The genomic and transcriptomic landscape of advanced renal cell cancer for individualized treatment strategies

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 Biology Center, Erasmus MC Cancer Institute, Erasmus University Medical Center Rotterdam, Department of Immunology and Department of Urology, Internal Postal Address NA-1218, 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

Keywords: non-clear cell renal cell carcinoma; personalized medicine; renal cell carcinoma; RNA-sequencing; whole-genome sequencing.

Abstract

Background

Differences in the clinical course and treatment responses in individual patients with advanced renal cell carcinoma (RCC) can largely be explained by the different genomics of this disease. To improve the personalized treatment strategy and survival outcomes for patients with advanced RCC, the genomic make-up in patients with advanced RCC was investigated to identify putative actionable mutations and signatures.

Methods

In this prospective multicenter study (NCT01855477), whole-genome sequencing (WGS) data of locally advanced and metastatic tissue biopsies and matched whole-blood samples were collected from 91 patients with histopathologically confirmed RCC. WGS data were analyzed for small somatic variants, copy-number alterations and structural variants. For a subgroup of patients, RNA sequencing (RNA-Seq) data could be analyzed. RNA-Seq data were clustered on immunogenic and angiogenic gene expression patterns according to a previously developed angio-immunogenic gene signature.

Results

For papillary and clear cell RCC, putative actionable drug targets were detected by WGS in 100% of the patients. RNA-Seq data of clear cell and papillary RCC were clustered using a previously developed angio-immunogenic gene signature. Analyses of driver mutations and RNA-Seq data revealed clear differences among different RCC subtypes, showing the added value of WGS and RNA-Seq over clinicopathological data.

Conclusions

By improving both histological subtyping and the selection of treatment according to actionable targets and immune signatures, 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 data. Prospective clinical trials are needed to evaluate the impact of genomic and transcriptomic diagnostics on survival outcome for advanced RCC patients.

Background

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 a slowly progressive disease over the years to a rapidly progressive disease with fast clinical deterioration. Therefore, the management of advanced ccRCC varies from active surveillance to systemic therapy.

The therapeutic landscape for patients with advanced ccRCC has changed significantly, in recent years. 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 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, non-clear cell RCC (nccRCC) is a heterogeneous group of different histological subtypes, including papillary and chromophobe RCC (18). As holds for advanced ccRCC, the course of nccRCC differs significantly between patients (19, 20). Since the nccRCC subtypes are considered rare diseases, randomized phase three clinical trials are still lacking for nccRCC (21). As a result, no clear standard of care for patients with advanced nccRCC has been defined (20, 22).

The development of RCC, including its metastatic potential and response to treatment, could largely 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). To improve the individualized treatment strategy and 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

111 with anti-angiogenic drugs, ICIs or a combination of these therapies and contribute to the next

112 steps in personalized medicine in patients with RCC.

Methods

Patient cohort, study procedures, sample collection, clinical data

All patients provided written informed consent for participation in the prospective multicenter Center for Personalized Cancer Treatment (CPCT-02) study (NCT01855477) by the Declaration of Helsinki. The CPCT-02 trial was approved by the medical ethical committee of the University Medical Center Utrecht and 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 (26). In summary, for this analysis, core needle biopsies from the tumor lesion, peripheral whole blood samples and clinical data were collected across 24 hospitals in the Netherlands. Response to treatment was determined according to RECIST v1.1 (27). 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. After checking the informed consents and the pathological reports, we selected 91 WGS samples (51 previously described by Priestley *et al. (26),* supplementary data file) and 28 completely new matching RNA-Seq 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 (28). 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. 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. 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* (26). Briefly, read pair mapping was performed using BWA-mem (29) 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) (30) was used for structural variant (SV) calling and LINX (v1.11) (30) for gene fusion event calling. Computational ploidy estimation and copy-number (CN) assessment was performed using the PURPLE (PURity & PLoidy Estimator) pipeline (30), estimating tumor purity and CN profile by combining B-allele frequency (BAF), read depth, and SVs.

Somatic variant annotation and filtering

As part of the data request, somatic variants were determined using Strelka and provided by the HMF. Variant Call Format (VCF) files with somatic variants were annotated based on GRCh37 with HUGO gene symbols, HGVS notations, gnomAD (31) frequencies using VEP (32) (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) (31): gnomAD exome (ALL) allele frequency ≥ 0.001; and gnomAD genome (ALL) ≥ 0.005. Variants specific to 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 (33) (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. The 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

176 homozygous loss, -1 = shallow, possibly heterozygous loss), one for diploid (0 = diploid) and two
177 for amplifications (1 = few additional copies, often broad gain, 2 = more copies, often focal gain).
178 Annotation of GISTIC2.0 peaks was performed as follows: A) Wide peaks were annotated with all
179 overlapping canonical UCSC genes within the genomic limits of the said peak. B) Focal peaks were
180 annotated based on overlapping genomic coordinates, using custom R scripts and UCSC gene
181 annotations. As a separate analysis, GISTIC2.0 was executed with `brlen` set to 0.5, calling arm-
182 level events if at least 50% of the chromosomal arm was affected by broad CN events.

183 **Structural variant analysis**

184 SVs affecting genes were imported using custom R scripts, overlapping genes on at least one
185 breakpoint, using GRCh37 genomic coordinates. SVs with an upstream or downstream Tumor
186 Allele frequency (TAF) below 0.1 as determined by PURPLE and GRIDSS (30) were discarded along
187 with SVs that affected all gene exons. In the case of both (multiple) mutations and/or SVs in the
188 same gene, these were annotated as 'multiple mutations'.

189 **Fusion gene analysis**

190 WGS-based LINX TSV files were imported using R and overlapped with the three pillars of
191 ChimerDB (34); deep sequencing data (ChimerSeq), text mining of PubMed publications
192 (ChimerPub), with extensive manual annotations (ChimerKB). Events not present in any pillar of
193 ChimerDB and intra-gene fusions were filtered out. RNA-Seq based fusion genes detected with
194 Isofox (https://github.com/hartwigmedical/hmftools/tree/master/isofox) were imported using R
195 and overlapped with the fusion events detected in the DNA sequencing.

196

Recurrent non-coding (and coding) somatic variants

All filtered non-coding somatic variants, occurring in at least four samples were selected as recurrent. Since none of these non-coding recurrent variants were located in regulatory elements (except for the *TERT* promoter) of genes known to be relevant in RCC, we chose to only communicate the *TERT* hotspot. Using the same methodology for coding variants, selecting those arising in at least four samples, did not result in pertinent findings.

Somatic Driver Genes Analysis

On the ccRCC samples we utilized the dN/dS model (192 Poisson rate parameters; under the full trinucleotide model) to identify genes undergoing mutational selection in ccRCC using the R package dndscv (35) (v0.0.1.0). The substitution and InDel models were used and corrected for sequence composition, gene length and mutational signatures (COSMIC database, see supplementary files). These models test the ratio between nonsynonymous (missense, nonsense and essential splice site) and background (synonymous) mutations. A q-value < 0.05 (including and excluding the InDel model) was used to identify genes that drive selection.

Mutational signatures analysis

Mutational signatures analysis was performed using the MutationalPatterns R package (v3.2.0)(36). The mutational signatures based on single base substitutions (N = 90 v3 signatures) were downloaded from COSMIC (37). 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). Subsequently, a constrained linear combination of the ninety mutational signatures was constructed per sample using non-

218 negative least squares regression implemented in the R package pracma (v2.2.9). Mutational
219 signatures were bootstrapped (*N* = 100) with MutationalPatterns and argument 'method' set to
220 "strict" to assess calling stability. Signature contribution for each sample was determined per 100
221 samplings, per signature.

222 **Chromothripsis**

223 Chromothripsis (CT), also known as chromosomal shattering, followed by seemingly random re-
224 ligation, was detected using Shatterseek (38) (v0.4) with default settings. The following definition
225 of CT was employed: (1) ≥25 intrachromosomal SVs involved in the event; (2) ≥7 oscillating CN
226 segments (2 CN states) or ≥14 oscillating CN segments (3 CN states); (3) CT event involving ≥20
227 Mb; (4) satisfying the test of equal distribution of SV types (p-value > 0.05); and (5) satisfying the
228 test of nonrandom SV distribution within the cluster region or chromosome (p-value ≤ 0.05).

229 **Actionable targets**

230 iClusion (https://iclusion.com) data, which connects specific or gene-level aberrations to clinical
231 cancer studies, were provided by HMF. This integrates clinical interpretations from Precision
232 Oncology Knowledge Base (OncoKB) (39), Clinical Interpretation of Variants in Cancer (CIViC) (40)
233 and Cancer Genome Interpreter (CGI) (41). All targets and biomarkers were overlapped with
234 filtered molecular data to verify presence. No RCC-specific effectiveness is taken into
235 consideration with this approach, purely resting the link on the molecular basis of the target to
236 broaden the horizon of actionable targets. Studies included in the databases mentioned can
237 conceivably show associations between drugs and targets in other cancer types than RCC or the
238 subtypes. 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 into the following three categories: on-label drugs for RCC, off-label available, and 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") (42). This evaluation board considers previous approvals by the U.S. Food and Drug Administration (FDA) and/or European Medicines Agency (EMA).

Germline analysis

Known pathogenic germline variants (GRCh37) related to cancer and/or Von Hippel-Lindau syndrome were retrieved from ClinVar (43) 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 the VCF files of our cohort. Variants with at least two reads and passing variant caller quality control were included. Furthermore, variants 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 were aligned using STAR (44) 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 (45) (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 (46) with 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 the 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 (47) (Monte Carlo approach with Adaptive Multilevel Splitting) with MSigDB (48) Hallmarks and Reactome pathways (49) 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 were 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 the WGS data are available at https://github.com/hartwigmedical/ and/or can be provided by authors upon request.

Results

Patient selection

In total, WGS data of 91 patients with histopathologically confirmed RCC from 24 hospitals in the Netherlands were included in the analyses (**Figure 1A**). Matched 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. The RCC subtypes of the remaining seven patients could not be further defined pathologically. 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 time of the 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 the 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%).

Whole-genome sequencing analyses

The tumor and blood biopsies were whole genome sequenced (WGS) with a median coverage of 102X and 38X (Supplementary figure 1A). The number and type of somatic changes in the WGS data were analyzed on both small mutations and structural variations (supplementary figure 2), showing similar numbers as previous RCC cohorts (7,224 substitutions, Q1–Q3: [5648-8581] *vs*. 7,050, Q1–Q3: [6434-9504], *p*-value = 0.45, Mann-Whitney U test (5)).

Somatic small variants, copy-number alterations (CNAs) and structural variants (SVs) were identified as described previously (26). 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 an undefined subtype of RCC) had a TMB > 10 bases. In the two 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 detected in a sample from a patient with tubulocystic RCC (*N* = 1) which had minimal genomic aberrations in general, and only two detectable SVs (**Figure 2C/D;** tandem duplication and break-end). SBS40 (unknown etiology) was the dominant mutational signature, with a mean relative contribution of 74% in this cohort. To assess the reliability of the SBS40 calling, the mutational signature calling was bootstrapped, which revealed a very high variance (median difference in the assignment of 37%) in the relative contribution of SBS40 (Supplementary figure 3).

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). For 57% of ccRCC, the genome-wide ploidy was 2, while 32% had a genome doubling with a ploidy of 3 or higher (Supplementary figure 2D). Differences in DNA ploidy have been associated with tumor differentiation and diploidy has been related to well-differentiated RCC (50). Considering SVs, a total of 3,121 deletions, 941 translocations, 2,196 tandem duplications, four insertions, 2,714 inversions and

641 break-ends were detected (Supplementary Fig 2F). The number of patients with events considered chromothripsis was limited and present in only five patients with ccRCC, two patients with pRCC and absent in the other subtypes. Furthermore, these chromothripsis events primarily 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 4), while these specific translocations were more frequently detected in the cohort of patients without chromothripsis events (29.7%, Supplementary figure 5). While clinical data concerning pre-treatment (**Figure 2H**), first treatment after biopsy (**Figure 2I**), and accompanying RECIST score (**Figure 2J**) was available, no correlative conclusions were drawn due to heterogeneity which would undermine any reliability attached to the data (Supplementary figure 6).

In patients with ccRCC, both previously described and novel amplifications and deletions were detected. Statistically significant CNA peaks (A) and arm-level copy-number alterations (B) in ccRCC are presented in Supplementary figure 7. Arm-level CNAs covering more than 50% were detected for amplifications of 1q, 5q, 7q, 8q, 12p and 20q, and deletions in 3p, 9p and 14q. All these CNAs have been described previously(51). Furthermore, in the current cohort, amplifications of 5p, 7p, 12q, 16p and 20p, and deletions of 4p, 4q, 6p, 8p, 9q, 14p, 18p and 18q were also statistically significant.

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. The driver gene analyses revealed that most driver genes in ccRCC encompassed variations in the chromosome 3p region: *PBRM1* (94.4%), *VHL* (93.1%), *SETD2* (90.3%), and *BAP1* (87.5%) together with events in tumor suppressor genes on other

chromosomes, such as *CDKN2A*/*B* (66.7%) and *PTEN* (30.6%) (**Figure 3,** supplementary data file). In addition, many patients with ccRCC had focal deletions in genes described as putative tumor suppressor genes, e.g. *PTPRD* (65.3%) and *NEGR1* (44.4%) (4, 52). Deletions of *PTPRD* have been described as a potential risk factor for the development of ccRCC (52, 53). Moreover, amplifications were present in genes associated with cell proliferation and angiogenesis, such as *CDK6* (55.6%, also prevalent in pRCC at 77.8%) and *CCNE1* (16.7%) (54, 55). Pathogenic germline mutations related to cancer or Von Hippel-Lindau syndrome were found in nine patients in different RCC subtypes and included *ATM, FLCN, CHEK2, FH, SDHA* and *MITF* (**Figure 3**).

In addition, genes that were previously described as frequently mutated somatically (q-value < 0.05) in (cc)RCC by Braun *et al*. (56) and in pRCC by Turajlic *et al*. (57), that were not statistically significant in our driver gene analysis, were included to extend our analysis. The commonly most affected and added genes in ccRCC were *MITF* (3p), *TERT* (5p), *RADIL* (7p), and *MET* (7q), predominantly as bystanders of arm-level events. The *TERT* promoter hotspot variant (C228T) (5) was detected in both ccRCC (N = 10) and pRCC (N = 2). *MET* and *RADIL* were also frequently affected in pRCC (both in 77.8%), along with *MED13* and *USP32* (both in 66.7%). Genes associated with poor prognosis in ccRCC showed a low mutational frequency in our cohort (4), such as the Krebs cycle genes (e.g. *SDHA*, *FH*).

Furthermore, previously validated fusion events (34) were detected, with *CLTC-VMP1* and *SFPQ-TFE3* both occurring once in the ccRCC group, along with a fusion event of *ASPSCR1-TFE3* in one patient with pRCC. The histopathological diagnosis had to be reconsidered for both patients with a detected *TFE3* fusion. As a result, these patients were reallocated in a different subcategory of

RCC, i.e. MiT family translocation renal cell carcinomas (1). Overall, characteristic clear cell driver gene events — such as somatic *VHL* mutations/deletions — were infrequent in nccRCC, except for *CDKN2A* deletions in pRCC.

Next, we investigated if the tumors contained targetable variants (**Figure 4**), that are known drivers in RCC (**Figure 3**) but also proteins which are considered targetable in other cancer indications. For example, specific drugs have been developed for target-specific variants encoded by *CDK4/6* or *EGFR* (58-60). This may result in off-label availability of these drugs for patients with similar aberrations in RCC, e.g. in the context of a clinical trial (61). Furthermore, somatic aberrations in cancer genes, e.g. *TP53*, are also considered biomarkers for targeted treatments (62). Lastly, some variants leading to specific mechanisms were also considered 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 (63). Overall, for the majority (90 out of 91) 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 (Supplementary figure 1B) 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 8), 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* (64) and *MAPKAPK3* (65) correlate with poor survival in RCC. In addition, various other genes were differentially expressed and have been described in other malignancies, such as *TUSC2* (66), *CAPN1* (67), *PCSK6* (68), and *CD2* (69). The differential expression of these genes confirmed a clear distinction between pRCC and ccRCC at the transcriptomic level.

Second, the pathway analysis (**Figure 5B**) of the 1,546 significant differentially expressed genes revealed cancer hallmarks such as oxidative phosphorylation and epithelial-mesenchymal transition (EMT), among others (70). Significant Reactome pathways (49) 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 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 9).

The 66 gene-signature has been based on previous data from the IMmotion150 trial (25). High expression of 'angiogenic' genes and specific '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 likely response to ICIs. According to this gene signature, ccRCC cases in the current cohort could be classified as either immunogenic or angiogenic (**Figure 5C**).

418 Moreover, all patients with pRCC had low expression of these genes, except for one individual
419 patient with expression of *EDNRB (71)*.

420

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* alterations (93.1%), most common driver gene mutations in ccRCC included alterations in tumor suppressor genes of different pathways such as *SETD2* (90.3%) and *PTEN* (30.6%). 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 90 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 generated. RNA-Seq could be applied to distinguish ccRCC, pRCC and histologically undefined RCC based on the differential gene expression. Using 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 (*N* = 4) using these signatures was not feasible.

At the genomic level, the findings for ccRCC mainly corresponded with previous findings (4, 5, 72). Previously described statistically significant arm-level events were also found in this cohort, i.e., amplifications of 1q, 5q, 7q, 8q, 12p and 20q, and deletions in 3p, 9p and 14q. However, a striking difference in the frequency of some events is caused by a distinct definition of 'arm-level'. For instance, both full loss of 3p and focal loss of events along 3p21-p25 (*VHL*, *PBRM1*, *BAP1*, *SETD2*) are counted as 'arm-level' in the ccRCC characterization from the Cancer Genome Atlas Research Network (4). In contrast, we defined these (and other large) events as covering >50%

of the chromosome arm (Supplementary Figure 7). 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 (73, 74), this is certainly not a result of a bias in the sequencing or our workflows. In contrast to other ccRCC studies, the number of chromothriptic events was limited in our study and occurred in only five patients. In previous studies, chromothripsis was defined as the combination of a chromothripsis event and a translocation event with concurrent 3p loss and 5q gain, which were called "t(3;5) chromothripsis events" (5). Although both chromothripsis and translocation (t(3;5)) events occurred in our cohort, for most of the cases, these events were independent of each other (χ^2 test, p-value = 0.6193, events visualized in Supplementary figures 4 and 5).

At the transcriptomic level, assigning ccRCC biopsies to 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 (75, 76), resulting in a clinical need to select patients who will benefit from angiogenic or immunogenic treatment (17). RNA-seq 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. However,

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, 77). However, further research in a prospective setting is still warranted (78).

The distinctive mutational gene pattern between ccRCC and nccRCC 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* alterations and 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 their downstream consequences. For instance, the frequently mutated *VHL* gene in ccRCC has been the basis for the developing angiogenesis inhibitors. The loss-of-function 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 (63, 79). Several TKIs that have been approved for the treatment of advanced ccRCC (6, 7, 63, 80) and interfere at different levels in this cascade (63, 79). More recently, Hypoxia Inducible Factor (HIF) inhibition has also shown proven efficacy in patients with *VHL* alterations (81). However, as patients with nccRCC in our cohort showed no mutations in *VHL* or other angiogenesis related pathways, it is questionable whether treatment directed against the *VHL* pathway would be the most effective therapy for this group of patients. On the other hand, germline mutations associated with specific genetic syndromes were detected in several patients with nccRCC. For instance, the patient with tRCC had a germline *FH* mutation, which makes it likely that this patient suffered from hereditary leiomyomatosis renal cell carcinoma syndrome (HLRCC) (82). Moreover, patients with a germline *FLCN* mutation may suffer from the Birt-Hogg-

Dubé syndrome, often associated with chromophobe renal cell carcinoma or oncocytoma (83). Our germline analysis shows that the number of germline mutation is relatively high in the more rare subtypes of RCC, demonstrating the importance of germline mutation analysis for (rare forms of) nccRCC. Since germline mutations play an important role in the development of RCC and are often associated with particular types of RCC, detecting these mutations allows for personalized and targeted treatment of the affected 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 (84, 85). In our cohort, nearly 8% of RCC cases could not be sub-classified through histopathological assessment. Thereby, for two patients, a fusion gene was predicted using WGS which led to the 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 a significant clinical impact. Since different subtypes and growth patterns of RCC are driven by gene expression (86), 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 8 and 9, 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 and therefore it was not possible to reliably correlate genomic and

transcriptomic findings to clinical data. A correlation with clinical data and outcome could have confirmed whether patients with certain gene signatures indeed had benefit from a specific treatment. Therefore, validation in a prospective trial is needed before 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 less common and heterogeneous subtypes, 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, with a broad spectrum of pre-treatment schedule, different treatment strategies after biopsy (**Supplementary figure 6**). All patients had metastatic or locally advanced disease but biopsies taken from metastatic and the kidney (primary site) were both included in the analyses. This strategy could have significantly impacted the analysis. However, previous studies have shown clear consistencies between primary tumor biopsies and their metastasis counterpart [5]. In addition, the heterogeneity in this cohort reflects the 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.

Conclusions

In conclusion, there are evident genomic and transcriptomic differences between RCC subtypes. The analysis of driver mutations, in combination with the 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 individualize treatment strategies in most 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 contribute to improved survival outcomes in individual patients with advanced RCC.

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 systemic 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 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 biopsies from soft tissue, muscle biopsies and biopsies from subcutaneous tissue. Figure **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 the base level. Track **G** shows the presence of chromothripsis. Track **H** shows whether patients were treatment naive at the 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 the 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)

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

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

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

598
599

Declarations

Ethics approval and consent to participate

All patients provided written informed consent for participation in the prospective multicenter Center for Personalized Cancer Treatment (CPCT-02) study (NCT01855477). The CPCT-02 trial was approved by the medical ethical committee of the University Medical Center Utrecht and the Netherlands Cancer Institute. Local approval was provided for each participating site.

Consent for publication

Not applicable.

Availability of data and materials

Data were 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 the WGS data are available at https://github.com/hartwigmedical/ and/or can be provided by authors upon request.

Competing interests

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

Funding

This study has been supported by the Erasmus Trustfonds and the Erasmus Foundation.

Authors' 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.

Acknowledgements

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

References

1. Moch H, Cubilla AL, Humphrey PA, Reuter VE, Ulbright TM. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. Eur Urol. 2016;70(1):93-105.
2. Cancer Genome Atlas Research N, Linehan WM, Spellman PT, Ricketts CJ, Creighton CJ, Fei SS, et al. Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma. N Engl J Med. 2016;374(2):135-45.
3. Padala SA, Barsouk A, Thandra KC, Saginala K, Mohammed A, Vakiti A, et al. Epidemiology of Renal Cell Carcinoma. World J Oncol. 2020;11(3):79-87.
4. Cancer Genome Atlas Research N. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature. 2013;499(7456):43-9.
5. Mitchell TJ, Turajlic S, Rowan A, Nicol D, Farmery JHR, O'Brien T, et al. Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. Cell. 2018;173(3):611-23 e17.
6. Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Rixe O, et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. N Engl J Med. 2007;356(2):115-24.
7. Mendez-Vidal MJ, Molina A, Anido U, Chirivella I, Etxaniz O, Fernandez-Parra E, et al. Pazopanib: Evidence review and clinical practice in the management of advanced renal cell carcinoma. BMC Pharmacol Toxicol. 2018;19(1):77.
8. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. N Engl J Med. 2015;373(19):1803-13.
9. Motzer RJ, Rini BI, McDermott DF, Aren Frontera O, Hammers HJ, Carducci MA, 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 RJ, Escudier B, Oudard S, Hutson TE, Porta C, Bracarda S, et al. Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. Lancet. 2008;372(9637):449-56.
11. Powles T, Plimack ER, Soulieres D, Waddell T, Stus V, Gafanov R, 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. 2020;21(12):1563-73.
12. Choueiri TK, Motzer RJ, Rini BI, Haanen J, Campbell MT, Venugopal B, 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. 2020;31(8):1030-9.
13. Rini BI, Powles T, Atkins MB, Escudier B, McDermott DF, Suarez C, 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. 2019;393(10189):2404-15.
14. Flippot R, Escudier B, Albiges L. Immune Checkpoint Inhibitors: Toward New Paradigms in Renal Cell Carcinoma. Drugs. 2018;78(14):1443-57.
15. Heng DY, Xie W, Regan MM, Harshman LC, Bjarnason GA, Vaishampayan UN, 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. 2013;14(2):141-8.
16. Guida A, Le Teuff G, Alves C, Colomba E, Di Nunno V, Derosa L, 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. 2020;11(49):4582-92.

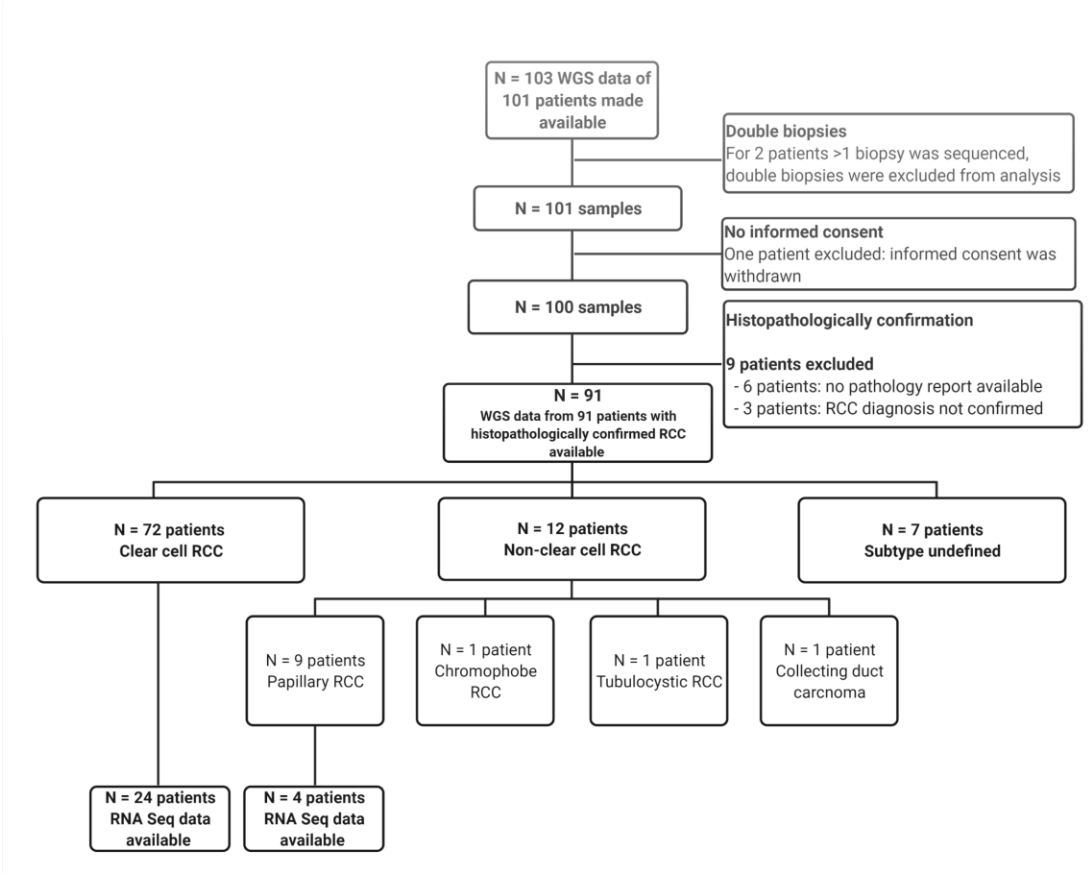
17. Vano YA, Ladoire S, Elaidi R, Dermeche S, Eymard JC, Falkowski S, et al. First-Line Treatment of Metastatic Clear Cell Renal Cell Carcinoma: What Are the Most Appropriate Combination Therapies? Cancers (Basel). 2021;13(21).
18. Truong LD, Shen SS. Immunohistochemical diagnosis of renal neoplasms. Arch Pathol Lab Med. 2011;135(1):92-109.
19. Hong B, Hou H, Chen L, Li Z, Zhang Z, Zhao Q, et al. The Clinicopathological Features and Prognosis in Patients With Papillary Renal Cell Carcinoma: A Multicenter Retrospective Study in Chinese Population. Front Oncol. 2021;11:753690.
20. Sepe P, Ottini A, Pircher CC, Franza A, Claps M, Guadalupi V, et al. Characteristics and Treatment Challenges of Non-Clear Cell Renal Cell Carcinoma. Cancers (Basel). 2021;13(15).
21. Powles T, Albiges L, Bex A, Grünwald V, Porta C, Procopio G, et al. ESMO Clinical Practice Guideline update on the use of immunotherapy in early stage and advanced renal cell carcinoma. Annals of Oncology. 2021;32(12):1511-9.
22. Osterman CK, Rose TL. A Systematic Review of Systemic Treatment Options for Advanced Non-Clear Cell Renal Cell Carcinoma. Kidney Cancer. 2020;4(1):15-27.
23. Turajlic S, Xu H, Litchfield K, Rowan A, Chambers T, Lopez JI, et al. Tracking Cancer Evolution Reveals Constrained Routes to Metastases: TRACERx Renal. Cell. 2018;173(3):581-94 e12.
24. Ricketts CJ, De Cubas AA, Fan H, Smith CC, Lang M, Reznik E, et al. The Cancer Genome Atlas Comprehensive Molecular Characterization of Renal Cell Carcinoma. Cell Rep. 2018;23(1):313-26 e5.
25. D'Costa NM, Cina D, Shrestha R, Bell RH, Lin YY, Asghari H, et al. Identification of gene signature for treatment response to guide precision oncology in clear-cell renal cell carcinoma. Sci Rep. 2020;10(1):2026.
26. Priestley P, Baber J, Lolkema MP, Steeghs N, de Bruijn E, Shale C, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. Nature. 2019;575(7781):210-6.
27. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.
28. Casparie M, Tiebosch AT, Burger G, Blauwgeers H, van de Pol A, van Krieken JH, 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. 2007;29(1):19-24.
29. 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) 2019:314-24.
30. Cameron DL, Baber J, Shale C, Papenfuss AT, Valle-Inclan JE, Besselink N, et al. GRIDSS, PURPLE, LINX: Unscrambling the tumor genome via integrated analysis of structural variation and copy number. bioRxiv. 2019:781013.
31. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434-43.
32. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biology. 2016;17(1):122.
33. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biology. 2011;12(4):R41.
34. Jang YE, Jang I, Kim S, Cho S, Kim D, Kim K, et al. ChimerDB 4.0: an updated and expanded database of fusion genes. Nucleic Acids Research. 2019;48(D1):D817-D24.
35. Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P, et al. Universal Patterns of Selection in Cancer and Somatic Tissues. Cell. 2017;171(5):1029-41 e21.

747 36. Blokzijl F, Janssen R, van Boxtel R, Cuppen E. MutationalPatterns: comprehensive genome-wide
748 analysis of mutational processes. Genome Medicine. 2018;10(1):33.
749 37. Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al. The repertoire of
750 mutational signatures in human cancer. Nature. 2020;578(7793):94-101.
751 38. Cortés-Ciriano I, Lee JJ-K, Xi R, Jain D, Jung YL, Yang L, et al. Comprehensive analysis of
752 chromothripsis in 2,658 human cancers using whole-genome sequencing. Nature Genetics.
753 2020;52(3):331-41.
754 39. Debyani C, Jianjiong G, Sarah P, Ritika K, Hongxin Z, Jiaojiao W, et al. OncoKB: A Precision
755 Oncology Knowledge Base. JCO Precision Oncology. 2017(1):1-16.
756 40. Griffith M, Spies NC, Krysiak K, McMichael JF, Coffman AC, Danos AM, et al. CIViC is a community
757 knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. Nature
758 Genetics. 2017;49(2):170-4.
759 41. Tamborero D, Rubio-Perez C, Deu-Pons J, Schroeder MP, Vivancos A, Rovira A, et al. Cancer
760 Genome Interpreter annotates the biological and clinical relevance of tumor alterations. Genome
761 Medicine. 2018;10(1):25.
762 42. College ter Beoordeling van Geneesmiddelen. https://www.cbg-meb.nl/ [Available from:
763 https://www.cbg-meb.nl/.
764 43. Landrum MJ, Chitipiralla S, Brown GR, Chen C, Gu B, Hart J, et al. ClinVar: improvements to
765 accessing data. Nucleic Acids Research. 2019;48(D1):D835-D44.
766 44. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-
767 seq aligner. Bioinformatics. 2012;29(1):15-21.
768 45. Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, et al. Ensembl 2020. Nucleic
769 Acids Research. 2019;48(D1):D682-D8.
770 46. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq
771 data with DESeq2. Genome Biology. 2014;15(12):550.
772 47. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set
773 enrichment analysis. bioRxiv. 2021:060012.
774 48. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular
775 Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417-25.
776 49. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome pathway
777 knowledgebase. Nucleic Acids Res. 2020;48(D1):D498-D503.
778 50. Abou-Rebyeh H, Borgmann V, Nagel R, Al-Abadi H. DNA ploidy is a valuable predictor for
779 prognosis of patients with resected renal cell carcinoma. Cancer. 2001;92(9):2280-5.
780 51. Beroukhim R, Brunet JP, Di Napoli A, Mertz KD, Seeley A, Pires MM, et al. Patterns of gene
781 expression and copy-number alterations in von-hippel lindau disease-associated and sporadic clear cell
782 carcinoma of the kidney. Cancer Res. 2009;69(11):4674-81.
783 52. Ortiz B, White JR, Wu WH, Chan TA. Deletion of Ptprd and Cdkn2a cooperate to accelerate
784 tumorigenesis. Oncotarget. 2014;5(16):6976-82.
785 53. Du Y, Su T, Tan X, Li X, Xie J, Wang G, et al. Polymorphism in protein tyrosine phosphatase
786 receptor delta is associated with the risk of clear cell renal cell carcinoma. Gene. 2013;512(1):64-9.
787 54. Tadesse S, Yu M, Kumarasiri M, Le BT, Wang S. Targeting CDK6 in cancer: State of the art and
788 new insights. Cell Cycle. 2015;14(20):3220-30.
789 55. Li Y, Shen Y, Zhu Z, Wen H, Feng C. Comprehensive analysis of copy number variance and
790 sensitivity to common targeted therapy in clear cell renal cell carcinoma: In silico analysis with in vitro
791 validation. Cancer Med. 2020;9(16):6020-9.
792 56. Braun DA, Hou Y, Bakouny Z, Ficial M, Sant' Angelo M, Forman J, et al. Interplay of somatic
793 alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal
794 cell carcinoma. Nat Med. 2020;26(6):909-18.

795 57. Turajlic S, Larkin J, Swanton C. SnapShot: Renal Cell Carcinoma. Cell. 2015;163(6):1556- e1.
796 58. Du Z, Brown BP, Kim S, Ferguson D, Pavlick DC, Jayakumaran G, et al. Structure-function analysis
797 of oncogenic EGFR Kinase Domain Duplication reveals insights into activation and a potential approach
798 for therapeutic targeting. Nat Commun. 2021;12(1):1382.
799 59. Ho GF, Chai CS, Alip A, Wahid MIA, Abdullah MM, Foo YC, et al. Real-world experience of first-
800 line afatinib in patients with EGFR-mutant advanced NSCLC: a multicenter observational study. BMC
801 Cancer. 2019;19(1):896.
802 60. Braal CL, Jongbloed EM, Wilting SM, Mathijssen RHJ, Koolen SLW, Jager A. Inhibiting CDK4/6 in
803 Breast Cancer with Palbociclib, Ribociclib, and Abemaciclib: Similarities and Differences. Drugs.
804 2021;81(3):317-31.
805 61. van der Velden DL, Hoes LR, van der Wijngaart H, van Berge Henegouwen JM, van Werkhoven E,
806 Roepman P, et al. The Drug Rediscovery protocol facilitates the expanded use of existing anticancer
807 drugs. Nature. 2019;574(7776):127-31.
808 62. Zhu G, Pan C, Bei JX, Li B, Liang C, Xu Y, et al. Mutant p53 in Cancer Progression and Targeted
809 Therapies. Front Oncol. 2020;10:595187.
810 63. Alonso-Gordoa T, Garcia-Bermejo ML, Grande E, Garrido P, Carrato A, Molina-Cerrillo J.
811 Targeting Tyrosine kinases in Renal Cell Carcinoma: "New Bullets against Old Guys". Int J Mol Sci.
812 2019;20(8).
813 64. Lin S, Zheng L, Lu Y, Xia Q, Zhou P, Liu Z. Comprehensive analysis on the expression levels and
814 prognostic values of LOX family genes in kidney renal clear cell carcinoma. Cancer Med.
815 2020;9(22):8624-38.
816 65. Oh E, Kim JH, Um J, Jung DW, Williams DR, Lee H. Genome-Wide Transcriptomic Analysis of Non-
817 Tumorigenic Tissues Reveals Aging-Related Prognostic Markers and Drug Targets in Renal Cell
818 Carcinoma. Cancers (Basel). 2021;13(12).
819 66. Rimkus T, Sirkisoon S, Harrison A, Lo HW. Tumor suppressor candidate 2 (TUSC2, FUS-1) and
820 human cancers. Discov Med. 2017;23(128):325-30.
821 67. Pan XQ, Huang W, Jin LW, Lin HZ, Xu XY. A Novel Pyroptosis-Related Prognostic Signature for
822 Risk Stratification and Clinical Prognosis in Clear Cell Renal Cell Carcinoma. Dis Markers.
823 2022;2022:8093837.
824 68. Wang P, Wang F, Wang L, Pan J. Proprotein convertase subtilisin/kexin type 6 activates the
825 extracellular signal-regulated kinase 1/2 and Wnt family member 3A pathways and promotes in vitro
826 proliferation, migration and invasion of breast cancer MDA-MB-231 cells. Oncol Lett. 2018;16(1):145-50.
827 69. Han J, Choi YL, Kim H, Choi JY, Lee SK, Lee JE, et al. MMP11 and CD2 as novel prognostic factors
828 in hormone receptor-negative, HER2-positive breast cancer. Breast Cancer Res Treat. 2017;164(1):41-56.
829 70. Ding XF, Zhou J, Hu QY, Liu SC, Chen G. The tumor suppressor pVHL down-regulates never-in-
830 mitosis A-related kinase 8 via hypoxia-inducible factors to maintain cilia in human renal cancer cells. J
831 Biol Chem. 2015;290(3):1389-94.
832 71. Wuttig D, Zastrow S, Fussel S, Toma MI, Meinhardt M, Kalman K, et al. CD31, EDNRB and
833 TSPAN7 are promising prognostic markers in clear-cell renal cell carcinoma revealed by genome-wide
834 expression analyses of primary tumors and metastases. Int J Cancer. 2012;131(5):E693-704.
835 72. Carlo MI, Mukherjee S, Mandelker D, Vijai J, Kemel Y, Zhang L, et al. Prevalence of Germline
836 Mutations in Cancer Susceptibility Genes in Patients With Advanced Renal Cell Carcinoma. JAMA Oncol.
837 2018;4(9):1228-35.
838 73. van Riet J, van de Werken HJG, Cuppen E, Eskens F, Tesselaar M, van Veenendaal LM, et al. The
839 genomic landscape of 85 advanced neuroendocrine neoplasms reveals subtype-heterogeneity and
840 potential therapeutic targets. Nat Commun. 2021;12(1):4612.

74. Angus L, Smid M, Wilting SM, van Riet J, Van Hoeck A, Nguyen L, et al. The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. Nat Genet. 2019;51(10):1450-8.
75. Motzer RJ, Tannir NM, McDermott DF, Aren Frontera O, Melichar B, Choueiri TK, et al. Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma. N Engl J Med. 2018;378(14):1277-90.
76. Motzer RJ, Escudier B, McDermott DF, Aren Frontera O, Melichar B, Powles T, 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. 2020;8(2).
77. Motzer RJ, Powles T, Atkins MB, Escudier B, McDermott DF, Alekseev BY, 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.
78. Epaillard N, Simonaggio A, Elaidi R, Azzouz F, Braychenko E, Thibault C, 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. 2020;107(5S):eS22-eS7.
79. Choueiri TK, Kaelin WG, Jr. Targeting the HIF2-VEGF axis in renal cell carcinoma. Nat Med. 2020;26(10):1519-30.
80. van der Veldt AA, Meijerink MR, van den Eertwegh AJ, Bex A, de Gast G, Haanen JB, et al. Sunitinib for treatment of advanced renal cell cancer: primary tumor response. Clin Cancer Res. 2008;14(8):2431-6.
81. Jonasch E, Donskov F, Iliopoulos O, Rathmell WK, Narayan VK, Maughan BL, et al. Belzutifan for Renal Cell Carcinoma in von Hippel-Lindau Disease. N Engl J Med. 2021;385(22):2036-46.
82. Merino MJ, Torres-Cabala C, Pinto P, Linehan WM. The morphologic spectrum of kidney tumors in hereditary leiomyomatosis and renal cell carcinoma (HLRCC) syndrome. Am J Surg Pathol. 2007;31(10):1578-85.
83. Benusiglio PR, Giraud S, Deveaux S, Mejean A, Correas JM, Joly D, et al. Renal cell tumour characteristics in patients with the Birt-Hogg-Dube cancer susceptibility syndrome: a retrospective, multicentre study. Orphanet J Rare Dis. 2014;9:163.
84. Nyk L, Malewski W, Kaczmarek K, Kryst P, Pyzlak M, Andrychowicz A, et al. Interobserver Variability in Assessment of Renal Mass Biopsies. Urol J. 2020;18(4):400-3.
85. Haas M. Donor kidney biopsies: pathology matters, and so does the pathologist. Kidney Int. 2014;85(5):1016-9.
86. Osunkoya AO, Young AN, Wang W, Netto GJ, Epstein JI. Comparison of gene expression profiles in tubulocystic carcinoma and collecting duct carcinoma of the kidney. Am J Surg Pathol. 2009;33(7):1103-6.

Figure 1



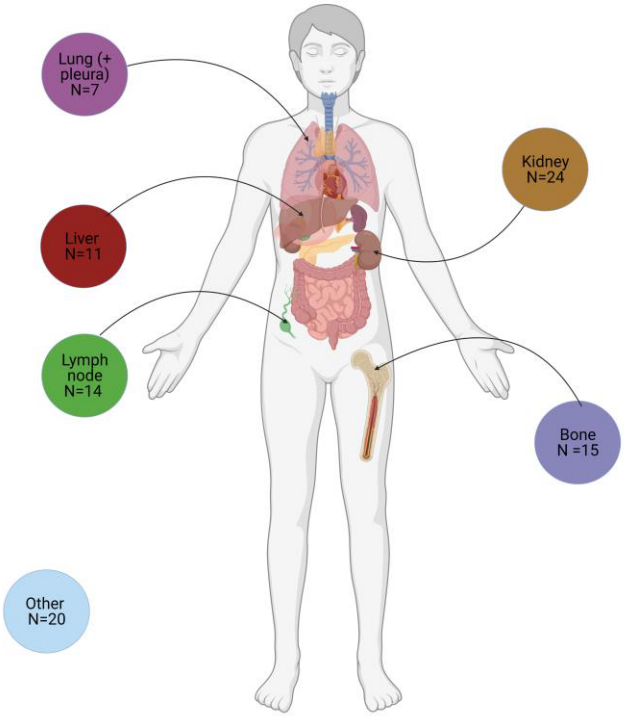
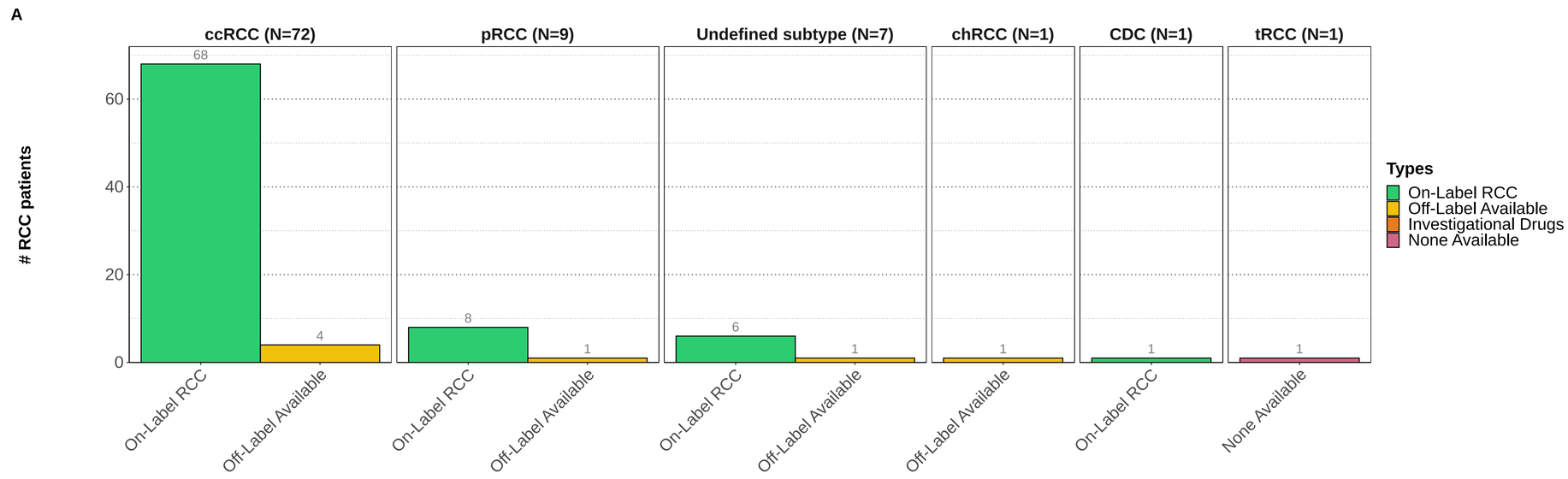


Figure 2

Figure 3

Figure 4



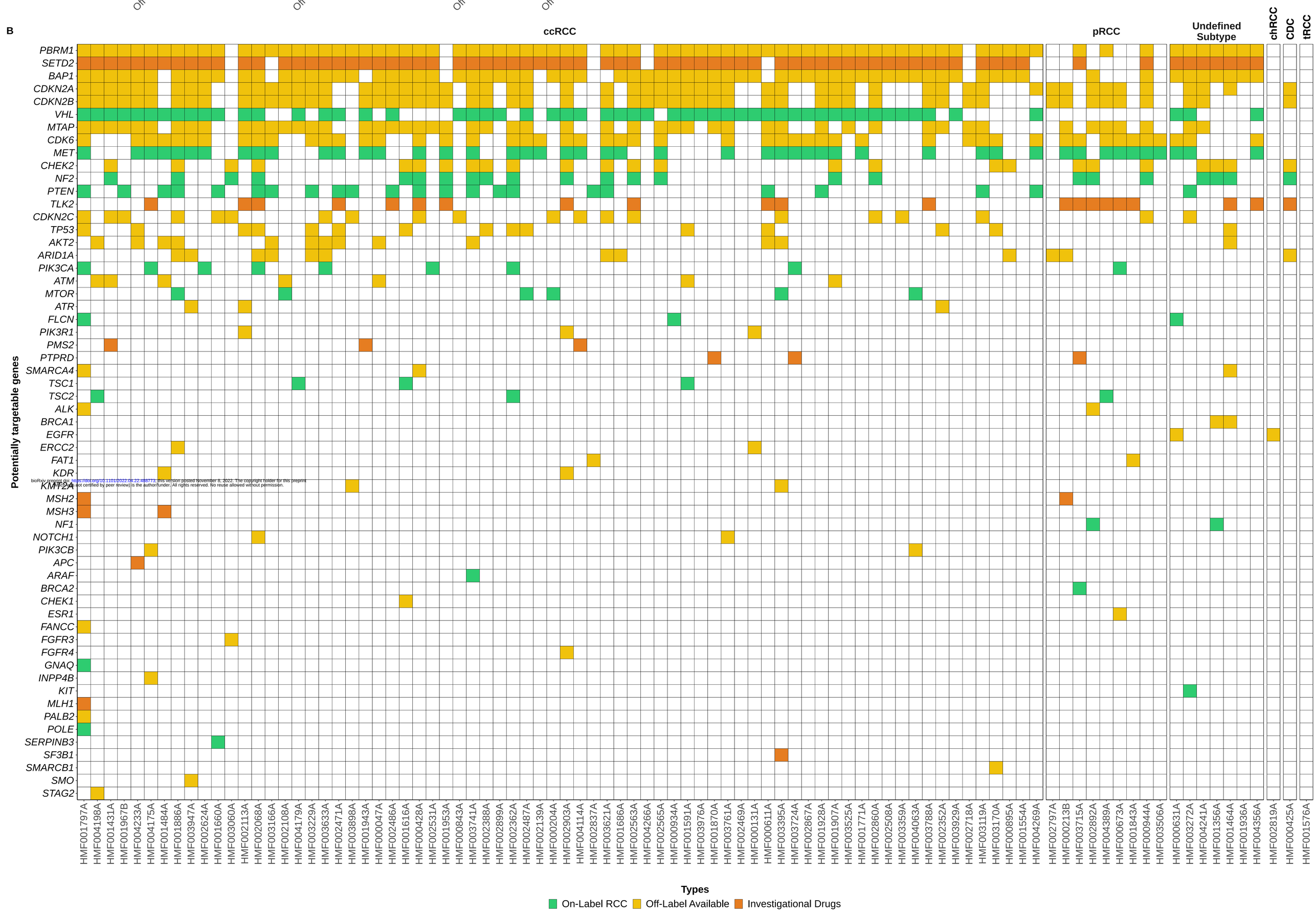


Figure 5