation. Biotinylated RNA baits were hybridized to the synthesized and purified cDNA targets, and the bait-target complexes were amplified in a post-capture PCR reaction. The resultant libraries were quantified and normalized, and the pooled libraries were denatured, diluted, and sequenced. Raw data was demultiplexed using the Illumina DRAGEN FFPE accelerator. FASTQ files were aligned with STAR aligner (Alex Dobin, release 2.7.4a github). A full 22,948-gene dataset of expression data was produced by the Salmon, which provides fast and bias-aware quantification of transcript expression(43) BAM files from STAR aligner were further processed for RNA variants using a proprietary custom detection pipeline. The reference genome used was GRCh37/hg19, and analytical validation of this test demonstrated ≥ 97% Positive Percent Agreement (PPA), ≥ 99% Negative Percent Agreement (NPA), and ≥ 99% Overall Percent Agreement (OPA) with a validated comparator method. Identified fusion transcripts were further evaluated to determine breakpoint positions and functional domains retained from fused genes.

RNA expression analyses

Previously described gene sets that represent key molecular pathways among transcriptionally distinct RCC subpopulations were evaluated.(17) Gene expression values were log-transformed and standardized to z-scores, with a composite signature score calculated as the mean z-score of the gene set for each sample. To assess the relative abundance of immune and stromal cell populations in the tumor microenvironment, gene expression values were analyzed using the Microenvironment Cell Populations (MCP)-counter tool.(21)

Immunohistochemistry (IHC)

IHC was performed on full formalin-fixed paraffin-embedded (FFPE) sections of glass slides. Slides were stained using the Agilent DAKO Link 48 (Santa Clara, CA) automated platform and staining techniques, per the manufacturer’s instructions, and were optimized and validated per CLIA/CAP and ISO requirements. Staining was scored for intensity (0 = no staining; 1+ = weak staining; 2+ = moderate staining; 3+ = strong staining) and staining percentage (0-100%).
PDL1 (SP142) staining results were categorized as positive (≥2+ and ≥5% tumor cells) or negative (0 or 0%).

**Tumor Mutational Burden (TMB)**

TMB was measured by counting all non-synonymous missense, nonsense, in-frame insertion/deletion, and frameshift mutations found per tumor that had not been previously described as germline alterations in dbSNP151, Genome Aggregation Database (gnomAD) databases, or benign variants identified by Caris's geneticists. A cutoff point of ≥ 10 mutations per megabase (mt/MB) was used based on the KEYNOTE-158 pembrolizumab trial. (44)

**Statistical analysis**

All statistical analyses were performed with JMP V13.2.1 (SAS Institute) or R Version 3.6.1 (https://www.R-project.org). Continuous data were assessed using Mann-Whitney U test, and categorical data was evaluated using Chi-square or Fisher’s exact test, where appropriate.

**Study approval**

The present study was conducted in accordance with the guidelines of the Declaration of Helsinki, Belmont Report, and US Common Rule. With compliance to policy 45 CFR 46.101(b), this study was conducted using retrospective, de-identified clinical data, and patient consent was not required.

**Data availability statement**

The datasets generated during and/or analyzed during the current study (including the figures in the manuscript and supplement) are available from the corresponding author on reasonable request. The deidentified sequencing data are owned by Caris Life Sciences, and qualified researchers can apply for access to these summarized data by contacting Andrew Elliott, PhD and signing a data usage agreement.

**References**


Table 1. Study cohort characteristics by RCC histological subtype.  

Mixed tumors included samples with histologic features of more than one subtype, most commonly papillary with clear cell changes, or unspecific features. *P<0.05, **P<0.01 when compared to clear cell subtype.

<table>
<thead>
<tr>
<th>Histologic subtype</th>
<th>Tumors N (%)</th>
<th>Male N (%)</th>
<th>Female N (%)</th>
<th>Median Age at Tissue Collection (Range)</th>
<th>Primary N (%)</th>
<th>Metastatic N (%)</th>
<th>Sarcomatoid/ Rhabdoid features (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear cell</td>
<td>508 (77.3%)</td>
<td>355 (69.9%)</td>
<td>153 (30.1%)</td>
<td>62 (19-90+)</td>
<td>250 (49.2%)</td>
<td>258 (50.8%)</td>
<td>41 (8.1%)</td>
</tr>
<tr>
<td>Papillary</td>
<td>63 (9.6%)</td>
<td>50 (79.4%)</td>
<td>13 (20.6%)</td>
<td>66* (21-87)</td>
<td>39 (61.9%)</td>
<td>24 (38.1%)</td>
<td>5 (7.9%)</td>
</tr>
<tr>
<td>Mixed</td>
<td>34 (5.2%)</td>
<td>26 (76.5%)</td>
<td>8 (23.5%)</td>
<td>63 (48-81)</td>
<td>15 (44.1%)</td>
<td>19 (55.9%)</td>
<td>8 (23.5%)**</td>
</tr>
<tr>
<td>Chromophobe</td>
<td>30 (4.6%)</td>
<td>21 (70.0%)</td>
<td>9 (30.0%)</td>
<td>63 (24-77)</td>
<td>17 (56.7%)</td>
<td>13 (43.3%)</td>
<td>6 (20.0%)*</td>
</tr>
<tr>
<td>MiT Translocation</td>
<td>8 (1.2%)</td>
<td>1 (12.5%)**</td>
<td>7 (87.5%)</td>
<td>54 (30-72)</td>
<td>6 (75.0%)</td>
<td>2 (25.0%)</td>
<td>1 (12.5%)</td>
</tr>
<tr>
<td>Medullary</td>
<td>8 (1.2%)</td>
<td>7 (87.5%)</td>
<td>1 (12.5%)</td>
<td>23.5** (14-41)</td>
<td>5 (50.0%)</td>
<td>5 (50.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Collecting duct</td>
<td>6 (0.9%)</td>
<td>4 (66.7%)</td>
<td>2 (33.3%)</td>
<td>63.5 (61-75)</td>
<td>5 (83.3%)</td>
<td>1 (16.7%)</td>
<td>1 (16.7%)</td>
</tr>
</tbody>
</table>
Table 2 – Study cohort characteristics by the presence of sarcomatoid/rhabdoid features

<table>
<thead>
<tr>
<th>Histologic subtype</th>
<th>Sarc/Rhab (</th>
<th>Tumors N (%)</th>
<th>Male N (%)</th>
<th>Female N (%)</th>
<th>Median Age (Range)</th>
<th>Primary N (%)</th>
<th>Metastatic N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear cell</td>
<td>+</td>
<td>41 (8.1%)</td>
<td>24 (58.5%)</td>
<td>17 (41.5%)</td>
<td>57 (19-82)</td>
<td>34 (82.9%)</td>
<td>7 (17.1%)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>467 (91.9%)</td>
<td>311 (70.9%)</td>
<td>136 (29.1%)</td>
<td>62 (28-90+)</td>
<td>216 (46.3%)</td>
<td>251 (53.7%)</td>
</tr>
<tr>
<td>Non-clear cell</td>
<td>+</td>
<td>21 (14.1%)</td>
<td>16 (71.4%)</td>
<td>6 (28.6%)</td>
<td>63 (49-83)</td>
<td>13 (61.9%)</td>
<td>8 (38.1%)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>128 (85.9%)</td>
<td>94 (73.4%)</td>
<td>34 (26.6%)</td>
<td>63 (14-87)</td>
<td>73 (57.0%)</td>
<td>55 (43.0%)</td>
</tr>
</tbody>
</table>
Figure 1 - Consort diagram of study inclusion process.

1031 RCC patient samples submitted to Caris Life Sciences for molecular profiling

Excluded 348 samples “Carcinoma, Not otherwise specified (NOS)” histology

26 samples with rare histology (e.g. Angiomyolipoma, Non-small cell, Oncocytoma, Wilms tumor, etc.)

Central review of 149 non-ccRCC samples to confirm histology

Final cohort of 657 total samples (508 ccRCC and 149 non-ccRCC)
**Figure 2** - RCC subtypes exhibit distinct gene expression profiles. Differential expression of 10 gene sets representing key molecular pathways by RCC subtype (A). Radial plots of the median gene signature expression level by RCC subtype (B). Mann-Whitney U test: *P<0.05, **P<0.01, ***P<0.001 when compared to ccRCC.
Figure 3 - Genomic alterations associated with gene signatures across RCC histologies. Oncoprint of the most commonly altered genes, with heatmap indicating the difference in gene signature score differences between biomarker-positive (i.e. mutated) and -negative tumors, in clear cell (A), chromophobe (B), collecting duct (C), papillary (D), and mixed tumors (E). Note: Genes with < 2 altered samples were excluded. *P<0.05.
Figure 4 - Association of gene scores with unique tumor microenvironments. (A) Heatmap of immunotherapy (IO)-related biomarkers, relative abundance of immune and stromal cell population estimated from RNA expression, and expression of key immune checkpoint genes across all RCC samples, with adjacent heatmap indicating the Spearman correlation strength with gene scores. (B) Radial plot of the median relative abundance of cell types by RCC subtype. (C) Prevalence of IO-related biomarkers by RCC subtype. *P<0.05, **P<0.01 when compared to ccRCC.
**Figure 5** - Sarcomatoid/rhabdoid features are associated with distinct expression profiles. Radial plot of the median gene signature expression level by subgroups.
Supplemental Figure S1 - Radial plots of the median gene signature expression level by patient demographics.
**Supplemental Figure S2** - Heatmap of gene signature score differences between biomarker-positive (i.e. mutated) and -negative tumors. Note: Genes with < 2 altered samples were excluded. Mann-Whitney U test: *P<0.05.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Overall N</th>
<th>Clear cell +Sarco/Rhab</th>
<th>Non-Clear cell +Sarco/Rhab</th>
</tr>
</thead>
<tbody>
<tr>
<td>VHL</td>
<td>65.9%</td>
<td>27</td>
<td>6</td>
</tr>
<tr>
<td>PBRM1</td>
<td>31.7%</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>SETD2</td>
<td>27.5%</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>BAP1</td>
<td>15.0%</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>TP53</td>
<td>15.0%</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>PTEN</td>
<td>10.0%</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>KDM5C</td>
<td>9.7%</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>ATM</td>
<td>7.3%</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>ARID1A</td>
<td>4.9%</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>MTOR</td>
<td>4.9%</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>SMARCB1</td>
<td>4.9%</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>TSC1</td>
<td>4.9%</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Clear cell -Sarco/Rhab

<table>
<thead>
<tr>
<th>Gene</th>
<th>Overall N</th>
<th>Clear cell -Sarco/Rhab</th>
<th>Non-Clear cell -Sarco/Rhab</th>
</tr>
</thead>
<tbody>
<tr>
<td>VHL</td>
<td>79.2%</td>
<td>369</td>
<td>9</td>
</tr>
<tr>
<td>PBRM1</td>
<td>49.2%</td>
<td>227</td>
<td>9</td>
</tr>
<tr>
<td>SETD2</td>
<td>23.3%</td>
<td>105</td>
<td>9</td>
</tr>
<tr>
<td>KDM5C</td>
<td>17.4%</td>
<td>61</td>
<td>9</td>
</tr>
<tr>
<td>BAP1</td>
<td>11.9%</td>
<td>53</td>
<td>9</td>
</tr>
<tr>
<td>TP53</td>
<td>8.2%</td>
<td>37</td>
<td>9</td>
</tr>
<tr>
<td>PTEN</td>
<td>7.5%</td>
<td>34</td>
<td>9</td>
</tr>
<tr>
<td>MTOR</td>
<td>5.0%</td>
<td>23</td>
<td>9</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>5.0%</td>
<td>23</td>
<td>9</td>
</tr>
<tr>
<td>TSC1</td>
<td>4.1%</td>
<td>19</td>
<td>9</td>
</tr>
<tr>
<td>ARID1A</td>
<td>3.9%</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>NF2</td>
<td>3.7%</td>
<td>17</td>
<td>9</td>
</tr>
</tbody>
</table>

Signature Score Difference

Biomarker Positive - Negative
Supplemental Figure S3 - Radial plots of the median gene signature expression level for each RCC subtype, including clear cell (A), chromophobe (B), collecting duct (C), medullary (D), MiT family translocation (E), mixed (F), and papillary (G). Black dotted line represents the overall study cohort median expression level.