Genomic and transcriptomic landscape of advanced renal cell cancer to individualize treatment strategy

K. de Joode[¥1], W.S. van de Geer[¥1,2], G.J.L.H. van Leenders[3], P. Hamberg[4], H.M. Westgeest[5], A. Beeker[6], S.F. Oosting[7], J.M. van Rooijen[8], L.V. Beerepoot[9], M. Labots[10], R.H.J. Mathijssen[1], M.P. Lolkema[1,11], E. Cuppen[12,13], S. Sleijfer[1,11], H.J.G. van de Werken[x*2,14,15], A.A.M. van der Veldt[x*1,16]

[¥]*Both authors contributed equally and are considered first author*

[x]*Both authors contributed equally and are considered last author*

**Corresponding author*

[1]Department of Medical Oncology, Erasmus MC Cancer Institute, University Medical Center, Rotterdam, The Netherlands;
[2]Cancer Computational Biology Center, Erasmus MC Cancer Institute, University Medical Center, Rotterdam, The Netherlands;
[3]Department of Pathology, Erasmus MC, Rotterdam, the Netherlands;
[4]Department of Internal Medicine, Franciscus Gasthuis & Vlietland, Rotterdam, The Netherlands;
[5]Department of Internal Medicine, Amphia Hospital, Breda, The Netherlands;
[6]Department of Internal Medicine, Spaarne Gasthuis, Hoofddorp, The Netherlands;
[7]Department of Medical Oncology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands;
[8]Department of Internal Medicine, Martini Hospital, Groningen, The Netherlands;
[9]Department of Internal Medicine, Elisabeth-Tweesteden hospital, Tilburg, The Netherlands;
[10]Department of Medical Oncology, Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Medical Oncology, Cancer Center Amsterdam, The Netherlands;
[11]Center for Personalized Cancer Treatment, Rotterdam, The Netherlands;
[12]Center for Molecular Medicine and Oncode Institute, University Medical Center Utrecht, Utrecht, The Netherlands;
[13]Hartwig Medical Foundation, Amsterdam, The Netherlands;
[14]Department of Urology, Erasmus MC Cancer Institute, University Medical Center, Rotterdam, The Netherlands;
[15]Department of Immunology, Erasmus MC Cancer Institute, University Medical Center, Rotterdam, The Netherlands;
[16]Departments of Radiology & Nuclear Medicine, Erasmus MC, Rotterdam, The Netherlands

*Corresponding author(s)

Astrid A.M. van der Veldt, Departments of Medical Oncology and Radiology & Nuclear Medicine, Erasmus MC Cancer Institute, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands.
Tel: +31 (0)10 704 0 704. E-mail address: a.vanderveldt@erasmusmc.nl

Harmen J.G. van de Werken, Cancer Computational Bioldatogy Center, Erasmus MC Cancer Institute, Erasmus University Medical Center Rotterdam, Department of Immunology and Department of Urology, Internal Postal Address NA-1218-335, PO Box 2040, 3000 CA Rotterdam, the Netherlands.
Tel: +31 (0)10 703 1 829. Fax: +31 (0)10 704 4 731. E-mail address: h.vandewerken@erasmusmc.nl

Abstract

The genomic landscape of advanced renal cell cancer (RCC) was characterized for 91 patients to identify actionable targets and signatures by combining whole-genome sequencing (WGS) with matched RNA sequencing (RNA-Seq). WGS data were analyzed for somatic small variants, copy-number alterations (CNAs) and structural variants. Somatic aberrations were analyzed for driver genes, CNA drivers, mutational signatures, catastrophic events and fusion genes. For papillary and clear cell RCC, potential actionable drug targets were detected by WGS in 89% and 100% of the patients, respectively. RNA-Seq data of clear cell and papillary RCC were clustered according to a previously developed angio-immunogenic gene signature. WGS and RNA-Seq may improve therapeutic decision making for most patients with advanced RCC, including patients with non-clear cell RCC for whom no standard treatment is available to date. Prospective clinical trials are needed to evaluate the impact of genomic and transcriptomic diagnostics on survival outcome for advanced RCC patients.

Introduction

Renal cell carcinoma (RCC) consists of different histological subtypes [1,2]. The most common histological subtype is clear cell RCC (ccRCC), accounting for approximately 75% of the RCC cases [3]. The vast majority of ccRCC is characterized by the loss of the short arm of chromosome 3 (3p)[4], which harbors several tumor suppressor genes. The function of these genes - *VHL*, *BAP1*, *PBRM1*, and *SETD2* - is frequently inactivated due to additional somatic mutations or epigenetic changes of these genes on the other allele [4,5]. Although these genetic aberrations can be observed in most patients with ccRCC, the clinical behavior in individual patients differs significantly, from slowly progressive disease over years to rapidly progressive disease with fast clinical deterioration. Therefore, the management of advanced ccRCC varies from active surveillance to systemic therapy.

In the past years, the therapeutic landscape for patients with advanced ccRCC has changed significantly. The introduction of tyrosine kinase inhibitors (TKIs) [6,7], immune checkpoint inhibitors (ICIs) [8,9], mammalian target of rapamycin (mTOR) inhibitors [10], and combinations of these anti-cancer therapies [11-13], has significantly improved the outcome for patients with advanced ccRCC. However, there are considerable interindividual differences in outcome, and only a minority of patients experience durable responses [14]. For patients with advanced ccRCC, treatment decision making is nowadays guided by the International Metastatic RCC Database Consortium (IMDC) criteria [15-17]. These criteria include only clinical patient characteristics (i.e. hemoglobin level, time from diagnosis to start of systemic therapy, Karnofsky performance state, calcium level and neutrophil and platelets count).

Moreover, RCC with non-clear cell histology is a heterogeneous group of different subtypes, such as papillary and chromophobe RCC [18]. As holds true for advanced ccRCC, the course of non-clear cell RCC (nccRCC) differs significantly between patients [19,20]. Since the different nccRCC subtypes are considered rare diseases, randomized phase three clinical trials are lacking for nccRCC [21]. As a result, there is no clearly defined standard of care for patients with advanced nccRCC [20,22].

The development of RCC, including its metastatic potential and response to treatment, could mainly be explained by the different genomics [2,4] and evolutionary pathways [5,23] of this disease. Previous studies have focused on the molecular characterization of primary RCC [2,4,24] and the genomic evolution of ccRCC [4,5,23]. For example, RNA expression analysis in ccRCC has identified different immunogenic and angiogenic gene expression signatures [24,25], however, predictive value for treatment efficacy has not been validated. To improve the individualized treatment strategy and the survival outcomes for patients with ccRCC and nccRCC, more insight into the genomic make-up of advanced RCC is required.

The objective of this study was to describe the genomic landscape of advanced RCC, by combining whole-genome sequencing (WGS) with matched RNA sequencing (RNA-Seq) data. First, WGS was applied to characterize the genomic make-up of RCC and to identify potential actionable targets for systemic treatment in individual patients with ccRCC and nccRCC. Next, both WGS and matched RNA-Seq data were combined for patients with ccRCC and papillary RCC (pRCC). The RNA-Seq data were applied to cluster RCC based on immunogenic and angiogenic gene expression patterns, aiming to identify those patients who could benefit from either treatment with anti-angiogenic drugs, immunogenic drugs or a combination of these therapies.

Results

Patient selection

In total, WGS data from 91 patients with histopathologically confirmed RCC were included in the analyses (**Figure 1A**). Additional RNA-Seq data were available for 28 patients (**Figure 1A**). Overall, 72 patients were diagnosed with ccRCC, nine patients with pRCC, one with chromophobe RCC, one with tubulocystic RCC, and one with collecting duct carcinoma. For the remaining seven patients with RCC, the subtype could not be further defined. The main biopsy sites were the kidney (*N* = 24), bone (*N* = 15), and lymph nodes (*N* = 14) (**Figure 1B**). The median age of patients at the time of biopsy was 65 years (range 40-83), 79% of the patients were male, and 78% of the patients did not receive any systemic treatment before biopsy (**Table 1**). Most patients (84/91, 92%) were treated with systemic therapy after biopsy was collected. This treatment consisted mostly of TKIs (61/84, 73%) or ICIs (19/84, 23%), whereas the remaining patients received combination treatment (4/84, 5%). For 68 out of 84 (81%) patients, the first tumor response (RECIST v1.1 [26]) to treatment (post tumor biopsy) could be established. Most patients had stable disease (SD) (46%), followed by progressive disease (PD) (14%) and partial response (PR) (14%).

Whole-genome sequencing analyses

Somatic small variants, copy-number alterations (CNAs) and structural variants (SVs) were identified as described previously [27]. The median tumor mutational burden (TMB) of patients with ccRCC was 2.8 [interquartile range (IQR) 1.1] (**Figure 2A**). Only two patients (one with ccRCC and one with undefined subtype of RCC) had a TMB > 10 [28]. In both these samples with high TMB, mutational signatures were related to defective DNA mismatch repair, covering more than 25% of their single-nucleotide variants (SNVs) (**Figure 2A/E**). Interestingly, the lowest TMB (i.e. 0.16) was found in a patient with tubulocystic RCC (*N* = 1) who had very little genomic aberrations in general, and only two detectable SVs (**Figure 2C/D;** tandem duplication and break-end). In the total cohort, SBS40 (unknown etiology) was the dominant mutational signature, with a mean relative contribution of 74%. To assess the reliability of the SBS40 calling, the mutational signature calling was bootstrapped, which revealed a very high variance (median difference in assignment of 37%) in the relative contribution of SBS40 (Supplementary figure 1).

The frequency of genomic SNVs, multi-nucleotide variants (MNVs), InDels, and the collective coding mutations showed a similar pattern across the different RCC subtypes (Supplementary figure 2, supplementary data file). In total, 713,077 somatically acquired SNVs, 173,579 InDels and 9,964 MNVs were detected in the RCC genomes. Transversions were more frequently found than transitions in ccRCC, pRCC and undefined subtypes (Supplementary figure 2B and 2E). Missense variants were the most dominant protein variant type and accounted for > 60% of the small variants in all subtypes (Supplementary figure 2G). The genome ploidy was mostly diploid, and genome duplications were more frequent in ccRCC than in other subtypes (Supplementary figure 2D). For 57% of ccRCC, the genome-wide ploidy was 2, while 32% had a genome doubling

with a ploidy of 3 or higher. Differences in DNA ploidy have been associated with tumor differentiation and diploidy has been related to well-differentiated RCC [29]. Considering SVs, a total of 3,121 deletions, 941 translocations, 2,196 tandem duplications, 4 insertions, 2,714 inversions and 641 break-ends were detected (Supplementary Fig 2F). When comparing the mutational frequencies of ccRCC in our cohort to those in the TRACERx WGS cohort, a multicenter prospective study analyzing the evolutionary features of ccRCC [5], the number of substitutions was similar in both cohorts (7,224, Q1–Q3: [5648-8581] *vs*. 7,050, Q1–Q3: [6434-9504], *p*-value = 0.45). The number of patients with events considered as chromothripsis was limited and present in only five patients with ccRCC, two patients with pRCC and absent in the other subtypes. Furthermore, these chromothriptic events mostly did not involve the classic t(3;5) event. Only one of the samples with chromothripsis did have a t(3;5) translocation (Supplementary figure 3), while these specific translocations were more frequently detected in the cohort of patients without chromothriptic events (29.7%, Supplementary figure 4).

Next, the WGS data were analyzed on driver genes of the ccRCC samples using the dN/dS algorithm and on CNAs by GISTIC 2.0. These driver gene analyses revealed that most driver genes in ccRCC encompassed variations in the chromosome 3p region: *SETD2* (88.9%), *VHL* (73.6%), *PBRM1* (61.1%), and *BAP1* (16.7%) together with mutations in tumor suppressor genes on other chromosomes, such as *CDKN2A* (68.1%) and *PTEN* (16.7%) (**Figure 3,** supplementary data file). In addition, a substantial number of patients with ccRCC had focal deletions in genes described as possible tumor suppressor genes, e.g. *PTPRD* (59.7%) and *NEGR1* (36.1%) [4,30]. Deletions of *PTPRD* have been described as a possible risk factor for the development of ccRCC [30,31]. Moreover, amplifications were present in genes associated with cell proliferation and angiogenesis, such as

158 *CDK6* (55.6%) and *CCNE1* (19.4%) [32,33]. In total, pathogenic germline mutations related to cancer
159 or Von Hippel-Lindau syndrome were found in nine patients in different RCC subtypes and
160 included *ATM, FLCN, CHEK2, FH, SDHA* and *MITF* (**Figure 3**).

161 In addition, genes which were previously described as frequently mutated somatically (*q*-value <
162 0.05) in (cc)RCC by Braun *et al*. [34] and in pRCC by Turajlic *et al*. [35], that were not statistically
163 significant in our driver gene analysis, were included to extend our analysis. Nevertheless, many
164 of these added genes showed a low mutational frequency in our cohort, similarly for genes
165 associated with poor prognosis ccRCC [4], such as the Krebs cycle genes (e.g. *SDHA*, *FH*).

166 Furthermore, previously validated fusion events [36] were detected, with *CLTC-VMP1* and *SFPQ-*
167 *TFE3* both occurring once in the ccRCC group, along with a fusion event of *ASPSCR1-TFE3* in one
168 patient with pRCC. For both patients with a detected *TFE3* fusion, the histopathological diagnosis
169 had to be reconsidered. As a result, these patients were re-allocated in a different subcategory
170 of RCC, i.e. MiT family translocation renal cell carcinomas [1]. The previously described *TERT*
171 promoter hotspot variant (C228T) [5] was found in both ccRCC (*N* = 10) and pRCC (*N* = 2). Overall,
172 characteristic clear cell driver gene events — such as somatic *VHL* mutations/deletions — were
173 absent in nccRCC, except for *CDKN2A* deletions in pRCC.

174 In patients with ccRCC, both previously described and novel amplifications and deletions were
175 detected. Statistically significant CNA peaks and arm-level copy-number alterations in ccRCC are
176 presented in Supplementary figure 5. Previously described arm-level CNAs [37] included
177 amplifications of 1q, 5q, 7q, 8q, 12p and 20q, and deletions in 3p, 9p and 14q. Furthermore, in
178 the current cohort, amplifications of 5p, 7p, 12q, 16p and 20p, and deletions of 4p, 6p, 9q, 14p,

18p and 18q were also statistically significant. Next, we investigated if the tumors contained targetable variants (**Figure 4**), which showed that most patients with RCC had one or more variants indicated as potential targets or biomarkers for treatment. For example, for target-specific variants encoded by *CDK4/6* or *EGFR*, specific drugs have been developed for other cancer types [38-40]. This may result in off-label availability of these drugs for patients with similar aberrations in RCC, e.g. in context of a clinical trial [41]. Furthermore, somatic aberrations in cancer genes, e.g. *TP53*, are also considered biomarkers for targeted treatments [42]. Lastly, some variants leading to specific mechanisms were also considered as targetable for treatment. For instance, TKIs targeting VEGF signaling are known to be effective in patients with *VHL* mutations and are on-label available for patients with RCC [43]. Overall, for the majority (96%) of patients in this cohort actionable targets were detected, even for patients with nccRCC for whom no standard treatment is available to date.

Transcriptome analyses of advanced RCC

Differential Expression Analysis (DEA) was performed on RNA-Seq data to discriminate the two most frequently diagnosed histological subtypes, ccRCC (*N* = 24) and pRCC (*N* = 4). Next to the *t*-distributed stochastic neighbor embedding (*t*-SNE), which showed a clear separation between the two subtypes (Supplementary figure 6), the DEA resulted in 1,546 significantly (adjusted *p*-value < 0.05) differentially expressed genes. The hundred genes with the smallest adjusted *p*-value are shown in **Figure 5A.** In this top hundred list, several genes are known to be associated with the development or course of RCC. For instance, *LOX* [44] and *MAPKAPK3* [45] correlate with poor survival in RCC. In addition, various other genes were differentially expressed and have been described in other malignancies, such as *TUSC2* [46], *CAPN1* [47], *PCSK6* [48], and *CD2* [49]. The differential

expression of these genes confirmed a clear distinction of pRCC and ccRCC at the transcriptomic level.

Second, the pathway analysis (**Figure 5B**) of the significant differentially expressed genes revealed cancer hallmarks such as oxidative phosphorylation and epithelial-mesenchymal transition (EMT), among others [50]. Significant Reactome pathways [51] of differentially expressed genes in ccRCC compared to pRCC samples, mainly showed pathways related to VEGF and collagen formation. Pathways related to poor prognosis [4], such as the AMPK complex, the Krebs cycle genes, the pentose phosphate pathway and fatty acid synthesis, were not found to be differentially expressed between ccRCC and pRCC. A heatmap of the top differentially expressed genes between ccRCC and pRCC and a *t*-SNE plot show that RCC samples of undefined subtype cluster with either ccRCC or pRCC samples based on the differential gene expression of these subtypes (Supplementary figure 7).

The 66 gene-signature has been based on previous data from the IMmotion150 trial [25]. High expression of 'angiogenic' genes and certain 'invasion' genes is applied to sub-classify RCC as 'angiogenic', which would be predictive for response to TKIs targeting VEGF signaling. In case of high expression in 'Ca^{2+}-Flux', 'T-Effector', and other 'invasion' genes, RCC is sub-classified as 'immunogenic', indicating a response to ICIs is likely. According to this gene signature, ccRCC cases in the current cohort could be classified as either immunogenic or angiogenic (**Figure 5C**). Moreover, all patients with pRCC had low expression of these genes, except for one individual patient with expression in *EDNRB* [52].

Discussion

In this study, the genomic and transcriptomic landscape of advanced RCC was characterized for 91 individual patients. First, genomic data showed that, next to *VHL* mutations (73.6%), most common driver gene mutations in ccRCC included alterations in tumor suppressor genes of different pathways such as *SETD2* (88.9%) and *PTEN* (16.7%). While TMB was comparable amongst the different subtypes of RCC, the driver gene analyses showed a distinctive pattern between patients with ccRCC and nccRCC. Furthermore, WGS revealed potential actionable targets for 87 out of 91 patients and WGS might therefore contribute to a more individualized treatment strategy for patients with advanced RCC.

For a subgroup of patients (*N* = 28), transcriptomic data were also available. RNA-Seq could be applied to distinguish ccRCC, pRCC and histologically undefined RCC based on the differential gene expression. The application of the 66-gene signature [25] on the RNA-Seq data, made it possible to sub-categorize ccRCC into immunogenic or angiogenic signatures, whereas classification in pRCC using these signatures was not feasible.

At genomic level, the findings for ccRCC corresponded mostly with previous findings [4,5,53]. The massive contribution of SBS40 to the mutational landscape of nearly all RCC subtypes in this cohort is remarkable. However, bootstrapping showed that SBS40 was the least robust signature, indicating that this signature could act as a sink for mutations that are difficult to fit. Since CPCT-02 cohorts with other tumor types did not show the high contribution of SBS40 [54,55], this is most certainly not a result of a bias in the sequencing or in our workflows. In contrast to other studies, the number of chromothriptic events was limited in this study and occurred in only five patients

with ccRCC. In previous studies, chromothripsis was defined as the combination of a chromothriptic event together with a translocation event with concurrent 3p loss and 5q gain, which were called "t(3;5) chromothripsis events" [5]. Although both chromothriptic and translocation (t(3;5)) events occurred in our cohort, for most of the cases these events were independent of each other (Supplementary figure 3 and 4).

At transcriptomic level, assigning ccRCC biopsies into either immunogenic or angiogenic signatures may indicate which treatment could be most beneficial for individual patients. The introduction of ICIs has significantly changed the therapeutic landscape for patients with advanced ccRCC [56,57], resulting in a clinical need to select patients who will benefit either from angiogenic or immunogenic treatment [17]. These data could assist clinical decision making when choosing the optimal treatment strategy for the individual patient with advanced ccRCC. For patients with high expression of genes annotated as immunogenic, first–line treatment with ICIs should be considered, whereas for patients with high expression in angiogenic genes treatment with a TKI should be taken into consideration. For those patients with low expression throughout all these genes, combination treatment with TKI/ICI may be considered, although treatment based on actionable targets identified by WGS could be the most effective option. Treatment selection based on gene expression has already shown promising results for patients with RCC [12,58], however, further research in a prospective setting is still warranted [59].

The distinctive mutational gene pattern between ccRCC and nccRCC clearly showed that these tumors are different entities, while differences within the nccRCC subtypes were also evident. For example, none of the patients with nccRCC had somatic *VHL* mutations and also other ccRCC

driver genes hardly showed mutations in nccRCC. This is of great importance, as the development of targeted drugs is based on driver mutations or consequences downstream. For instance, the frequently mutated *VHL* gene in ccRCC has been the basis for the development of angiogenesis inhibitors for this disease. The mutations in this tumor-suppressor gene result in the accumulation of *HIF1α/2α*, eventually leading to overexpression of *VEGF/PDGF*, *AXL*, and *MET,* among others [43,60]. Several TKIs that have been approved for the treatment of advanced ccRCC [6,7,43,61] interfere at different levels in this cascade [43,60]. More recently, Hypoxia Inducible Factor (HIF) inhibition has also shown proven efficacy in patients with *VHL* mutations [62]. However, as patients with nccRCC in our cohort showed no mutations in *VHL* or other related pathways, it is questionable whether treatment directed against the *VHL* pathway would be the most effective therapy for this particular group of patients.

In clinical practice, RCC is usually defined histopathologically. As a result, there is a large dependency on experienced pathologists. However, discrepancies among these experts remain [63,64]. In our cohort, nearly 8% of RCC cases could not be sub-classified through histopathological assessment. Thereby, for two patients a fusion gene was detected by WGS which led to revision of the original histological diagnosis and allocation to a different subgroup, i.e. MiT family translocation renal cell carcinoma [1]. As the histopathological classification defines the treatment strategy, this could have significant clinical impact. Since different subtypes and growth patterns of RCC are driven by gene expression [65], a next generation sequencing-based classifier could be feasible. Here, we showed that analyses of driver mutations (*VHL*, *PBRM1*, *SETD2)* and RNA-Seq data reveal clear differences among the different RCC subtypes. As shown in Supplementary

Figures 6 and 7, clustering of the undefined RCC subtypes is feasible and could be useful to clarify the histological subtype in clinical practice.

This study has some important limitations. First, the collected clinical data within the CPCT-02 study were limited. For instance, only the first tumor response (according to RECIST v1.1.) after biopsy was available and the response rates in this cohort were relatively low when compared to the clinical trials. On one hand, these lower tumor response rates could be due to the unavailability of ICIs, as patients were included between 2016 and 2019, when ICIs were not yet available in first-line setting for RCC in the Netherlands. In addition, patients within the CPCT-02 study were treated in a real-world setting, which is known for its lower response rates compared to response rates of clinical trials. In addition, due to the limited clinical data collection, it was not possible to reliably correlate genomic and transcriptomic findings to clinical data. A correlation with clinical data could have confirmed whether patients with certain gene signatures indeed had benefit from a specific treatment. Therefore, validation in a prospective trial is needed prior to clinical implementation.

Second, the limited number of patients with nccRCC made it challenging to run separate analyses for this group. Since the subgroup of patients with nccRCC consists of different less common and heterogeneous subtypes, and very little is known about the genomics of nccRCC. Therefore, we decided to include all patients with nccRCC, even subgroups containing only a single patient. Finally, the collected data were heterogeneous. For instance, not only biopsies of metastases, but also biopsies from the kidney (including primary tumors) were included for the analyses. However, it is not conceivable that this has significantly impacted the analysis, since

previous studies have shown clear consistencies between primary tumors and their metastasis [5]. In addition, the heterogeneity in this cohort reflects daily clinical practice of patients who present with advanced RCC, including primary metastatic disease. Despite this clinical heterogeneity, a clear genomic and transcriptomic signal could be extracted, indicating that the genomic and transcriptomic analyses are feasible for clinical implementation.

In conclusion, there are evident genomic and transcriptomic differences between RCC subtypes. The analysis of driver mutations, in combination with clustering of RNA-Seq data, could assist the histopathological subtyping of RCCs in clinical practice. In addition, RNA-Seq data could identify patients with ccRCC who may benefit more from treatment with either ICIs, TKIs or a combination of these drugs. Genomic and transcriptomic analyses are promising to identify actionable targets and to individualize treatment strategies in the majority of patients with RCC, even for patients with nccRCC. Although these results are promising, prospective clinical trials are still needed to evaluate whether genomic and transcriptomic diagnostics indeed contribute to improved survival outcomes in individual patients with advanced RCC.

Materials and Methods

Center for Personalized Medicine: Patient cohort, study procedures, sample collection, clinical data

In accordance with the Declaration of Helsinki, all patients within this study provided written informed consent for participation in the Center for Personalized Cancer Treatment (CPCT-02) study (NCT01855477) before study procedures started. The CPCT-02 study was approved by the medical ethical committee of the University Medical Center Utrecht, as well as the Netherlands Cancer Institute and local approval was provided for each participating site. Details regarding inclusion criteria, the study protocol, sampling, and sequencing have been previously described [27]. In summary, core needle biopsies from the tumor lesion, peripheral whole blood samples and clinical data were collected across hospitals in the Netherlands. The response to treatment was determined according to RECIST v1.1 [26]. WGS data from 103 biopsies of 101 patients with advanced RCC were made available. Only one sample per patient was selected for the genomic analyses. When multiple biopsies of one patient were available, the sample covering the most clinical information and/or the sample with the highest estimated tumor cell percentage was selected. The selection resulted in 91 WGS samples (51 previously described by Priestley *et al.* [27]) and 35 RNA samples.

Pathological diagnosis

To confirm the histopathological diagnosis of RCC, pathology reports were requested via PALGA, the nationwide network and registry of histo- and cytopathology in the Netherlands [66]. Slides and tissue blocks were not available for pathological revision. Alternatively, the pathology reports were reviewed by a genito-urinary pathologist (GvL) to determine whether the microscopic

description and immunohistochemistry were compatible with the original diagnosis. The following subtypes were annotated: clear cell, papillary, chromophobe, tubulocystic, and collecting duct carcinoma. Histopathologically confirmed RCC of which the subtype remained unclear was categorized as undefined subtype.

Whole genome sequencing and preprocessing

Between the 8th of August 2016 and the 3rd of October 2019, tumor and whole-blood pairs were whole-genome sequenced at the Hartwig Medical Foundation (HMF) central sequencing center. A HiSeqX system was applied and 2 × 150 base read pairs were generated using standard settings (Illumina, San Diego, CA, USA). Preprocessing was performed as described by Priestley *et al* [27]. Briefly, read pair mapping was performed using BWA-mem [67] to the reference genome GRCh37 (human) with subsequent systematic variant calling and several quality control and/or correction steps. The Genome Rearrangement IDentification Software Suite (GRIDSS) [68] was used for structural variant (SV) calling and LINX (v1.11) [68] for gene fusion event calling. Computational ploidy estimation and copy-number (CN) assessment was performed using the PURPLE (PURity & PLoidy Estimator) pipeline [68], estimating tumor purity and CN profile by combining B-allele frequency (BAF), read depth, and SVs.

Somatic variant annotation and filtering

Somatic variants were determined using Strelka and provided by the HMF as part of the data request. Variant Call Format (VCF) files with somatic variants were annotated based on GRCh37 with HUGO gene symbols, HGVS notations, gnomAD [69] frequencies using VEP [70] (database release 95, merged cache), with setting "--per_gene". Exclusively somatic single-nucleotide variants

(SNVs), small InDels, multi-nucleotide variants (MNVs) with ≥ 3 alternative read observations and passing variant caller quality control were included in the analyses. Furthermore, population variants were removed to prevent germline leakage, based on the gnomAD database (v2.0.2) [69]: gnomAD exome (ALL) allele frequency ≥ 0.001; and gnomAD genome (ALL) ≥ 0.005. Variants specific for the Dutch CPCT cohort were removed based on a panel-of-normals from 1,762 representative normal blood HMF samples. The most deleterious mutation was used to annotate the overlapping gene for each sample.

Tumor Mutational Burden calculation

The number of mutations per megabase pair was calculated as the amount of somatic genome-wide SNVs, MNVs, and InDels divided by the number of callable nucleotides in the human reference genome (GRCh37) FASTA file:

$$TMB = \frac{\left(SNVs_g + MNVs_g + InDels_g\right)}{\left(\frac{2858674662}{10^6}\right)}$$

Ploidy and copy-number analysis

Broad and focal somatic CN alterations in ccRCC were identified by GISTIC2.0 [71] (v2.0.23), using the following parameters: genegistic 1, gcm extreme, maxseg 4000, broad 1, brlen 0.98, conf 0.95, rx 0, cap 3, saveseg 0, armpeel 1, smallmem 0, res 0.01, ta 0.1, td 0.1, savedata 0, savegene 1, qvt 0.1. Distinction between shallow and deep CN events per region was based on thresholding performed by GISTIC2.0. The alterations were assigned a score taking both the amplitude and the frequency of its occurrence across samples into account (G-score). Thresholding was divided into five CN categories; two for deletions (-2 = deep, possibly homozygous loss, -1 = shallow,

possibly heterozygous loss), one for diploid (0 = diploid) and two for amplifications (1 = few additional copies, often broad gain, 2 = more copies, often focal gain). Annotation of GISTIC2.0 peaks was performed as follows: A) Wide peaks were annotated with all overlapping canonical UCSC genes within the genomic limits of said peak. B) Focal peaks were annotated based on overlapping genomic coordinates, using custom R scripts and UCSC gene annotations.

Structural variant analysis

SVs affecting genes were imported using custom R scripts, overlapping genes on at least one breakpoint, using GRCh37 genomic coordinates. SVs with an upstream or downstream Tumor Allele frequency (TAF) below 0.1 as determined by PURPLE and GRIDSS [68] were discarded along with SVs that affected all exons of a gene. In the case of both (multiple) mutations and/or SVs in the same gene, these were annotated as 'multiple mutations'.

Fusion gene analysis

WGS-based LINX TSV files were imported using R and overlapped with the three pillars of ChimerDB [36]; deep sequencing data (ChimerSeq), text mining of PubMed publications (ChimerPub), with extensive manual annotations (ChimerKB). Events that were not present in any pillar of ChimerDB and intra-gene fusions were filtered out. RNA-Seq based fusion genes detected with Isofox (https://github.com/hartwigmedical/hmftools/tree/master/isofox) were imported using R and overlapped with the fusion events detected in the DNA sequencing.

Somatic Driver Genes Analysis

We utilized the dN/dS model (192 Poisson rate parameters; under the full trinucleotide model) to identify genes undergoing mutational selection in the ccRCC patients with the R package

dndscv [72] (v0.0.1.0). Both the substitution model and InDel model were used and were corrected for sequence composition, gene length and mutational signatures. These models test the ratio between nonsynonymous (missense, nonsense and essential splice site) and background (synonymous) mutations. To identify genes that drive selection, a *q*-value < 0.05 (both including and excluding the InDel model) was used.

Mutational signatures analysis

Mutational signatures analysis was performed using the MutationalPatterns R package (v3.2.0)[73]. The mutational signatures based on single base substitutions (*N* = 90 v3 signatures) were downloaded from COSMIC [74]. SNVs were categorized according to their respective trinucleotide context (GRCh37) into a mutational spectrum matrix M_{ij} (where *i* represents 1:96 trinucleotide contexts and *j* represents the number of 1:91 samples) and subsequently, a constrained linear combination of the ninety mutational signatures was constructed per sample using non-negative least squares regression implemented in the R package pracma (v2.2.9). Mutational signatures were bootstrapped (*N* = 100) with MutationalPatterns and argument 'method' set to "strict" to assess calling stability. Signature contribution for each sample was determined per 100 samplings, per signature.

Chromothripsis

Chromothripsis (CT), also known as chromosomal shattering, followed by seemingly random re-ligation, was detected using Shatterseek [75] (v0.4) with default settings. The following definition of CT was employed: (1) ≥25 intrachromosomal SVs involved in the event; (2) ≥7 oscillating CN segments (2 CN states) or ≥14 oscillating CN segments (3 CN states); (3) CT event involving ≥20

Mb; (4) satisfying the test of equal distribution of SV types (p-value > 0.05); and (5) satisfying the test of nonrandom SV distribution within the cluster region or chromosome (p-value ≤ 0.05).

Actionable targets

iClusion (https://iclusion.com) data, which connects specific or gene-level aberrations to clinical cancer studies, were provided by HMF. This integrates clinical interpretations from Precision Oncology Knowledge Base (OncoKB) [76], Clinical Interpretation of Variants in Cancer (CIViC) [77] and Cancer Genome Interpreter (CGI) [78]. All targets and biomarkers were overlapped with filtered molecular data to verify presence. Targets marked as "gene-level" were generalized for other variation in those genes, not listed in the iClusion data. The identified targets were assessed and manually categorized in the following three categories: on-label drugs for RCC, off-label available, investigational drugs. Drugs were considered on-label when approval was given for any subtype of RCC in the Netherlands. Whether drugs were on- or off-label available in the Netherlands is defined by the Dutch Medicines Evaluation Board ("College ter Beoordeling van Geneesmiddelen") [79]. This evaluation board takes previous approvals by the U.S. Food and Drug Administration (FDA) and/or European Medicines Agency (EMA) in consideration.

Germline analysis

Known pathogenic germline variants (GRCh37) related to cancer and/or Von Hippel-Lindau syndrome were retrieved from ClinVar [80] that were less than 51 bp long, with a review status of "practice guideline", "expert panel", "multiple submitters" or "at least one star". These ClinVar variants were used as filter for the import of germline variants from VCF files of our cohort. Variants with at least 2 reads and passing variant caller quality control were included.

Furthermore, variants that were annotated with "high" impact, in genes with known germline variation in RCC (*FH*, *SDHA*, *SDHB*, *SDHC*, *SDHD*, *TCEB1*, *FLCN*, *CHEK2*) were included.

RNA sequencing

RNA was isolated from biopsy using the QIAsymphony RNA Kit (Qiagen, Hilden, Germany) for tissue and quantified on the Qubit. Between 50 and 100 ng of RNA was used as input for the KAPA RNA HyperPrep Kit with RiboErase (Human/Mouse/Rat) library preparation (Roche) on an automated liquid handling platform (Beckman Coulter). RNA was fragmented (high temperature in the presence of magnesium) to a target length of 300 bp. Barcoded libraries were sequenced as pools on either a NextSeq 500 (V2.5 reagents) generating 2 x 75 base read pairs or on a NovaSeq 6000 generating 2 x 150 base read pairs using standard settings (Illumina, San Diego, CA, USA). BCL output from the sequencing platform was converted to FASTQ using Illumina's bcl2fastq tool (versions 2.17 to 2.20) using default parameters. RNA-Seq data was aligned using STAR [81] to GRCh37 resulting in unsorted BAMs including chimeric reads as output. Gene and transcript counts were generated and used for subsequent fusion detection using Isofox (https://github.com/hartwigmedical/hmftools/tree/master/isofox).

RNA sequencing analyses

Raw read counts were imported in R and filtered on protein coding genes based on Ensembl GTF file [82] (Homo sapiens GRCh37, version 87). *t*-distributed stochastic neighbor embedding (*t*-SNE) was performed on variance stabilized read counts (generated by DESeq2 varianceStabilizingTransformation) of all protein coding genes. Differential expression analysis between ccRCC and pRCC was performed on raw read counts using DESeq2 [83] and the Wald-test.

Statistical significant results with Benjamini-Hochberg adjusted *p*-value < 0.05 were further filtered to base mean > 100 counts and absolute $\log_2$ fold change ≥ 1. The heatmap with the top most significantly differentially expressed genes (based on lowest adjusted *p*-value) was made using variance stabilized read counts and euclidean distances on scaled data. Gene signature heatmap was produced with centered Z-Scores with Euclidean distances. Gene set enrichment analyses were performed using fgsea [84] (Monte Carlo approach with Adaptive Multilevel Splitting) with MSigDB [85] Hallmarks and Reactome pathways [51] as gene sets. Reproduction of the D'Costa *et al.* gene signature [25] was done using 65 of 66 original genes, since *PECAM1* was on a genome patch not included in the RNA-Seq mapping supplied by HMF. Heatmaps were produced using pheatmap with Ward.D clustering.

Data and material availability

Data was provided by HMF, which were used under data request number DR-088 for the current study. Both WGS, RNA-Seq and clinical data are freely available for academic use from the HMF through standardized procedures and request forms can be found at https://www.hartwigmedicalfoundation.nl. All tools and scripts used for processing of the WGS data are available at https://github.com/hartwigmedical/ and/or can be provided by authors upon request.

Tables

Table 1. Overview of patients' characteristics

	Frequency	Percentage
Sex		
Male	**72**	**79%**
Female	**19**	**21%**
Median age 65 years [range 40-83]	---------	---------
Histological subtype		
Clear cell RCC (ccRCC)	**72**	**79%**
Non-clear cell RCC (nccRCC)	**12**	**13%**
Papillary RCC (pRCC)	9	75%
Chromophobe RCC (chRCC)	1	8%
Collecting duct carcinoma (CDC)	1	8%
Tubulocystic RCC (tRCC)	1	8%
Undefined subtype	**7**	**8%**
Prior treatment (n=number of lines)		
No	**71**	**78%**
Yes (1)	**10**	**11%**
Yes (≥2)	**10**	**11%**
Treatment after biopsy (N = 84)		
Tyrosine kinase inhibitors (TKIs)	**61**	**73%**
Pazopanib	36	59%
Sunitinib	23	38%
Cabozantinib	1	1.6%
Lenvatinib	1	1.6%
Immune checkpoint inhibitors (ICIs)	**19**	**23%**
Nivolumab monotherapy	13	68%
Nivolumab + ipilimumab	6	32%
Combination treatment	**4**	**5%**
Avelumab + axitinib	4	100%

Figure legends

Figure 1 Overview of sample selection and biopsy sites

1A illustrates the selection of samples for both the WGS (*N* = 91) and matched RNA-Seq (*N* = 28) analyses. In **1B**, the main biopsy sites are shown. Next to the illustrated biopsy sites, other biopsy sites (*N* = 20) include for instance biopsies from subcutaneous tissue. **1B** was created with biorender.

Figure 2 Overview of genomic characteristics of whole-genome sequenced advanced RCC cohort (*N* = 91)

Track **A** shows the tumor mutational burden (mutations per Mb; yellow for low (0-5), orange for medium (5-10) and red for high (> 10)). Track **B** shows the mean genome-wide ploidy, with white representing diploidy. Tracks **C** and **D** illustrate the abundance of structural variants and the relative frequency of the types of these variants. Tracks **E** and **F** show the relative mutational signature contribution (COSMIC signatures v3) and the relative frequency of mutational changes at base level. Track **G** shows the presence of chromothripsis. Track **H** shows whether patients were treatment naive at time of biopsy. Tracks **I** and **J** indicate the first treatment given after biopsy (if any) and the first tumor response according to RECIST v1.1, respectively. On the x-axis the figure is arranged in descending order by tumor mutational burden per RCC subtype. ccRCC = clear cell renal cell carcinoma. pRCC = papillary renal cell carcinoma. Undefined subtype = renal cell carcinoma, with undefined subtype. chRCC = chromophobe renal cell carcinoma. CDC = collecting duct carcinoma. tRCC = tubulocystic renal cell carcinoma. NA = not available.

Figure 3 Overview of coding mutations and copy-number alterations in driver genes in whole-genome sequenced advanced renal cell carcinoma cohort (*N* = 91)

The oncoplot in track **A** shows mutations (filled center) and copy-number alterations (grid cell background) of driver genes determined by dN/dS and GISTIC2.0. Track **B** also shows an oncoplot, but on selected genes, not passing any statistical threshold. Germline pathogenic mutations are indicated with a capital letter 'G' (and red border), utilizing the same color coding as the somatic mutations. Consequential fusion genes are indicated in yellow, with a red border. Track **C** shows the tumor mutational burden (mutations per Mb; yellow for low (0-5), orange for medium (5-10)

524 and red for high (> 10)). Tracks **D**, **E** and **F** show whether patients were treatment naive at time of
525 biopsy, if systemic treatment was given after time of biopsy, and the first tumor response after
526 systemic treatment according to RECIST v1.1, respectively. Bold sample names with asterisks
527 indicate MiT family translocation RCC. Figure is arranged in descending order by tumor mutational
528 burden per RCC subtype on the x-axis.

529 **Figure 4 Overview of DNA-based biomarkers and potential treatment options in the whole-**
530 **genome sequenced advanced renal cell carcinoma cohort (*N* = 91)**
531 Track **A** Percentage of potential available treatment options based on genomic characteristics.
532 Treatment options are categorized according to the highest level of drug availability in clinical
533 practice (on label available – off label available – investigational drugs). Track **B** Potentially
534 actionable alterations at gene-level with each column representing a sample, ordered
535 descendingly by tumor mutational burden per subtype on the x-axis. Detailed description of
536 actionable targets can be found in Supplementary Data file 1.

537 **Figure 5 RNA sequencing cohort and differential expression analysis between clear cell renal**
538 **cell carcinoma (ccRCC) and papillary RCC (pRCC) with classification according to gene**
539 **signatures[25]**
540 Track **A** shows a heatmap of Z-scores of variance stabilized values with unsupervised clustering of
541 the top 100 transcripts based on smallest adjusted *p*-value and colored according to Z-scores.
542 Tracks **B** shows gene set enrichments based on sets (y-axis) from the molecular signatures
543 database hallmarks and Reactome pathways, with the normalized enrichment score (NES) on the
544 x-axis. Bar charts are visualized with ccRCC taken as reference (*N* = 24) (positive NES equals
545 expression up in ccRCC and down in pRCC (*N* = 4)). Track **C** shows unsupervised clustering on the
546 rows (patients) with color coding indicating the RCC subtype (purple for ccRCC and pink for pRCC)
547 and colored according to Z-scores. The x-axis has been cut into several gene groups related to
548 angiogenesis, invasion, Ca^{2+}-flux and T-effector cells, as defined by D'Costa *et al*. [25] and by their
549 stated order.

550
551

Acknowledgements

This publication and the underlying study have been made possible partly on the basis of the data that Hartwig Medical Foundation and the Center of Personalized Cancer Treatment (CPCT) have made available to the study. Furthermore, we would like to thank the nationwide network and registry of histo- and cytopathology in the Netherlands (PALGA) for their assistance on retrieving and storing the pathological records of the included patients with RCC. Figure 1 was created with BioRender (https://biorender.com/).

Author contributions

K.J., W.S.G., H.J.G.W. and A.A.M.V. wrote the manuscript which all authors have reviewed critically. W.S.G. and H.J.G.W. performed the bioinformatical analyses. K.J. and A.A.M.V. assessed the clinical data. G.J.L.H.L. reviewed and assessed the histopathological data. P.H., H.M.W., A.B., S.F.O, J.M.R., L.V.B., M.L. and R.H.J.M. are clinical contributors. M.P.L. and S.S. participated in the board of the CPCT-02 study. E.C. coordinated the sequencing of samples and contributed to the bioinformatical analyses.

Conflicts of interest

P.H. declares consultancy roles for Astellas, MSD, Ipsen, Pfizer, AstraZeneca, and Bristol-Myers Squibb, all outside the submitted work; H.M.W. declares honoraria from Roche and Astellas and travel expenses from Ipsen and Astellas, all outside the submitted work; S.F.O declares research grants from Novartis, Pfizer and Celldex Therapeutics and advisory board for Bristol Myers Squibb (all paid to the institution); M.L. declares speakers fee of BMS and advisory board of MSD, both paid to institution; R.H.J.M. declares speakers fee from Novartis and advisory role for Servier, patency from Pamgene, and investigator-initiated research (paid to institution) from Astellas, Bayer, Boehringer-Ingelheim, Cristal Therapeutics, Pamgene, Pfizer, Novartis, Roche, Sanofi, Servier, all outside the submitted work; M.P.L. declares advisory board for Amgen, Astellas, Astra Zeneca, Bayer, INCa, Janssen Cilag BV, MSD, Novartis, Pfizer, Roche, Sanofi, Servier, consulting role for Julius Clinical and Research Grants (paid to institution) from Astellas, Janssen, MSD, Sanofi, all outside the submitted work; H.J.G.W. declares speakers honoraria from Bayer,

580 Depositary receipts for shares from Cergentis B.V. all outside the submitted work; A.A.M.V.
581 reports advisory board (all paid to institution) of BMS, MSD, Merck, Pfizer, Ipsen, Eisai, Pierre
582 Fabre, Roche, Novartis, Sanofi, all outside the submitted work. All other authors declare no
583 competing interests.

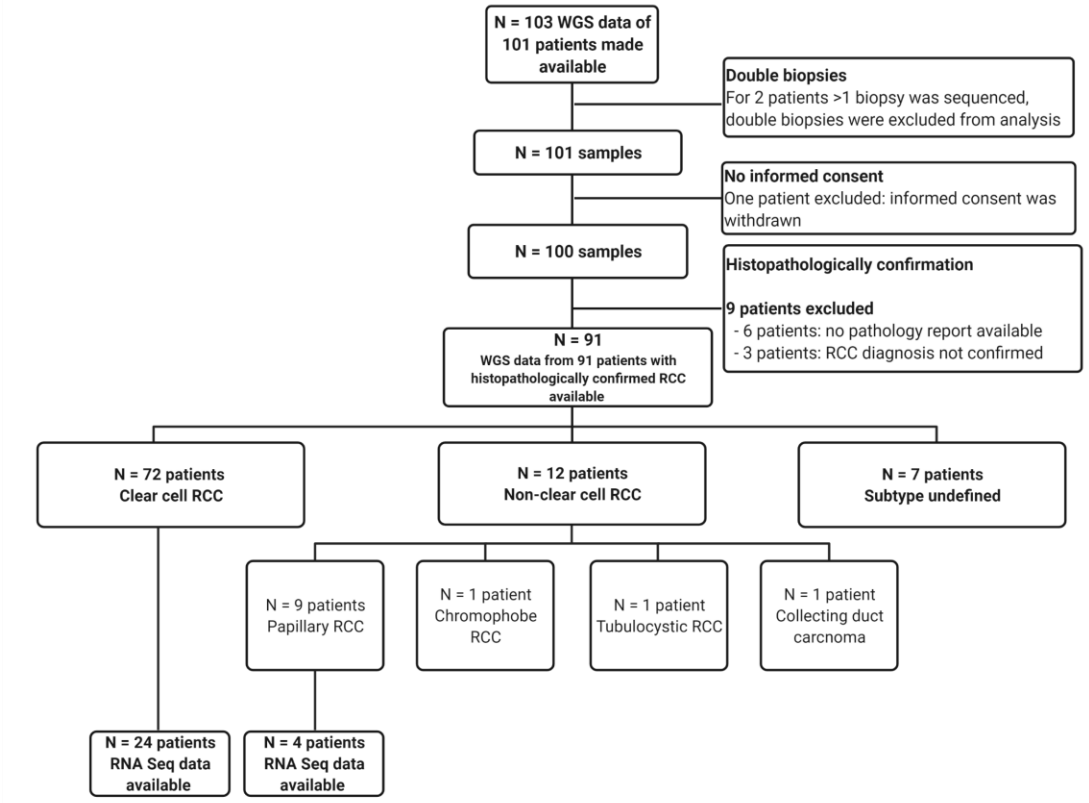
References

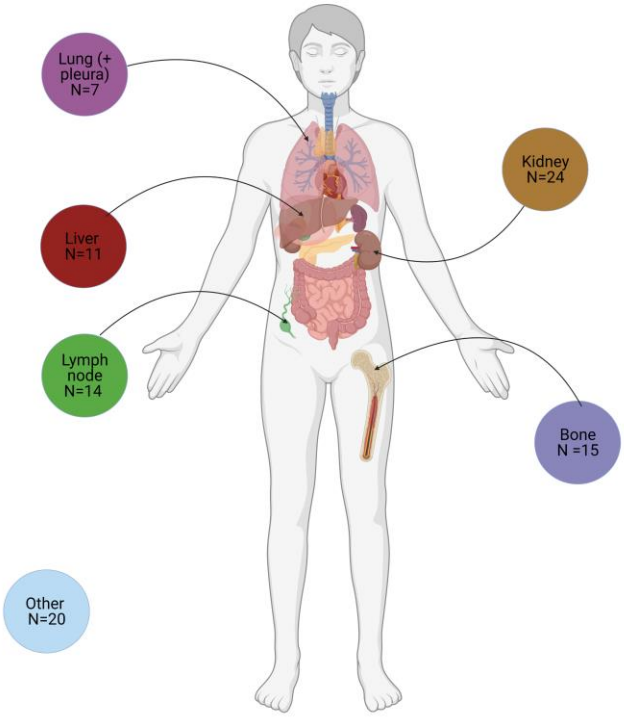
1. Moch, H., Cubilla, A.L., Humphrey, P.A., Reuter, V.E. & Ulbright, T.M. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. *Eur Urol* **70**, 93-105 (2016).
2. Cancer Genome Atlas Research, N. *et al.* Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma. *N Engl J Med* **374**, 135-45 (2016).
3. Padala, S.A. *et al.* Epidemiology of Renal Cell Carcinoma. *World J Oncol* **11**, 79-87 (2020).
4. Cancer Genome Atlas Research, N. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **499**, 43-9 (2013).
5. Mitchell, T.J. *et al.* Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. *Cell* **173**, 611-623 e17 (2018).
6. Motzer, R.J. *et al.* Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med* **356**, 115-24 (2007).
7. Mendez-Vidal, M.J. *et al.* Pazopanib: Evidence review and clinical practice in the management of advanced renal cell carcinoma. *BMC Pharmacol Toxicol* **19**, 77 (2018).
8. Motzer, R.J. *et al.* Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *N Engl J Med* **373**, 1803-13 (2015).
9. Motzer, R.J. *et al.* Nivolumab plus ipilimumab versus sunitinib in first-line treatment for advanced renal cell carcinoma: extended follow-up of efficacy and safety results from a randomised, controlled, phase 3 trial. *Lancet Oncol* (2019).
10. Motzer, R.J. *et al.* Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. *Lancet* **372**, 449-56 (2008).
11. Powles, T. *et al.* Pembrolizumab plus axitinib versus sunitinib monotherapy as first-line treatment of advanced renal cell carcinoma (KEYNOTE-426): extended follow-up from a randomised, open-label, phase 3 trial. *Lancet Oncol* **21**, 1563-1573 (2020).
12. Choueiri, T.K. *et al.* Updated efficacy results from the JAVELIN Renal 101 trial: first-line avelumab plus axitinib versus sunitinib in patients with advanced renal cell carcinoma. *Ann Oncol* **31**, 1030-1039 (2020).
13. Rini, B.I. *et al.* Atezolizumab plus bevacizumab versus sunitinib in patients with previously untreated metastatic renal cell carcinoma (IMmotion151): a multicentre, open-label, phase 3, randomised controlled trial. *Lancet* **393**, 2404-2415 (2019).
14. Flippot, R., Escudier, B. & Albiges, L. Immune Checkpoint Inhibitors: Toward New Paradigms in Renal Cell Carcinoma. *Drugs* **78**, 1443-1457 (2018).
15. Heng, D.Y. *et al.* External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: a population-based study. *Lancet Oncol* **14**, 141-8 (2013).
16. Guida, A. *et al.* Identification of international metastatic renal cell carcinoma database consortium (IMDC) intermediate-risk subgroups in patients with metastatic clear-cell renal cell carcinoma. *Oncotarget* **11**, 4582-4592 (2020).
17. Vano, Y.A. *et al.* First-Line Treatment of Metastatic Clear Cell Renal Cell Carcinoma: What Are the Most Appropriate Combination Therapies? *Cancers (Basel)* **13**(2021).
18. Truong, L.D. & Shen, S.S. Immunohistochemical diagnosis of renal neoplasms. *Arch Pathol Lab Med* **135**, 92-109 (2011).
19. Hong, B. *et al.* The Clinicopathological Features and Prognosis in Patients With Papillary Renal Cell Carcinoma: A Multicenter Retrospective Study in Chinese Population. *Front Oncol* **11**, 753690 (2021).

20. Sepe, P. *et al.* Characteristics and Treatment Challenges of Non-Clear Cell Renal Cell Carcinoma. *Cancers (Basel)* **13**(2021).
21. Powles, T. *et al.* ESMO Clinical Practice Guideline update on the use of immunotherapy in early stage and advanced renal cell carcinoma. *Annals of Oncology* **32**, 1511-1519 (2021).
22. Osterman, C.K. & Rose, T.L. A Systematic Review of Systemic Treatment Options for Advanced Non-Clear Cell Renal Cell Carcinoma. *Kidney Cancer* **4**, 15-27 (2020).
23. Turajlic, S. *et al.* Tracking Cancer Evolution Reveals Constrained Routes to Metastases: TRACERx Renal. *Cell* **173**, 581-594 e12 (2018).
24. Ricketts, C.J. *et al.* The Cancer Genome Atlas Comprehensive Molecular Characterization of Renal Cell Carcinoma. *Cell Rep* **23**, 313-326 e5 (2018).
25. D'Costa, N.M. *et al.* Identification of gene signature for treatment response to guide precision oncology in clear-cell renal cell carcinoma. *Sci Rep* **10**, 2026 (2020).
26. Eisenhauer, E.A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* **45**, 228-47 (2009).
27. Priestley, P. *et al.* Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* **575**, 210-216 (2019).
28. Alexandrov, L.B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415-21 (2013).
29. Abou-Rebyeh, H., Borgmann, V., Nagel, R. & Al-Abadi, H. DNA ploidy is a valuable predictor for prognosis of patients with resected renal cell carcinoma. *Cancer* **92**, 2280-5 (2001).
30. Ortiz, B., White, J.R., Wu, W.H. & Chan, T.A. Deletion of Ptprd and Cdkn2a cooperate to accelerate tumorigenesis. *Oncotarget* **5**, 6976-82 (2014).
31. Du, Y. *et al.* Polymorphism in protein tyrosine phosphatase receptor delta is associated with the risk of clear cell renal cell carcinoma. *Gene* **512**, 64-9 (2013).
32. Tadesse, S., Yu, M., Kumarasiri, M., Le, B.T. & Wang, S. Targeting CDK6 in cancer: State of the art and new insights. *Cell Cycle* **14**, 3220-30 (2015).
33. Li, Y., Shen, Y., Zhu, Z., Wen, H. & Feng, C. Comprehensive analysis of copy number variance and sensitivity to common targeted therapy in clear cell renal cell carcinoma: In silico analysis with in vitro validation. *Cancer Med* **9**, 6020-6029 (2020).
34. Braun, D.A. *et al.* Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nat Med* **26**, 909-918 (2020).
35. Turajlic, S., Larkin, J. & Swanton, C. SnapShot: Renal Cell Carcinoma. *Cell* **163**, 1556-1556 e1 (2015).
36. Jang, Y.E. *et al.* ChimerDB 4.0: an updated and expanded database of fusion genes. *Nucleic Acids Research* **48**, D817-D824 (2019).
37. Beroukhim, R. *et al.* Patterns of gene expression and copy-number alterations in von-hippel lindau disease-associated and sporadic clear cell carcinoma of the kidney. *Cancer Res* **69**, 4674-81 (2009).
38. Du, Z. *et al.* Structure-function analysis of oncogenic EGFR Kinase Domain Duplication reveals insights into activation and a potential approach for therapeutic targeting. *Nat Commun* **12**, 1382 (2021).
39. Ho, G.F. *et al.* Real-world experience of first-line afatinib in patients with EGFR-mutant advanced NSCLC: a multicenter observational study. *BMC Cancer* **19**, 896 (2019).
40. Braal, C.L. *et al.* Inhibiting CDK4/6 in Breast Cancer with Palbociclib, Ribociclib, and Abemaciclib: Similarities and Differences. *Drugs* **81**, 317-331 (2021).
41. van der Velden, D.L. *et al.* The Drug Rediscovery protocol facilitates the expanded use of existing anticancer drugs. *Nature* **574**, 127-131 (2019).

42. Zhu, G. *et al.* Mutant p53 in Cancer Progression and Targeted Therapies. *Front Oncol* **10**, 595187 (2020).
43. Alonso-Gordoa, T. *et al.* Targeting Tyrosine kinases in Renal Cell Carcinoma: "New Bullets against Old Guys". *Int J Mol Sci* **20**(2019).
44. Lin, S. *et al.* Comprehensive analysis on the expression levels and prognostic values of LOX family genes in kidney renal clear cell carcinoma. *Cancer Med* **9**, 8624-8638 (2020).
45. Oh, E. *et al.* Genome-Wide Transcriptomic Analysis of Non-Tumorigenic Tissues Reveals Aging-Related Prognostic Markers and Drug Targets in Renal Cell Carcinoma. *Cancers (Basel)* **13**(2021).
46. Rimkus, T., Sirkisoon, S., Harrison, A. & Lo, H.W. Tumor suppressor candidate 2 (TUSC2, FUS-1) and human cancers. *Discov Med* **23**, 325-330 (2017).
47. Pan, X.Q., Huang, W., Jin, L.W., Lin, H.Z. & Xu, X.Y. A Novel Pyroptosis-Related Prognostic Signature for Risk Stratification and Clinical Prognosis in Clear Cell Renal Cell Carcinoma. *Dis Markers* **2022**, 8093837 (2022).
48. Wang, P., Wang, F., Wang, L. & Pan, J. Proprotein convertase subtilisin/kexin type 6 activates the extracellular signal-regulated kinase 1/2 and Wnt family member 3A pathways and promotes in vitro proliferation, migration and invasion of breast cancer MDA-MB-231 cells. *Oncol Lett* **16**, 145-150 (2018).
49. Han, J. *et al.* MMP11 and CD2 as novel prognostic factors in hormone receptor-negative, HER2-positive breast cancer. *Breast Cancer Res Treat* **164**, 41-56 (2017).
50. Ding, X.F., Zhou, J., Hu, Q.Y., Liu, S.C. & Chen, G. The tumor suppressor pVHL down-regulates never-in-mitosis A-related kinase 8 via hypoxia-inducible factors to maintain cilia in human renal cancer cells. *J Biol Chem* **290**, 1389-94 (2015).
51. Jassal, B. *et al.* The reactome pathway knowledgebase. *Nucleic Acids Res* **48**, D498-D503 (2020).
52. Wuttig, D. *et al.* CD31, EDNRB and TSPAN7 are promising prognostic markers in clear-cell renal cell carcinoma revealed by genome-wide expression analyses of primary tumors and metastases. *Int J Cancer* **131**, E693-704 (2012).
53. Carlo, M.I. *et al.* Prevalence of Germline Mutations in Cancer Susceptibility Genes in Patients With Advanced Renal Cell Carcinoma. *JAMA Oncol* **4**, 1228-1235 (2018).
54. van Riet, J. *et al.* The genomic landscape of 85 advanced neuroendocrine neoplasms reveals subtype-heterogeneity and potential therapeutic targets. *Nat Commun* **12**, 4612 (2021).
55. Angus, L. *et al.* The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. *Nat Genet* **51**, 1450-1458 (2019).
56. Motzer, R.J. *et al.* Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma. *N Engl J Med* **378**, 1277-1290 (2018).
57. Motzer, R.J. *et al.* Survival outcomes and independent response assessment with nivolumab plus ipilimumab versus sunitinib in patients with advanced renal cell carcinoma: 42-month follow-up of a randomized phase 3 clinical trial. *J Immunother Cancer* **8**(2020).
58. Motzer, R.J. *et al.* Final Overall Survival and Molecular Analysis in IMmotion151, a Phase 3 Trial Comparing Atezolizumab Plus Bevacizumab vs Sunitinib in Patients With Previously Untreated Metastatic Renal Cell Carcinoma. *JAMA Oncol* (2021).
59. Epaillard, N. *et al.* BIONIKK: A phase 2 biomarker driven trial with nivolumab and ipilimumab or VEGFR tyrosine kinase inhibitor (TKI) in naive metastatic kidney cancer. *Bull Cancer* **107**, eS22-eS27 (2020).
60. Choueiri, T.K. & Kaelin, W.G., Jr. Targeting the HIF2-VEGF axis in renal cell carcinoma. *Nat Med* **26**, 1519-1530 (2020).
61. van der Veldt, A.A. *et al.* Sunitinib for treatment of advanced renal cell cancer: primary tumor response. *Clin Cancer Res* **14**, 2431-6 (2008).

62. Jonasch, E. *et al.* Belzutifan for Renal Cell Carcinoma in von Hippel-Lindau Disease. *N Engl J Med* **385**, 2036-2046 (2021).
63. Nyk, L. *et al.* Interobserver Variability in Assessment of Renal Mass Biopsies. *Urol J* **18**, 400-403 (2020).
64. Haas, M. Donor kidney biopsies: pathology matters, and so does the pathologist. *Kidney Int* **85**, 1016-9 (2014).
65. Osunkoya, A.O., Young, A.N., Wang, W., Netto, G.J. & Epstein, J.I. Comparison of gene expression profiles in tubulocystic carcinoma and collecting duct carcinoma of the kidney. *Am J Surg Pathol* **33**, 1103-6 (2009).
66. Casparie, M. *et al.* Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol* **29**, 19-24 (2007).
67. Vasimuddin, M., Misra, S., Li, H. & Aluru, S. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems,. *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)* 314-324 (2019).
68. Cameron, D.L. *et al.* GRIDSS, PURPLE, LINX: Unscrambling the tumor genome via integrated analysis of structural variation and copy number. *bioRxiv*, 781013 (2019).
69. Karczewski, K.J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434-443 (2020).
70. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biology* **17**, 122 (2016).
71. Mermel, C.H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biology* **12**, R41 (2011).
72. Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**, 1029-1041 e21 (2017).
73. Blokzijl, F., Janssen, R., van Boxtel, R. & Cuppen, E. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. *Genome Medicine* **10**, 33 (2018).
74. Alexandrov, L.B. *et al.* The repertoire of mutational signatures in human cancer. *Nature* **578**, 94-101 (2020).
75. Cortés-Ciriano, I. *et al.* Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *Nature Genetics* **52**, 331-341 (2020).
76. Debyani, C. *et al.* OncoKB: A Precision Oncology Knowledge Base. *JCO Precision Oncology*, 1-16 (2017).
77. Griffith, M. *et al.* CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nature Genetics* **49**, 170-174 (2017).
78. Tamborero, D. *et al.* Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Medicine* **10**, 25 (2018).
79. College ter Beoordeling van Geneesmiddelen. https://www.cbg-meb.nl/.
80. Landrum, M.J. *et al.* ClinVar: improvements to accessing data. *Nucleic Acids Research* **48**, D835-D844 (2019).
81. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2012).
82. Yates, A.D. *et al.* Ensembl 2020. *Nucleic Acids Research* **48**, D682-D688 (2019).
83. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).
84. Korotkevich, G. *et al.* Fast gene set enrichment analysis. *bioRxiv*, 060012 (2021).
85. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425 (2015).





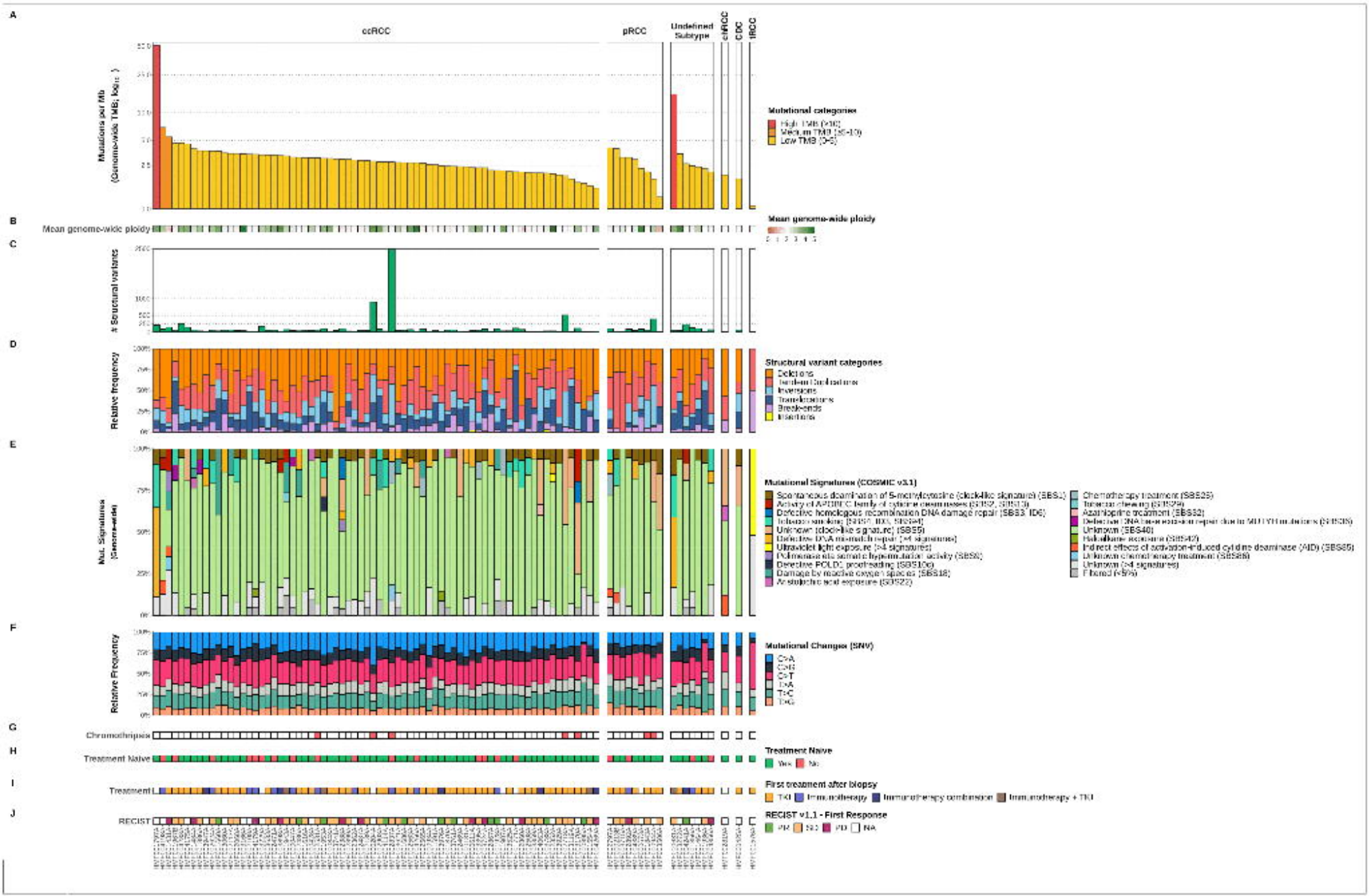
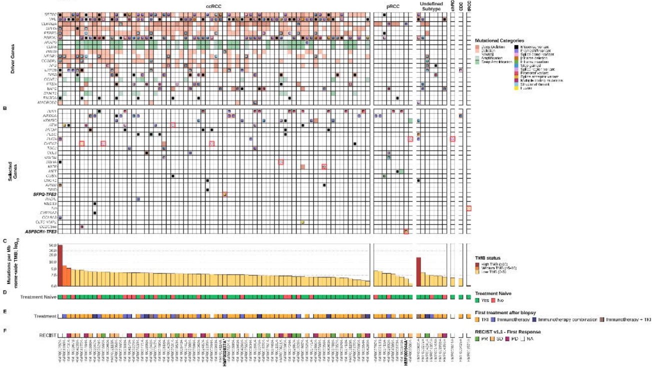

A
ccRCC
pRCC
Undefined Subtype
cHRCC
CDC
tRCC
Mutations per Mb (Genome-wide TMB; log10)
Mutational categories
High TMB (>10)
Medium TMB (3-10)
Low TMB (0-3)
B
Mean genome-wide ploidy
Mean genome-wide ploidy
C
Structural variants
D
Relative frequency
Structural variant categories
Deletions
Tandem Duplications
Inversions
Translocations
Break-ends
Insertions
E
Mut. Signatures (Genome-wide)
Mutational Signatures (COSMIC v3.1)
Spontaneous deamination of 5-methylcytosine (clock-like signature) (SBS1)
Activity of APOBEC family of cytosine deaminases (SBS2, SBS13)
Defective homologous recombination DNA damage repair (SBS3, ID6)
Tobacco smoking (SBS4, ID3, SBS92)
Unknown (clock-like signature) (SBS5)
Defective DNA mismatch repair (>4 signatures)
Ultraviolet light exposure (>4 signatures)
Polymerase eta somatic hypermutation activity (SBS9)
Defective POLD1 proofreading (SBS10b)
Damage by reactive oxygen species (SBS18)
Aristolochic acid exposure (SBS22)
Chemotherapy treatment (SBS25)
Tobacco chewing (SBS29)
Azathioprine treatment (SBS32)
Defective DNA base excision repair due to MUTYH mutations (SBS36)
Unknown (SBS40)
Haloalkane exposure (SBS42)
Indirect effects of activation-induced cytidine deaminase (AID) (SBS85)
Unknown chemotherapy treatment (SBS86)
Unknown (>4 signatures)
Filtered (<5%)
F
Relative Frequency
Mutational Changes (SNV)
C>A
C>G
C>T
T>A
T>C
T>G
G
Chromothripsis
H
Treatment Naive
Treatment Naive
Yes
No
I
Treatment I
First treatment after biopsy
TKI
Immunotherapy
Immunotherapy combination
Immunotherapy + TKI
J
RECIST
RECIST v1.1 - First Response
PR
SD
PD
NA

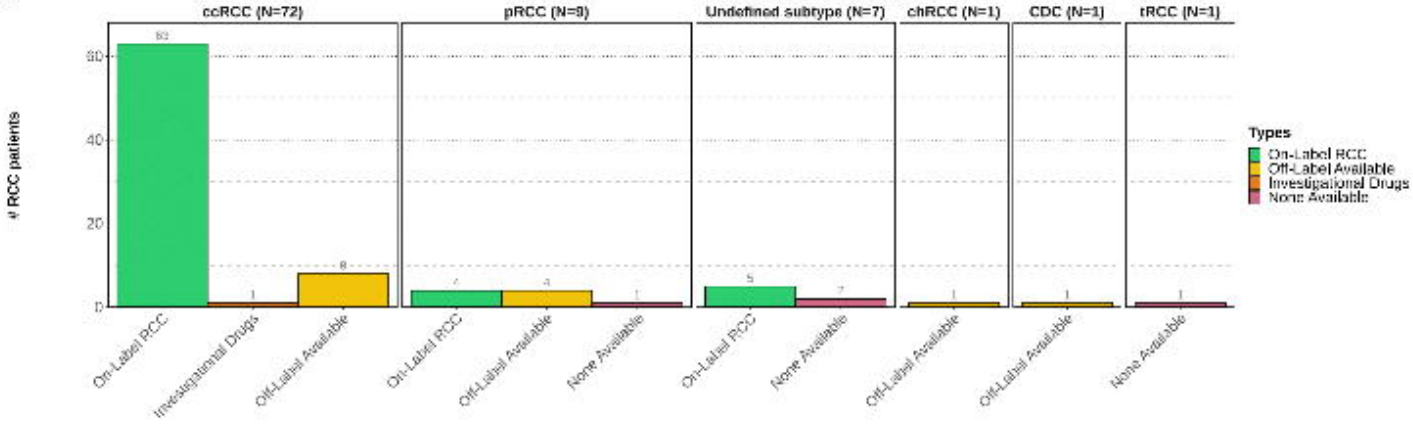
Figure 3



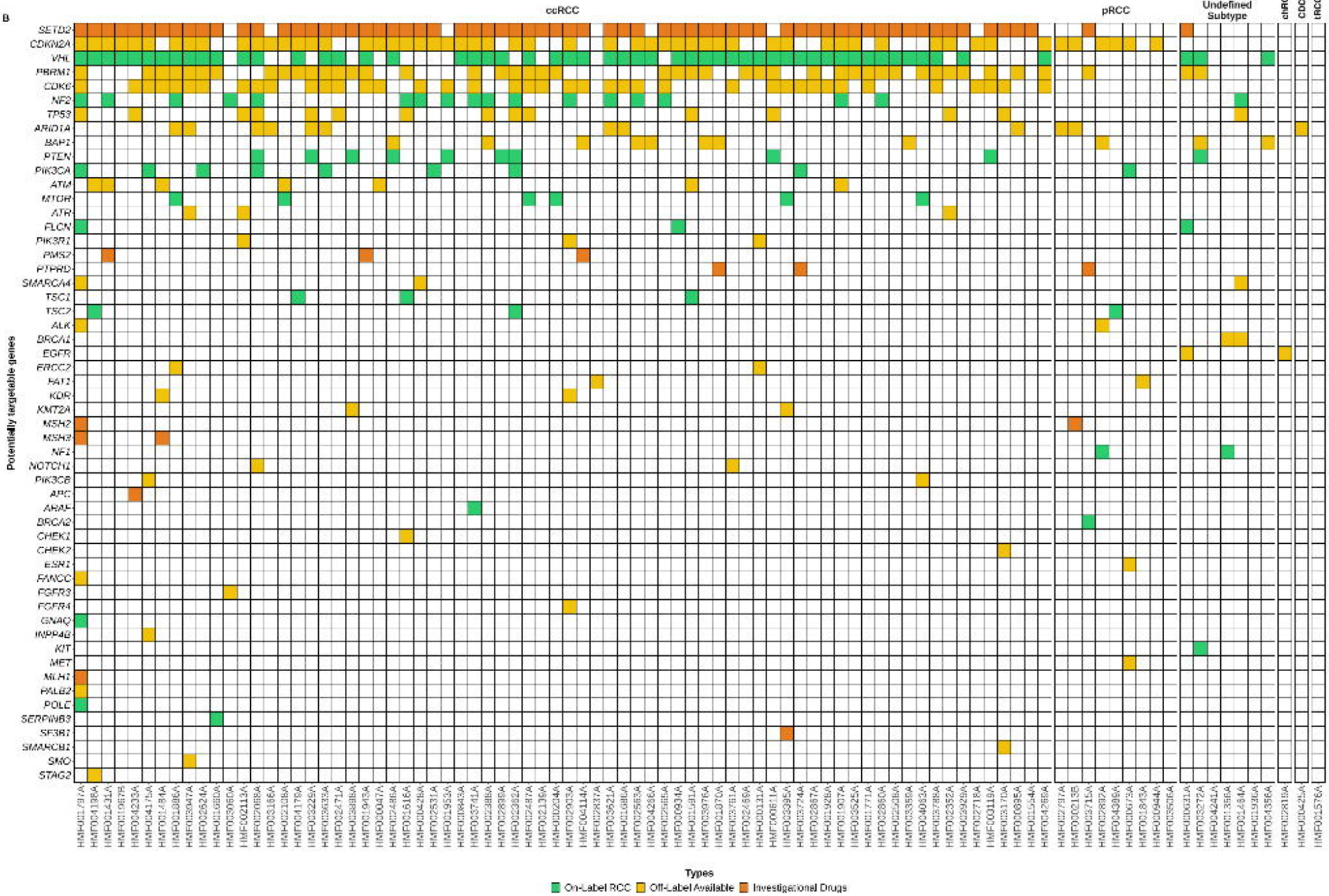
A
ccRCC
pRCC
Undefined Subtype
chRCC
CDC
tRCC
Driver Genes
Mutational Categories
Deep Deletion
Deletion
Neutral
Amplification
Deep Amplification
B
Selected Genes
SFPQ-TFE3
ASPSCR1-TFE3
C
TMB status
D
Treatment Naive
Yes
No
E
Treatment
First treatment after biopsy
TKI
Immunotherapy
Immunotherapy combination
Immunotherapy + TKI
F
RECIST
RECIST v1.1 - First Response
PR
SD
PD
NA

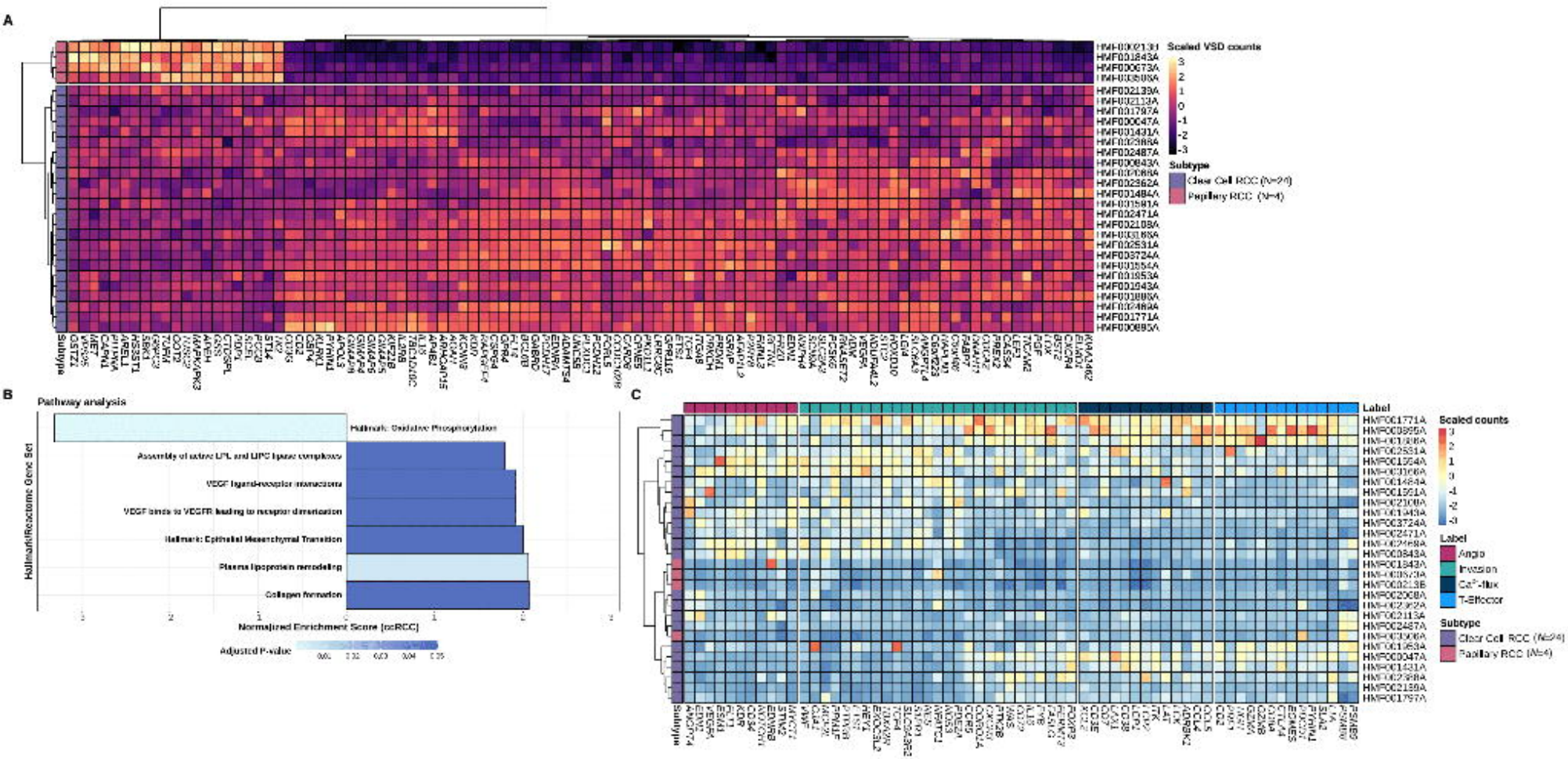
Actionable targets
iClusion Data

A



B




A
Scaled VSD counts
3
2
1
0
-1
-2
-3
Subtype
Clear Cell RCC (N=20)
Papillary RCC (N=4)
B
Pathway analysis
Hallmark:Reactome Gene Set
Hallmark: Oxidative Phosphorylation
Assembly of active LPL and LIPC lipase complexes
VEGF ligand-receptor interactions
VEGF binds to VEGFR leading to receptor dimerization
Hallmark: Epithelial Mesenchymal Transition
Plasma lipoprotein remodeling
Collagen formation
Normalized Enrichment Score (ccRCC)
Adjusted P-value
C
Label
Scaled counts
Angio
Invasion
Ca2+-flux
T-Effector
Subtype
Clear Cell RCC (N=20)
Papillary RCC (N=4)