Dissertation

Minimal-invasive therapy monitoring in plasma DNA of patients with metastasized renal cell carcinoma

submitted by

Tina MOSER, MSc., BSc.

for the Academic Degree of

Doctor of Philosophy (PhD)

at the

Medical University of Graz
Diagnostic & Research Institute of Human Genetics

under the Supervision of

Assoz. Prof.in Priv.-Doz.in Mag.a Dr.in rer.nat.
Ellen HEITZER

2020

STATUTORY DECLARATION

I hereby declare that this dissertation is my own original work and that I have fully acknowledged by name all of those individuals and organizations that have contributed to the research for this dissertation. Due acknowledgement has been made in the text to all other material used. Throughout this dissertation and in all related publications I followed the "Guidelines of the Medical University of Graz on Good Scientific Practice".

Tina Moser; November 2020

DISCLOSURES

Part of this thesis has been published in:

Moser T[1], Ulz P[1,2], Zhou Q[1], Perakis S[1], Geigl JB[1], Speicher MR[1] & Heitzer E[1,3]. Single-Stranded DNA Library Preparation Does Not Preferentially Enrich Circulating Tumor DNA. Clin Chem 2017 Oct;63(10):1656-1659.

and

Smith CG[4,*], Moser T[1,*], Mouliere F[5], Field-Rayner J[6], Eldridge M[4], Riediger AL[4], Chandrananda D[4], Heider K[4], Wan JCM[4], Warren AY[7], Morris J[4], Hudecova I[4], Cooper WN[4], Mitchell TJ[7,8], Gale D[4], Ruiz-Valdepenas A[4], Klatte T[8,9], Ursprung S[4,10], Sala E[4,7,10], Riddick ACP[7], Aho TF[7], Armitage JN[7], Perakis S[1], Pichler M[11], Seles M[12], Wcislo G[13], Welsh SJ[7], Matakidou A[7,14], Eisen T[7,14,15], Massie CE[16], Rosenfeld N[4,+], Heitzer E[1,3,+] & Stewart GD[4,7,8,+]. Comprehensive characterization of cell-free tumor DNA in plasma and urine of patients with renal tumors. Genome Med 2020 Feb 28;12(1):23-020-00723-8.

[1]Diagnostic and Research Center for Molecular Biomedicine, Institute of Human Genetics, Medical University of Graz, Graz, Austria

[2]Freenome, South San Francisco, CA, USA

[3]Christian Doppler Laboratory for Liquid Biopsies for Early Detection of Cancer, Graz, Austria

[4]Cancer Research UK Cambridge Institute, CRUK Major Centre, Cambridge, United Kingdom

[5]Department of Pathology, Cancer Center Amsterdam, Amsterdam UMC, The Netherlands

[6]Cambridge Urology Translational Research and Clinical Trials Office, University of Cambridge, Cambridge, United Kingdom

[7]Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom

[8]Department of Surgery, University of Cambridge, Cambridge, United Kingdom

[9]Department of Urology, Royal Bournemouth Hospital, Bournemouth, United Kingdom

[10]Department of Radiology, University of Cambridge, Cambridge, United Kingdom

[11]Department of Internal Medicine Graz, Division of Oncology, Medical University of Graz, Graz, Austria

[12]Department of Urology, Medical University of Graz, Graz, Austria

[13]Department of Oncology, Military Institute of Medicine, Warsaw, Poland

[14]Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Cambridge, United Kingdom

[15]Department of Oncology, University of Cambridge, Cambridge, United Kingdom

[16]Hutchison/MRC Research Centre, University of Cambridge, Cambridge, United Kingdom

[*]Christopher G. Smith and Tina Moser contributed equally to this work.

[+]Nitzan Rosenfeld, Ellen Heitzer, and Grant D. Stewart are joint senior authors.

All the co-authors have agreed to the inclusion of their published data in the dissertation and permission to reproduce illustrations and figures from own or third-party publications has been granted.

ACKNOWLEDGMENTS

At this point, I would like to thank all the people who have supported me professionally and personally throughout my time as a student at the Diagnostic and Research Institute of Human Genetics.

First and foremost, I would like to thank my supervisor **Prof. Ellen Heitzer** for giving me the opportunity to conduct my PhD in her research team and more importantly, for providing me with her expertise and for her tireless support and patience. I am genuinely grateful for her supervision, advice and guidance from the very beginning of my research.

I would like to express my appreciation to **Prof. Michael Speicher**, who has always supported me with valuable advice, new ideas, and inspiring discussions. Thank you for involving me in other fascinating projects, which made me recognize the tremendous potential of Liquid Biopsies for precision medicine.

I would like to especially thank **Dr. Christopher G. Smith**, **Dr. Nitzan Rosenfeld**, and **Dr. Grant D. Stewart**. Our collaboration made it possible to develop new hypotheses and concepts that have enriched our study and, subsequently, our joint publication. I truly enjoyed working with all of you!

Special thanks go to all my **colleagues** and my former co-worker **Qing** for their help and support during the last four years. **Samantha**, thank you for your friendship. Even in the most challenging situations, your ability to always make me laugh encouraged me to keep going. Additionally, I want to thank **Tom** for always having an open door for my essential (and sometimes not at all essential) questions and strong personal support.

Finally, I would like to mention my friends and family, on whom I can rely in every imaginable situation. Words fail me to express my appreciation to **Philipp** – thank you for always having my back! I am deeply grateful to **Hetsch** and **Sabine**, **Matthias**, **Nici** and all the "**Feldbacher**" and "**Kopfinger**" **family members** for their endless caring support, encouragement and persistent confidence in me. I am so happy to have you all in my life!

Tina Moser received funding from the Medical University of Graz through the PhD program Molecular Medicine (MolMed). This work was supported by the Austrian Science Fund FWF [P28949-B28], and by the Austrian Federal Ministry for Digital and Economic Affairs (Christian Doppler Research Fund for Liquid Biopsies for Early Detection of Cancer).

TABLE OF CONTENTS

ABBREVIATIONS

°C	Degree Celsius
BAP1	BRCA1 Associated Protein 1
AKT	AKT Serine/Threonine Kinase
ARE	Antioxidant Responsive Element
ARID1A	AT-Rich Interaction Domain 1A
bp	Base Pair
ccRCC	Clear Cell Renal Cell Carcinoma
CDKN2A	Cyclin Dependent Kinase Inhibitor 2A
cfDNA	Cell-free DNA
chRCC	Chromophobe Renal Cell Carcinoma
CRC	Colorectal Cancer
CT	Computerized Tomography
CTC	Circulating Tumor Cell
ctDNA	Circulating tumor DNA
CTLA-4	Cytotoxic T-Lymphocyte-Associated Protein 4
DNA	Deoxyribonucleic Acid
dNTPs	Deoxynucleotide Triphosphate
dsDNA	Double-Stranded Deoxyribonucleic acid
EGFR	Epidermal Growth Factor Receptor
FDA	Food and Drug Administration
FFPE	Formalin Fixed Paraffin Embedded
gmAF	Global Mutant Allele Fraction
HIF	Hypoxia-Inducible Factor

IMDC	International Metastatic Renal Cell Carcinoma Database Consortium
INVAR-TAPAS	Integration of Variant Reads From Targeted Panel Sequencing
IR	Informative Reads
ITH	Intratumor Heterogeneity
KDM5C	Lysine Demethylase 5C
LINE-1	Long Interspersed Nucleotide Element-1
LOD	Limit of Detection
M	Molar
mAF	Mutant Allele Frequency
MET	MET Proto-Oncogene, Receptor Tyrosine Kinase
mFast-SeqS	modified Fast Aneuploidy Screening Test-Sequencing System
ml	Milliliter
mRCC	Metastatic Renal Cell Carcinoma
MRI	Magnet Resonance Imaging
MSKCC	Memorial Sloan Kettering Cancer Center
MTOR	Mechanistic Target Of Rapamycin
mTORC1	Mechanistic Target Of Rapamycin Complex 1
NF1	Neurofibromin 1
ng	Nanogram
NGS	Next-Generation Sequencing
OS	Overall Survival
PBRM1	Polybromo 1
PCR	Polymerase Chain Reaction
PD	Progressive Disease
PD-1	Programmed Cell Death Protein 1

PDGF	Platelet-Derived Growth Factor
PD-L1	Programmed Cell Death 1 Ligand 1
PFS	Progress-free Survival
pg	Picogram
PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
Plasma-Seq	Whole-genome Sequencing of Plasma DNA
pRCC	Papillary Renal Cell Carcinoma
PTEN	Phosphatase and Tensin Homolog
RCC	Renal Cell Carcinoma
RECIST	Response Evaluation Criteria in Solid Tumors
RNA	Ribonucleic acid
RR	Response Rate
SCNA	Somatic Copy Number Alteration
SD	Stable Disease
SETD2	SET Domain Containing 2, Histone Lysine Methyltransferase
SNV	Somatic Nucleotide Variant
ssDNA	Single-Stranded Deoxyribonucleic acid
sWGS	Shallow Whole-Genome Sequencing
TCEB1/ELOC	Transcription Elongation Factor B/ Elongin C
TE	Tris-EDTA
TERT	Telomerase Reverse Transcriptase
TKI	Tyrosine Kinase Inhibitor
tMAD	Trimmed Median Absolute Deviation from Copy Number Neutrality
TP53	Tumor Protein p53
TSC1/2	TSC Complex Subunit 1/2

TSG	Tumor Suppressor Gene
U	Unit
UMI	Unique Molecular Identifier
VEGF	Vascular Endothelial Growth Factor
VEGFR	VEGF Receptor
VHL	Von Hippel-Lindau Tumor Suppressor
WES	Whole-Exome Sequencing
wt/vol	Weight/Volume
µg	Microgram
µl	Microliter
µM	Micromolar

LIST OF FIGURES

LIST OF TABLES

ZUSAMMENFASSUNG

Die sogenannte Liquid Biopsy („Flüssigbiopsie“), die sich auf die minimal-invasive Analyse von zirkulierenden Tumormarkern aus Blut und anderen Körperflüssigkeiten bezieht, stellt eine innovative Methodik zur Erkennung und Überwachung krebsspezifischer Veränderungen dar. Insbesondere die Analyse der zirkulierenden Tumor-DNA (engl.: circulating tumor DNA; ctDNA) hat ein großes klinisches Potenzial und wurde bereits bei vielen soliden Tumoren zur Überwachung des Behandlungserfolges während einer anti-tumoralen Therapie eingesetzt. Über einen potenziellen Nutzen von ctDNA bei Nierenzellkarzinomen (engl.: renal cell carcinoma; RCC) ist allerdings nur wenig bekannt. Dem Nierenzellkarzinom ist eine ausgeprägte pathologische und genetische Heterogenität inhärent, die Behandlungsentscheidungen erschwert und eine immense Hürde für die personalisiere Krebsmedizin darstellt.

Im Rahmen dieser Arbeit wurde das Vorhandensein und die potenziellen klinischen Anwendungen der zirkulierenden Tumor-DNA in den Köperflüssigkeiten von PatientInnen mit Nierentumoren bestimmt. Ein besonderer Schwerpunkt lag auf jenen PatientInnen, die sog. zielgerichtete Therapien erhielten, um die Fähigkeit der ctDNA zur longitudinalen Überwachung des Krankheitsverlaufs zu validieren.

Zunächst konzentrierten wir uns auf die Optimierung der Nachweisstrategien von ctDNA. Zu diesem Zweck fokussierten wir uns darauf, ctDNA-Fragmente auf der Grundlage der Fragmentlänge anzureichern, da diese Berichten zufolge kürzer sind als jene von zellfreier DNA (cfDNA) aus hämatopoetischen Zellen. Eine *in silico* Größenselektion von Fragmenten im Bereich zwischen 90 und 150 bp verbesserte die ctDNA-Detektion maßgeblich, mit mehr als zweifacher medianer Anreicherung in bestimmten Fällen. Weiters führten wir eine detaillierte Analyse von Baseline und longitudinalen Liquid Biopsy Proben durch, wobei wir eine Reihe hochmoderner Sequenzierungsansätze in zwei unabhängigen prospektiven klinischen Kohorten einsetzten, die das gesamte klinische Spektrum von Nierentumoren repräsentieren. Wir wandten zunächst eine Kombination aus genomweiten Methoden mit *in silico* Größenselektion auf Plasmaproben von PatientInnen mit metastasiertem Nierenzellkarzinom an. Unsere Daten zeigten niedrigere ctDNA-Konzentrationen in RCC im Vergleich zu anderen soliden Krebsarten mit ähnlicher Größe und ähnlichem Stadium, wobei bei insgesamt 32,6% der PatientInnen Tumor-DNA im Plasma nachgewiesen werden konnte.

Zusätzlich konnte die Verwendung einer gezielten, hochauflösenden Methode, die sich auf die Analyse von zehn RCC-assoziierten Genen beschränkt, nicht die Detektion der genomweiten Methoden übertreffen (18,6%). Darüber hinaus zeigten genomweite Sequenzierungsverfahren, die auf Plasma- und Urinproben der zweiten Patientenkohorte ähnlich niedrige Nachweisraten (39,6%). Hochauflösende, personalisierte Assays verbesserten jedoch die Detektion mit zunehmendem Erfolg (~50%). Insgesamt konnten wir zeigen, dass der ctDNA-Nachweis im Plasma bei PatientInnen mit größeren Tumoren und jenen mit venösen Tumorthromben wahrscheinlicher war. Trotz der relativ niedrigen Nachweisraten ergaben unsere Daten neue Einblicke in die Möglichkeiten der ctDNA für die minimal-invasive Beurteilung von RCC-Tumoren. Insbesondere die Analyse von > 200 longitudinalen Plasmaproben zeigte, dass die ctDNA-Dynamik in Abhängigkeit vom Krankheitsverlauf stark schwankte und daher für die Therapieüberwachung bei RCC verwendet werden kann. Bei den meisten dieser PatientInnen waren die ctDNA-Spiegel zu Beginn der Behandlung und beim Fortschreiten der Erkrankung erhöht, während diese mit dem Therapieansprechen abnahmen. Schließlich konnte gezeigt werden, dass ctDNA aus dem Plasma und erstmals auch aus dem Urin in der Lage ist, die genetische Heterogenität besser aufzuzeigen als eine einzelne Gewebebiopsie.

Zusammenfassend zeigen diese Daten, dass in Köperflüssigkeiten der betroffenen PatientInnen mit Nierenzellkarzinom nur geringe Mengen von ctDNA zu finden sind. Diese umfassenden Ergebnisse geben einen Einblick in die Biologie von Nierenzellkarzinomen, und leisten einen wichtigen Beitrag zur Realisierung der personalisieren Krebsmedizin, die als hoffnungsvolles Instrument in der Onkologie gilt.

ABSTRACT

Liquid biopsy, which refers to the minimally invasive analysis of circulating tumor markers from blood and other bodily fluids, represents an innovative tool for detecting and monitoring cancer-specific alterations. In particular, the analysis of circulating tumor DNA (ctDNA) has great clinical potential and has already been applied in many solid tumors to monitor treatment response and therapeutic effectiveness. However, the utility of ctDNA in renal cell carcinoma (RCC) is largely lagging behind. Extensive pathological and genetic heterogeneity is inherent to RCC, which confounds treatment decisions and is a significant hurdle in precision medicine.

Here, we aimed to establish the potential utility of ctDNA in the clinical management of RCC by characterizing the levels and composition of ctDNA in patient body fluids. A major focus was on those patients who received targeted therapies to validate the ability of ctDNA to track the disease course longitudinally.

Firstly, we focused on the optimization of ctDNA detection strategies. To this end, we aimed to enrich ctDNA fragments based on fragment length, which is reported to be shorter than cell-free DNA (cfDNA) from hematopoietic cells. *In silico* size selection of fragments in the size range between 90 and 150 bp improved ctDNA detection, with a more than two-fold median enrichment in particular cases. Secondly, we undertook a detailed analysis of baseline and longitudinal plasma and urine samples using a series of state-of-the-art approaches across two independent prospective clinical cohorts representing the full clinical spectrum of renal tumors. We initially applied a combination of untargeted methods and *in silico* size selection to plasma samples derived from metastatic RCC (mRCC) patients. We observed lower ctDNA levels in RCC relative to other solid cancers of similar size and stage, detecting plasma ctDNA in 32.6% of patients. The use of a high-resolution *de novo* mutation calling assay by targeting the ten commonly mutated genes in RCC did not improve detection beyond that obtained by untargeted methods (18.6%). Moreover, broad untargeted sequencing approaches applied to plasma and urine samples from the second cohort of patients indicated similarly low detection rates (39.6%). However, high-resolution personalized assays improved detection rates with incremental success (~50%). Overall, ctDNA detection in plasma was more likely amongst patients with larger tumors and in those with venous tumor thrombus. Despite the relatively low detection rates, our data revealed novel insights into the capabilities of ctDNA for minimally invasive assessment of RCC tumors. In particular, analysis of > 200 longitudinal

plasma samples revealed that ctDNA dynamics widely fluctuated in accordance with disease course and thus may be used for therapy monitoring in RCC. For most of these patients, ctDNA levels were elevated at treatment initiation and disease progression but decreased with response. Finally, further interrogation of those patients with detected ctDNA revealed that plasma and, for the first time, urine ctDNA may better represent genetic heterogeneity than a single tissue biopsy.

In summary, our findings confirm RCC as a ctDNA-low malignancy. Nevertheless, these comprehensive results provide an important insight into the biology of renal cell carcinoma and contribute to the realization of precision cancer medicine, which is considered a promising tool in oncology.

1. INTRODUCTION

1.1. Renal cell carcinoma

Malignant kidney tumors, of which more than 90% are renal cell carcinomas (RCCs), are the 7th and 10th most common cancer type in men and women, respectively (1). RCC accounts for ~2% of all cancer diagnoses and deaths worldwide, and the incidence continues to rise (2). The disease is often detected incidentally after radiological imaging, which is usually performed for a variety of non-specific symptoms and abdominal complaints (3). Despite the increased detection of early-stage renal masses that are potentially curable, approximately 25%-30% of patients present with synchronous metastatic disease (mRCC, metastatic RCC) at the time of diagnosis (4). Even after nephrectomy with curative intent for clinically localized disease, 20-40% of patients will eventually experience disease relapse and develop metastases. Although the therapeutic strategies for advanced-stage RCC patients have evolved remarkably over the last decade, they still fail to provide complete or long-term responses (5). Overall, the prognosis of patients with metastatic disease is poor, with a 5-year survival rate of <10% (6,7).

1.1.1. Disease heterogeneity

RCC arises from the renal tubular epithelial cells and represents an extremely complex group of malignancies with different histological subtypes characterized by distinct clinical courses, each having its own biological and genetic characteristics (3,8,9). According to the latest WHO (World Health Organization) histological classification, the disease comprises 16 distinct tumor subtypes (10). The clear-cell type (ccRCC) represents the most common histology of RCC, accounting for 75% of cases; papillary renal cell carcinoma (pRCC) comprises 15-20% of all RCCs and is subdivided into type 1 and type 2. Chromophobe renal cell carcinoma (chRCC), the third most frequently observed subtype, accounts for 5% of all kidney cancers (11). Clear-cell and papillary RCCs originate from proximal tubule or parietal cells, whereas chRCC arises from intercalated cells in the distal convoluted tubule (12). Minor subtypes represent less than 1% of all kidney tumors, which include medullary RCC, collecting duct RCC, MiT family translocation RCC, and oncocytomas (3).

1.1.2. Mutational landscape, essential genes, and pathways

1.1.2.1. Clear-cell RCC

Large-scale sequencing studies have enabled the identification of recurrent cancer loci that characterize RCC tumorigenesis, including *VHL*, the mutated gene underlying the von Hippel-Lindau disease. Chromosome 3p loss is a near-ubiquitous early event in the development of ccRCC (>90% patients) (13,14). It is responsible for the loss of heterozygosity of four well-known tumor suppressor genes (TSGs), i.e., *VHL*, *PBRM1*, *SETD2*, and *BAP1* (3,7,14,15). The biallelic inactivation of the *VHL* gene, through non-synonymous mutations (60%–70%) or epigenetic silencing (5%–10%; promoter methylation) represents the most frequent mutational event seen in ccRCC (13-15). Disruption of the *VHL* complex results in the accumulation of the transcription factors hypoxia-inducible factor (HIF) 1α and 2α, with subsequent downstream activation of HIF-dependent factors, including vascular endothelial growth factor (VEGF), platelet-derived growth factors (PDGFs), and other genes involved in angiogenesis, tumor cell survival, and cell proliferation (16,17). Further alterations in genes encoding other parts of the *VHL* complex, such as transcription elongation factor B (*TCEB1*), are found in 1-5% of cases and are associated with distinct clinical courses. Alterations in *TCEB1*, along with the loss of chromosome 8, are mutually exclusive with *VHL* and 3p loss. These findings illustrate the importance of the VHL-HIF pathway in ccRCC and represent an important target for therapeutic agents (8).

Nevertheless, *VHL* loss alone is insufficient for ccRCC tumorigenesis and additional prevalent mutations have been discovered in chromatin- and histone remodeling genes, including *PBRM1* (30-40%), *SETD2* (~12%), *BAP1* (~10%), and *KDM5C* (~5%) (13-15,18,19). On the backbone of 3p loss, mutations result in a complete loss of function of the TSGs *PBRM1*, *SETD2*, and *BAP1*. *BAP1* mutations are typically mutually exclusive with *PRBM1* or *SETD2* alterations, whereas alterations in *PRBM1* and *SETD2* have the tendency for clonal co-occurrence (12,20). RCC patients whose tumors carry *BAP1* mutations have been characterized by an aggressive disease and shorter survival compared to those with a *PBRM1* alteration (21). Other recently published studies have identified additional pathways and components that are recurrently mutated in RCC, including the PI3K–AKT–mTOR signaling pathway, i.e., *PIK3CA*, *MTOR*, *TSC1/2*, and *PTEN* (3,9,14,15). In a study by Sato et al. (15), 26% of patients had mutations that involved this pathway, which provides a rationale for the effectiveness of mTOR inhibitors in RCC.

1.1.2.2. Non-clear cell RCC

In addition to their common tissue of origin, pRCCs and ccRCCs share some molecular features. While alterations in the *MET* oncogene are enriched in familial and sporadic forms (13%) of type I pRCC (22,23), pRCC type II tumors are associated with *CDKN2A* silencing, *TERT* mutations, and alterations of the NRF2/antioxidant response element (ARE) pathway. Type II tumors are also characterized by mutations in *BAP1*, *SETD2*, and *PRBM1*, albeit with and without chromosomal loss of 3p (3,9,22).

Compared to ccRCC and pRCC tumors, chRCCs are associated with a relatively low somatic mutational burden. The most prevalent mutations associated with this tumor subtype include two TSGs, i.e., *TP53* (32%) on chromosome 17 and *PTEN* (9%) on chromosome 10 (24). In particular, *PTEN*-mutated tumors are associated with poor survival (11).

1.1.3. Somatic copy number alterations

Besides mutations, somatic copy number alterations (SCNAs) comprise a significant driving force of the disease (Figure 1).

A hallmark event present in >90% of ccRCCs is the chromosome 3p loss, which is mainly caused by an unbalanced translocation between chromosome 3 and chromosome 5 with chromothripsis (in 43% of samples), typically resulting in chromosome 3p loss and 5p gain. Thus, chromosome 5p gain represents the second-most frequent SCNA event, found in 65-70% of these patients (13,14,25). The arm-level loss of chromosome 3p occurs in a handful of cells many years to decades before diagnosis, mainly during childhood or adolescence. In contrast, the biallelic inactivation of the *VHL* gene usually happens much later in life (13). Additional recurrent focal and arm-level copy number alterations are associated with ccRCC, including chromosomal losses of chromosomes 1p, 4q, 6q, 8p, 9p, and 14q, and gains on chromosomes 7q, 8q, and 12p (14).

Type I pRCC tumors are associated with nearly universal copy number events, including the gain of chromosomes 7 (where *MET* is located) and 17, while type II tumors are characterized by frequent loss of chromosome 9p (22).

Despite the overall low mutational burden, chRCC tumors are characterized by a high frequency of heterozygous chromosomal losses, including chromosomes 1, 2, 6, 10, 13, 17, and 21 (80% of all cases). On the backbone of *PTEN* and *TP53* mutations, chromosomal losses of 10 and 17 lead to complete loss of function of these genes (24).

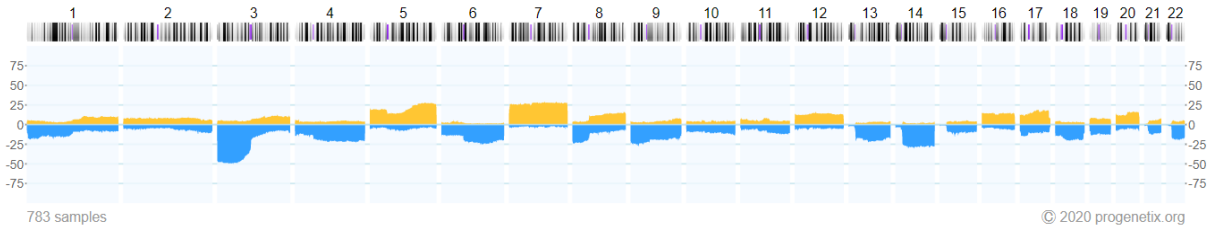


Figure 1: Frequency of RCC-specific somatic copy number alterations.
Taken from the Progenetix database (26) (CC BY 4.0).

1.1.4. Treatment landscape of RCC

Early diagnosis and surgical removal are still the only long-term curative strategy for patients with RCC (5). While nephrectomy and subsequent follow-up to track relapse is the recommended strategy for patients with localized disease (1), the management of advanced or metastatic RCC patients is more challenging. This disease is characterized by distinct clinical courses, which can range from indolent to rapidly progressing, lethal disease (16). Therefore, clinical nomograms such as the Memorial Sloan Kettering Cancer Center (MSKCC) system (27,28) and the International Metastatic Renal Cell Carcinoma Database Consortium (IMDC) model (29) have been developed to stratify patients according to their prognosis and guide decision-making (1,9).

It is widely known that RCCs are a group of chemotherapy-resistant diseases (3). In contrast to other solid cancers, surgery plays a central role in the management of patients with mRCC (30). For instance, complete resection of single-site lung metastasis or adrenal metastasis is associated with an impressive 5-year overall survival rate of around 40% or 60%, respectively (5,31). Previous systemic treatment options were limited to cytokine-based therapies (i.e., interferon-alpha and interleukin-2), which have been replaced by therapies with higher response rates (RR) and longer progression-free (PFS) and/or overall survival (OS) (32). Over the last decade, knowledge of the tumor biology and specific molecular pathways that are altered in RCC has dramatically expanded the available therapeutic armamentarium. The treatment landscape now includes targeted therapeutics, immune-checkpoint inhibitors, and combination strategies (Figure 2) (9,12).

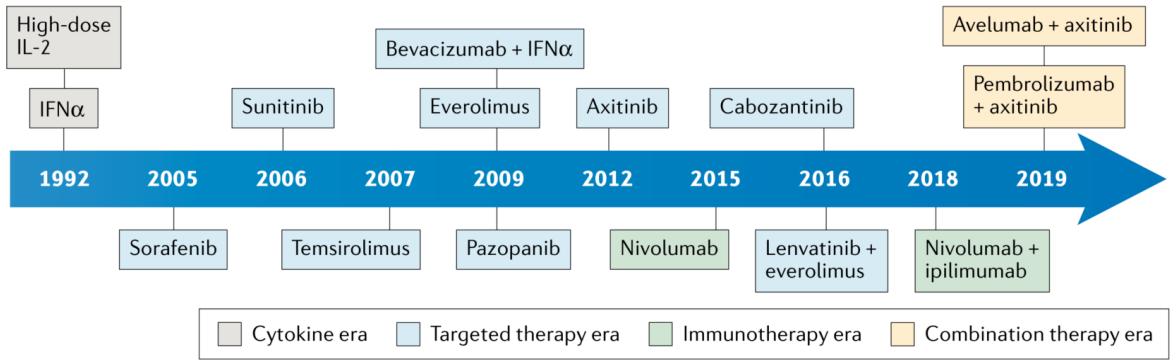


Figure 2: Evolution of the treatment landscape for patients with metastatic RCC.
Reprinted from Dizman et al. (12) with permission from Springer Nature (license number 4886051328019).

Due to the highly vascular nature of RCC tumors (33), agents targeting pro-angiogenic factors such as VEGF remain the preferred frontline option. These drugs are blocking either the circulating VEGF using monoclonal antibodies (e.g., bevacizumab in combination with interferon-alpha) or the tyrosine kinase activity of the VEGF receptor (VEGFR) using tyrosine kinase inhibitors (TKIs, e.g., cabozantinib, pazopanib, sorafenib, and sunitinib) (1,7,32). However, selected patients, in particular those with the poor-risk advanced disease, may instead benefit from the inhibition of the mechanistic target of rapamycin (mTOR) pathway using the mTOR inhibitor temsirolimus (34). Following one prior anti-angiogenetic therapy, the TKI axitinib or the mTOR inhibitor everolimus, either as monotherapy or in combination with a VEGFR inhibitor (i.e., lenvatinib), are also approved treatment options (1). Most of the TKIs are multi-kinase inhibitors with targets that include VEGFR1-3 and fibroblast growth factor receptors (FGFR, e.g., lenvatinib), or platelet-derived growth factor receptors (PDGFR, e.g., axitinib, sunitinib, pazopanib), or MET and AXL (e.g., cabozantinib) (7,12). As increased expression and activation of the tyrosine kinases MET and AXL has been implicated in the clinical resistance to anti-VEGF therapies, cabozantinib demonstrated significant improvement in PFS, OS, and RR after progression on initial VEGF-targeted agents (35,36).

New treatment strategies include the immune-checkpoint inhibitors that block programmed cell death protein 1 (PD-1) (e.g., nivolumab, pembrolizumab), programmed cell death 1 ligand 1 (PD-L1) (e.g., avelumab) or cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) (e.g., ipilimumab). These therapeutics administered either alone or in combination with other drugs such as VEGF-targeted therapies (e.g., axitinib and pembrolizumab) will further revolutionize the cancer care of mRCC patients (9,12,17). For instance, the combination of ipilimumab and

nivolumab has already demonstrated clinical efficacy, with significantly improved OS as well as overall RR over sunitinib among patients with untreated intermediate- and poor-risk RCC disease (37). Therefore, it has emerged as the preferred treatment option in this patient population (1).

Despite significant improvement in clinical outcomes through targeted therapies, primary or acquired resistance is seen in the vast majority of patients (38). For instance, 30% of all patients showed primary resistance to sunitinib, and the remaining 70% who initially respond will develop acquired resistance within 6-15 months (39). The mechanisms of VEGF resistance are multifactorial, including restoration of angiogenesis through the activation of VEGF-independent pathways and reduced bioavailability either through increased efflux (40) or lysosomal sequestration of TKIs (41). It is suggested that tumors activate an "angiogenic switch," leading to either upregulation of the existing VEGF pathway or activation of alternative non-VEGF pathways responsible for tumor revascularization (40). Similar to TKIs, inhibition of mTORC1 with mTOR-inhibitors are generally not durable, and all patients will eventually experience disease progression while on treatment (42).

1.1.4.1. Challenges of targeted therapies

Clinical outcomes of many tumor types might be improved by the understanding and detection of the molecular events that underlie their pathogenesis (43). Targeted therapies tailored to the genetic composition of tumors are administered. Currently, tumor biopsies are still the gold standard for categorizing tumors and informing clinical decision-making. Despite significant improvement in clinical outcomes through targeted therapies, several new issues and challenges have been introduced, including costs, lack of effective drugs against most genomic alterations, and technical limitations of molecular tests. Most significantly, intratumor heterogeneity (ITH), which describes the presence of genetically distinct subclones within tumors, has long been recognized to play a central role in cancer progression, treatment failure, and drug resistance (31,38,44,45).

Gerlinger et al. (46) initially reported that ITH might be a pervasive feature of ccRCC. Multi-region sequencing of primary ccRCC tumors and the associated metastatic sites revealed that *VHL* and chromosome 3p loss were the only ubiquitous (clonal) driver events. Conversely, other driver mutations, including alterations in *TP53,* in the PI3K–AKT–mTOR signaling pathway (*PIK3CA, PTEN*, *mTOR*) and in the histone-modifying enzymes and chromatin

remodeling complexes (*SETD2*, *BAP1*, *KDM5C*) were subclonal, i.e., these alterations were only present in some regions of the tumor (47,48). Overall, only the minority of somatic mutations were present across all primary tumor regions, whereas up to 63-69% of all nonsynonymous mutations were not detectable across all biopsy specimens (46). These data highlighted that a single tumor-biopsy specimen is not sufficient to capture the full spectrum of genetic diversity and that ITH hampers the validation of oncology biomarkers owing to sampling bias (46,47), and furthermore, contributes to intrinsic resistance by positive selection of preexisting drug-resistant subclones (44). In addition to existing heterogeneity, tumors evolve in space and time. Notably, under the selective pressure of sequential lines of therapy, subclones with acquired resistance may arise *de novo* during disease progression, which may result in differences in the proportion and pattern of specific alterations between the primary tumor and metastases or relapses (44). Furthermore, it has been shown that metastatic sites were significantly more homogeneous and were associated with fewer driver somatic alterations compared to primary tumors (49). However, in most cases, no information about the genetic make-up of metastases is available (45).

Due to the issues mentioned above, the interpretation of tumor biopsies that are often not accessible may be confounded by ITH (46) and therefore are potentially sub-optimally to inform treatment decisions (50). ITH should be evaluated in every aspect of the patient's course of the disease, primarily to provide a tailored therapy that best matches the molecular alterations of the patient’s tumor (45).

1.1.5. The clinical need for molecular profiling with a particular focus on therapy monitoring in mRCC patients

The lack of validated biomarkers for RCC in clinical practice poses several challenges for management of the disease. These include the following: early detection; histological classification of RCC subtypes, i.e., chRCC from clear-cell type or benign oncocytomas; and early diagnosis of minimal residual disease following surgery with curative intent, which would enable effective stratification of patients for adjuvant therapy trials (51). At advanced stages, especially in the context of metastatic disease, useful predictive markers for targeted therapy selection, prognosis or treatment response, and resistance monitoring have not yet been established (4). Although serial sampling of the tumor genome to monitor treatment response should represent a prerequisite for personalized therapy, it is currently almost impossible since consecutive biopsies are associated with patient burden due to the invasiveness of the

procedure. Therefore, the therapeutic transition to another drug is solely based on unacceptable toxicity and progressive disease based on clinical parameters, i.e., radiological imaging and/or blood measurements (52), which are often not very accurate. The conventional follow-up evaluation strategy to determine response and resistance to therapy is mostly based on computed tomography (CT) or magnet resonance imaging (MRI) (1). Although drug efficacy is mainly assessed by RECIST (Response Evaluation Criteria in Solid Tumours) guidelines (53), these criteria remain debatable (1).

For this reason, there is an increased desire to use molecular stratification to forecast the course of individual renal cancers and to unravel the diversity of outcome. Most significantly, methods are clearly required for rapid, cost-effective, and minimally invasive means to predict and monitor the response to targeted therapies throughout the disease. To this end, liquid biopsies (see section 1.2) are emerging as a minimally invasive alternative to traditional tumor sampling techniques (51).

1.2. Circulating tumor DNA and liquid biopsy

Liquid biopsies refer to the analyses of tumor-derived material such as circulating tumor DNA (ctDNA), circulating tumor cells (CTCs), and tumor-derived extracellular vesicles, which enable tumor genome characterization by minimally invasive means, i.e., a blood draw (54-59). Besides peripheral blood, other biological fluids such as urine and saliva are also suitable for the analysis of circulating tumor-derived analytes (59,60). In particular, the field of cell-free DNA (cfDNA), which in patients with cancer contains ctDNA, has made significant progress over the past decade and is now evolving into a widely used prognostic and predictive biomarker (see recent reviews (57,58)).

ctDNA can be released from multiple tumor lesions into the bloodstream. Hence, liquid biopsies might reflect both ITH and spatially separated metastatic deposits and thus potentially overcome the limitations of conventional tissue biopsies (57,59,61). Recent studies reported the utility of ctDNA as a reliable biomarker to detect minimal residual disease and to predict early relapse (62-65). Also, due to the short half-life, ctDNA provides a real-time snapshot of the entire tumor burden (66,67) and can be used for serial monitoring the treatment response. Previous studies reported that ctDNA dynamics correlate with the patient's response, and more importantly, progression can be detected before it is clinically obvious (68-72). Liquid biopsies also find application in the monitoring of evolutionary dynamics of a patient's tumor and in the

detection of resistance mechanisms that inevitably arise for almost all targeted therapies across all tumor entities (73-76). In particular, the potentiality of ctDNA to elucidate actionable targets (77,78) could considerably support clinical decision-making in precision medicine.

More recently, the field of liquid biopsy is moving forward into the early diagnosis of pre-symptomatic patients (79,80), i.e., shifting patient management to early cancer detection, where treatments are more likely to be successful.

1.2.1. Origin and fragmentation patterns of cell-free DNA

The understanding of the origin and the mechanisms of release is not necessarily complete. Previous studies strongly indicate that the mechanisms of cfDNA release into the bloodstream include apoptosis, necrosis, and active secretion (Figure 3A) (81-84).

However, necrotic cell death causes DNA of varying sizes, including high-molecular weight DNA (>10 kb), which is rarely seen in plasma samples. Thus, the relative contribution of necrosis to cfDNA release is still unknown (81). Alternatively, living non-neoplastic and tumor cells can actively secrete extracellular vesicles (e.g., exosomes) that contain nucleic acids into the bloodstream (Figure 3) (85,86). For example, Kahlert et al. (87) observed that exosome-associated DNA (> 10 kb) reflects the genomic landscape and mutational status of the parental tumor cells and could therefore serve as a promising disease biomarker.

Genome-wide sequencing approaches have revealed that the fragmentation pattern of cfDNA derived from pregnant women was reminiscent of nuclease-cleaved nucleosomes (88), as the modal distribution of cfDNA fragment lengths has a modal value near 166 bp, which corresponds to the DNA wrapped around a nucleosome (~147 bp) plus a linker region (~20 bp) (88,89). These data suggest that cfDNA molecules are mainly associated with the release of DNA from dying, apoptotic cells after enzymatic processing (90). However, the size of ctDNA fragments has been a matter of debate for a long time, with studies reporting contradictory evidence for increased or decreased integrity of ctDNA in plasma from cancer patients (91,92). For example, our group observed that the fragment size distribution of plasma DNA in metastasized cancer patients was associated with multiples of 166 bp (i.e., oligonucleosomes) (93). Conversely, other studies analyzing fragment size distribution of plasma DNA from cancer patients and healthy controls indicated that ctDNA fragments are even more fragmented than cfDNA molecules (Figure 3B), with a modal distribution between 132 and 145 bp (94,95). These observations are consistent with the data of Jiang et al. (91), who reported that shorter

plasma DNA molecules derived from hepatocellular carcinoma patients preferentially carry the tumor-associated copy number alterations. In addition, it was found that overall size profiles correlate with the amount of tumor DNA in plasma, i.e., shorter DNA fragments were enriched in patients with high ctDNA fractions (91).

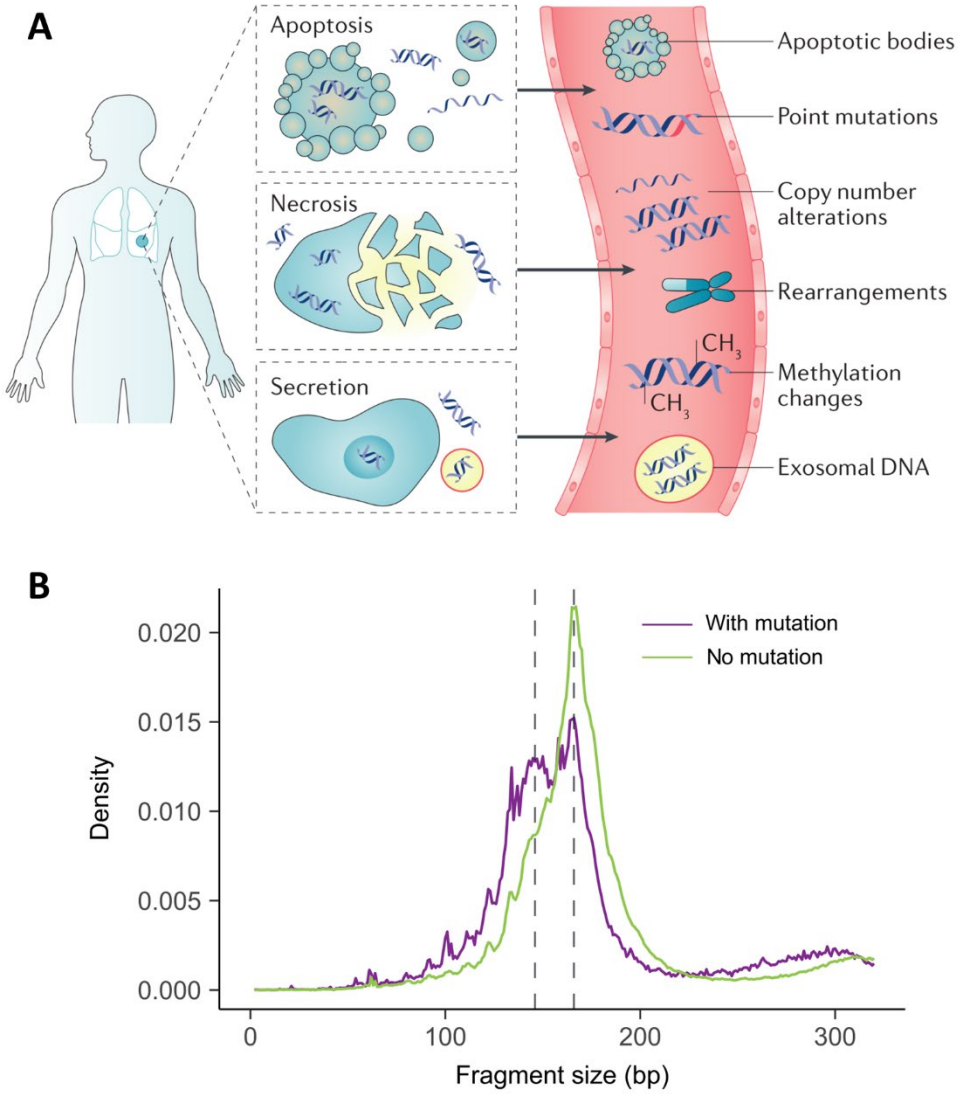


Figure 3: Origin and fragmentation patterns of circulating tumor DNA.

(A) Circulating tumor DNA can be released into the blood by various mechanisms, including apoptosis, necrosis, and secretion. Taken from Wan et al. (57). Reprinted with permission from Springer Nature (license number 4886041080426).

(B) The size distribution of cell-free (purple) and circulating tumor DNA (green). Adapted from Mouliere et al. (96). Reprinted with permission from AAAS (license number 4886050962104).

1.2.2. Detection technologies for ctDNA analysis

It has been shown that the proportion of cfDNA that is tumor-derived has a variable contribution ranging from <0.1% to >10% of the total cfDNA concentration (67). Despite the assumption that ctDNA levels correlate with the metabolic tumor volume (63,68), occasionally patients with late-stage cancers harbored very low ctDNA levels in their circulation (97,98). Bettegowda et al. (67) showed that the fraction of tumor-derived DNA can vary quite substantially and that, in addition to the stage of the disease, the cancer type also affects the ctDNA concentration in the circulation.

Overall, the low ctDNA fraction within the high background signal of cfDNA derived from cells of the hematopoietic lineage represents one of the major challenges for liquid biopsy-based assays (99). Owing to technological advances, targeted high-resolution and untargeted approaches are now available, which differ mainly in their analytical sensitivity and genomic coverage for the analysis of cell-free analytes (Figure 4) (100,101).

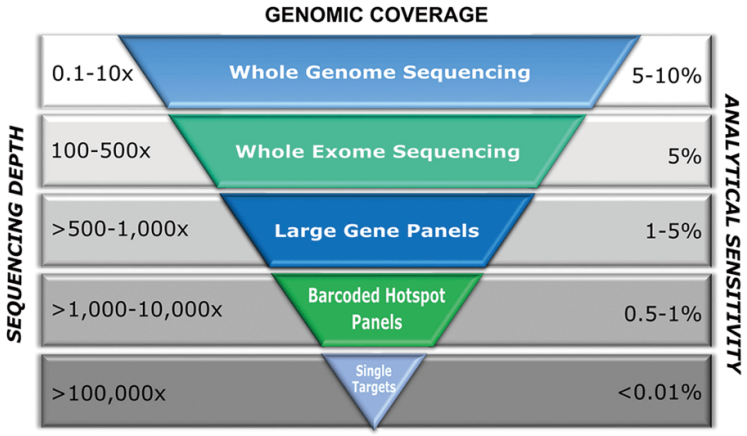


Figure 4: Methods of ctDNA detection.
The schema illustrates the genomic coverage and the achieved analytical sensitivity of the approaches. In short, increasing genomic coverages comes with decreasing analytical sensitivity. Reprinted from Zhou & Moser et al. (100) (CC BY-NC-ND 4.0).

The majority of ctDNA-based assays thus far have focused on the detection of cancer-specific somatic single nucleotide variants (SNVs), SCNAs, and rearrangements (102). In addition to these genetic alterations, cfDNA molecules exhibit epigenetic marks (e.g., methylation patterns and fragment size) from its cell of origin and prove fruitful to detect tumor-derived molecules (79,96,102,103) (see 1.2.2.3).

1.2.2.1. Targeted approaches

The first strategy involves targeted sequencing technologies that are tailored to specific genes and genomic loci of interest and is used to quantify the mutant allele frequency (mAF) of mutations in the plasma of cancer patients (99). Only the emergence of massively parallel next-generation sequencing technologies (NGS) such as amplicon-based sequencing (69) or hybrid capture sequencing (104,105) revealed the full potential of mutation-based analysis and enabled the simultaneous analysis of a larger number of cancer-specific mutations with high analytical sensitivity (102). Current strategies apply molecular barcodes (also called unique molecular identifier (UMIs)) to reduce background errors that occur during library preparation and/or sequencing, thereby improving the detection of low-frequency variants (57,58). Besides, previous data have demonstrated that tumor-informed patient-specific analysis, which includes prior tumor genotyping and custom panel design, further improves the sensitivity of ctDNA-based assays (63,106). Owing to technological progress, the analytical sensitivity of ctDNA detection improved dramatically in recent years, and today, targeted sequencing technologies allow the identification of mAFs far below 1% (101).

1.2.2.2. Untargeted approaches

Whole-genome sequencing (WGS) of plasma DNA provides a comprehensive view of the genomic landscape and enables the analysis of somatic copy number alterations (SCNAs) and structural variants (99). Heitzer et al. (93) developed a genome-wide approach called plasma-Seq, which allows the evaluation of SCNAs and therapeutically relevant focal amplifications of plasma-derived cfDNA using WGS with low coverage (0.1–0.2x) (75,90,107). Similar approaches such as tMAD (trimmed median absolute deviation from copy number neutrality) metric (96) and the ichorCNA algorithm (108) enable the quantification of copy number alterations and estimation of ctDNA levels. Moreover, whole-exome sequencing (WES) has been demonstrated to be useful to identify *de novo* mutations and potentially clinically actionable targets for tracking tumor evolution throughout treatment (73). Overall, untargeted sequencing approaches do not depend on *a priori* information of the tumor genome and best capture tumor heterogeneity. Nevertheless, drawbacks of genome-wide profiling of ctDNA compared to targeted methods include the presence of higher tumor fraction in the patients' blood, i.e., >3-10%, in order to provide informative results (100).

1.2.2.3. Beyond SCNAs and SNVs

Beyond genetic features, the size distribution of plasma DNA is an essential accessible parameter to enrich for tumor-derived molecules, given the size differences between cfDNA and ctDNA (shorter than 150 bp) (57,109) (Figure 3B). Since the analysis of cfDNA fragmentation patterns has attracted considerable attention, this area of research is referred to as "fragmentomics" (110). Currently, extensive efforts are being made to target short plasma DNA molecules. Recently, it became clear that the cfDNA fragmentation patterns are biased by the sequencing library preparation applied prior to sequencing (102). Developed to capture highly degraded DNA derived from ancient specimens, recent single-stranded DNA (ssDNA) library preparation approaches enable the recovery of both single-stranded DNA and double-stranded DNA, as well as shorter and damaged DNA fragments (111,112). Therefore, ssDNA library preparation might bypass the limits of the conventional double-stranded DNA library preparation method that tends to recover mostly non-degraded fragments greater than 100 bp in length (112). Since previous studies have indeed shown that ssDNA libraries could recover smaller DNA fragments in plasma samples collected from transplant recipients (112) and healthy individuals (113), in this study, we tested their use in cancer patients. An alternative to the enrichment of small fragment sizes *in vitro* is to select sequence reads within particular DNA fragment sizes *in silico* (96). Size selection of specific fragment sizes *in silico* (e.g., 90-150 bp) have enabled improved ctDNA detection, with a more than two-fold median enrichment of tumor fraction in selected cases (96).

In summary, fragment-specific analyses of plasma DNA seem to be promising. However, more work is clearly required for further implementation.

1.2.3. ctDNA in renal cancer

It has been previously published that liquid biopsy-based assays show great promise in various tumor entities, such as breast (66), colorectal (74), lung (63), ovarian (114) and prostate cancer (75). However, the clinical utility of ctDNA in the management of RCC is largely lagging behind. To date, there are only limited and often contradictory reports available investigating the quantitative information (i.e., the reflection of tumor burden) and qualitative alterations (i.e., genetic information) (59) of ctDNA in the locally advanced and metastatic RCC setting. Initial data (67,115) have suggested very low ctDNA levels and/or detection rates in this malignancy (51). In contrast, two extensive studies identified genomic alterations in cfDNA in about 80%

of patients with advanced RCC (98,116). Likewise, the utility of ctDNA-based assays in early-stage kidney cancers seems to be contradictory (79,117) and remains to be established. Furthermore, previous reports have suggested that ctDNA in urine may be informative, especially in genitourinary cancers (118,119). However, there is still no such study yet that has examined the presence of urine ctDNA in RCC patients (51).

Overall, there is still an unmet clinical need to provide insight into ctDNA levels, composition, and potential clinical utility of liquid biopsy-based approaches in the management of renal cancers of differing stage and subtype (51).

2. AIMS

The major aim of this thesis was the comprehensive characterization of ctDNA in liquid biopsy fluids of patients with RCC. Firstly, we focused on the optimization of ctDNA detection strategies on the basis of the biological features of plasma DNA, given the difference in size between cfDNA and ctDNA. To this end, we compared different library preparation protocols (i.e., single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA) library protocols) for the recovery of varying plasma DNA size fractions. In addition, we applied *in silico* size selection, which may provide substantial improvements in ctDNA detection. Secondly, we used several state-of-the-art sequencing approaches to provide insight into the presence, levels, composition, and potential clinical utility of ctDNA in mRCC patient management, i.e., predicting and monitoring response to targeted therapies. Thirdly, we teamed up with the Rosenfeld group at the University of Cambridge and validated our initial observations with an independent prospective clinical cohort, which additionally enabled us to determine the genetic heterogeneity and potential applications of ctDNA from plasma and urine.

To this end, the following aspects were investigated in particular:

- Evaluation of the performance of ssDNA and dsDNA library preparation protocols for ctDNA detection from plasma DNA
- Assessment of the utility of *in silico* size selection of fragments between 90 and 150 bp from shallow whole-genome sequencing (sWGS) to enrich ctDNA
- Assessment of ctDNA levels and detection rates in plasma and urine of patients with RCC using a series of well-established approaches across two independent prospective clinical cohorts
- Longitudinal monitoring of mRCC patients to validate the ability of ctDNA to track the disease course
- Assessment of intratumoral heterogeneity in liquid biopsy fluids

3. MATERIAL AND METHODS

3.1. Ethics statement

The MonReC (Monitoring Renal Cancer) study was approved by the Ethics Committee of the Medical University of Graz, Austria (approval number 27-210 ex 14/15) and by the Ethics Committee of the Military Institute of Medicine, Warsaw, Poland (approval number 33/WIM/2015) and conducted in accordance with the Declaration of Helsinki. DIAMOND patients were recruited according to the local ethical guidelines (REC ID 03/018). Written informed consent was obtained from all patients (51). Further studies which included other solid cancer and control subjects were approved by the Ethics Committee of the Medical University of Graz, Austria (approval number 21-229 ex 09/10, 21-228 ex 09/10, and 21-227 ex 09/10 for the collection and analysis of colorectal, prostate, and breast cancer patients, and 26-288 ex 14/14 for the study involving healthy controls).

3.2. Evaluation of fragment-specific analyses

3.2.1. Patient cohort

To evaluate the genomic differences and performance of double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA) libraries, we included plasma samples with a high tumor content derived from other ongoing studies. To this end, we prepared dsDNA and ssDNA libraries from a set of six plasma samples derived from patients with late-stage metastatic breast (B7_2, B13_1), colorectal (CRC; C177_2, C177_3), and prostate (P226_4, P226_6) cancer (120).

In addition, we applied *in silico* selection of sequence reads within particular DNA fragment size ranges to existing low pass WGS data sets of RCC patients (n=177 plasma samples) and an independent cohort of CRC patients (n=26 plasma samples). As a control set, we included 22 individuals free of malignant disease. All cancer patients (except for RCC patients, see page 22) were recruited and treated at the Department of Internal Medicine, Division of Oncology at the Medical University of Graz.

3.2.2. Sample collection and processing

For each patient, 9ml of whole blood was collected into BD Vacutainer EDTA tubes containing 10% NBF (BD Biosciences, Franklin Lakes, USA) or Streck tubes. For plasma collection, samples were centrifuged at 200g for 10 minutes with 0 brake and acceleration powers,

followed by an additional centrifugation of 1,600g for 10 minutes. Plasma was transferred into a new 15ml tube and centrifuged at 1,600g for 10 minutes. All aliquots were stored at -80°C before cell-free DNA isolation.

3.2.3. Cell-free DNA isolation

Plasma DNA was isolated using the QIAamp Circulating Nucleic Acid Kit (Qiagen, Hilden, Germany) from 2ml plasma according to the manufacturer's protocol. Following extraction, eluted plasma DNA concentration was quantified using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Vienna, Austria).

3.2.4. Double-stranded DNA sequencing library preparation

Shallow-whole genome sequencing libraries from plasma samples were prepared and sequenced as previously described (74,75,93,107). Briefly, shotgun libraries were generated from 5 to 10ng of plasma DNA using the TruSeq DNA Nano Sample Preparation Kit (Illumina, San Diego, USA) according to the manufacturer's instructions but with several modifications. Since plasma DNA is already highly fragmented, we omitted the fragmentation step. In addition, 25 PCR cycles were used for the selective amplification of the libraries. The Bioanalyzer DNA 7500 Kit (Agilent Technologies, Santa Clara, USA) was used to determine the fragment size distribution. Final libraries were pooled in an equimolar fashion, and 75 bp paired-end sequencing was performed on the Illumina MiSeq or NextSeq platform (Illumina, San Diego, USA) (51). On average, we obtained ~10 million reads per sample, which accounts for 0.1-0.2× coverage of the whole genome.

3.2.5. Single-stranded DNA sequencing library preparation

In this study, we compared our well-established TruSeq Nano Library protocol (Illumina, San Diego, USA) with a single-stranded DNA library preparation protocol (Figure 5) published by Burnham et al. (112). This method was adapted to the protocol described in Nature Protocols by Gansauge and Meyer (111).

Figure 5: Schematic diagram of single-stranded and double-stranded preparation protocol.
In contrast to dsDNA library preparation, in which multiple adapters are ligated to dsDNA fragments, ssDNA library preparation first generates ssDNA fragments to which a biotinylated linker (Btn) is attached. Ligation of the first and second adaptor is carried out while DNA is bound to streptavidin beads.

In brief, DNA fragments were dephosphorylated and denatured by mixing 2X CircLigase II buffer (Epicentre, London, UK), MnCl2, and 1U FastAP (Thermo Fisher Scientific, Waltham, USA) with 10ng plasma DNA in 43µl reaction volume. The reaction was incubated at 37°C for 10 minutes, followed by an incubation time of 2 minutes at 95°C, and was immediately quenched in an ice batch. Fragments were then ligated to a biotinylated adapter oligonucleotide CL78 (Table 1) using CircLigase II and 50% PEG-4000 solution (wt/v) (Sigma-Aldrich/Merck KGaA, Darmstadt, Germany) for a sample volume of 80µl, and was incubated for 1 hour at 60°C. Stop solution (0.5M EDTA, Tween-20; Sigma-Aldrich/Merck KGaA, Darmstadt, Germany) was added, and the ligation product was incubated for 1 minute at 95°C. After two

washing steps with the bead binding buffer (5M NaCl, 1M Tris-HCl [pH 8], 0.5M EDTA, Tween-20, and SDS [20% wt/vol]), Dynabeads MyOne Streptavidin C1 beads (Thermo Fisher Scientific, Waltham, USA) was resuspended using the same buffer. Adapter-ligated fragments were immobilized on streptavidin beads by rotating for at least 20 minutes at room temperature. Collected beads were washed with wash buffer A (5M NaCl, 1M Tris-HCl [pH 8], 0.5M EDTA, Tween-20, SDS [20% wt/vol]) and once with wash buffer B (5M NaCl, 1M Tris-HCl [pH 8], 0.5M EDTA, Tween-20). For the primer annealing and extension, beads were mixed with 10X Isothermal Amplification Buffer (New England Biolabs, Ipswich, USA), dNTPs (10mM each; Thermo Fisher Scientific, Waltham, USA), and oligonucleotide CL9 (Table 1) for 2 minutes at 65°C. After incubation for 1 minute in an ice bath, 24U Bst 2.0 DNA polymerase (New England Biolabs, Ipswich, USA) was added, and the reaction was incubated from 15°C to 37°C by ramping the temperature up at 1°C/minute, and incubation was held at 37°C for 5 minutes. Beads were washed once with wash buffer A and incubated with stringency wash buffer (SDS [20% wt/vol], 20×SSC buffer) for 3 minutes at 45°C. After these steps, beads were once washed with wash buffer B. After clean-up, beads were combined with 10× Buffer Tango (Thermo Fisher Scientific, Waltham, USA), dNTP mix (10mM each), Tween 20 (1%), and 5U T4 DNA polymerase (Thermo Fisher Scientific, Waltham, USA), and incubated by gentle shaking for 15 minutes at room temperature. EDTA (0.5 M) was added to each reaction mixture, and beads were rewashed with each wash buffer A, stringency wash buffer, and wash buffer B, as described above. A double-stranded adapter was prepared by mixing oligonucleotide CL53, oligonucleotide CL73 (both Table 1), TE (pH 8), and 5M NaCl, and the reaction mix was incubated for 10 seconds at 95°C, and decreasing the temperature to 14°C at a rate of 1°C/second. After the removal of 3′ overhangs, a second adapter (double-stranded adapter, see above) was added by blunt-end ligation (1-hour incubation at room temperature with gentle shaking) with T4 DNA ligase (Thermo Fisher Scientific, Waltham, USA), T4 DNA ligase buffer, PEG-4000 solution and Tween-20. After washing steps as previously described (wash buffer A, stringency wash buffer, and wash buffer B), beads were resuspended with TET buffer (1M Tris-HCl [pH 8], 0.5M EDTA, Tween-20), and library molecules were eluted from the beads by heat denaturation (95°C for 1 minute).

Library amplification and indexing were performed with 11 PCR cycles, followed by one purification step with AMPure XP beads (Beckman Coulter, Brea, USA). Libraries were quantified by qPCR, and the quality was assessed using a Bioanalyzer High Sensitivity Kit

(Agilent Technologies, Santa Clara, USA). Final libraries were sequenced on an Illumina MiSeq platform (Illumina, San Diego, USA) in a 76 bp paired-end run. On average, we obtained ~10 million reads per sample.

Table 1: Synthetic oligos used for single-stranded sequencing library preparation.

Name	Description	Sequence (5'-3')
CL9	Extension primer	GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC T*N*N* N*N
CL53	Double-stranded adapter 1	ACA CGA CGC TCT TC/3ddC/
CL57	Double-stranded adapter 2	/5Phos/GGA AGA GCG TCG TGT AGG GAA AGA G*T*G* T*A*
CL78	Single-stranded adapter	/5Phos/AGA TCG GAA GTT TTT TTT TT/3BioTEG/

*, phosphorothioate bond, /ddC/, dideoxycytidine, /5Phos/, 5' phosphorylation /3BioTEG/, 3' biotin-TEG

3.2.6. Somatic copy number analysis and tumor fraction estimation

Single-stranded and double-stranded sWGS data were analyzed using an in-house script, i.e., plasma-Seq pipeline as described previously (75,90,93,121). Genome-wide copy number plots were generated. Estimation of tumor fraction from cfDNA and additional visualization of copy number changes were performed using the publicly available ichorCNA algorithm, which uses a probabilistic model for the simultaneous prediction of large-scale CNAs and quantification of tumor content (108). Since this method has a lower limit of tumor detection (LOD) of 0.03 (108), all samples with tumor fractions below this detection threshold were considered to be ctDNA-negative.

3.2.7. *In silico* size selection

In order to enrich for tumor-derived signal, we performed *in silico* size selection for sequencing reads of 90-150 bp in length as previously published by Mouliere et al. (96). To quantify the impact of *in silico* size selection, we compared SCNA profiles generated with and without size selection.

Briefly, the paired-end sequence reads were aligned to the human hg19 genome using bwa-mem. PCR duplicates were marked and removed using MarkDuplicates (122). The fragment lengths were obtained by the mapping locations of the alignments. Samtools software was used

to select reads, which correspond to fragment lengths in a specific size range (90-150 bp) (96). To assess the effect of size selection and the accompanying increase in noise on the calling of SCNAs, we generated subsampled BAM files to retain an equivalent number of sequencing reads of samples after size selection. We used Picard's DownsampleSam function to downsample a given BAM file to a specific proportion of reads. Finally, we analyzed the original, downsampled, and *in silico* size selected data using the plasma-Seq pipeline (75,90) or the ichorCNA algorithm (108). SCNA profiles generated from these data sets were compared. In particular, samples with less than 2 million sequencing reads were only classified as ctDNA-positive if RCC- or CRC-specific SCNAs were found in the ProGenetix database (26).

3.2.8. Fragment length analysis

We analyzed the fragment size distribution of plasma DNA samples using paired-end sequencing sWGS data. After alignment to the human hg19 genome, insert sizes were extracted from the BAM file using Picard's CollectInsertSizeMetrics function (122). Insert size distributions were plotted in R (version 4.0.2).

3.3. Comprehensive characterization of ctDNA in patients with renal tumors

3.3.1. Study design

Here, we aimed to present the most comprehensive assessment of ctDNA in plasma of patients with mRCC (MonReC cohort). A primary focus was on those patients who received targeted therapies to validate the ability of ctDNA to track the disease course longitudinally.

To strengthen the impact of our initial observation, we collaborated with the research group of Nitzan Rosenfeld, Christopher G. Smith (both Cancer Research UK Cambridge Institute, Cambridge, UK), and Grant D. Stewart (Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK), who had recruited 48 patients with a full range of renal subtypes and masses from benign to locally advanced patients (DIAMOND cohort). This cooperation made it possible to analyze ctDNA using a series of well-established approaches across two independent prospective clinical cohorts (Figure 6) and further enabled the determination of genetic heterogeneity and potential applications of ctDNA from plasma and urine.

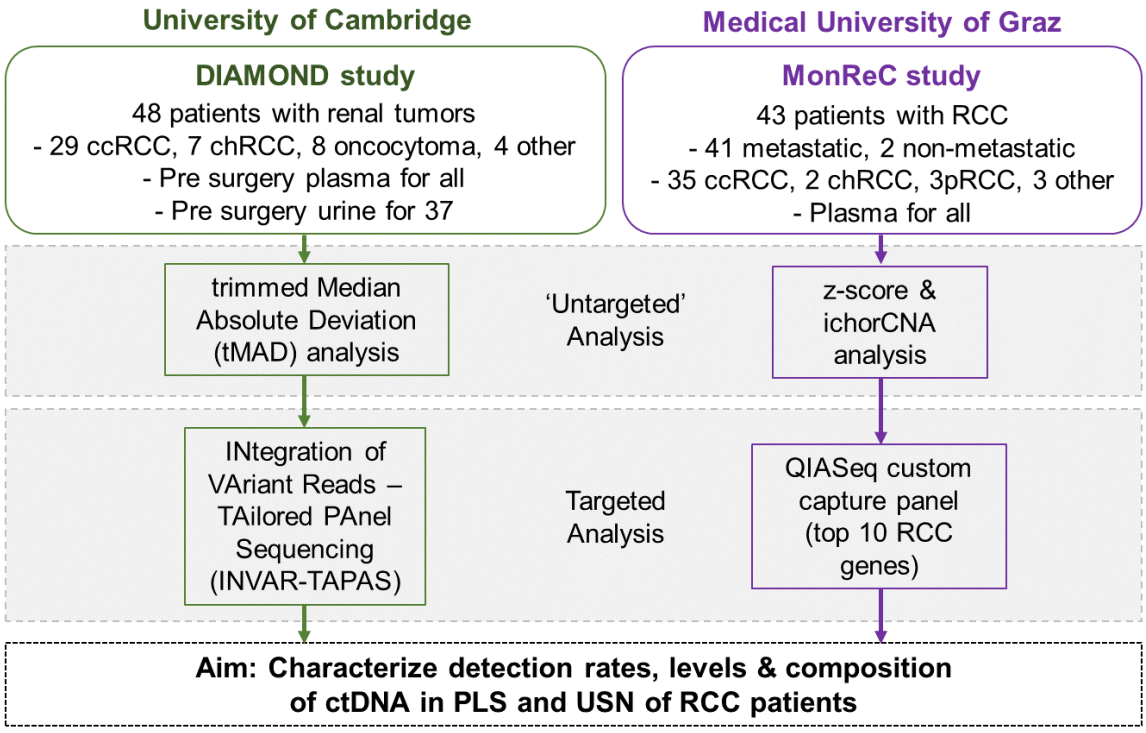


Figure 6: ctDNA analysis pipeline.

[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

For MonReC (right panel), we applied a combination of modified Fast Aneuploidy Screening Test-Sequencing System (mFAST-SeqS) and ichorCNA analysis of shallow whole-genome sequencing (sWGS) to infer tumor fraction. For DIAMOND (left panel), we utilized the trimmed Median Absolute Deviation (tMAD) score of shallow whole-genome sequencing (sWGS) data for ctDNA detection. Targeted methods, including Qiagen's QIASeq custom capture panel and a novel analysis algorithm, termed INtegration of VAriant Reads from TArgeted PAnel Sequencing (INVAR-TAPAS), were subsequently utilized.

3.3.2. MonReC patient cohort

At our site, 43 patients with a range of tumor stages and subtypes, including clear-cell, chromophobe, and papillary RCC, were recruited into the MonReC study. At the time of the first blood draw, 41 patients were diagnosed with metastasized RCC, and two patients had localized cancer (2/43). The majority of patients (35/43 patients) had their primary tumor resected through radical nephrectomy. 18/43 RCC patients (mean age 62.5 years, range 46-81 years) were recruited and treated at the Department of Urology or the Division of Oncology, Department of Internal Medicine, at the Medical University of Graz (Graz, Austria). 25/43 RCC patients (mean age 58.9 years, range 41-68 years) were recruited and treated at the Department of Oncology at Military Institute of Medicine (Warsaw, Poland) (51). We analyzed patients'

plasma samples using state-of-the-art sequencing technologies (Figure 6). Patient characteristics are outlined in Figure 7 and Supplementary Table 1.

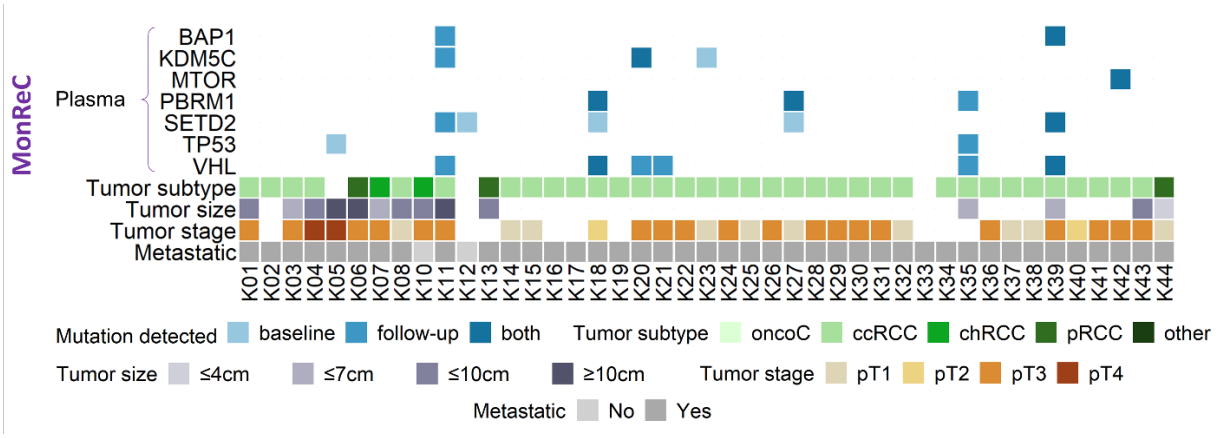


Figure 7: MonReC patient's characteristics and tumor genomic profile.

[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

Plasma samples were collected from 41 patients with mRCC and two with localized RCC (total n=43). The plasma mutation status using QIASeq of frequently mutated RCC genes is shown at the top. White cubes indicate that no mutation in these genes was observed. Blue, medium blue, and dark blue cubes display that a mutation was detected at baseline, during follow-up, or at both time points, respectively. Shown in descending order are tumor -subtype, -size, and metastatic at the time of sampling.

3.3.3. MonReC sample collection, processing, and plasma isolation

For this study, peripheral blood samples were collected at multiple time points, i.e., at first diagnosis of metastases, during several lines of treatment, and/or when the disease progressed along with the introduction of a new line of treatment (mean time between first blood draw and surgery 48 months, range 0-269). In total, we collected 252 blood samples from 43 RCC patients. Within this study, 9 ml of blood were collected into BD Vacutainer® EDTA tubes containing 10% NBF (BD Biosciences, Franklin Lakes, USA) or Streck tubes. Blood samples collected at the Medical University of Graz (n=18) were processed at the Institute of Human Genetics. Plasma samples (204/252 samples) collected from patients recruited at the Department of Oncology at the Military Institute of Medicine (Warsaw, Poland) were stored at -80°C before shipping to Graz. Plasma DNA was extracted from 2 ml of plasma using the QIAamp Circulating Nucleic Acid Kit (Qiagen, Hilden, Germany) (51).

3.3.4. MonReC sample pre-screening

In order to assess the tumor content in patients samples prior to higher-resolution genome-wide and targeted approaches, we analyzed all plasma samples (n=252) using our mFAST-SeqS

approach (123). Briefly, this method is based on the selective amplification of *LINE-1* sequences throughout the genome. Amplicon libraries were prepared as previously described (123), using 0.5 - 1ng of plasma DNA. *LINE-1* libraries were sequenced on the Illumina MiSeq or NextSeq platform (Illumina, San Diego, USA) to generate a minimum of 100,000 single-end reads for each amplicon library. Z-scores were calculated by subtracting the mean and dividing by the standard deviation of normalized read counts for the respective chromosome arm from healthy controls (n=35) to identify over- and underrepresentation of each chromosome arm (51).

3.3.5. MonReC shallow whole-genome sequencing and ichorCNA analysis

sWGS libraries were prepared and sequenced from plasma samples using our conventional dsDNA method (93). For detailed information, see section 3.2.4. Libraries were sequenced in paired-end mode (2x 75 bp) on either an Illumina MiSeq or NextSeq instrument (Illumina, San Diego, USA). On average, we obtained 31.3 million reads per sample (range 10.2 – 41.9 M reads). We additionally applied *in silico* size selection to enrich for ctDNA signal (see 3.2.7). All plasma samples were analyzed by sWGS to map SCNAs and to quantify tumor fraction using the ichorCNA algorithm (108). To improve the parameter estimation in samples with low tumor fractions, we used the ichorCNA algorithm with optimized settings (124).

3.3.6. MonReC mutation profile generation using QIAseq custom panel

To assess the mutational spectrum of genes that are frequently mutated in RCC, we performed high-resolution mutation profiling using the customized QIAseq Targeted DNA Panel (CDHS-11685Z-538; Qiagen, Hilden, Germany). This panel was designed to cover the full coding sequence of ten genes (*BAP1*, *KDM5C*, *MET*, *MTOR*, *PBRM1*, *PIK3CA*, *PTEN*, *SETD2*, *TP53*, *VHL*) involved in RCC tumorigenesis and progression. Library preparation was performed using 10ng input DNA according to the QIAseq Targeted DNA Panel Handbook (R2). The workflow includes the following steps: DNA is first enzymatically fragmented, end-repaired, and A-tailed. During a ligation process, each DNA molecule is assigned an Illumina-specific adapter harboring a unique molecular barcode (UMI) and sample index. Post-UMI assignment, genomic targets are enriched by a single pool of gene-specific primers and one universal primer. The library was amplified, and Illumina-specific adapter sequences and additional samples indices were added within a universal PCR using 18 PCR cycles. The Bioanalyzer High Sensitivity DNA Kit (Agilent, Santa Clara, USA) was used to determine the quality and

fragment sizes from all libraries. Further, we quantified the libraries using the QIAseq Library Quant Assay (QIAGEN). Libraries were pooled in equimolar amounts and sequenced on an Illumina NextSeq platform (Illumina, San Diego, USA) to generate 150 bq paired-end reads. We obtained an average amount of 3.47 million reads (ranging from 1.33 to 5.7 M) per sample (51).

3.3.6.1. MonReC mutation calling

Raw sequencing data were analyzed using the cloud-based QIAseq data analysis pipeline, which utilizes a barcode-aware variant caller (smCounter V1) to distinguish between true variants from sequencing artifacts (125). We only considered variants which passed the predefined quality control from the smCounter pipeline. In addition, we excluded synonymous variants and common polymorphisms, which were characterized as variants present with minor allele frequencies of >1% in population frequency databases (ExAC, TOPMED, 1000 Genomes, GnomAD). To improve the precision of mutation calling, we prepared and analyzed all plasma samples in duplicate and only variants called in both replicates were considered as true variants. All called variants were visually inspected using the Integrative Genomics Viewer (IGV) (version 2.3.58) (51).

3.3.7. MonReC establishment of the limit of detection for QIAseq custom panel

We evaluated the assay performance using a commercially available reference material (Seraseq ctDNA reference material v2; SeraCare Life Sciences, Milford, USA), which contains multiple mutations (n=40) at defined variant allele frequencies (VAFs). The plasma-like material includes a wild-type sample and an additional five samples with VAFs of 0.125%, 0.25%, 0.5%, 1%, and 2%. The QIASeq Custom panel covered three mutations in *PIK3CA*, two variants in *PTEN,* and five mutations in the *TP53* gene. To assess the sensitivity of our panel, we prepared and sequenced all samples in duplicates. On average, 3.4 million reads (range 3.0 – 3.8 million) per sample were obtained. The assay enabled detection rates of 90% and 100% for 0.02 mAF, 70% for 0.01, and 30% for 5.0×10^{-3}. None of the variants could be detected at a VAF of 2.5×10^{-3}. The QIAseq panel failed to identify 8 and 10 out of 10 variants at 1.25×10^{-3} VAF. Moreover, the wild type sample clearly showed a negative result. The maximum achievable limit of detection (LOD) was set to 5.0×10^{-3} (Supplementary Figure 1). In order to increase variant calling stringency, we analyzed all plasma samples in duplicate and

only considered variants to be true if detected in both replicates and if they had a VAF higher than the LOD (51).

3.3.8. DIAMOND- authors contributions

Since the results of the DIAMOND cohort are briefly discussed in this thesis, the methods used are briefly described here. Christopher G. Smith and colleagues performed the experiments and data collection presented in this section. In addition, the dry lab, including tMAD analysis, mutation calling on WES data, and INVAR, were also carried out by the group members of the Rosenfeld's lab. The data was compiled by Christopher G. Smith, Nitzan Rosenfeld, and Grant D. Stewart.

3.3.9. DIAMOND patient cohort

The DIAMOND cohort included 48 patients with the full ranges of renal subtypes and masses from benign to locally advanced patients. In total, 29 patients with stage I (11/29), stage II (1/29) and stage III (17/29) ccRCC, seven patients with stage I (2/7), stage II (2/7) and stage III (3/7) chrRCC, eight patients with oncocytomas, one patient with stage III pRCC, one patient with stage II MiT family translocation RCC and two patients with oncocytic renal neoplasm (51). General clinical characteristics for each patient are outlined in Figure 8.

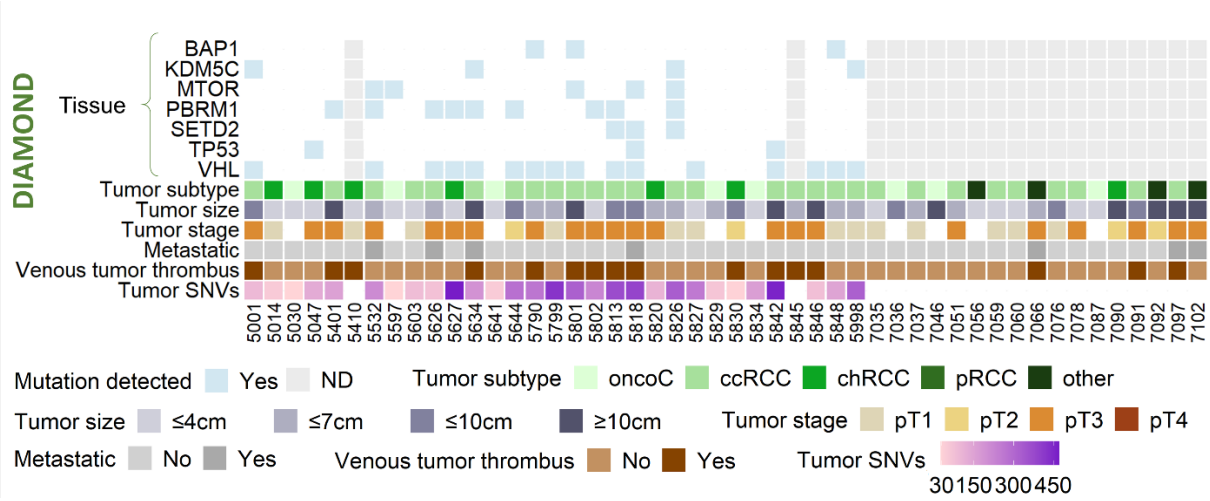


Figure 8: Patient characteristics and tumor genomic profile of the DIAMOND cohort.

[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

For the DIAMOND cohort, plasma (n = 48) and urine (n = 37) were collected from patients with a range of tumor subtypes and stages. The tumor tissue mutation status of frequently mutated RCC genes is shown at the top. White cubes indicate that no mutation in these particular genes was observed. Grey cubes display that no tumor tissue was available. Shown in descending order are tumor subtype, tumor size, tumor stage, metastatic at baseline,

evidence of venous tumor thrombus, and the number of tumor SNVs (necessary for the tumor-informed INVAR-TAPAS assay).

3.3.10. DIAMOND tumor collection and extraction

DIAMOND patients underwent nephrectomy (partial or total) as part of curative treatment or cytoreductive surgery. Tumor tissue samples (n=29) were obtained from either fresh frozen (FF) or formalin-fixed paraffin-embedded (FFPE) material. FF samples were extracted from a tissue < 20mg, and DNA was isolated using a DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany). DNA extraction from FFPE samples, which were obtained from 2-mm-diameter and 3-mm^3-deep cores, was performed using the GeneRead DNA FFPE kit (Qiagen, Hilden, Germany), as recommended by the manufacturer's instructions. Only one modification was made; instead of the 1 hour incubation at 56°C, the incubation was performed overnight. DNA concentrations of FFPE and FF DNA was quantified using the PheraStar FSX platform (BMG LabTech, Allmendgrün, Germany) (51).

3.3.11. DIAMOND body fluid collection, processing, and isolation

Plasma was collected from all DIAMOND patients (n=48) prior to surgery (n=37 patients; mean 8.9, range 0-35 days pre-surgery). Within this cohort, the approaches for sample collection and processing varied. For samples collected prior to April 2016 (21/32 patients), plasma from 8ml peripheral blood, collected in EDTA tubes, was extracted using one centrifugation step at 4,000rpm for 20 minutes. For samples collected after April 2016 (11/32 patients), 12ml blood were collected and samples were centrifuged at 1,600g for 10 minutes with an additional centrifugation of 13,300rpm for 10 minutes. Plasma and buffy coat were transferred into separate tubes and stored at -80°C before DNA isolation (51).

For a subset of DIAMOND patients (n=37), urine samples were collected prior to surgery (mean 8.9, range 0–35 days pre-surgery). 30-50ml urine was collected into a 50ml falcon tube containing 0.5M EDTA and centrifuged at 2,400g for 10 minutes. Urine supernatant (USN) was stored at -80°C prior to analysis. In addition, the urine cell pellet (UCP) was collected by resuspending the pellet with 1ml supernatant and an additional centrifugation step at 13,300rpm for 10 minutes. The dry pellet was stored at -80°C before DNA extraction. Cell-free DNA and buffy coat samples were extracted from bodily fluid samples using the QIAsymphony platform (Qiagen, Hilden, Germany) (51).

3.3.12. DIAMOND tumor DNA sequencing library preparation

Whole exome sequencing (WES) of all tumor specimens and germline buffy coat DNA samples was performed to detect patient-specific somatic variants. Briefly, 50ng of DNA was sheared using the Covaris LE 220 (Covaris, Woburn, USA) according to the manufacturer's protocol. Sequencing libraries were prepared using the ThruPLEX DNA-seq kit (Rubicon Genomics, Ann Arbor, USA), and 5x PCR cycles during library preparation were used. Library quantification was performed by qPCR using the KAPA library quantification kits (Roche, Basel, Swiss), and the library size was checked using a Bioanalyzer or a Tapestation (Agilent Technologies, Santa Clara, USA). After library preparation, exome capture was carried out using the TruSeq Exome Library protocol (Illumina, San Diego, USA) with the use of i5 and i7 TruSeq HT xGen universal blocking oligos (IDT) during the hybridization steps. After capture, 8 PCR cycles were carried out. Multiplexed libraries were sequencing using an Illumina platform (HiSeq 4000) (51).

3.3.13. DIAMOND tumor mutation calling

Mutation calling was performed on genomic regions with a >20x depth in the matching sequencing data from the buffy coat using the CallableLoci tool from GATK (126). Mutect2 (127) was used for somatic mutation calling. The filter settings are described in detail in the corresponding publication (51). Also, we filtered variants present with minor allele frequencies (AFs) of >0.02 in the 1000 Genomes project global database and excluded variants with an AF >0 observed in normal adjacent tissue samples. Moreover, we observed a large number of C:G>A:T mutations in FFPE-based tissue data, often preceded by C or T and G or A, respectively. Thus, we filtered all C:G>A:T calls in order to suppress artifacts. As a result, we developed patient-specific SNV lists for a subset of patients with renal tumors. On average, 297.36 unique SNVs per patient were identified. These mutation lists were used to design the custom capture sequencing panels (51).

3.3.14. DIAMOND body fluid shallow whole-genome sequencing

Shallow whole-genome sequencing libraries were prepared using all tissue (described above) and body fluid samples, i.e., plasma and urine (USNA and UCP). Indexed sWGS libraries were prepared from 15-30 ng plasma, and USN derived DNA using the ThruPLEX Tag-Seq kits (Rubicon Genomics, Ann Arbor, USA). The number of PCR amplification cycles varied between 8 and 11 cycles, dependent on DNA input. UCP DNA was fragmented using acoustic

shearing before library preparation. Library concentration was determined by qPCR. WGS libraries prepared from tissue and all biofluids were pooled for each sample type equimolarly and sequenced in a paired-end mode (150bp) using an Illumina HiSeq 4000. On average, 16.4 million reads per sample were obtained (51).

3.3.14.1. DIAMOND trimmed median absolute deviation (tMAD) analysis

WGS data were analyzed using a bioinformatics pipeline published previously by Mouliere et al. (96). Tissue samples were analyzed for CNAs in R using the QDNAseq pipeline (128).

For all body fluids, CNAs were quantified from sWGS data using the CNAclinic pipeline (129), as described by Mouliere et al. (96). The trimmed median absolute deviation (tMAD) from the copy number neutrality was calculated (130). Briefly, we randomly downsampled sequencing reads to 10 million reads per dataset and performed the analysis using a bin size of 30 kbp. Read counts in plasma and USN cases were normalized by the read counts from an identically analyzed and processed cohort of healthy control samples, and the median absolute deviation from log2R = 0 of segmented bins was quantified. To set the detection threshold, we calculated the tMAD score for sWGS data from 46 healthy controls and used the maximal value (median=0.01, range 0.004–0.015). The sensitivity of the tMAD metric was set at 0.3% based on previous determinations (96). The tMAD metric was also applied to UCP samples for the first time. Due to unavailable matched healthy control samples, normalization of this dataset was carried out using the samples' own mean logR, and ctDNA positivity was determined if a signal that deviated from the copy number neutral state was detected (51).

In silico size selection of sequencing reads within a particular fragment size range was performed to enrich for tumor-derived cfDNA fragments. Briefly, sequencing BAM files were downsampled (~2 million reads), and tMAD analysis was carried out using a bin size of 500 kb. Samples with <2 million reads were considered unsuitable for tMAD analysis. Normalization was performed with a similarly downsampled cohort of healthy controls (51).

3.3.15. DIAMOND Tailored Panel Sequencing (TAPAS)- custom hybrid-capture panel design and plasma sequencing

Following mutation calling from WES data, custom personalized capture panels (2.077-Mb; 57,306 probes) were designed using the Agilent SureDesign software (Agilent Technologies, Santa Clara, USA). In addition to the introduction of SNVs detected by WES, we also included

open reading frames of RCC-specific oncogenes and tumor suppressor genes, see corresponding publication (51).

ThruPLEX libraries generated from plasma and USN derived DNA (see above) were hybridized in a 3-plex (333.3ng each; 1000ng total) using the SureSelect XTHS panel (Agilent Technologies, Santa Clara, USA), with the addition of i5 and i7 specific blockers (IDT), as recommended by the manufacturer's instructions. We used 13 PCR cycles for post-capture library amplification. Library fragment sizes were checked, and library concentrations were then quantified using the KAPA library quantification kit. Quantified libraries were sequenced on an Illumina HiSeq4000 platform, generating paired-end 150 bp reads. For each sample, we aimed to obtain at least 30 million sequencing reads (51).

3.3.15.1. DIAMOND INVAR-TAPAS pipeline

ThruPlex Tag-seq library BAM files were collapsed with the CONNOR tool (131) using a minimum family size of 2 and a consensus frequency threshold of 90%.

All customized capture sequencing data were analyzed using the INtegration of VAriant Reads - TAilored PAnel Sequencing (INVAR-TAPAS) pipeline, published by Wan and colleagues (106). Briefly, patient-specific mutation lists are generated by tumor genotyping, which were subsequently used to design personalized hybrid-capture sequencing panels. This pipeline combines custom error-suppression approaches with the use of UMIs and consideration of mutation classes, signal-enrichment methods based on biological features of tumor-derived fragments, and tumor mutant allele fraction to reduce background noise and to enrich for the mutant signal. An INVAR score per sample is generated by aggregating the signal across all loci, considering all parameters mentioned above. An INVAR score threshold is determined based on sequencing data derived from patients using non-matched mutation lists. Finally, to quantify the ctDNA signal across multiple mutated loci, a global mutant allele fraction (gmAF) is obtained by calculating a background-subtracted, depth-weighted mean allele fraction across the patient-specific loci. The sensitivity of this assay relies on the total number of informative reads (IR) that cover the relevant loci. The ability to detect ctDNA in samples with less than 20,000 IR was limited. The limit of detection of this approach was set to 1.0×10^{-4} mAF, i.e., samples with ctDNA levels lower than this threshold would need to be re-analyzed (51).

3.3.16. DIAMOND association between ctDNA discovery and clinicopathological features

Abdominal CT scans were used to determine the maximum diameter of the primary tumor. In addition, cross-section imaging was used to determine whether the patients had advanced local disease, i.e., an extension of a tumor thrombus into the renal vein or inferior vena cava. To assess the cellular proliferation rate, immunohistochemical (IHC) staining of FFPE specimens was performed utilizing an anti-Ki67 monoclonal antibody (MIB-1 clone at 1:100 dilution; DAKO Agilent Technologies LDA) (51).

3.3.17. DIAMOND representation of intratumor heterogeneity in body fluids

Tumor heterogeneity, which is a characteristic feature of RCC (46), was assessed by multi-tumor region profiling. For DIAMOND patient 5842, ten spatially distinct tumor regions were obtained after surgery. WES on all tissue samples and subsequent mutation calling was performed as previously described. mAF for each site was extracted by the SAMtools (version 1.3.1) (126) function mpileup. The heatmap function in R was used to visualize the dataset. Hierarchical clustering was done by variants (columns) and not by samples (rows) according to Euclidean distance (51).

3.4. Data analysis

Statistical analyses and plotting of data were performed using R (version 4.0.2) and GraphPad Prism 6.0 (GraphPad Software, San Diego, USA). A p-value of < 0.050 was considered statistically significant.

4. RESULTS

4.1. ctDNA detection can be improved by fragment-specific analysis

4.1.1. Single-stranded DNA library preparation does not preferentially enrich short ctDNA

Recently, it was demonstrated that different library preparation methods significantly affect cfDNA yield and fragment sizes (111,112). In this study, we addressed whether the single-stranded DNA (ssDNA) library preparation approach indeed results in a substantial recovery of short plasma DNA fragments and whether it can achieve a higher ctDNA fraction content compared to the conventional double-stranded DNA (dsDNA) approach in which such shorter fragments are lost through size selection. To compare our conventional dsDNA library protocol (93) with an ssDNA adapter-ligation method (112) (Figure 5), both types of libraries were prepared from plasma samples derived from six patients with metastatic breast (B7_2, B13_1), colorectal (C177_2, C177_3), and prostate (P226_4, P226_6) cancer (120). First, we tested for differences in fragment length distribution by paired-end sWGS and observed distinct DNA size profiles from dsDNA and ssDNA libraries (Figure 9A) and substantial enrichment of small DNA fragments (≤ 150 bp) after ssDNA library preparation ($p < 0.050$, paired Wilcoxon test; Figure 9B). In contrast to dsDNA libraries, ssDNA libraries showed a plateau in their ascending slope toward the peak at 166 bp; this was caused by short DNA species (≤ 150 bp). Short DNA fragments represented a mean of 9.9% (range 6.7%-13.9%) and 46.5% (range 40.5%-56.4%) of total DNA in dsDNA and ssDNA, respectively (120). Moreover, dsDNA did not exhibit the 10 bp periodicity typically found in ssDNA libraries. While the conventional library preparation recovered di-nucleosomal fragments in the range of 320 bp, these fragments were less prevalent in the ssDNA libraries.

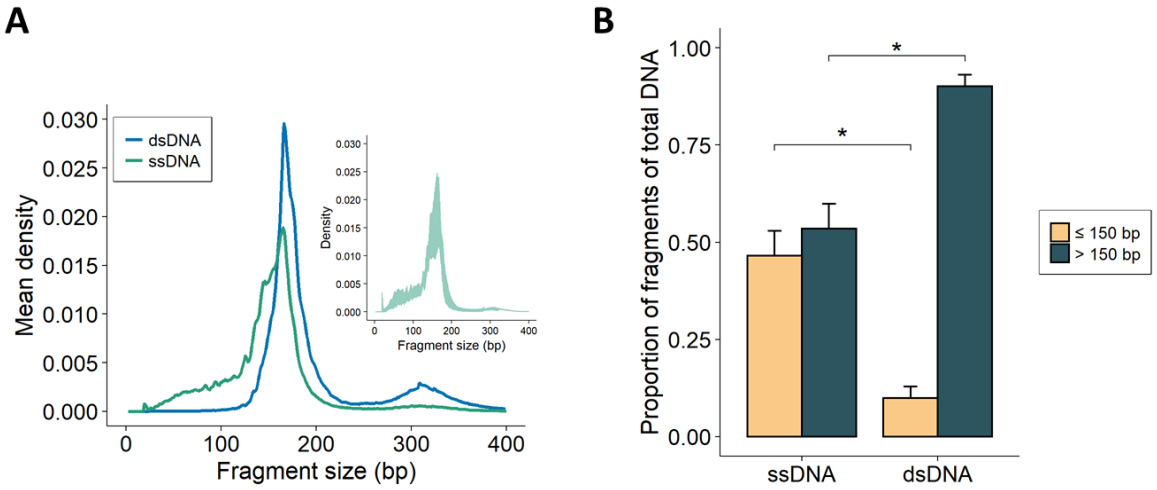


Figure 9: Fragmentation profiles and genomic characteristics of dsDNA and ssDNA libraries.

(A) Plasma DNA fragment length distribution. Mean density plot of plasma DNA from six cancer patients measured by paired-end sequencing after dsDNA (blue) and ssDNA library preparation (green). The sample-to-sample variability is shown in the small panel.

(B) Improved recovery of small plasma DNA fragments after ssDNA library preparation. Comparison between proportions of DNA fragments of specific fragment lengths, i.e., ≤ 150 bp and > 150 bp, revealed that ssDNA libraries significantly enriched smaller fragments (paired Wilcoxon test, $p < 0.050$) versus dsDNA libraries.

Since recent data suggested that tumor-derived fragments are shorter than those from normal cells (96) and ssDNA libraries showed greater efficiency in the recovery of small DNA species, we expected an improved tumor-derived signal. However, when we compared the respective tumor fractions calculated with ichorCNA, tumor levels were all within a narrow range (Supplementary Table 2; [dsDNA ichorCNA tumor fraction/ssDNA] B7_2: 0.2967/0.2857; B13_1: 0.8069/0.7966; C177_2: 0.0834/0.1084; C177_3: 0.4802/0.4678; P226_4: 0.0730/0.0754; P226_6: 0.2564/0.2482). Moreover, genome-wide copy number profiles generated from dsDNA and ssDNA libraries were nearly consistent within all libraries with no significant differences (Figure 10). In contrast, the extent of SCNAs between dsDNA and ssDNA libraries was significantly correlated ([log2 ratios dsDNA/ssDNA] B7_2: R=0.95; B13_1: R=0.92; C177_2: R=0.91; C177_3: R=0.99; P226_4: R=0.98; P226_6: R=0.95; Pearson's R, each $p < 2.2e\text{-}16$).

Overall, we confirmed that ssDNA libraries significantly recovered smaller plasma DNA size fractions; however, we did not find a preferential enrichment of ctDNA (120). Thus, these

findings were contradictory to our expectations, and we concluded that ssDNA libraries did not have any advantage over our typically employed dsDNA approach.

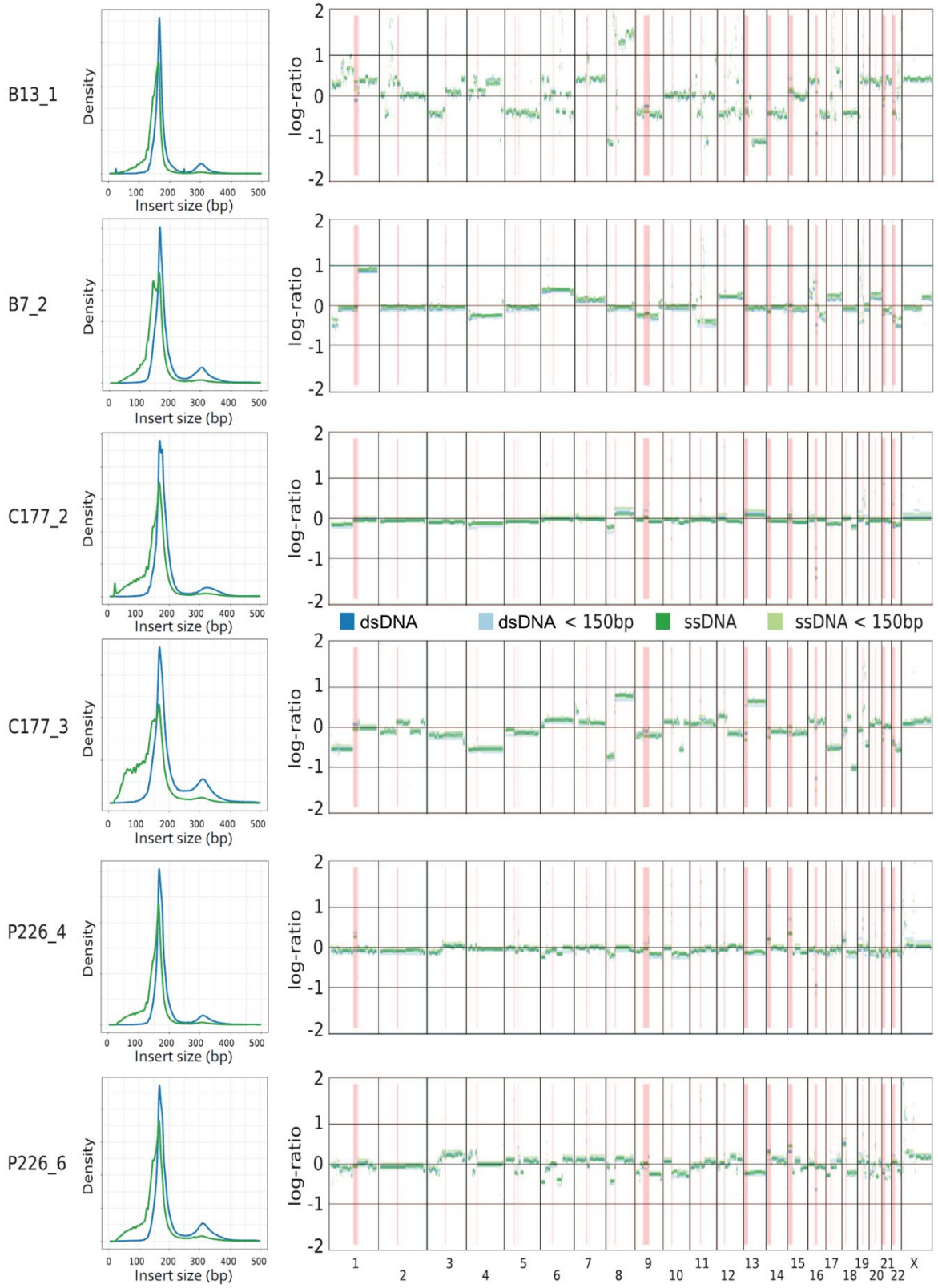


Figure 10: Comparison of dsDNA and ssDNA library preparation protocols.
[Figure and legend adapted from Clinical Chemistry (Moser et al., 2017)]
As depicted in Figure 9, DNA size profiles for both dsDNA (blue) and ssDNA (green) library preparations are shown for each patient next to the respective genome-wide copy number profiles for the dsDNA (blue) and ssDNA (green) library preparations. In addition, we performed *in silico* analysis of DNA fragments with a size of < 150 bp (light blue for dsDNA and light green for ssDNA; Supplementary Table 2 and section 4.1.2). Known problematic genome regions (> 1 Mb) that are affected by mappability issues and therefore prone to deviations in read counts as previously described (128) are highlighted in red. The x-axis defines the chromosomes; the y-axis shows log2 copy number ratios.

4.1.2. *In silico* size selection preferentially enriches short ctDNA

An alternative strategy to the enrichment of small DNA molecules by different sequencing library preparation methods is to *in silico* select specific fragment size ranges (96). We applied *in silico* selection of sequencing reads within particular DNA fragment size ranges (90-150 bp) to sWGS data sets (all dsDNA libraries) generated from RCC patients, a cancer type in which low ctDNA levels in the blood circulation is suspected (67). Analysis of SCNAs and the respective tumor fractions was carried out using ichorCNA (108) on all plasma samples after and without size selection. A lower limit of 0.03 ctDNA tumor fraction was previously established (108) and therefore samples with tumor fractions below that threshold were considered ctDNA-negative. Out of 177 plasma samples, only 16 samples (9.0%) showed RCC-specific SCNAs, such as loss of chromosome 3p or gain of 8q. *In silico* size selection further improved the detection rate and revealed 19 additional ctDNA-positive samples (35/177, 19.8%) (Figure 11 & Figure 12). On average, the ichorCNA tumor fraction increased 2.2-fold (range 0.9-6.8) from a median of 0.01 (range 0-0.18) to a median of 0.07 (range 0.03-0.23). Figure 12 shows two exemplary RCC cases with no detectable and low ctDNA; however, after *in silico* size selection, the tumor fraction significantly increased and a variety of SCNAs, including those typical for RCC, e.g., loss of 3p, gain of 5p and 7p or gain of 8p, could be observed.

Since fragments smaller than 150 bp comprised only a small fraction of the total reads (median 7.5%, range 2.9-26.6), size selection led to a significantly reduced number of sequencing reads, which in turn resulted in increased background noise. To confirm that the elevation of tumor fractions did not arise as a result of noise introduced by the lower read count, we randomly downsampled the initial sequence data to retain the same amount of reads after size selection (Figure 12 & Figure 13). We additionally carried out a comparison of sWGS data from healthy

control samples (Figure 13B) and again confirmed that size-based selection preferentially enriched ctDNA (51). While the proportion of smaller fragments (90-150 bp) in controls was substantially lower than in RCC (3.8% versus 7.5% of total reads), which subsequently led to higher background noise, none of the control samples harbored the RCC-specific SCNAs.

Finally, we aimed to validate these findings in an independent sWGS data set of CRC patients, and so we applied the same approach to existing data sets from CRC patients. We observed that the data set from CRC patients largely showed the same result; small fragments comprised 9.7% of the total reads and size selection increased the tumor fractions for nearly all plasma samples, with an average increase of 1.4-fold (range 0.4-2.6) (Figure 11B). One exemplary CRC case is depicted in Figure 13, where *in silico* size selection enriched the tumor-specific signal 2.6-fold.

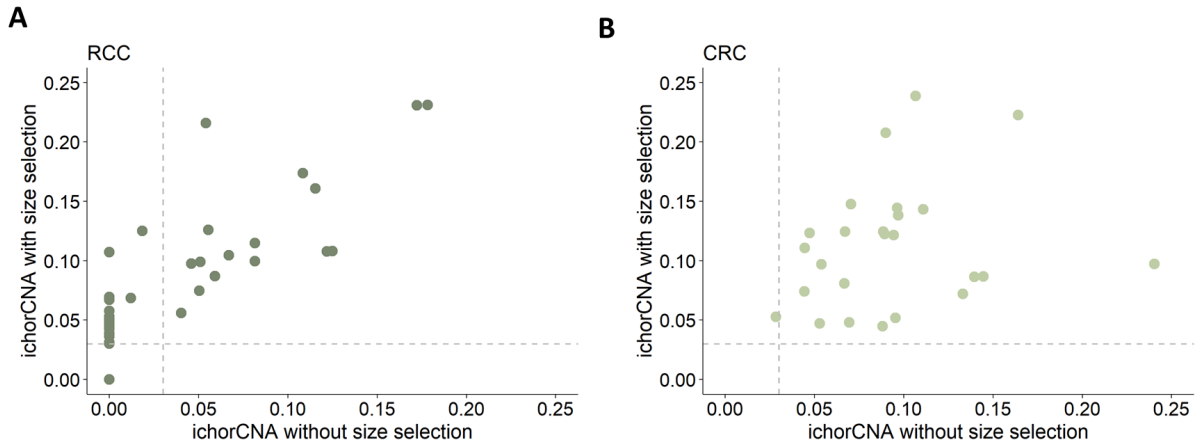


Figure 11: Enrichment of ctDNA using *in silico* size selection.

(A) IchorCNA tumor fractions of plasma samples from RCC patients with (y-axis) and without *in silico* size selection (x-axis). On average, *in silico* size selection increased the tumor fraction 2.2-fold from a median of 0.01 (range 0-0.18) to a median of 0.07 (range 0.03-0.23). Despite an increased ctDNA signal, size selection revealed 19 additional ctDNA-positive samples. Samples with ichorCNA tumor fractions > 0.03 (dashed lines) were considered ctDNA-positive.

(B) Size selection of fragments between 90-150 bp (y-axis) again improved ctDNA levels (1.4-fold, range 0.4-2.6) from a median of 0.09 (range 0.03-0.24) to a median of 0.10 (range 0.04-0.24) in CRC patients.

Taken together, our observations confirm previous reports (91,95) that ctDNA fragments may be shorter than cfDNA fragments. *In silico* size selection of fragments smaller than 150 bp significantly increased the tumor fraction, in particular in plasma samples with initially low ctDNA content (Figure 11). It is noteworthy that even the use of combination strategies for

enhancing tumor signal by combing ssDNA library preparation and *in silico* size selection (smaller than 150 bp) again did not lead to increased sensitivity (Figure 10, Supplementary Table 2), confirming the original results, i.e., ssDNA ligation libraries may not be the method of choice to improve ctDNA detection.

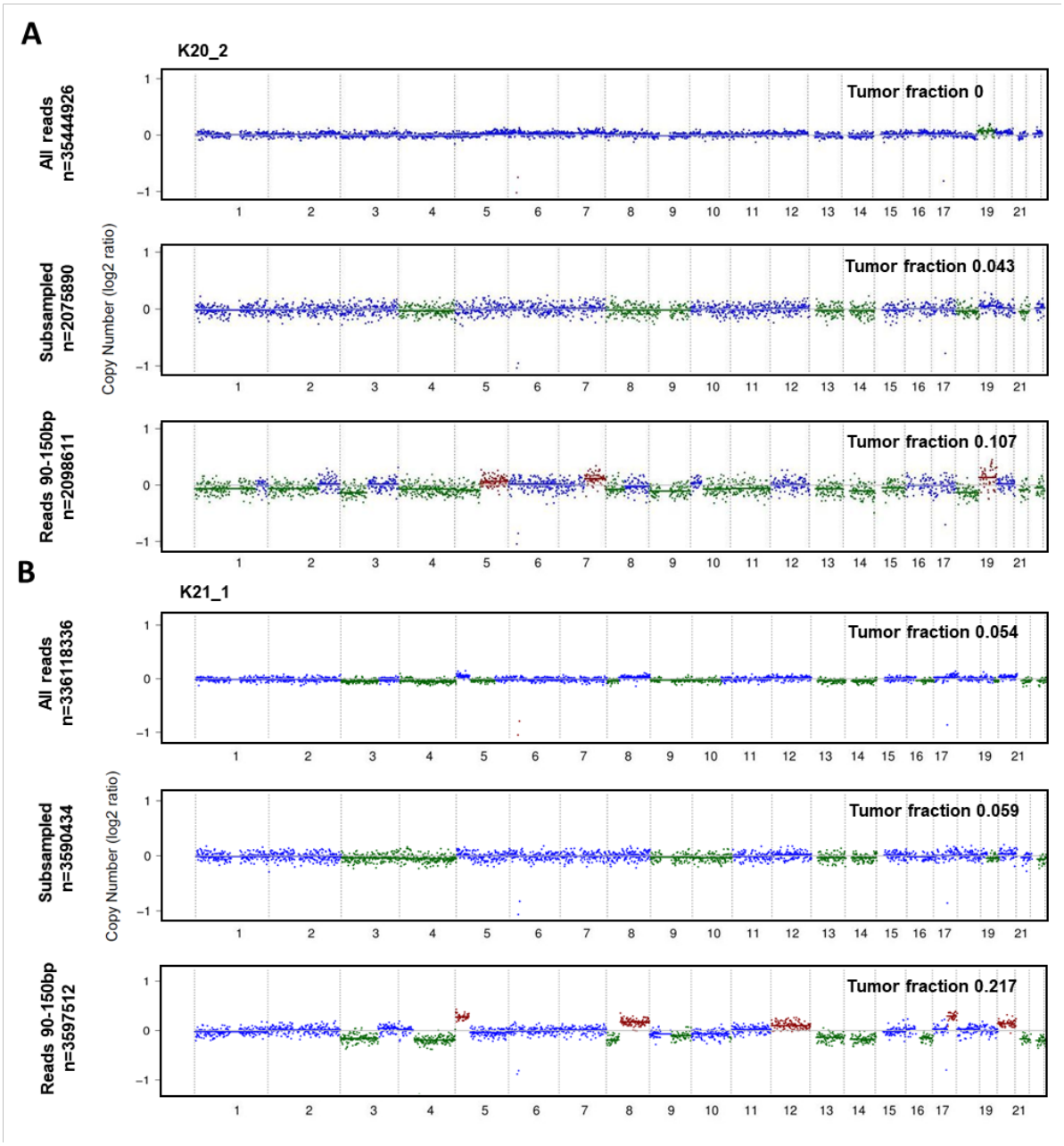


Figure 12: Analysis of the enrichment with and without size selection of plasma samples derived from two patients with RCC.

(A) Shown is one exemplary RCC case (K20_2), in which no SCNAs were detected using all available sequencing reads (top panel). *In silico* size selection of reads with fragment sizes of 90-150 bp led to an increase in the tumor fraction and detected RCC-specific alterations (lower panel).

(B) Shown is one RCC case (K21_1), in which *in silico* size selection improved the ctDNA signal more than fourfold (lower panel).

(A-B) We confirmed that improved SCNA calling did not simply arise as a result of noise introduced by the lower read count by randomly downsampling the initial sequence reads to a read count equivalent to that attained after size selection (middle panel). Blue indicates copy number neutral, red copy number gain, green copy number loss.

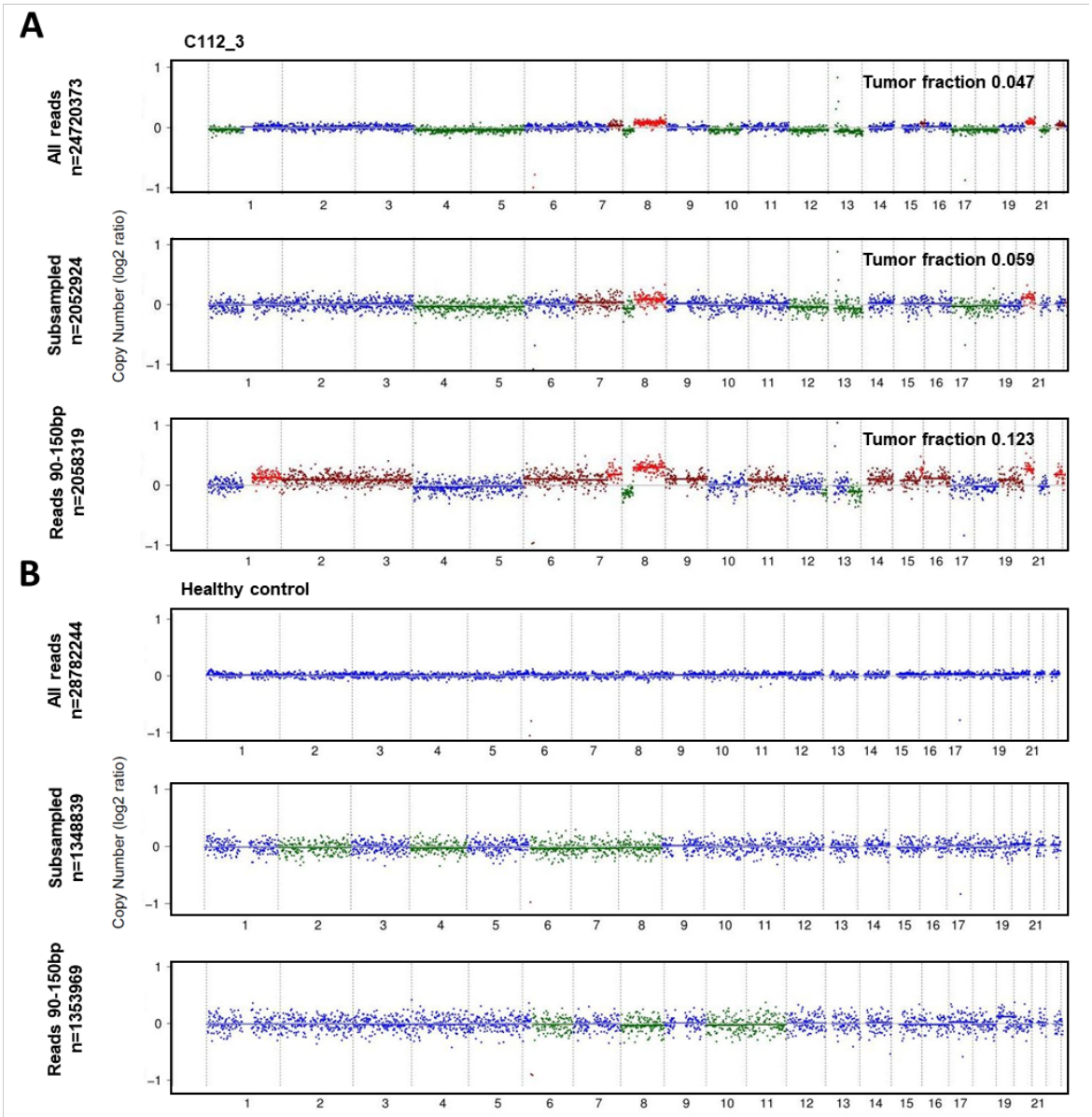


Figure 13: Analysis of the enrichment with and without size selection of plasma samples derived from one patient with CRC and one healthy control.

(A) Shown is one exemplary case of one CRC patient (C112_3), where *in silico* selection enriched the tumor-specific signal 2.6-fold and detected SCNAs commonly observed in CRC (lower panel). We confirmed that improved SCNA calling did not simply arise as a result of noise introduced by the lower read count by randomly

downsampling the initial sequence reads to a read count equivalent to that attained after size selection (middle panel).

(B) [Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

After (lower panel) and without (top panel) *in silico* size selection of one healthy control sample. Due to the lower proportions of smaller DNA fragments in healthy controls, size selection led to increased background noise. However, in the control sample, only whole chromosomes were called, and no RCC or CRC specific SCNAs were observed. Blue indicates copy number neutral, red copy number gain, green copy number loss.

4.2. Assessment of plasma from metastatic RCC patients suggest low ctDNA levels

4.2.1. ctDNA detection using untargeted assays

Within the MonReC cohort, we mainly recruited metastatic patients (41/43), most of whom underwent primary tumor resection (35/43 patients) (51). First, we started to investigate the general plasma DNA parameters in renal tumors such as cfDNA concentration and ctDNA fraction. Consistent with previous reports (102), we observed that all RCC cases (mean 29.8 ng/ml, range 4.5-336.6 ng/ml) compared to healthy subjects (mean 9.2 ng/ml, range 2.9-20.8 ng/ml) had significantly higher amounts of cfDNA per milliliter of plasma ($p < 0.001$, Wilcoxon; Figure 14A).

Given the paucity of data surrounding ctDNA levels and optimal assays in RCC, we applied a combination of cost-efficient and rapid untargeted sequencing approaches, albeit with limited sensitivity, to establish the first measure of ctDNA fractions. Initially, we applied our previously published mFAST-SeqS approach (123) and conducted sWGS to assess copy numbers at a genome-wide level and to calculate tumor fraction using the ichorCNA algorithm (108). mFAST-SeqS establishes a genome-wide z-score as an estimate of ctDNA levels in the plasma of cancer patients (123), with a genome-wide z-score ≥3 indicating a tumor fraction of >3-10% (132). At baseline, z-scores ranged from -0.622 to 3.674 (median: 0.654). We identified elevated ctDNA levels in less than 5% (2/43, 4.7%) of all baseline samples. Faced by this very low ctDNA detection rate in RCC, we applied *in silico* selection (see section 4.1.2) to the baseline sWGS data sets. *In silico* selection improved the sensitivity of the ichorCNA metric and the ctDNA detection rate increased from 18.6% to 32.6% (14/43 of patients), resulting in ctDNA positivity in six additional patients (Figure 14B). On average, the ichorCNA tumor fraction increased 2.2-fold (range 0.9-5.7) to a median of 0.08 (range 0.04-0.23) (51). In addition, we tested for correlation between tumor fraction and overall plasma DNA levels and

observed no association between these two parameters (R=0.21, $p < 0.050$, Pearson's R; Figure 14C). This is in agreement with previous reports that in patients with advanced cancer, there is a several-fold increase in the amount of cfDNA, but the proportion of ctDNA may vary widely (67).

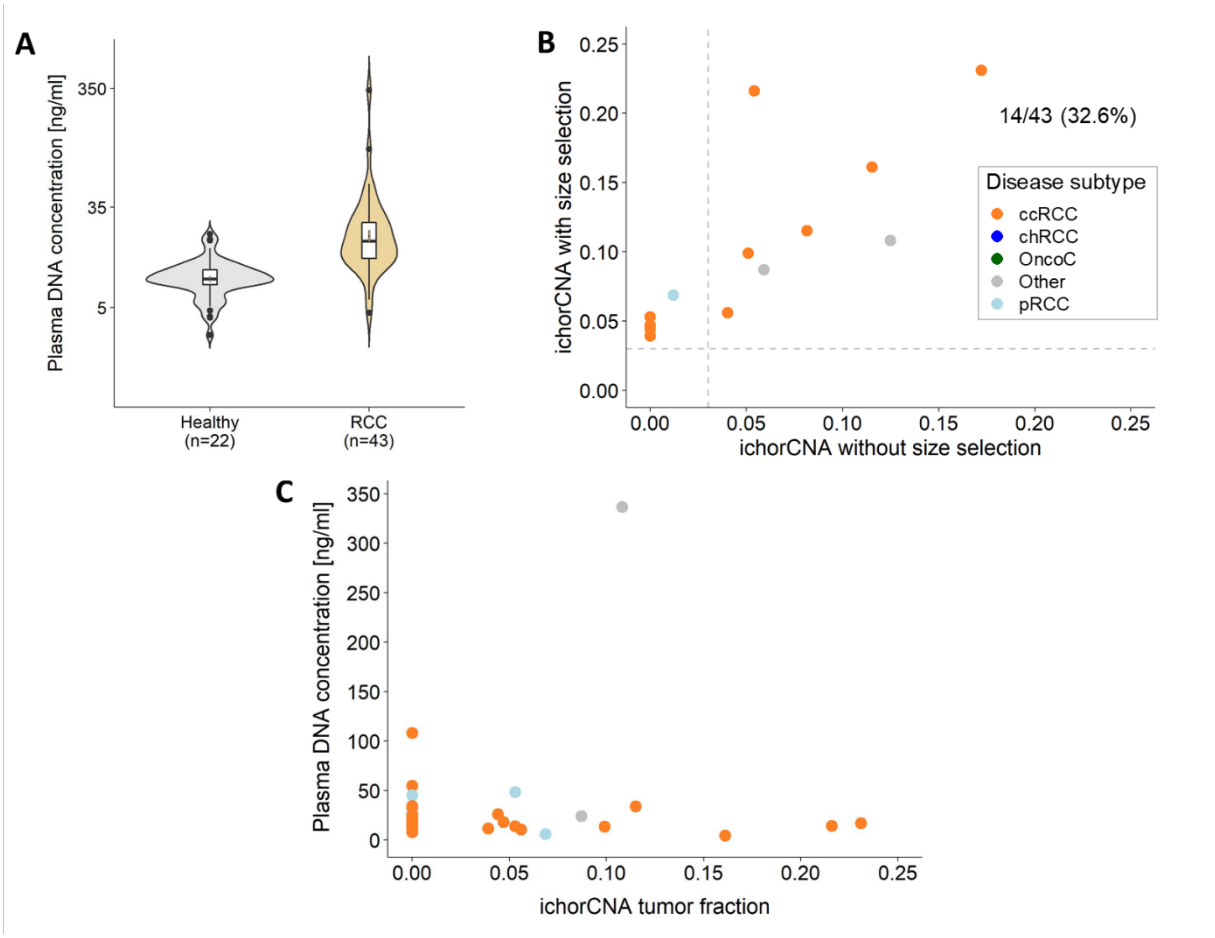


Figure 14: General plasma DNA parameters and ctDNA detection using untargeted assays.

(A) Plasma DNA quantities (ng/ml) in all RCC baseline cases (n=43) versus healthy controls (n=22). cfDNA concentrations were significantly higher in cancer patients than in healthy individuals (Wilcoxon, $p < 0.001$).

(B) [Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

ichorCNA tumor fractions of ctDNA-positive plasma samples (n = 14), before (x-axis), and after *in silico* size selection for sequencing reads 90–150 bp in length (y-axis). On average, the tumor fraction increased 2.2-fold (range 0.9–5.7) and revealed six additional patients with detectable ctDNA, in addition to the eight patient samples detected without size selection. The dotted lines indicate the defined ichorCNA threshold. Data points are colored according to disease subtype.

(C) Comparison of ctDNA fractions estimated using ichorCNA and size selection (x-axis) and plasma DNA concentration in ng/ml (y-axis) for plasma samples from RCC patients revealed no significant difference (Pearson's R=0.21, $p < 0.050$).

Overall, the detection rates of ctDNA in patients with RCC using broad, untargeted sequencing approaches were low; only 31.7% (13/41) of metastatic patients were ctDNA-positive (51). To independently validate these low ctDNA detection rates, we applied the previously published trimmed Median Absolute Deviation (tMAD) metric (96) to all sWGS data from the baseline samples. We observed largely the same result (Figure 15). A summary of all ctDNA-positive baseline samples is shown in Supplementary Table 3.

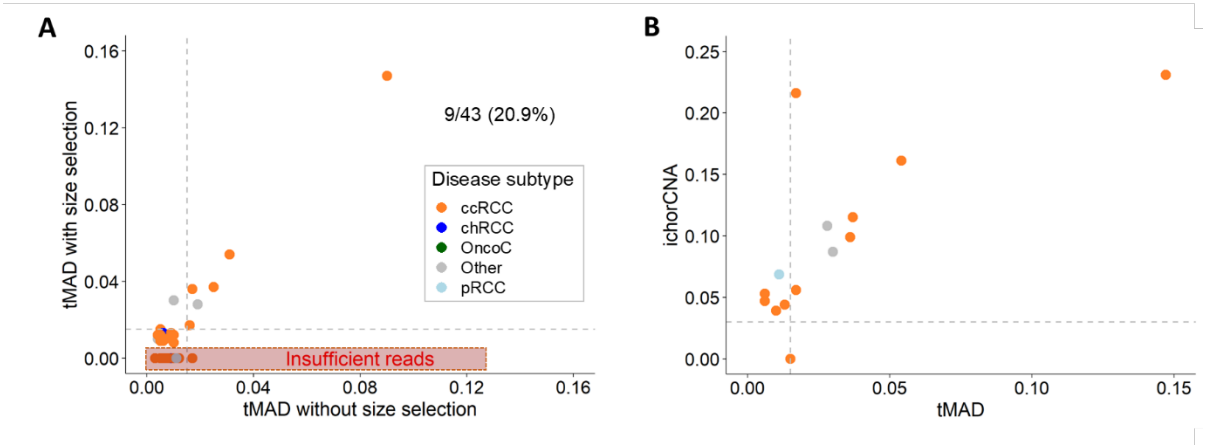


Figure 15: ctDNA detection using the tMAD metric.

[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

(A) Size selection (x-axis) improved the ctDNA detection rate using the tMAD metric, ctDNA positivity was detected in two additional patients (9/43, 20.9%). Data points are colored according to disease subtype.

(B) Comparison of tMAD (x-axis) and ichorCNA tumor fraction (y-axis) distribution. In eight patients, ctDNA was detected using both metrics (Supplementary Table 3). SCNAs were detected when the samples surpassed the tMAD threshold of > 0.015 and ichorCNA threshold of > 0.03 (dashed lines).

4.2.1.1. Comparison of ctDNA detection rates in various advanced cancers

In addition to the assessment of ctDNA levels in RCC, we included a pan-cancer analysis comparison to reveal ctDNA levels in different solid cancers. To this end, we assessed the ctDNA fraction, as quantified by the mFAST-SeqS assay (123) and the ichorCNA metric (108), from a large set of cancer types of similar stage, i.e., patients with metastatic renal, sarcoma, breast, lung, colorectal and prostate cancer. As shown in Figure 16, the frequency of cases with detectable ctDNA varied with tumor type. We observed that 20%-~40% of patients with metastasized lung, colorectal, breast, and prostate cancers had detectable ctDNA levels. In contrast, less than 8% of patients with RCC and none of the samples from patients with sarcoma harbored a genome-wide z-score exceeding the threshold and therefore, ctDNA detection using untargeted approaches was less likely in these cancer types (Figure 16A). This was also

confirmed by tumor fractions calculated with ichorCNA, i.e., plasma samples from patients with RCC showed considerably lower ctDNA levels than patients with other solid cancers (Figure 16B) (51). Sorting the cancer types according to their ctDNA proportion revealed a similar pattern to that shown by Bettegowda et al. (67) (based on targeted mutation analysis) and Mouliere and colleagues (96) (based on cfDNA fragment sizes). Our findings match with an independent sequencing attempt carried out at the University of Cambridge using their established tMAD score (section 4.4).

Overall, the pan-cancer analysis comparison revealed extremely low ctDNA yields in RCC versus other solid tumors (51); thus, our findings illustrated the quantitative limitation associated with the use of ctDNA in RCC and underlined the need to use higher-resolution methods.

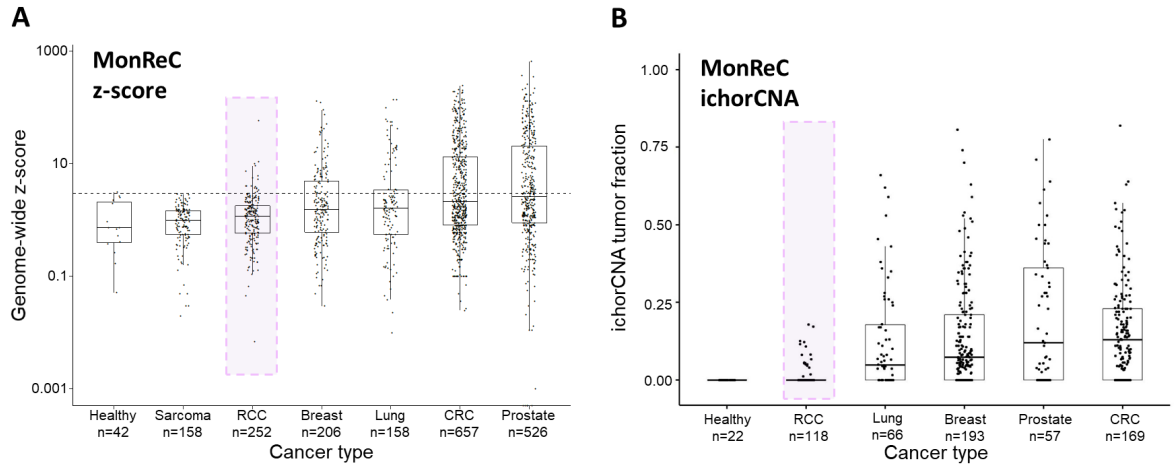


Figure 16: ctDNA detection rates using untargeted sequencing approaches in different types of cancers.
[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]
(A) Genome-wide z-scores distribution and **(B)** ichorCNA tumor fractions across different tumor types. RCC plasma samples from the MonReC cohort were compared to samples from other cancer types and clearly highlighted RCC as ctDNA-low malignancy. The detection threshold is indicated by a grey dashed line. RCC samples are highlighted. A similar comparison was carried out, utilizing the tMAD metric (Figure 21).

4.2.2. ctDNA detection using a high-sensitivity targeted assay

Despite the challenges mentioned above in ctDNA detection, it should be noted that the sensitivity of the untargeted sequencing approaches employed here is limited to a tumor fraction of ~3-10% (132), i.e., >3-10% of the DNA fragments within a sample originate from tumor cells. Therefore, we hypothesized that techniques with higher analytical sensitivity are required to thoroughly quantify the ctDNA concentrations in RCC. For the MonReC cohort, no tissue

was available; therefore, we established the mutational spectrum within baseline plasma samples using a *de novo* approach, i.e., the QIAseq Targeted DNA Custom Panel (Qiagen, Hilden, Germany), targeting ten genes that are frequently mutated in RCC (*BAP1*, *KDM5C*, *MET*, *MTOR*, *PBRM1*, *PIK3CA*, *PTEN*, *SETD2*, *TP53*, *VHL*). As a first step, we evaluated the performance of the QIAseq assay and reported a maximal achievable sensitivity of 5×10^{-3} mAF (Supplementary Figure 1) (51). Based on existing mutation analysis data from RCC (15), we expected a somatic mutation to be available in more than 80% of metastatic ccRCC patients in at least one of these genes (51).

However, ctDNA was observed in only eight baseline samples (8/43; 18.6%, Figure 17), with an average mAF of 8.3×10^{-2} (range 3.5×10^{-2}–0.18). We observed two or more mutations in three patients (K18, K27, K39). Alterations in the *SETD2* gene appeared to be the most commonly seen mutation in this dataset, with a mutation being found in four of the eight ctDNA-positive patients (50%) (Table 2). Alterations in *KDM5C*, *PBRM1,* and *VHL* constituted the next most prevalent SNV, being observed in 25% of all baseline cases (2/8). Somatic mutations in *BAP1*, *MTOR,* and *TP53* were identified in one case each (1/8). Overall, mutations were seen in 70% of all genes analyzed. Moreover, most of these patients, except for patient K42, had detectable SCNAs before and/or after size selection (51). A summary of various assays targeting tumor-derived DNA molecules in all baseline samples is depicted in Supplementary Table 3.

These data illustrated that, especially in ctDNA-low malignancies such as RCC, a gene panel, targeting only a small number of cancer-specific mutations, may not enhance ctDNA detection beyond that attained by untargeted methods, highlighting that in some settings, untargeted, low-cost, and rapid methods can be as effective for ctDNA quantification as the more expensive and time-consuming targeted approaches.

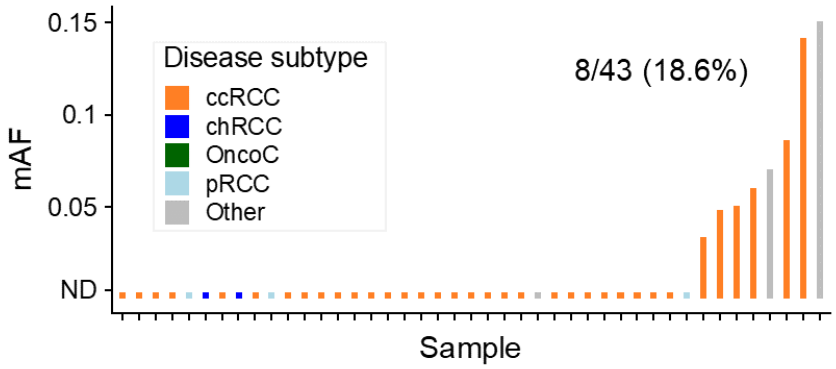


Figure 17: ctDNA detection using a *de novo* mutation calling approach.

[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

Summary of targeted ctDNA profiling using a *de novo* mutation calling approach. At baseline, mutations were detected in 8/43 (18.6%) plasma samples. The y-axis indicates mAF, which ranged from 3.5×10^{-2}–0.15 (if two or more mutations were identified, the mean was calculated). The disease subtype is indicated by bar color; see insert for figure legend.

Table 2: Mutation detected at baseline using a *de novo* mutation calling approach, targeting ten frequently altered genes in RCC.

[Table and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

Patient	Disease subtype	Gene	Mutation according to HGVS (RefSeq No.: cds:protein)	mAF [%]	Consequence	Impact
K05	Unknown	TP53	NM_000546.5:c.701A>G:p.Tyr234Cys	7.1	Missense	Moderate
K12	Unknown	SETD2	NM_014159.6:c.6391C>T:p.Gln2131*	15.1	Stop Gain	High
K18	ccRCC	VHL	NM_000551.3:c.473T>C:p.Leu158Pro	6.3	Missense	Moderate
		SETD2	NM_014159.6:c.4652delA:p.Asp1551Valfs*14	5.9	Frameshift	High
		PBRM1	NM_018313.4:c.1846_1876del:p.Lys616Alafs*16	4.3	Frameshift	High
K20	ccRCC	KDM5C	NM_004187.3:c.4336C>A:p.His1446Asn	8.7	Missense	Moderate
K23	ccRCC	KDM5C	NM_004187.3:c.137delT:p.Ile46Thrfs*27	5.1	Frameshift	High
K27	ccRCC	SETD2	NM_014159.6:c.7537A>C:p.Thr2513Pro	5.3	Missense	Moderate
		PBRM1	NM_018313.4:c.127dupA:p.Thr43Asnfs*10	4.4	Frameshift	High
K39	ccRCC	VHL	NM_000551.3: c.606_607insT:p.Gln203Serfs*53	13.3	Frameshift	High
		SETD2	NM_014159.6:c.4562T>G:p.Leu1521Arg	17.5	Missense	Moderate
		BAP1	NM_004656.3:c.947_948insTC:p.Ala317Argfs*19	11.9	Frameshift	High
K42	ccRCC	MTOR	NM_004958.3:c.3239G>A:p.Arg1080His	3.5	Missense	Moderate

4.3. Longitudinal monitoring of ctDNA in metastatic RCC patients treated with multiple systemic therapies

One of the primary aims of the MonReC study was to determine the potential of plasma DNA to monitor treatment response in mRCC. Therefore, we collected serial plasma samples for 37/43 (86%) patients. Longitudinal blood samples (median 5, range 2-21) were obtained before and during treatment during a median follow-up period of 6 months (range 0.4-19.2) (51).

4.3.1. Serial monitoring of genome-wide z-score of ctDNA

It has been previously published that the mFAST-SeqS approach provides a promising monitoring tool for treatment response (132). To this end, as mentioned above, we used mFAST-SeqS as an initial measure of tumor content (123,132). We observed a median z-score of 1.0 (range -0.9-58.0) and elevated ctDNA levels in 19/252 samples (7.5%, Figure 16) from nine patients. The median time between the first and last genome-wide z-score assessment was 3.6 months (25th-75th percentile: 0.95-7.60). Using a linear mixed model with a random intercept at the patient level, the mean estimates for genome-wide z-scores at baseline was 0.40 (95% CI 0.20-0.60), 0.79 (95% CI 0.63-0.95) during treatment and 0.72 (95% CI 0.44-0.99) at progression. Pairwise differences (Figure 18) revealed significant higher z-scores during treatment compared to baseline ($p < 0.001$, Wald test) and also at progression versus baseline ($p = 0.029$). Surprisingly, z-scores did not change significantly between treatment and progression ($p = 0.583$) (51).

Overall, in the majority of patients, the mFAST-SeqS-based ctDNA levels were below the detection limit of this approach. Therefore, the overall low ctDNA levels within this cohort and the analytical sensitivity may have hindered the reliable assessment of quantitative changes in the mFAST-SeqS z scores.

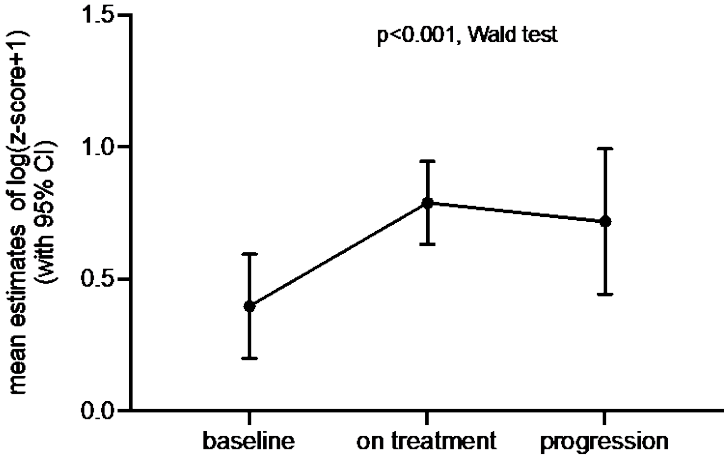


Figure 18: Longitudinal monitoring of genome-wide z-scores.

[Figure and legend originally published in Genome Medicine (Smith & Moser et al., 2020)]

Comparison of z-score distribution using a linear mixed model with a random intercept at the patient level revealed significantly higher z-scores during treatment (Wald test, $p < 0.001$) and at progression (Wald test, $p=0.029$) versus baseline. No significant difference between z-scores between treatment and progression (Wald test, $p=0.583$) was observed.

4.3.2. Serial monitoring of ctDNA mAF dynamics reflect treatment response

We performed longitudinal mutation analysis and ichorCNA analysis of sWGS data on 14 RCC patients (Figure 19). In addition to follow-up samples from patients with detectable ctDNA at baseline (K18, K20, K23, K27, K39, K42), we also performed mutation analyses from patients with clinical signs of disease progression. Therefore, in three additional patients (K11, K21, K35), we observed detectable ctDNA in at least one follow-up sample (Figure 19, Figure 20). In addition, we analyzed follow-up samples derived from five patients with detectable SCNAs (K08, K13, K19, K40, K44) at baseline using the ichorCNA algorithm (Figure 19, Supplementary Table 3) (51).

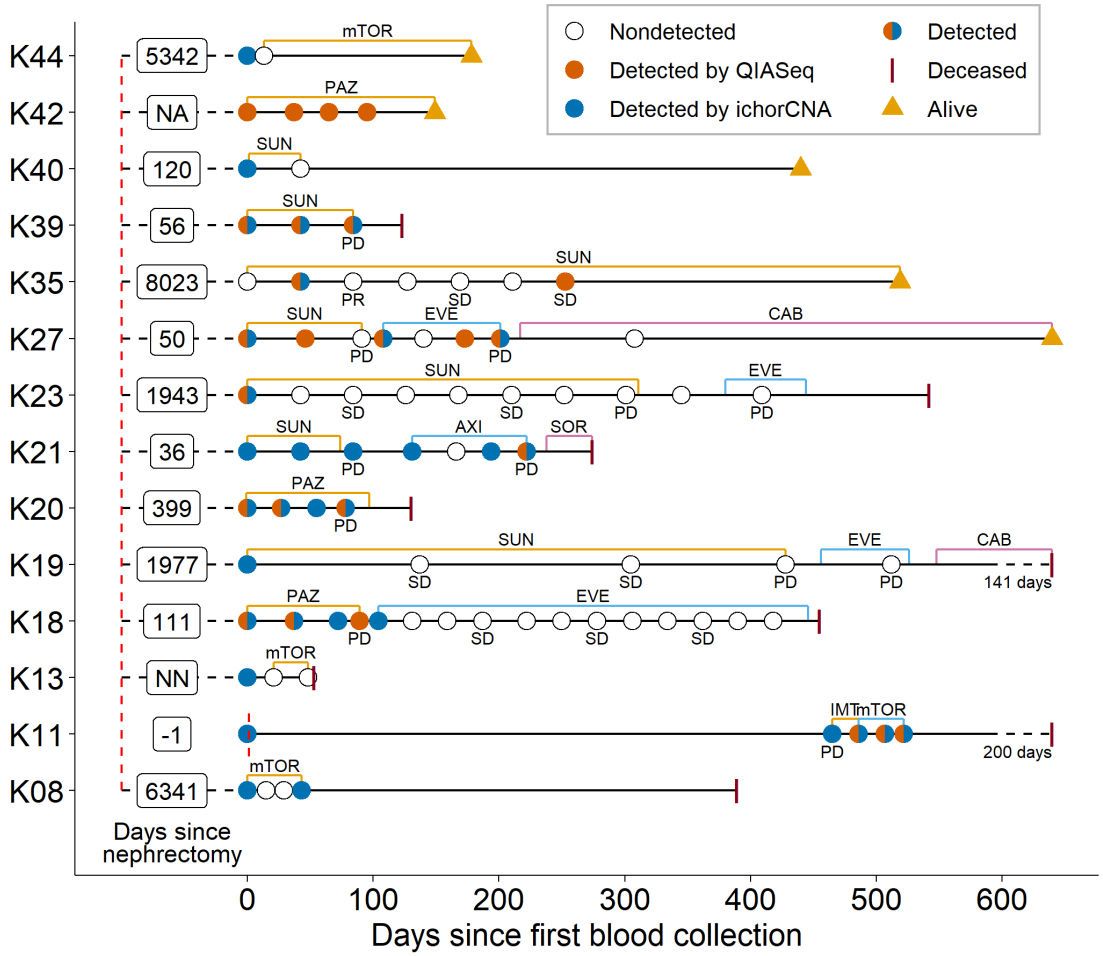


Figure 19: Summary of longitudinal ctDNA analysis of mutations detected by the QIASeq panel and ichorCNA tumor fractions.

[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

ctDNA positivity was assessed with the customized QIAseq panel and/or ichorCNA algorithm. At progression (PD), ctDNA positivity was observed in the majority of patients, whereas during stable disease (SD) or partial response (PR), patients were ctDNA-negative. The time between the first blood draw and patients' nephrectomy is shown in days (NA, not available; NN, no nephrectomy). Colored lines indicate type and duration of treatment (mTOR=mTOR inhibitor; PAZ=pazopanib, SUN=sunitinib; EVE=everolimus; CAB=cabozantinib; AXI=axitinib; SOR=sorafenib; IMT=immune therapy).

We invariably detected similar ctDNA mAF patterns assessed by the QIASeq panel in almost all patients' plasma, i.e., elevated ctDNA levels at treatment initiation (Figure 19, Figure 20, and Supplementary Figure 2). In patients who responded to therapy, we observed a decline in the mAF of all variants (Figure 20A-B, Supplementary Figure 2). In contrast, at progression or when no treatment response was gained (Figure 20C-E, Supplementary Figure 2), the ctDNA levels increased or remained elevated. In the plasma from two patients (K18 and K23) who

were ctDNA-positive before treatment start, the mAFs decreased below our detection limit for almost all time points; even when the disease progressed, the samples remained ctDNA-negative (Supplementary Figure 2C-D). In addition, SCNA-derived mAFs estimated with ichorCNA revealed similar ctDNA patterns as mutation-based assays (Figure 20, Supplementary Figure 2), except patient K42 (Supplementary Figure 2B), for whom no SCNAs were observed (51).

Figure 20 shows patient K39, diagnosed with stage III ccRCC, with the highest mAFs and ichorCNA tumor fractions in this cohort. Nephrectomy and resection of lymph node metastases were carried out fifty-six days before the first blood draw. At baseline, when treatment with sunitinib was initiated and across all three time points, *BAP1*, *SETD2*, *VHL* mutations (Figure 20E), and the RCC-specific SCNAs such as loss of 3p, 9p, and loss of whole chromosome 14 were detected. After 12 weeks, the patient presented with clinically progressive disease and died eight weeks later, mirroring the rapidly progressing disease. In contrast, after ~22 years following nephrectomy with curative intent, patient K35 progressed, i.e., multiple metastatic osteolytic lesions were detected on imaging (Figure 20B). After two years of receiving targeted therapy, the patient still responded, and the therapy schedule was continued (51).

Taken together, ctDNA dynamics mostly corresponded to treatment response and revealed the extensive heterogeneous genetic landscape, diverse clinical presentations, and clinical outcomes of metastatic RCC patients. Nevertheless, the accurate detection of response-specific ctDNA mAF dynamics was affected by the overall low ctDNA fractions in patients with RCC and the performance of the variant caller (Supplementary Figure 3). More work is clearly required to determine the clinical utility of ctDNA and validate any prognostic or predictive utility.

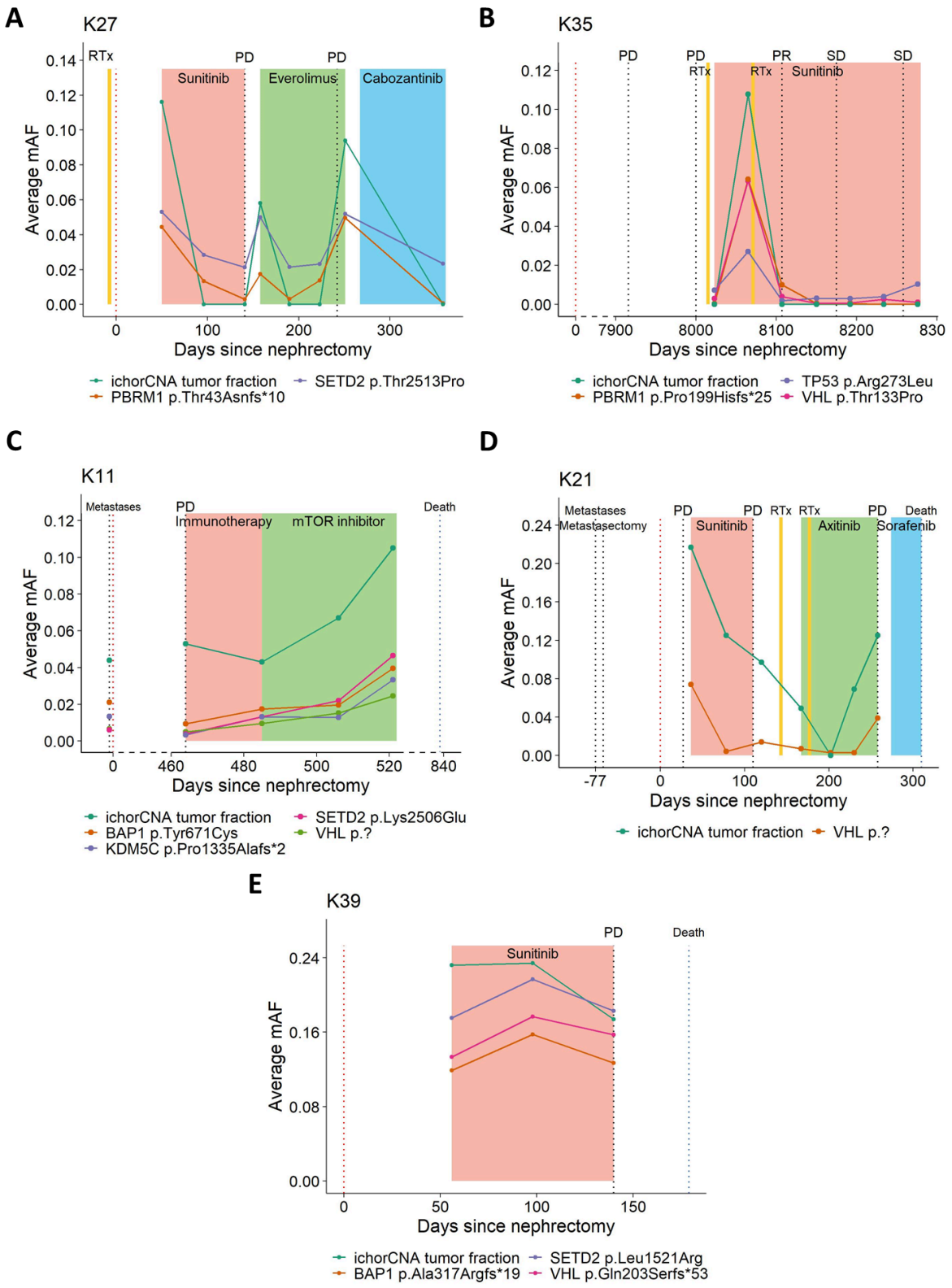


Figure 20: Dynamic changes in ctDNA mAF during systemic therapy for metastatic RCC.

[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

(A-D) Plots illustrating dynamic changes in ctDNA in longitudinal plasma from mRCC patients K27, K35, K11, K21, and K39, respectively.

(A) At baseline, when 1st line treatment with sunitinib was initiated, the patient presented high ctDNA levels. Due to progressive disease, the treatment strategy was changed. After treatment switch to a 3rd line therapy (cabozantinib), patient K27 responded, and a decline in ctDNA levels was registered.

(B) Patient K35 progressed several years after nephrectomy. After the increase of ctDNA levels at the second time point, ctDNA remained low, mirroring patients' treatment response.

(C-E) At progression or when no treatment response was gained, ctDNA increased or remained elevated. **(E)** Patient K39 displayed a high ctDNA load in the plasma at all analyzed time points. The patient progressed rapidly through one line of targeted therapy (sunitinib) and died six months after the diagnosis of ccRCC.

In each panel, the time of diagnosis (red) and the time of patients' death (blue) are indicated as vertical dotted lines. SD denotes stable disease, PD progressive disease, and PR partial response. The ichorCNA tumor fraction is shown in green, and the respective somatic mutations detected by the QIAseq panel are displayed in various colors. Therapeutic interventions are shown in multiple colors (RTx= radiotherapy). Further patient-specific plots are presented in Supplementary Figure 2.

4.4. Assessment of ctDNA in plasma and urine of patients with the full range of renal tumors confirms RCC as ctDNA-low malignancy

We aimed to verify our initial findings from the MonReC study with a second prospective clinical cohort recruited and analyzed at the Cancer Research UK Cambridge Institute. Thus, samples from patients in the DIAMOND cohort were analyzed using a combination of genome-wide and more sensitive sequencing approaches. In addition to plasma samples, urine samples were available for a subset of DIAMOND patients (37/48), as previous studies described that proximal sampling, i.e., sampling of biofluids collected proximal to the tumor site, may yield a higher concentration of tumor-derived DNA (118,133). Urine allows the analysis of urine supernatant (USN) and also the urine cell pellet (UCP) DNA, which is not cell-free but has already been successfully used to non-invasively identify tumor-derived alterations (118).

First, liquid biopsy fluid samples, taken before partial or radical nephrectomy, were analyzed for SCNAs and their respective tumor levels utilizing the tMAD score (96). Applying tMAD to plasma samples, ctDNA was detected in only three of 48 (6.3%) analyzed samples (Figure 21). Consistent with data from our cohort (4.2.1), *in silico* size selection improved ctDNA detection and increased the tMAD score in this cohort (Figure 21A). On average, the tMAD scores increased 2.2-fold (range 1.25–4.83) and led to ctDNA detection in eight additional patients (11/48, 22.9%), including one patient with oncocytoma. Parallel analysis of ctDNA in USN samples indicated similar low ctDNA levels and revealed detectable ctDNA in only eight

patients (8/37, 21.6%) (Figure 21B, ccRCC n=6, oncocytoma n=1, MiT family translocation n=1). ctDNA positivity was observed in the UCP of three patients (3/20, 15%) (Figure 21B), including one ccRCC patient with the largest tumor of the DIAMOND cohort (tumor diameter 23cm). As shown in recent studies (118,134), we observed similar ctDNA levels and detection rates with minimal overlap between plasma and urine (Figure 21B) (51).

Overall, these data again confirmed that ctDNA levels are lower in RCC compared to cancer types of similar size and stage (Figure 21C).

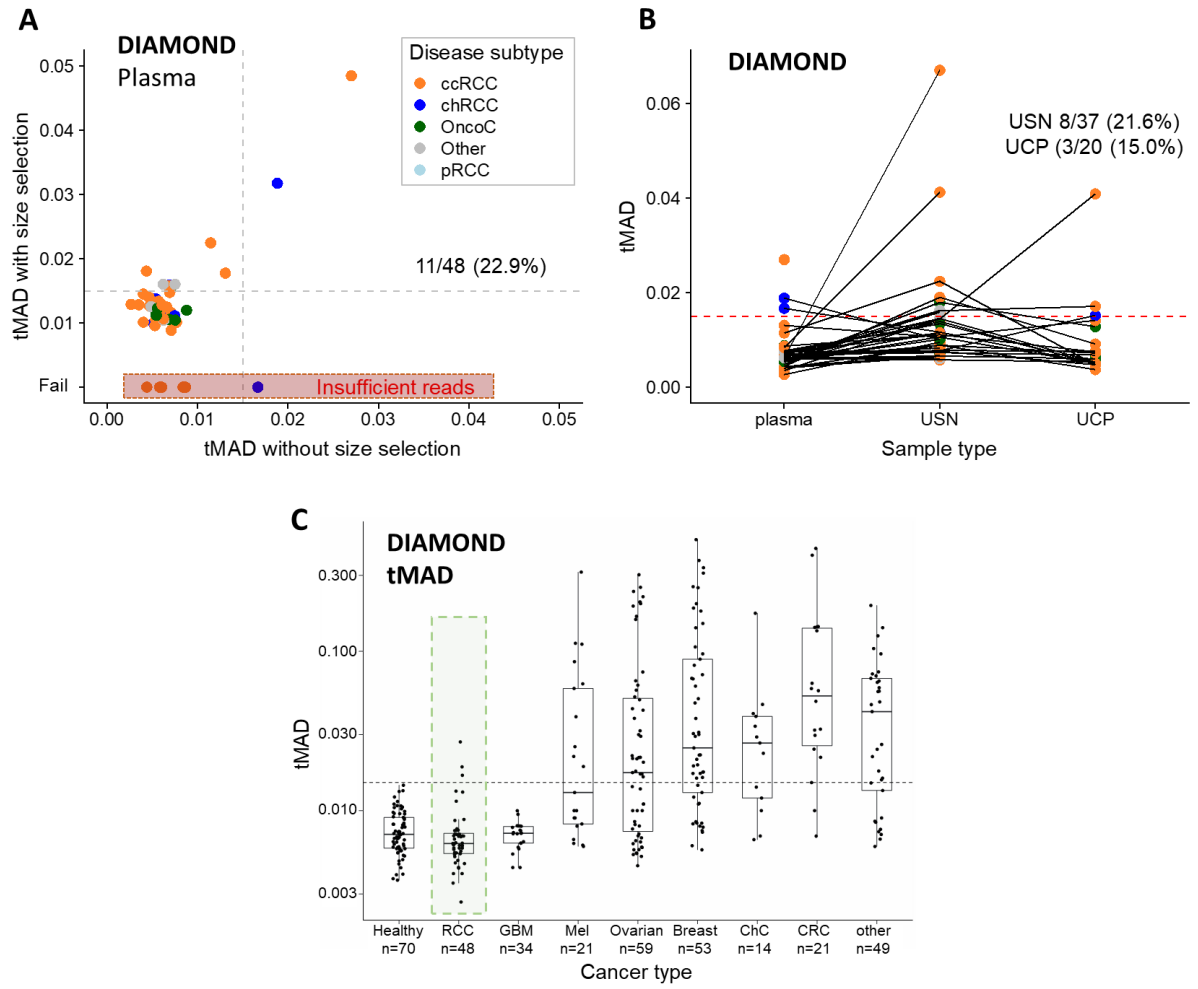


Figure 21: Plasma and urine ctDNA detection using untargeted assays in the DIAMOND cohort.
[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]
(A) Distribution of tMAD scores across plasma samples before (x-axis) and after *in silico* size selection for sequencing reads 90–150 bp in length (y-axis). On average, the tMAD score increased 2.2-fold (range 1.25–4.83), and revealed detectable ctDNA in eight additional patient samples, resulting in ctDNA detection in 11/48 (22.9%) patients. A tMAD score of > 0.015 (gray dashed line) indicates SCNA, and thus ctDNA. Data points are colored

according to disease subtype. Four patient samples had insufficient sequencing reads after size selection for tMAD analysis (red highlight).

(B) tMAD metrics showed similar low ctDNA levels in urine as observed in plasma. Plot showing the difference in tMAD scores across DIAMOND plasma (without size selection), USN and UCP samples. Sample data points from the same patient are connected by lines. The detection threshold is indicated by a red dashed line.

(C) Comparison of tMAD scores across different tumor types. RCC plasma samples from the DIAMOND cohort were compared to samples from other cancer types and highlight clearly RCC as a low ctDNA tumor type. (GBM = glioblastoma, Mel = melanoma, ChC = cholangiocarcinoma). A similar comparison is shown in Figure 16.

Second, the previously described INtegration of VAriant Reads – TAilored PAnel Sequencing (INVAR-TAPAS) approach (106) was applied to a subset of DIAMOND (29/48) and urine (20/48) samples. Seven DIAMOND plasma and six urine samples were unsuitable for the INVAR pipeline due to insufficient informative reads and thus were excluded from the analysis. Of the remaining analyzed plasma samples, 12/22 (54.5%) samples were ctDNA-positive (Figure 22A). As demonstrated in Figure 22, plasma ctDNA detection rates and levels varied amongst patients with renal tumors, i.e., 83.3% (10/12) of these patients were diagnosed with ccRCC. In addition to ctDNA detection in large ccRCC tumors (7.4cm-23cm; 9/12 patients), it was also possible in patients with small ccRCC (2.8 cm; 1/12 patient) and chRCC tumors (2 cm; 1/12 patient) with global mutant allele fractions (gMAF) of 6.4×10^{-5} and gmAF 1.8×10^{-4}, respectively. Even one patient with a benign oncocytoma (3.9 cm) revealed detectable ctDNA levels (gmAF of 2.7×10^{-4}).

As in the plasma, the detection rate has improved with the use of patient-specific assays, leading to seven patients (7/14, 50%) with detectable ctDNA in USN (Figure 22B; ccRCC n=5, chRCC n=1, oncocytoma n=1). In addition, we tested for a correlation between gMAFs in plasma and urine, but there was no relationship between these two parameters (Spearman's rho = 0.28, *p=0.300*; Figure 22C). Nevertheless, in most patients, the tumor fraction was too low for accurate quantification (51).

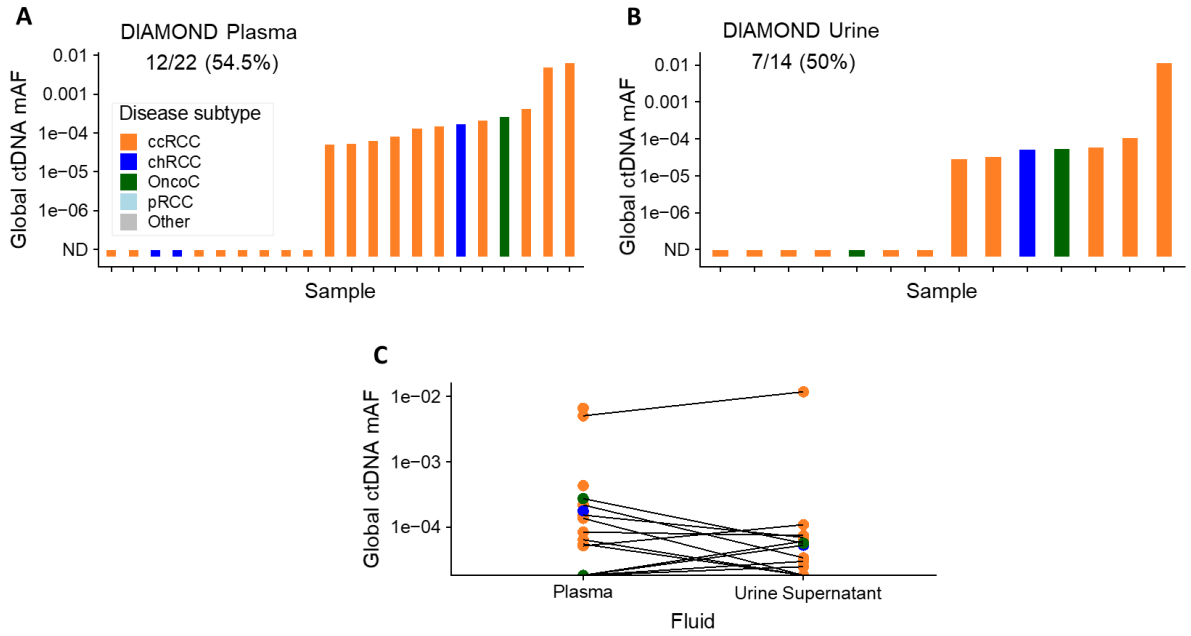


Figure 22: Tumor-informed profiling yields improved ctDNA detection in the DIAMOND cohort.

[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

(A) Application of INVAR-TAPAS to DIAMOND plasma samples. ctDNA was detected in the plasma of 12/22 (54.5%) patients, with global ctDNA mAF (gmAF) shown on the y-axis. The disease subtype is indicated by bar color.

(B) INVAR-TAPAS improved ctDNA detection rates in USN samples. 7/14 patients (50%) harbored detectable ctDNA.

(C) Comparison of gmAF of plasma and USN samples. In patients for whom we had access to both fluids, lines connect data points (Spearman's rho = 0.28, *p=0.300)*.

In summary, the detection rates and levels of tumor-derived DNA in both body fluids varied amongst patients with renal tumors. Surprisingly, we demonstrated that detection was also possible in patients with early-stage disease and even in lesions deemed to be benign. Targeted personalized high-sensitivity assays improved ctDNA detection in plasma and urine (~50%) beyond that attained by untargeted approaches (39.6% of patients had detectable in plasma and/or urine). In total, 27 of 48 (56.3%) DIAMOND patients harbored detectable ctDNA levels, whether in plasma or urine, using non-targeted and targeted methods. Nevertheless, not all samples were analyzed with the most sensitive sequencing method or had both liquid biopsy fluids available. In patients where both liquid biopsy fluids were available, 11/13 (84.6%) harbored detectable ctDNA. In contrast, the overall ctDNA detection rate of 34.5% was substantially lower in the MonReC cohort, most likely due to the use of the less sensitive *de*

novo mutation calling panel (51). A summary of ctDNA detection in all patients and biofluids is shown in Figure 23.

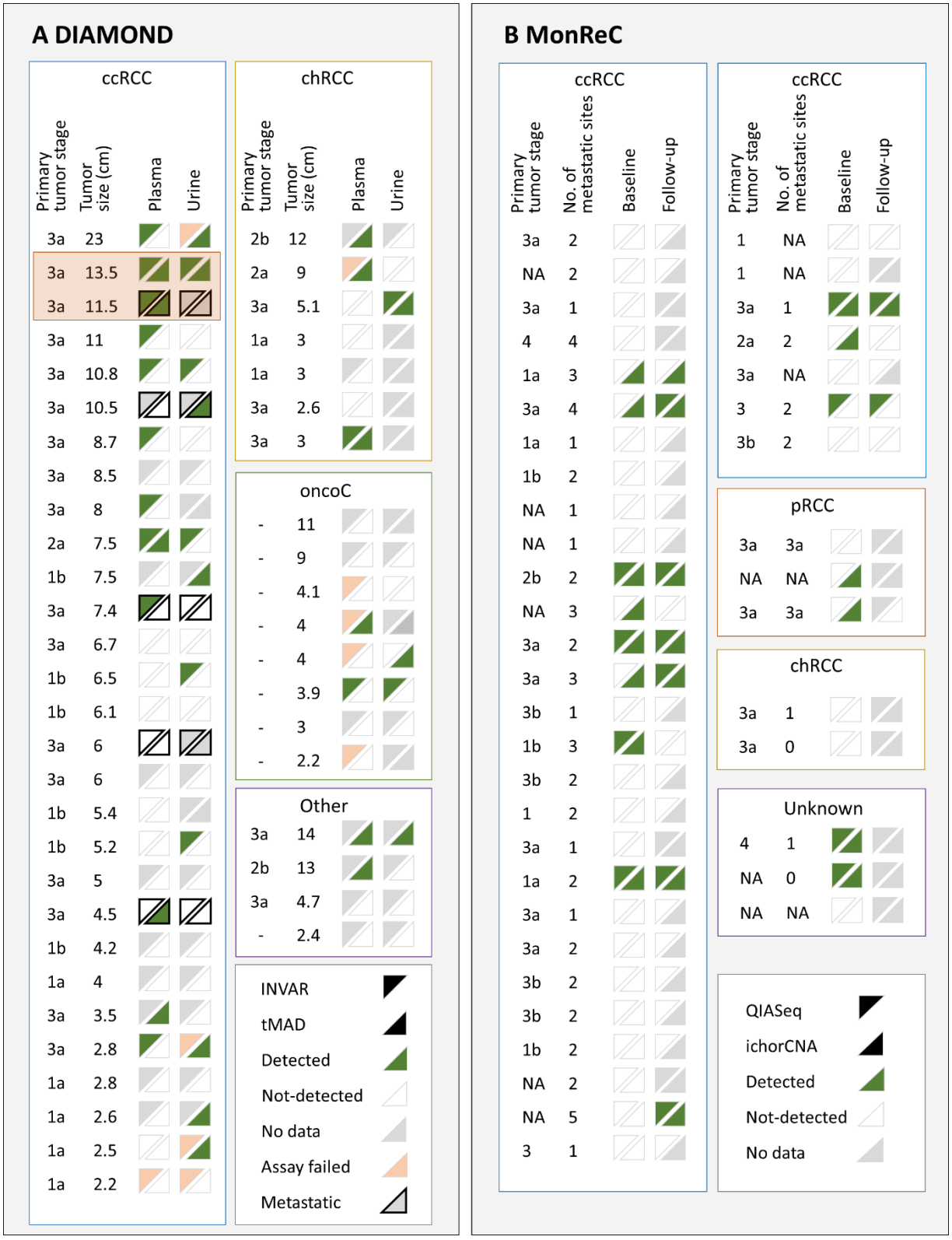


Figure 23: ctDNA detection in patients with the full clinical spectrum of renal tumors.

[Figure and legend originally published in Genome Medicine (Smith & Moser et al., 2020)]

(A) ctDNA discovery in baseline plasma (left of triangle box pair) and urine (right of pair) of DIAMOND ccRCC (left), chRCC (top right), and oncocytoma (oncoC, middle right) patients. Four patients with "other" disease subtypes are also presented ("other," bottom right). Samples are ranked in descending order according to tumor size (cm). For each data point, the upper left triangle displays the results of INVAR-TAPAS analysis and the bottom right of the results of tMAD analysis. Green triangles indicate ctDNA positivity, white triangles highlight samples in which ctDNA was undetected, gray triangles indicate no data available (because the assay was not applied to that sample or no sample was available), and pink triangles indicate failed assay. Data points with a black outline indicate patients with metastatic disease at the time of sampling. Two DIAMOND patients 5842 and 5634 (intratumoral heterogeneity section 4.6) are highlighted with an orange box.

(B) Summary of ctDNA detection in baseline (left of triangle box pair) and follow-up (right of pair) plasma of MonReC patients. Each subtype is shown in a separate box (ccRCC, clear cell; pRCC, papillary; chRCC, chromophobe; NA, unknown). The upper left triangle indicates the results of Qiagen's mutation-based assay, the bottom right shows the results of ichorCNA analysis. Triangle color, as above. Forty-one patients had metastatic disease and, where data were available, the number of metastatic sites is indicated.

4.5. Clinicopathological features are associated with ctDNA discovery

Based on the above findings, the association between clinical parameters and detectability of ctDNA was explored by analyzing plasma and urine samples in DIAMOND patients. First, we observed a significant ($p < 0.020$, Mann-Whitney's U test) correlation between ctDNA discovery and tumor size across all tumor subtypes (Figure 24A). This relationship was driven by plasma with no apparent association in urine. In addition, plasma ctDNA was more likely detected in locally advanced RCC patients denoted by venous tumor thrombus or inferior vena cava tumor thrombus as compared to those without (detection by INVAR-TAPAS +/− tMAD metric; $p < 0.050$, Fisher's exact test, Figure 24B). Surprisingly, we revealed no association in both body fluids between cellular proliferation rate as measured by ki67 staining and ctDNA discovery ($p=0.600$, T-test) (Figure 24C) (51). As the patients' tumors were resected many years before the first blood draw, no conclusions could be drawn from the MonReC cohort.

In summary, ctDNA detection in both liquid biopsy fluids remains challenging. We have identified predictors of ctDNA detection in renal cancer associated with larger tumors and the presence of venous tumor thrombus (51).

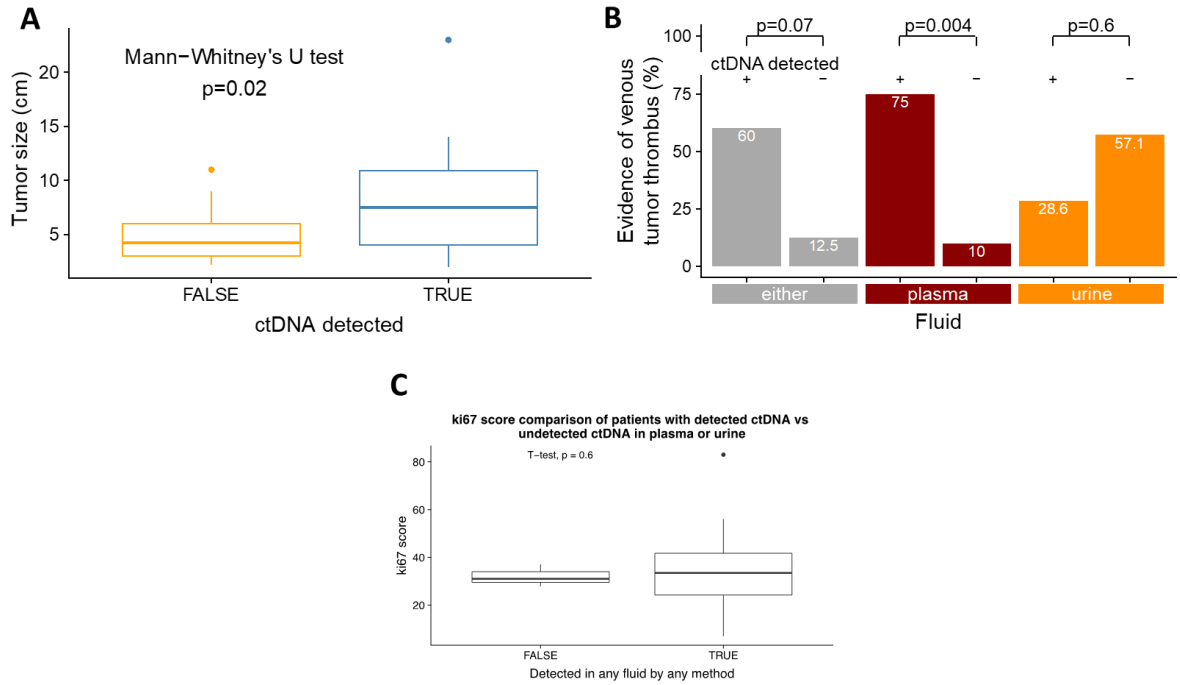


Figure 24: Plasma and urine ctDNA detection rates varied amongst patients with renal tumors.

[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

(A) DIAMOND patients with detected ctDNA had larger tumors. A comparison of tumor size (diameter, cm) between patients with and without detected ctDNA revealed that the former had larger tumors (Mann-Whitney's U test, *p=0.020*). Detection was via tMAD and/or INVAR-TAPAS, and in either fluid. The association was driven by plasma, with no apparent relationship in urine.

(B) ctDNA discovery in plasma by INVAR-TAPAS was significantly more frequent amongst patients with venous tumor thrombus (Fisher's exact test, $p < 0.050$) as compared to those without. This was not the case when considering ctDNA in urine or ctDNA in either fluid.

(C) ctDNA detection was not associated with tumor cell proliferation. Comparison of cellular proliferation rate (as measured by ki67 staining) between patients with and without detected ctDNA (by targeted and untargeted methods) revealed no association in plasma and urine (T-test, *p=0.600*).

4.6. Liquid biopsies capture intratumoral heterogeneity

An aim of the DIAMOND study was to determine the representation of tumor heterogeneity in fluid samples, as intratumor heterogeneity appears to be a consistent feature of RCC (46), which poses a significant challenge for personalized medicine (51). While the levels of ctDNA in most samples were insufficient to allow in-depth analysis of the representation of tumor heterogeneity in liquid biopsy samples, two DIAMOND patients had sufficient tumor-derived molecules in their circulation. For DIAMOND patient 5842 (Figure 23), WES on DNA from ten spatially distinct tumor biopsies of the nephrectomy specimen was performed and revealed

somatic alterations with varying apparent clonality (51). Firstly, we compared baseline ctDNA mAFs in plasma and urine to the number of tumor regions after dense spatial sampling that a mutation was called in (Figure 25A), and reported an incrementally increasing mutation mAF as more tumor samples were considered ($p < 0.050$, Wilcoxon T-test). In addition, private mutation clusters from each tumor region were observed in matched plasma and urine samples, i.e., both body fluids overcame this apparent heterogeneity with 90% and 100% of regions represented by at least one mutation in plasma and urine, respectively (Figure 25B). There was no indication that one or more tumor regions are more strongly represented in fluids than others (Supplementary Figure 4). Taken together, exploration of data from one patient revealed that plasma ctDNA and, for the first time, urine ctDNA are capable of overcoming genetic heterogeneity. These observations were supported by another DIAMOND patient (5634) from whom multi-region sequencing data and matched baseline plasma samples were available (data not shown) (51).

Our results suggest that liquid biopsies may better reflect the genomic architecture of a patient's tumor than a single tissue biopsy, and thus, may help guide treatment decisions in RCC patients.

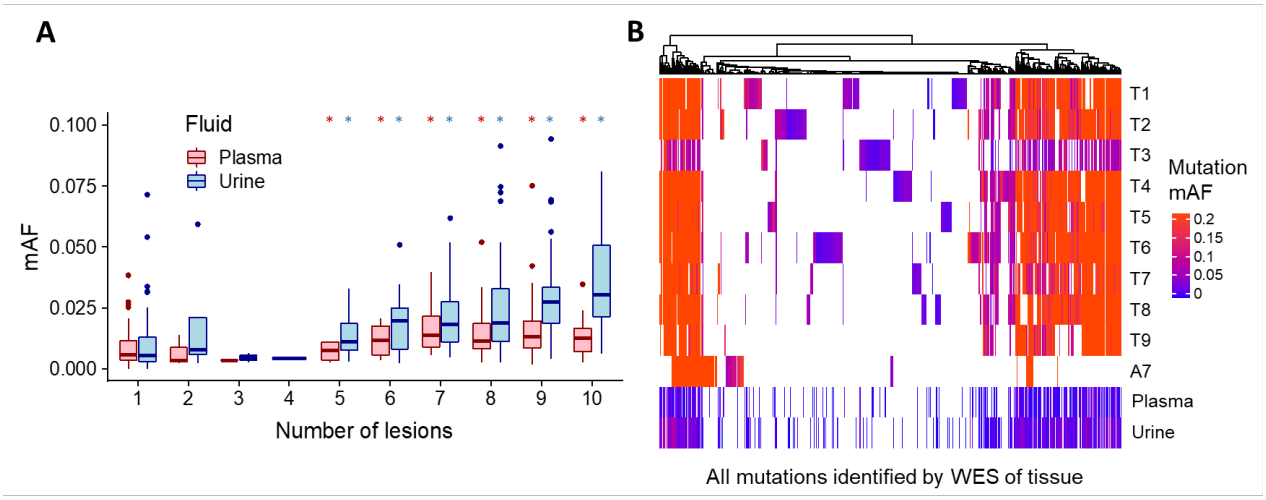


Figure 25: Representation of intratumoral heterogeneity in plasma and urine of DIAMOND patients.
[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]
We obtained multi-region (x10) sampled tumor tissue from one DIAMOND patient (5842).
(A) Comparison of baseline ctDNA mAF in plasma (red) and USN (blue) and the number of tumor regions that mutation was observed in after multi-region sampling of 5842. *illustrates significant difference as compared to mutations detected in just one region.
(B) mAF heatmap of mutations detected across ten tumor biopsies (T1–T9 = fresh frozen, A7 = FFPE) and baseline plasma and urine samples from 5842. All tumor regions were represented by at least one mutation in plasma and

urine. Individual SNVs are displayed in colored vertical lines. Hierarchical clustering was by mutation according to Euclidean distance. Color intensity corresponds to mutation mAF. All mutation clusters, even those private to individual regions, are represented by at least one mutation in plasma and urine (Supplementary Figure 4).

5. DISCUSSION

Approximately 50% of RCC patients succumb to the disease (135); thus, RCC represents the most lethal urological malignancy (136). Due to the mostly asymptomatic disease course, ~25%-30% of all patients have already developed metastatic RCC at the time of diagnosis (137). Moreover, RCC is associated with a wide range of histologically and molecularly distinct malignancies (8), characterized by distinct clinical presentations and clinical outcomes. Currently, it is still challenging to forecast the course of individual tumors, which is why many benign (e.g., oncocytomas) or indolent renal masses (e.g., pRCC type I, chRCC) are overtreated and aggressive tumor subtypes are undertreated (high-grade ccRCC) (52). Therefore, there is an increased desire to use molecular stratification approaches to identify subtype-specific genetic features to ensure appropriate diagnosis and disease-specific therapeutic options (11). Moreover, the characterization of the molecular profile of the tumor both at baseline and at disease progression is necessary to allow personalization of treatment options (3,138). Tracking the response to systemic therapies is of utmost importance in the metastatic RCC setting to continue selected therapy regimes if a response is observed or to change therapy at the earliest possible time if disease progression is detected (3). Nevertheless, treatment selection and the delivery of personalized therapies is complicated by the extensive genetic heterogeneity within RCC tumors (46,139).

Due to the invasiveness of the procedure and failure to recapitulate the genetic heterogeneity in space and time, tumor biopsies are potentially unsuitable for guiding clinical decision-making (140,141). Liquid biopsies, i.e., the analysis of tumor-specific circulating analytes (e.g., ctDNA) collected minimal-invasively or non-invasively through a blood draw or the sampling of other biofluids such as urine (57,58), potentially circumvent the hurdle of temporal and spatial heterogeneity (61). Since ctDNA can be released from multiple tumor regions, it is considered as a potential surrogate for the entire tumor genome (67). Numerous studies reported that it is possible to reconstruct tumor genomes from plasma DNA and indicated that ctDNA could be used to track tumor dynamics in real-time (93,97,142-144).

Although we and other groups have shown that tumor-specific alterations can be identified minimal-invasively via ctDNA in a variety of solid tumors, less is known about the relevance of ctDNA in locally advanced and metastatic RCC (51). To this end, our study is separated into three parts: Firstly, we focused on the optimization of ctDNA detection strategies based on the

biological properties of plasma DNA harnessing the difference in size between cfDNA and ctDNA. Secondly, we applied serval state-of-the-art approaches to provide insight into the presence, levels, composition, and potential clinical utility of ctDNA in mRCC patient management, i.e., predicting and monitoring response to targeted therapies. Thirdly, we aimed to verify our initial observations from the MonReC study with an independent second prospective clinical cohort (i.e., the DIAMOND cohort), which also enabled us to determine the genetic heterogeneity and potential applications of ctDNA from plasma and urine.

5.1. Fragmentomics can boost ctDNA detection

Over the past decade, the analytical sensitivity of ctDNA detection improved dramatically, owing to technological advances (e.g., molecular barcoding and background reduction by sophisticated bioinformatics algorithms) (58). Moreover, proximal sampling, i.e., the analysis of tumor-derived abnormalities in body fluids in close proximity to the tumor site, may improve the sensitivity of the analysis (57). More recently, novel ctDNA approaches have moved beyond genetic tumor-specific alterations and have begun to take advantage of the biological features of plasma DNA to further boost the potential of liquid biopsies (96,103). Despite various reports revealing contradictory findings regarding the size of plasma DNA fragments, several groups clearly suggested from paired-end whole-genome sequencing data that the fragment size of ctDNA is shorter than non-tumor cfDNA (91,94,95). Thus, there are extensive efforts underway to target small DNA fragments to enrich for ctDNA and reduce the cfDNA signal.

In our study, we compared two different library preparation protocols, i.e., ssDNA library preparation and our routinely employed dsDNA library approach, for the recovery of different cfDNA size fractions. Moreover, we assessed whether an alternative strategy, i.e., *in silico* size selection, could be used to improve ctDNA detection, as enhanced detection of somatic alterations by whole-exome and whole-genome sequencing using this approach was previously reported (96).

5.1.1. Comparison of library preparation protocols for the recovery of different cfDNA size fractions

It was recently shown that sequencing library preparation could have a significant effect on length profile measurements (111,112). Our results showed that ssDNA library preparation enriched for small DNA species; however, we did not demonstrate a preferential enrichment of ctDNA (51). Hence, our data support findings from Vong et al. (145), who reported that ssDNA

library preparation significantly improved the recovery of small DNA fragments, but did not result in an enrichment of shorter fetal DNA fragments from plasma DNA of pregnant females. The authors suggested that the biological differences of fetal and maternal plasma DNA fragments, such as the differential degree of DNA damage, may cause the failure of ssDNA libraries to preferentially enrich for fetal DNA (145). Recent reports suggest that tissue-specific nucleosome DNA wrapping, tissue-specific DNA releasing, and degradation mechanisms may be responsible for specific cfDNA fragment size subpopulations (102). For example, ctDNA molecules exhibit a more pronounced 10 bp ladder pattern compared to cfDNA, which could indicate these tissue-specific differences in nucleosome wrapping (95). Moreover, in cancer cells, genome-wide hypomethylation and cfDNA release during cell proliferation rather than apoptosis are discussed as a possible explanation for DNA fragments of various lengths (95,109,146).

We conclude that adaptor-ligation based ssDNA library construction may not be the method of choice for ctDNA enrichment in plasma samples. Nevertheless, we did not evaluate a template-switching library method (120).

5.1.2. Enhanced ctDNA detection using *in silico* size selection

In contrast, *in silico* size selection of fragments in the size range between 90 and 150 bp clearly enriched for ctDNA against the background signal of cfDNA. Thus, these findings are consistent with recent reports (95,96), indicating differences in size between cfDNA and ctDNA. After size selection, the tumor content increased in nearly all ctDNA-positive samples. Most significantly, this method improved ctDNA levels and detection rates, especially in plasma samples with initially low tumor content, e.g., patients with renal cancer. Contrary to our expectations, we did not observe an increased sensitivity by combining ssDNA library preparation and *in silico* size selection, again confirming that ssDNA libraries preferentially enriched heavily damaged cfDNA molecules (112,145). *In silico* size selection requires no additional step in the laboratory and enables re-analysis of existing sWGS datasets by applying different fragment size selection windows (133). However, the proportion of short DNA species (e.g., < 150 bp) is only about 10%; therefore, the majority of sequencing reads are not used. To this end, this method requires a higher sequencing depth to obtain sufficient reads for downstream analysis and is associated with higher costs.

Overall, our results highlight that examining biological features of plasma DNA with size-based methods can improve sensitivity for detecting ctDNA, and thus the evaluation of actionable genomic alterations. We conclude that in particular, the analysis of ctDNA in cancer types that shed low levels of tumor-derived DNA into the circulation could benefit from these fragment-specific approaches. However, size selection may result in loss of potential information, as it was recently shown that ctDNA fragments with low mutant allele frequencies were longer than cfDNA (147). Hence, further studies are urgently required to demonstrate how the enrichment of the full diversity of plasma fragment lengths may offer more information into the functional and biophysical features of plasma DNA (120).

5.2. Characterization of detection rates and levels of ctDNA in patients with renal tumors

To strengthen and independently validate our observations, we collaborated with Grant D. Stewart and Nitzan Rosenfeld's group. This cooperation made it possible to analyze liquid biopsy fluids and tissue samples using various well-established approaches and to include two independent, prospective clinical cohorts representing overlapping but distinct clinical phenotypes. While MonReC evaluated primarily metastatic RCC patients receiving multiple targeted therapies, DIAMOND included patients with the full spectrum of renal tumors from benign to locally advanced and metastatic patients (51).

5.2.1. Assessment of ctDNA in patient plasma

In this study, we aimed to present the most comprehensive dissection of ctDNA fractions in patients with renal tumors. We first applied a bottom-up approach to determine if inexpensive, broad, non-targeted sequencing approaches could be informative in RCC. Although these methods have already been successfully used in other cancer types such as breast (96,132), colorectal (74,107), and prostate cancer (75,121), they showed low plasma ctDNA detection rates in 18.6% and 6.3% of patients from the MonReC and DIAMOND cohorts, respectively. In this setting, the size selection of smaller fragments following sWGS improved the detection of SCNA. However, the ctDNA discovery was still moderate, with 33% and 23% ctDNA-positive patients in these two cohorts, respectively. When a non-personalized high-resolution assay targeting ten RCC-associated genes was applied to the MonReC cohort, detection did not improve beyond the broad, untargeted methods, revealing tumor-specific mutations in only 18.6% of patients. These findings highlighted that in some disease settings, untargeted, low-

cost, and rapid methods combined with fragment-specific analysis could be as effective for ctDNA quantification as more expensive and time-consuming targeted approaches. However, targeted personalized high-sensitivity approaches, which are substantially more costly, yielded improved ctDNA detection rates (~55%) (51).

Since RCC is often an aggressive, angiogenesis-driven disease in a highly vascularized organ with frequent cell necrosis, it is surprising and unclear why we found such low ctDNA levels (51). Previous reports suggested that renal tumors are frequently surrounded by a so-called pseudocapsule (148,149). Although the encapsulated tumors may represent a physical barrier to the release of ctDNA into the bloodstream and/or its filtration into urine, not all RCC disease settings (e.g., metastatic RCC) are likely to be associated with such a capsule. Our findings demonstrated that the detection of ctDNA in plasma, but not in urine, was more likely amongst patients with larger *in situ* tumors and those patients with locally advanced tumor growth (e.g., presence of a tumor thrombus in the renal vein or inferior vena cava). While Abbosh et al. (63) and Mair and colleagues (150) observed tumor cell proliferation as a predictor for ctDNA discovery, we did not observe an association in both body fluids (51).

Taken together, we reported that ctDNA detection rates and levels varied among patients with renal tumors. ctDNA discovery was limited in about a third of patients with mRCC recruited during the MonReC study, albeit with substantially higher tumor fractions than reported in the DIAMOND cohort (51). Beyond assays solely based on tumor-specific SNVs or SCNAs, *in vitro* and/or *in silico* size selection strategies for specific fragment size ranges (95,133,146) in combination with methods targeting multiple biomarkers, such as proteins (151) and cfDNA methylation patterns (79,117,152) may enhance the detection and potential of liquid biopsies in the RCC setting (51).

5.2.2. Assessment of ctDNA in patient urine

Recently, liquid biopsies have become an umbrella term for blood-based biomarkers, but it also applies to serial other body fluids. Sampling other body fluids in combination with or instead of plasma may be considered to improve the sensitivity for the evaluation of ctDNA (57,60). Previous reports have already shown that tumor-specific alterations can be detected in the urine of patients with bladder cancer (118), in the saliva of patients with head and neck squamous cell carcinomas (153), and cerebrospinal fluid of glioma patients (133). For a subset of DIAMOND patients, we quantified tumor fractions in urine, i.e., we analyzed cell-free and

cellular DNA and observed similar low ctDNA levels and detection rates as compared to plasma. In total, only seven patients were ctDNA-positive in both fluids. These findings suggest that potential differential mechanisms may control the release and levels of ctDNA in both bodily fluids, which requires further analysis. Furthermore, more work is highly needed for the evaluation of ctDNA in urine samples from patients with mRCC to compare levels of ctDNA in plasma versus urine in patients without their primary tumor present (51).

5.3. The potential of ctDNA in renal tumors

We focused on the analysis of ctDNA using sophisticated sequencing approaches in patients with renal tumors to validate the application of ctDNA-based assays for clinical decision-making. Among others, we were able to demonstrate that ctDNA detection is possible in patients with lesions deemed to be benign (assessed in DIAMOND patients), underlining the possibility of non-invasively differentiating small renal masses into histological subtypes in order to guide physicians towards the most appropriate clinical treatment (e.g., invasive surgery versus active monitoring) (51).

Here, we additionally focused on those patients who received targeted therapies to validate the ability of ctDNA to track the disease course longitudinally. To evaluate this usage, we analyzed > 200 longitudinal plasma samples collected from 37 RCC patients under targeted therapy. Furthermore, a major aim of the DIAMOND study was to investigate the potential of liquid biopsy fluids to assess the intratumoral heterogeneity; as such, we had access to multi-region sampled tumor tissue and matched plasma and urine samples.

5.3.1. Therapy monitoring via ctDNA in mRCC patients

Over the last decade, the therapeutic modalities for patients with advanced-stage RCC have evolved remarkably. With the development and introduction of immune-checkpoint inhibitors, targeted agents, and new combination strategies, the survival outcomes of RCC patients has improved dramatically (1,9). Whereas the clinical benefit of anti-angiogenic therapies such as VEGF-targeted TKIs has already been demonstrated, such treatment strategies rarely lead to complete response and disease progression is likely in almost all patients (41). Within this study, we have aimed to correlate ctDNA levels with disease course that included several rounds of a variety of treatments (51).

In contrast to the study by Suppan et al. (132), which showed that genome-wide z-scores of metastatic breast cancer patients were clearly associated with the response to therapy, we did not see a decrease of z-scores during treatment. Due to the overall low ctDNA levels and detection rates in the MonReC cohort, we applied sWGS in combination with size selection and the high-resolution *de novo* mutation calling approach. In the fourteen patients suitable for longitudinal analysis, ctDNA dynamics widely fluctuated in accordance with disease course. Most RCC patients harbored detectable ctDNA at treatment initiation, whereas during stable disease or response, levels of tumor-derived DNA decreased. ctDNA increased or remained elevated at disease progression or when a treatment response was not gained (51). Due to the limited studies characterizing ctDNA profiles and levels across RCC patients receiving systemic therapies, our study is of great importance. Recently, Yamamoto and colleagues (154) observed cancer-related mutations in 30% of patients with RCC (n=53). Although the authors reported that the course of the disease could be monitored using longitudinal samples, this was only investigated in a total of six patients (154). In a second study, Pal et al. (116) determined ctDNA changes across 220 mRCC patients receiving first-line and later lines of therapy using the commercial Guardant360 (Guardant Health) platform. Among those patients, 78.6% had detectable ctDNA, though with a median of one genomic alteration per patient. Moreover, they observed that the mutation rates for specific genes (*TP53*, *NF1*, *VHL*, *EGFR*, and *PIK3CA*) increased as patients progressed from the first line to salvage treatment, illustrating the potential mechanisms of therapeutic resistance during systemic therapy (116,155). Besides, the same group observed ctDNA positivity in 53% of mRCC patients (n=34), and as reflected in our study, showed an association between ctDNA discovery and disease volume (156). Unfortunately, these two studies did not indicate the range of detected mAFs, which impeded a direct comparison with the MonReC data set. Overall, we observed substantially lower detection rates with the MonReC cohort, possibly due to our smaller gene panel, which has a limit of detection of 5.0×10^{-3} as compared to 2.0×10^{-4} for Guardant360. Additionally, the commercial platform targeted 73 cancer-related genes, including *ARID1A*, *EGFR*, and *NF1* (116,156), many of which we did not analyze (51).

Taken together, larger gene panels and/or the application of personalized sequencing approaches such as INVAR-TAPAS offer the possibility to greatly improve the sensitivity of ctDNA-based assays (51). These approaches could be particularly useful for these "challenging" cancers that shed lower concentrations of tumor-derived molecules into the

bloodstream. Moreover, additional biological features of cfDNA, e.g., fragment size should perhaps be considered for disease monitoring (154). The current data show that no consensus on the levels and potentiality of ctDNA in RCC has yet been reached. In future work, the clinical utility of ctDNA in RCC, i.e., its prognostic or predictive value, needs to be further analyzed. Furthermore, no standardization of the optimal timing of the blood draws primarily in relation to systemic therapy dosing has yet been introduced and, therefore, needs further investigation (51).

5.3.2. Representation of intratumoral heterogeneity in liquid biopsy fluids

RCC is associated with an extensive intratumor heterogeneity (46), which is one of the leading causes of cancer progression and therapeutic failure. Furthermore, this genetic heterogeneity offers a particular challenge to the delivery of precision cancer medicine that best matches the genetic composition of a patient's tumor (31). Overall, it is an unmet clinical need to assess the tumor heterogeneity in every aspect during a patient's journey through cancer, i.e., at the time of diagnosis and at various time points during systemic treatments (45). To this end, we assessed the tumor heterogeneity through multi-region sampling and its representation in liquid biopsy-based fluids, i.e., plasma and urine. Due to the overall low ctDNA detection rates and levels, only two well-characterized DIAMOND patients were suitable for the analysis. Overall, our findings confirm that plasma ctDNA can capture tumor heterogeneity. In addition, for the first time, we demonstrated that even urine ctDNA is capable of the same (51). This ability to assess spatially separate tumor subclones through the use of urine represents a particularly attractive proposition given the quantities that can be collected at great ease.

In summary, our data reveal that liquid biopsy fluids may efficiently address the tumor heterogeneity in RCC patients in real-time, and, more importantly, that this is possible in a clinically relevant setting without burdening patients.

5.4. Limitations

There are multiple limitations in our study that must be acknowledged. Most significantly, low tumor content restricted the usage of liquid biopsy-based approaches in RCC patients. For example, longitudinal ctDNA dynamic monitoring was only informative in a subset of mRCC patients (n=14). Moreover, while the experimental design of the DIAMOND study had aimed for the assessment of heterogeneity in several RCC cases, we were hampered by the low levels

of ctDNA and sample availability. Thus, only two well-characterized DIAMOND patients were suitable for the exploration of intratumoral heterogeneity. Besides, multiple different ctDNA techniques were used for the heterogeneous patient populations, which in turn independently verified our initial observations that ctDNA detection in RCC is challenging (51). While we reported that enrichment strategies, e.g., *in silico* size selection of shorter sequencing reads (<150 bp), are required for ctDNA-low tumors, it may result in loss of potential informative plasma DNA sub-populations and introduce biases (146). Furthermore, the pre-analytical procedures varied not only within the DIAMOND study but also between the two prospective clinical cohorts (51). Overall, further studies are urgently needed to standardize the pre-analytical workflows, including blood collection, centrifugation speed, and plasma DNA extraction methods to ensure the validity and utility of liquid biopsy-based assays.

6. CONCLUSION

To the best of our knowledge, our study represents the most comprehensive characterization of ctDNA in RCC patients' plasma and urine samples, with a wide range of methods that could be applicable in different clinical scenarios.

Overall, our results revealed substantially lower levels of ctDNA in RCC than in other solid cancers of similar size and stage, with untargeted sequencing detection (without size selection) in 12.1% of patients from both cohorts (51). To this end, improved ctDNA detection methods, e.g., fragment-specific analyses, are clearly required for these "challenging" tumors. We were able to exploit the differences in fragment sizes using *in silico* size selection, which allowed a more sensitive evaluation of tumor-derived molecules. While single-stranded library constructions may not be the method of choice to enrich for ctDNA, *in silico* size selection enhanced ctDNA detection to 27.5% of patients from both cohorts. Surprisingly, even in the metastatic disease setting, the discovery of tumor-derived DNA was also limited to roughly a third of patients, again indicating generally low ctDNA (51).

In addition, a tumor-guided patient-specific assay, which is more expensive but analyzes in parallel a larger number of loci, was applied with incremental success (~50%) compared to a high-sensitivity *de novo* mutation calling approach (18.6%) targeting ten RCC associated genes. Despite the overall low ctDNA content, our data reveal novel insights into the capabilities of ctDNA for minimal-invasive assessment of renal tumors and, thus, may represent an essential contribution to the realization of personalized medicine for RCC patients. Firstly, while ctDNA was more likely amongst patients with larger tumors and in those patients with venous tumor thrombus, we reported that detection in plasma was also possible in patients with small (< 3cm) and even lesions deemed to be benign. Detection of lesions of this size and stage could improve patient outcomes by expediting treatment. Furthermore, we observed the potential of longitudinal analysis of blood-based samples to monitor disease course in some metastatic RCC patients, which in the future may enable early termination of ineffective therapy and a switch to an alternative treatment option. Finally, further interrogation of those patients with detected ctDNA revealed that plasma ctDNA and, for the first time, urine ctDNA are capable of overcoming genetic heterogeneity. Nevertheless, the utility of ctDNA as a prognostic or predictive biomarker in RCC patients has yet to be proven before its full potential can be recognized (51).

In summary, the results generated by our study provided valuable insight into the biology of renal cell carcinoma and contributed to a better understanding of the potentiality of liquid biopsy-based samples in these patients. Overall, our study highlighted RCC as a ctDNA-low malignancy. The biological reasons behind this phenomenon are still unknown and have yet to be clarified. Nevertheless, we illustrated potential clinical utility in the management of patients with RCC, which could contribute to the realization of minimally invasive precision oncology in the future. Further progress in pre-analytical and analytical approaches for liquid biopsy fluids is urgently needed before this can be realized (51).

REFERENCES

(1) Escudier B, Porta C, Schmidinger M, Rioux-Leclercq N, Bex A, Khoo V, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. Ann Oncol 2019 May 1;30(5):706-720.

(2) Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018 Nov;68(6):394-424.

(3) Hsieh JJ, Purdue MP, Signoretti S, Swanton C, Albiges L, Schmidinger M, et al. Renal cell carcinoma. Nat Rev Dis Primers 2017 Mar 9;3:17009.

(4) Maia MC, Salgia M, Pal SK. Harnessing cell-free DNA: plasma circulating tumour DNA for liquid biopsy in genitourinary cancers. Nat Rev Urol 2020 May;17(5):271-291.

(5) Krabbe LM, Bagrodia A, Margulis V, Wood CG. Surgical management of renal cell carcinoma. Semin Intervent Radiol 2014 Mar;31(1):27-32.

(6) Li P, Wong YN, Armstrong K, Haas N, Subedi P, Davis-Cerone M, et al. Survival among patients with advanced renal cell carcinoma in the pretargeted versus targeted therapy eras. Cancer Med 2016 Feb;5(2):169-181.

(7) Turajlic S, Swanton C, Boshoff C. Kidney cancer: The next decade. J Exp Med 2018 Oct 1;215(10):2477-2479.

(8) Turajlic S, Larkin J, Swanton C. SnapShot: Renal Cell Carcinoma. Cell 2015 Dec 3;163(6):1556-1556.e1.

(9) Kotecha RR, Motzer RJ, Voss MH. Towards individualized therapy for metastatic renal cell carcinoma. Nat Rev Clin Oncol 2019 Oct;16(10):621-633.

(10) 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 Jul;70(1):93-105.

(11) 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 Jun 19;23(12):3698.

(12) Dizman N, Philip EJ, Pal SK. Genomic profiling in renal cell carcinoma. Nat Rev Nephrol 2020 Aug;16(8):435-451.

(13) Mitchell T, Turajlic S, Rowan A, Nicol D, Farmery J, O'Brien T, et al. Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. Cell 2018 04/12;173.

(14) Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature 2013 Jul 4;499(7456):43-49.

(15) Sato Y, Yoshizato T, Shiraishi Y, Maekawa S, Okuno Y, Kamura T, et al. Integrated molecular analysis of clear-cell renal cell carcinoma. Nat Genet 2013 Aug;45(8):860-867.

(16) Lara PN, Jonasch E. Kidney Cancer: Principles and Practice. : Springer; 2012.

(17) Porta C, Cosmai L, Leibovich BC, Powles T, Gallieni M, Bex A. The adjuvant treatment of kidney cancer: a multidisciplinary outlook. Nat Rev Nephrol 2019 Jul;15(7):423-433.

(18) Dalgliesh GL, Furge K, Greenman C, Chen L, Bignell G, Butler A, et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature 2010 Jan 21;463(7279):360-363.

(19) Varela I, Tarpey P, Raine K, Huang D, Ong CK, Stephens P, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. Nature 2011 Jan 27;469(7331):539-542.

(20) Turajlic S, Xu H, Litchfield K, Rowan A, Horswell S, Chambers T, et al. Deterministic Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal. Cell 2018 Apr 19;173(3):595-610.e11.

(21) de Cubas AA, Rathmell WK. Epigenetic modifiers: activities in renal cell carcinoma. Nat Rev Urol 2018 Oct;15(10):599-614.

(22) Cancer Genome Atlas Research Network, 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 Jan 14;374(2):135-145.

(23) Ricketts CJ, Crooks DR, Sourbier C, Schmidt LS, Srinivasan R, Linehan WM. SnapShot: Renal Cell Carcinoma. Cancer Cell 2016 Apr 11;29(4):610-610.e1.

(24) Davis C, Ricketts C, Wang M, Yang L, Cherniack A, Shen H, et al. The Somatic Genomic Landscape of Chromophobe Renal Cell Carcinoma. Cancer Cell 2013 12/08;26:319-330.

(25) Shen C, Beroukhim R, Schumacher SE, Zhou J, Chang M, Signoretti S, et al. Genetic and functional studies implicate HIF1alpha as a 14q kidney cancer suppressor gene. Cancer Discov 2011 Aug;1(3):222-235.

(26) Baudis M. Progenetix. 2020; Available at: https://progenetix.org/. Accessed August, 2020.

(27) Motzer RJ, Mazumdar M, Bacik J, Berg W, Amsterdam A, Ferrara J. Survival and prognostic stratification of 670 patients with advanced renal cell carcinoma. J Clin Oncol 1999 Aug;17(8):2530-2540.

(28) Motzer RJ, Bacik J, Schwartz LH, Reuter V, Russo P, Marion S, et al. Prognostic factors for survival in previously treated patients with metastatic renal cell carcinoma. J Clin Oncol 2004 Feb 1;22(3):454-463.

(29) Ko JJ, Xie W, Kroeger N, Lee JL, Rini BI, Knox JJ, et al. The International Metastatic Renal Cell Carcinoma Database Consortium model as a prognostic tool in patients with metastatic renal cell carcinoma previously treated with first-line targeted therapy: a population-based study. Lancet Oncol 2015 Mar;16(3):293-300.

(30) Ljungberg B, Bensalah K, Canfield S, Dabestani S, Hofmann F, Hora M, et al. EAU guidelines on renal cell carcinoma: 2014 update. Eur Urol 2015 May;67(5):913-924.

(31) Iacobuzio-Donahue C, Litchfield K, Swanton C. Intratumor heterogeneity reflects clinical disease course. Nature Cancer 2020 01/01;1:3-6.

(32) Choueiri TK, Motzer RJ. Systemic Therapy for Metastatic Renal-Cell Carcinoma. N Engl J Med 2017 Jan 26;376(4):354-366.

(33) Lainakis G, Bamias A. Targeting angiogenesis in renal cell carcinoma. Curr Cancer Drug Targets 2008 Aug;8(5):349-358.

(34) Hudes G, Carducci M, Tomczak P, Dutcher J, Figlin R, Kapoor A, et al. Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. N Engl J Med 2007 May 31;356(22):2271-2281.

(35) Choueiri TK, Escudier B, Powles T, Mainwaring PN, Rini BI, Donskov F, et al. Cabozantinib versus Everolimus in Advanced Renal-Cell Carcinoma. N Engl J Med 2015 Nov 5;373(19):1814-1823.

(36) Choueiri TK, Escudier B, Powles T, Tannir NM, Mainwaring PN, Rini BI, et al. Cabozantinib versus everolimus in advanced renal cell carcinoma (METEOR): final results from a randomised, open-label, phase 3 trial. Lancet Oncol 2016 Jul;17(7):917-927.

(37) 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 Oct;20(10):1370-1385.

(38) Fisher R, Pusztai L, Swanton C. Cancer heterogeneity: implications for targeted therapeutics. Br J Cancer 2013 Feb 19;108(3):479-485.

(39) Morais C. Sunitinib resistance in renal cell carcinoma. J Kidney Cancer VHL 2014 Apr 22;1(1):1-11.

(40) Tamaskar I, Dhillon J, Pili R. Resistance to angiogenesis inhibitors in renal cell carcinoma. Clin Adv Hematol Oncol 2011 Feb;9(2):101-110.

(41) Makhov P, Joshi S, Ghatalia P, Kutikov A, Uzzo RG, Kolenko VM. Resistance to Systemic Therapies in Clear Cell Renal Cell Carcinoma: Mechanisms and Management Strategies. Mol Cancer Ther 2018 Jul;17(7):1355-1364.

(42) Rini BI, Atkins MB. Resistance to targeted therapy in renal-cell carcinoma. Lancet Oncol 2009 Oct;10(10):992-1000.

(43) Heitzer E, Ulz P, Geigl JB. Circulating tumor DNA as a liquid biopsy for cancer. Clin Chem 2015 Jan;61(1):112-123.

(44) Jamal-Hanjani M, Quezada SA, Larkin J, Swanton C. Translational implications of tumor heterogeneity. Clin Cancer Res 2015 Mar 15;21(6):1258-1266.

(45) Soultati A, Stares M, Swanton C, Larkin J, Turajlic S. How should clinicians address intratumour heterogeneity in clear cell renal cell carcinoma? Curr Opin Urol 2015 Sep;25(5):358-366.

(46) Gerlinger M, Rowan A, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. N Engl J Med 2012 03/08;366:883-92.

(47) Gerlinger M, Horswell S, Larkin J, Rowan A, Salm M, Varela I, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. Nat Genet 2014 02/02;46.

(48) Ricketts CJ, Linehan WM. Intratumoral heterogeneity in kidney cancer. Nat Genet 2014 Mar;46(3):214-215.

(49) 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 Apr 19;173(3):581-594.e12.

(50) Yong E. Cancer biomarkers: Written in blood. Nature 2014 Jul 31;511(7511):524-526.

(51) Smith CG, Moser T, Mouliere F, Field-Rayner J, Eldridge M, Riediger AL, et al. Comprehensive characterization of cell-free tumor DNA in plasma and urine of patients with renal tumors. Genome Med 2020 Feb 28;12(1):23-020-00723-8.

(52) Farber NJ, Kim CJ, Modi PK, Hon JD, Sadimin ET, Singer EA. Renal cell carcinoma: the search for a reliable biomarker. Transl Cancer Res 2017 Jun;6(3):620-632.

(53) 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 Jan;45(2):228-247.

(54) Alix-Panabières C, Pantel K. Clinical Applications of Circulating Tumor Cells and Circulating Tumor DNA as Liquid Biopsy. Cancer Discov 2016 May;6(5):479-491.

(55) Bardelli A, Pantel K. Liquid Biopsies, What We Do Not Know (Yet). Cancer Cell 2017 Feb 13;31(2):172-179.

(56) Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating liquid biopsies into the management of cancer. Nat Rev Clin Oncol 2017 Sep;14(9):531-548.

(57) Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. Nat Rev Cancer 2017 Apr;17(4):223-238.

(58) Heitzer E, Haque IS, Roberts CES, Speicher MR. Current and future perspectives of liquid biopsies in genomics-driven oncology. Nat Rev Genet 2019 Feb;20(2):71-88.

(59) De Rubis G, Rajeev Krishnan S, Bebawy M. Liquid Biopsies in Cancer Diagnosis, Monitoring, and Prognosis. Trends Pharmacol Sci 2019 Mar;40(3):172-186.

(60) Heitzer E, Perakis S, Geigl JB, Speicher MR. The potential of liquid biopsies for the early detection of cancer. NPJ Precis Oncol 2017 Oct 17;1(1):36-017-0039-5. eCollection 2017.

(61) Murtaza M, Dawson SJ, Pogrebniak K, Rueda OM, Provenzano E, Grant J, et al. Multifocal clonal evolution characterized using circulating tumour DNA in a case of metastatic breast cancer. Nat Commun 2015 Nov 4;6:8760.

(62) Tie J, Wang Y, Tomasetti C, Li L, Springer S, Kinde I, et al. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. Sci Transl Med 2016 Jul 6;8(346):346ra92.

(63) Abbosh C, Birkbak N, Wilson G, Jamal-Hanjani M, Constantin T, Salari R, et al. Phylogenetic ctDNA analysis depicts early stage lung cancer evolution. Nature 2017 04/26;545.

(64) Reinert T, Henriksen T, Christensen E, Sharma S, Salari R, Sethi H, et al. Analysis of Plasma Cell-Free DNA by Ultradeep Sequencing in Patients With Stages I to III Colorectal Cancer. JAMA Oncology 2019 05/09;5.

(65) McDonald BR, Contente-Cuomo T, Sammut SJ, Odenheimer-Bergman A, Ernst B, Perdigones N, et al. Personalized circulating tumor DNA analysis to detect residual disease after neoadjuvant therapy in breast cancer. Sci Transl Med 2019 Aug 7;11(504):eaax7392. doi: 10.1126/scitranslmed.aax7392.

(66) Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 2013 Mar 28;368(13):1199-1209.

(67) Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014 Feb 19;6(224):224ra24.

(68) Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med 2008 Sep;14(9):985-990.

(69) Forshew T, Murtaza M, Parkinson C, Gale D, Tsui DW, Kaper F, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 2012 May 30;4(136):136ra68.

(70) Gray E, Rizos H, Reid A, Boyd S, Pereira M, Lo J, et al. Circulating tumor DNA to monitor treatment response and detect acquired resistance in patients with metastatic melanoma. Oncotarget 2015 07/19;6.

(71) Montagut C, Siravegna G, Bardelli A. Liquid biopsies to evaluate early therapeutic response in colorectal cancer. Ann Oncol 2015 Aug;26(8):1525-1527.

(72) Imamura F, Uchida J, Kukita Y, Kumagai T, Nishino K, Inoue T, et al. Monitoring of treatment responses and clonal evolution of tumor cells by circulating tumor DNA of heterogeneous mutant EGFR genes in lung cancer. Lung Cancer 2016 Apr;94:68-73.

(73) Murtaza M, Dawson SJ, Tsui DW, Gale D, Forshew T, Piskorz AM, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature 2013 May 2;497(7447):108-112.

(74) Mohan S, Heitzer E, Ulz P, Lafer I, Lax S, Auer M, et al. Changes in colorectal carcinoma genomes under anti-EGFR therapy identified by whole-genome plasma DNA sequencing. PLoS Genet 2014 Mar 27;10(3):e1004271.

(75) Ulz P, Belic J, Graf R, Auer M, Lafer I, Fischereder K, et al. Whole-genome plasma sequencing reveals focal amplifications as a driving force in metastatic prostate cancer. Nat Commun 2016 Jun 22;7:12008.

(76) Zheng D, Ye X, Zhang MZ, Sun Y, Wang JY, Ni J, et al. Plasma EGFR T790M ctDNA status is associated with clinical outcome in advanced NSCLC patients with acquired EGFR-TKI resistance. Sci Rep 2016 Feb 12;6:20913.

(77) Ou SI, Nagasaka M, Zhu VW. Liquid Biopsy to Identify Actionable Genomic Alterations. Am Soc Clin Oncol Educ Book 2018 May 23;38:978-997.

(78) Oikkonen J, Zhang K, Salminen L, Schulman I, Lavikka K, Andersson N, et al. Prospective Longitudinal ctDNA Workflow Reveals Clinically Actionable Alterations in Ovarian Cancer. JCO Precision Oncology 2019 12/01; 2020/08(3):1-12.

(79) Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV, Liu MC, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Annals of Oncology 2020 June 2020;31(6):745-759.

(80) Lennon AM, Buchanan AH, Kinde I, Warren A, Honushefsky A, Cohain AT, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. Science 2020 Jul 3;369(6499):10.1126/science.abb9601. Epub 2020 Apr 28.

(81) Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer Res 2001 Feb 15;61(4):1659-1665.

(82) Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P. About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. Clin Chim Acta 2001 Nov;313(1-2):139-142.

(83) van der Vaart M, Pretorius PJ. The origin of circulating free DNA. Clin Chem 2007 Dec;53(12):2215.

(84) Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. Nat Rev Cancer 2011 Jun;11(6):426-437.

(85) Thakur BK, Zhang H, Becker A, Matei I, Huang Y, Costa-Silva B, et al. Double-stranded DNA in exosomes: a novel biomarker in cancer detection. Cell Res 2014 Jun;24(6):766-769.

(86) Thierry AR, El Messaoudi S, Gahan PB, Anker P, Stroun M. Origins, structures, and functions of circulating DNA in oncology. Cancer Metastasis Rev 2016 Sep;35(3):347-376.

(87) Kahlert C, Melo SA, Protopopov A, Tang J, Seth S, Koch M, et al. Identification of double-stranded genomic DNA spanning all chromosomes with mutated KRAS and p53 DNA in the serum exosomes of patients with pancreatic cancer. J Biol Chem 2014 Feb 14;289(7):3869-3875.

(88) Lo YM, Chan KC, Sun H, Chen EZ, Jiang P, Lun FM, et al. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. Sci Transl Med 2010 Dec 8;2(61):61ra91.

(89) Ramachandran S, Henikoff S. Replicating Nucleosomes. Sci Adv 2015 Aug 7;1(7):e1500587. doi: 10.1126/sciadv.1500587.

(90) Ulz P, Thallinger GG, Auer M, Graf R, Kashofer K, Jahn SW, et al. Inferring expressed genes by whole-genome sequencing of plasma DNA. Nat Genet 2016 Oct;48(10):1273-1278.

(91) Jiang P, Chan CW, Chan KC, Cheng SH, Wong J, Wong VW, et al. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. Proc Natl Acad Sci U S A 2015 Mar 17;112(11):E1317-25.

(92) Jiang P, Lo YMD. The Long and Short of Circulating Cell-Free DNA and the Ins and Outs of Molecular Diagnostics. Trends Genet 2016 Jun;32(6):360-371.

(93) Heitzer E, Auer M, Hoffmann EM, Pichler M, Gasch C, Ulz P, et al. Establishment of tumor-specific copy number alterations from plasma DNA of patients with cancer. Int J Cancer 2013 Jul 15;133(2):346-356.

(94) Mouliere F, Robert B, Arnau Peyrotte E, Del Rio M, Ychou M, Molina F, et al. High fragmentation characterizes tumour-derived circulating DNA. PLoS One 2011;6(9):e23418.

(95) Underhill HR, Kitzman JO, Hellwig S, Welker NC, Daza R, Baker DN, et al. Fragment Length of Circulating Tumor DNA. PLoS Genet 2016 Jul 18;12(7):e1006162.

(96) Mouliere F, Chandrananda D, Piskorz AM, Moore EK, Morris J, Ahlborn LB, et al. Enhanced detection of circulating tumor DNA by fragment size analysis. Sci Transl Med 2018 Nov 7;10(466):10.1126/scitranslmed.aat4921.

(97) Heidary M, Auer M, Ulz P, Heitzer E, Petru E, Gasch C, et al. The dynamic range of circulating tumor DNA in metastatic breast cancer. Breast Cancer Res 2014 Aug 9;16(4):421-014-0421-y.

(98) Zill OA, Banks KC, Fairclough SR, Mortimer SA, Vowles JV, Mokhtari R, et al. The Landscape of Actionable Genomic Alterations in Cell-Free Circulating Tumor DNA from 21,807 Advanced Cancer Patients. Clin Cancer Res 2018 Aug 1;24(15):3528-3538.

(99) Perakis S, Speicher MR. Emerging concepts in liquid biopsies. BMC Med 2017 Apr 6;15(1):75-017-0840-6.

(100) Zhou Q, Moser T, Perakis S, Heitzer E. Untargeted profiling of cell-free circulating DNA. Translational Cancer Research 2017;7(2).

(101) El-Heliebi A, Heitzer E. Chapter 5 - State of the Art and Future Direction for the Analysis of Cell-Free Circulating DNA. In: Filice M, Ruiz-Cabello J, editors. Nucleic Acid Nanotheranostics: Elsevier; 2019. p. 133-188.

(102) van der Pol Y, Mouliere F. Toward the Early Detection of Cancer by Decoding the Epigenetic and Environmental Fingerprints of Cell-Free DNA. Cancer Cell 2019 Oct 14;36(4):350-368.

(103) Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. Nature 2019 Jun;570(7761):385-389.

(104) Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med 2014 May;20(5):548-554.

(105) Newman AM, Lovejoy AF, Klass DM, Kurtz DM, Chabon JJ, Scherer F, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. Nat Biotechnol 2016 May;34(5):547-555.

(106) Wan JCM, Heider K, Gale D, Murphy S, Fisher E, Mouliere F, et al. ctDNA monitoring using patient-specific sequencing and integration of variant reads. Sci Transl Med 2020 Jun 17;12(548):10.1126/scitranslmed.aaz8084.

(107) Zhou Q, Perakis SO, Ulz P, Mohan S, Riedl JM, Talakic E, et al. Cell-free DNA analysis reveals POLR1D-mediated resistance to bevacizumab in colorectal cancer. Genome Med 2020 Feb 22;12(1):20-020-0719-6.

(108) Adalsteinsson VA, Ha G, Freeman SS, Choudhury AD, Stover DG, Parsons HA, et al. Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. Nat Commun 2017 Nov 6;8(1):1324-017-00965-y.

(109) Heitzer E, Auinger L, Speicher MR. Cell-Free DNA and Apoptosis: How Dead Cells Inform About the Living. Trends Mol Med 2020 May;26(5):519-528.

(110) Jiang P, Chan KCA, Lo YMD. Liver-derived cell-free nucleic acids in plasma: Biology and applications in liquid biopsies. J Hepatol 2019 Aug;71(2):409-421.

(111) Gansauge MT, Meyer M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nat Protoc 2013 Apr;8(4):737-748.

(112) Burnham P, Kim MS, Agbor-Enoh S, Luikart H, Valantine HA, Khush KK, et al. Single-stranded DNA library preparation uncovers the origin and diversity of ultrashort cell-free DNA in plasma. Sci Rep 2016 Jun 14;6:27859.

(113) Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. Cell 2016 Jan 14;164(1-2):57-68.

(114) Parkinson CA, Gale D, Piskorz AM, Biggs H, Hodgkin C, Addley H, et al. Exploratory Analysis of TP53 Mutations in Circulating Tumour DNA as Biomarkers of Treatment Response for Patients with Relapsed High-Grade Serous Ovarian Carcinoma: A Retrospective Study. PLoS Med 2016 Dec 20;13(12):e1002198.

(115) Ball MW, Gorin MA, Guner G, Pierorazio PM, Netto G, Paller CJ, et al. Circulating Tumor DNA as a Marker of Therapeutic Response in Patients With Renal Cell Carcinoma: A Pilot Study. Clin Genitourin Cancer 2016 Oct;14(5):e515-e520.

(116) Pal SK, Sonpavde G, Agarwal N, Vogelzang NJ, Srinivas S, Haas NB, et al. Evolution of Circulating Tumor DNA Profile from First-line to Subsequent Therapy in Metastatic Renal Cell Carcinoma. Eur Urol 2017 Oct;72(4):557-564.

(117) Nuzzo PV, Berchuck JE, Korthauer K, Spisak S, Nassar AH, Abou Alaiwi S, et al. Detection of renal cell carcinoma using plasma and urine cell-free DNA methylomes. Nat Med 2020 Jul;26(7):1041-1043.

(118) Patel KM, van der Vos KE, Smith CG, Mouliere F, Tsui D, Morris J, et al. Association Of Plasma And Urinary Mutant DNA With Clinical Outcomes In Muscle Invasive Bladder Cancer. Sci Rep 2017 Jul 17;7(1):5554-017-05623-3.

(119) Dudley JC, Schroers-Martin J, Lazzareschi DV, Shi WY, Chen SB, Esfahani MS, et al. Detection and Surveillance of Bladder Cancer Using Urine Tumor DNA. Cancer Discov 2019 Apr;9(4):500-509.

(120) Moser T, Ulz P, Zhou Q, Perakis S, Geigl JB, Speicher MR, et al. Single-Stranded DNA Library Preparation Does Not Preferentially Enrich Circulating Tumor DNA. Clin Chem 2017 Oct;63(10):1656-1659.

(121) Belic J, Graf R, Bauernhofer T, Cherkas Y, Ulz P, Waldispuehl-Geigl J, et al. Genomic alterations in plasma DNA from patients with metastasized prostate cancer receiving abiraterone or enzalutamide. Int J Cancer 2018 Sep 1;143(5):1236-1248.

(122) Broad Institute. Picard. n.d.; Available at: http://broadinstitute.github.io/picard/. Accessed August, 2020.

(123) Belic J, Koch M, Ulz P, Auer M, Gerhalter T, Mohan S, et al. Rapid Identification of Plasma DNA Samples with Increased ctDNA Levels by a Modified FAST-SeqS Approach. Clin Chem 2015 Jun;61(6):838-849.

(124) Ha G. ichorCNA - Parameter tuning and settings. 2018; Available at: https://github.com/broadinstitute/ichorCNA. Accessed August, 2020.

(125) Xu C, Nezami Ranjbar MR, Wu Z, DiCarlo J, Wang Y. Detecting very low allele fraction variants using targeted DNA sequencing and a novel molecular barcode-aware variant caller. BMC Genomics 2017 Jan 3;18(1):5-016-3425-4.

(126) Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009 Aug 15;25(16):2078-2079.

(127) Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013 Mar;31(3):213-219.

(128) Scheinin I, Sie D, Bengtsson H, van de Wiel MA, Olshen AB, van Thuijl HF, et al. DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. Genome Res 2014 Dec;24(12):2022-2032.

(129) Chandrananda D. CNAclinic. 2017; Available at: https://github.com/sdchandra/CNAclinic. Accessed August, 2020.

(130) Chandrananda D. tMAD. 2018; Available at: https://github.com/sdchandra/tMAD. Accessed August, 2020.

(131) Gates C. Connor. 2018; Available at: https://github.com/umich-brcf-bioinf/Connor. Accessed August, 2020.

(132) Suppan C, Brcic I, Tiran V, Mueller HD, Posch F, Auer M, et al. Untargeted Assessment of Tumor Fractions in Plasma for Monitoring and Prognostication from Metastatic Breast Cancer Patients Undergoing Systemic Treatment. Cancers (Basel) 2019 Aug 14;11(8):10.3390/cancers11081171.

(133) Mouliere F, Mair R, Chandrananda D, Marass F, Smith CG, Su J, et al. Detection of cell-free DNA fragmentation and copy number alterations in cerebrospinal fluid from glioma patients. EMBO Mol Med 2018 Dec;10(12):10.15252/emmm.201809323.

(134) Togneri FS, Ward DG, Foster JM, Devall AJ, Wojtowicz P, Alyas S, et al. Genomic complexity of urothelial bladder cancer revealed in urinary cfDNA. Eur J Hum Genet 2016 Aug;24(8):1167-1174.

(135) Cancer Research UK. Statistics by cancer type. 2017; Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/kidney-cancer.. Accessed August, 2020.

(136) Rossi SH, Klatte T, Usher-Smith J, Stewart GD. Epidemiology and screening for renal cancer. World J Urol 2018 Sep;36(9):1341-1353.

(137) Cohen HT, McGovern FJ. Renal-cell carcinoma. N Engl J Med 2005 Dec 8;353(23):2477-2490.

(138) Arsanious A, Bjarnason GA, Yousef GM. From bench to bedside: current and future applications of molecular profiling in renal cell carcinoma. Mol Cancer 2009 Mar 17;8:20-4598-8-20.

(139) Graham J, Heng DYC, Brugarolas J, Vaishampayan U. Personalized Management of Advanced Kidney Cancer. Am Soc Clin Oncol Educ Book 2018 May 23;38:330-341.

(140) Bedard PL, Hansen AR, Ratain MJ, Siu LL. Tumour heterogeneity in the clinic. Nature 2013 Sep 19;501(7467):355-364.

(141) Ramon Y Cajal S, Sese M, Capdevila C, Aasen T, De Mattos-Arruda L, Diaz-Cano SJ, et al. Clinical implications of intratumor heterogeneity: challenges and opportunities. J Mol Med (Berl) 2020 Feb;98(2):161-177.

(142) Leary RJ, Sausen M, Kinde I, Papadopoulos N, Carpten JD, Craig D, et al. Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing. Sci Transl Med 2012 Nov 28;4(162):162ra154.

(143) Heitzer E, Ulz P, Belic J, Gutschi S, Quehenberger F, Fischereder K, et al. Tumor-associated copy number changes in the circulation of patients with prostate cancer identified through whole-genome sequencing. Genome Med 2013 Apr 5;5(4):30.

(144) Diaz LA,Jr, Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. J Clin Oncol 2014 Feb 20;32(6):579-586.

(145) Vong JSL, Tsang JCH, Jiang P, Lee WS, Leung TY, Chan KCA, et al. Single-Stranded DNA Library Preparation Preferentially Enriches Short Maternal DNA in Maternal Plasma. Clin Chem 2017 May;63(5):1031-1037.

(146) Heitzer E, Speicher MR. One size does not fit all: Size-based plasma DNA diagnostics. Sci Transl Med 2018 Nov 7;10(466):10.1126/scitranslmed.aav3873.

(147) Liu X, Lang J, Li S, Wang Y, Peng L, Wang W, et al. Fragment Enrichment of Circulating Tumor DNA With Low-Frequency Mutations. Front Genet 2020 Feb 28;11:147.

(148) Ascenti G, Gaeta M, Magno C, Mazziotti S, Blandino A, Melloni D, et al. Contrast-enhanced second-harmonic sonography in the detection of pseudocapsule in renal cell carcinoma. AJR Am J Roentgenol 2004 Jun;182(6):1525-1530.

(149) Roy CS, El Ghali S, Buy X, Lindner V, Lang H, Saussine C, et al. Significance of the pseudocapsule on MRI of renal neoplasms and its potential application for local staging: a retrospective study. AJR Am J Roentgenol 2005 Jan;184(1):113-120.

(150) Mair R, Mouliere F, Smith CG, Chandrananda D, Gale D, Marass F, et al. Measurement of Plasma Cell-Free Mitochondrial Tumor DNA Improves Detection of Glioblastoma in Patient-Derived Orthotopic Xenograft Models. Cancer Res 2019 Jan 1;79(1):220-230.

(151) Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science 2018 Feb 23;359(6378):926-930.

(152) Shen SY, Singhania R, Fehringer G, Chakravarthy A, Roehrl MHA, Chadwick D, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. Nature 2018 Nov;563(7732):579-583.

(153) Wang Y, Springer S, Mulvey CL, Silliman N, Schaefer J, Sausen M, et al. Detection of somatic mutations and HPV in the saliva and plasma of patients with head and neck squamous cell carcinomas. Sci Transl Med 2015 Jun 24;7(293):293ra104.

(154) Yamamoto Y, Uemura M, Fujita M, Maejima K, Koh Y, Matsushita M, et al. Clinical significance of the mutational landscape and fragmentation of circulating tumor DNA in renal cell carcinoma. Cancer Sci 2019 Feb;110(2):617-628.

(155) Fenner A. Kidney cancer: Tracking and therapy selection using ctDNA. Nat Rev Urol 2017 Jul;14(7):389.

(156) Maia MC, Bergerot PG, Dizman N, Hsu J, Jones J, Lanman RB, et al. Association of Circulating Tumor DNA (ctDNA) Detection in Metastatic Renal Cell Carcinoma (mRCC) with Tumor Burden. Kidney Cancer 2017 Jul 26;1(1):65-70.

APPENDIX

Supplementary tables

Supplementary Table 1: Patient characteristics of the MonReC cohort.

[Table and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

MonReC ID	Surgery type	Days between plasma and surgery/diagnosis	Tumor Type	Tumor Stage	Tumor Size (cm)	Fuhrman Grade	Diagnosed with metastases	Status at the time of first blood draw
K01	RN	427	ccRCC	pT3aN1M1	8	G2	Yes	Metastases in bone and LN
K02	NN	12	ccRCC	NA	NA	NA	Yes	No nephrectomy; metastases in lung and LN
K03	NN	NA	ccRCC	pT3aN0M1	7	G2	Yes	No nephrectomy; bone metastases
K04	NN	1	ccRCC	pT4N0M0	8.5	G3	Yes	No nephrectomy; metastases in lungs, liver, bone and LN
K05	RN	-6	NA	pT4N0M1	15	G4	Yes	Prior nephrectomy; lung metastases
K06	RN	-32	pRCC	pT3aN0M1	11	G4	Yes	Prior nephrectomy; lung metastases
K07	RN	-12	chRCC	pT3aN0M1	6.5	G2	Yes	Prior nephrectomy; lung metastases
K08	RN	6341	ccRCC	pT1aN0M0	7.5	NA	No	Metastases in lung, LN and bone
K10	RN	2	chRCC	pT3aN0M0	8	G4	No	No metastases at first blood draw
K11	RN	-1	ccRCC	pT3aN1M1	10.5	G4	Yes	Prior nephrectomy; metastases in lungs, bone and LNs

K12	NN	NA	NA	NA	NA	NA	No	No metastases at first blood draw, surgery denied
K13	NN	NA	pRCC	NA	8	G2	Yes	CT punctured biopsy, metastases in lung, bone and LN
K14	RN	4441	ccRCC	pT1aN0M0	NA	G1	No	Metastases in lung
K15	RN	71	ccRCC	pT1bNxMx	NA	G2	NA	Metastases in the lungs and target metastases in the liver
K16	RN	4167	ccRCC	NA	NA	G2	NA	Relapsed tumors in the right kidney
K17	RN	8230	ccRCC	NA	NA	NA	NA	Metastases in the lungs
K18	RN	111	ccRCC	pT2bNxMx	NA	G4	NA	Metastases in lungs and liver
K19	RN	1977	ccRCC	NA	NA	G2	NA	Metastases in lungs, bones and sternum
K20	RN	399	ccRCC	pT3aN0Mx	NA	G3	NA	Metastases in lungs
K21	RN	36	ccRCC	pT3aNxM1	NA	G4	Yes	Metastasis in the brain, thoracic lymph nodes and retroperitoneum
K22	RN	1080	ccRCC	pT3bN1M1	NA	G2	Yes	Thoracic lymphadenopathy
K23	RN	1943	ccRCC	pT1bNxMx	NA	NA	NA	Lymph nodes
K24	RN	520	ccRCC	pT3bNxMx	NA	G3	NA	Metastases in lung and brain
K25	RN	5617	ccRCC	pT1NxMx	NA	G1	NA	Metastases in the lungs and pelvic bones
K26	RN	1008	ccRCC	pT3aNxMx	NA	G3	NA	Second tumor in left kidney (ccRCC, G2, pT1aNxMx), lung metastases

K27	RN	50	ccRCC	pT1aNxM1	NA	G4	Yes	Bone metastases
K28	RN	834	ccRCC	pT3aNxMx	NA	G3	NA	Cutaneous metastasis
K29	RN	367	ccRCC	pT3aNxMx	NA	G4	NA	Metastases in lungs and LN
K30	RN	659	ccRCC	pT3bNxMx	NA	G4	NA	Metastases in lungs, local relapse of the kidney
K31	RN	86	ccRCC	pT3bN0Mx	NA	G2	NA	Metastases in lungs and new metastases in the spine
K32	RN	3383	ccRCC	pT1bNxMx	NA	G2	NA	Metastases in pancreas and liver
K33	NA	NA	NA	NA	NA	NA	NA	NA
K34	NA	NA	ccRCC	NA	NA	NA	NA	Metastases in LN and bone
K35	RN	8023	ccRCC	NA	5.5	G1	No	Several metastases in bone, LN, lung, suprarenal gland, and recurrence in right kidney
K36	RN	229	ccRCC	pT3N0M1	NA	G4	NA	NA
K37	RN	333	ccRCC	pT1	NA	G1	NA	NA
K38	RN	3109	ccRCC	pT1	NA	G1	NA	NA
K39	RN	56	ccRCC	pT3a	7	G3	Yes	NA
K40	RN	120	ccRCC	pT2a	NA	G4	NA	Metastases in lung and omentum
K41	RN	768	ccRCC	pT3a	NA	G2	NA	NA

K42	NA	NA	ccRCC	pT3M1	NA	G2	Yes	Metastases in bone and lung, secondary tumor in prostate
K43	RN	129	ccRCC	pT3bN1M1	9	G3	Yes	Metastases in liver and lung
K44	RN	5342	pRCC	pT1aN0M0	1.3	G3	No	Metastases in lungs and LN

RN = radical nephrectomy, ccRCC = clear cell RCC, chRCC = chromophobe RCC, pRCC = papillary RCC, LN = lymph nodes, NA = not available, NN= no nephrectomy. Days between plasma and surgery indicate the time between fluid collection and nephrectomy, e.g., 36 indicates that the samples were collected 36 days after treatment (51).

Supplementary Table 2: Comparison of overall ctDNA fraction between dsDNA and ssDNA libraries.

Patient ID	Library construction	ichorCNA tumor fraction [%]	
		all fragments	*in silico* size selection < 150bp
B7_2	ssDNA	28.57	28.9
	dsDNA	29.67	30.2
B13_1	ssDNA	79.66	78.16
	dsDNA	80.69	80.35
C177_2	ssDNA	10.84	10.33
	dsDNA	8.34	12.25
C177_3	ssDNA	46.78	46.62
	dsDNA	48.02	46.2
P226_4	ssDNA	7.54	8.91
	dsDNA	7.30	7.20
P226_6	ssDNA	24.82	28.79
	dsDNA	25.64	26.82

Supplementary Table 3: Summary of MonReC baseline samples with ctDNA positivity.

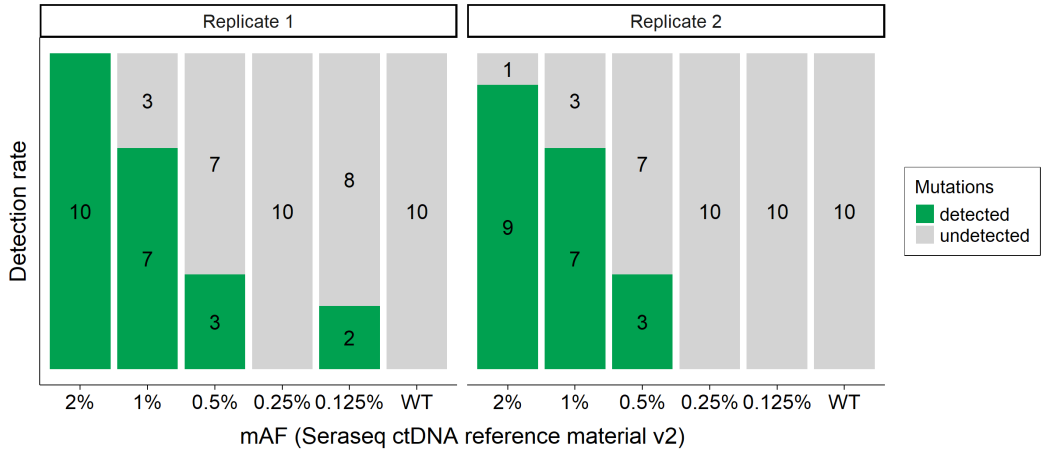
[Table and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

ID	Mutation [%]	ichorCNA tumor fraction [%]			tMAD			mFAST-SeqS
	Mean mAF	all fragments	subsampled	size selection 90-150bp	10M 30kbp vs Control	2M 500kbp vs Control	size selection 90-150bp	z-score
K02_1	ND	0.0	2.2	4.3	0.005	0.012	0.015	1.6
K05_1	7.1	5.5	5.9	8.7	0.010	0.015	0.030	0.3
K08_1	ND	0.0	2.1	4.7	0.006	0.011	NA	0.1
K11_1	ND	0.0	1.9	4.4	0.009	0.011	0.013	2.4
K12_1	15.1	12.5	2.6	10.8	0.019	0.021	0.028	3.7
K13_1	ND	0.0	1.7	4.3	0.006	0.012	NA	1.8
K18_1	5.1	11.5	11.1	16.1	0.031	0.033	0.054	1.5
K19_1	ND	0.0	0.0	3.9	0.006	0.011	0.010	0.7
K20_1	8.7	5.1	5.0	9.9	0.017	0.021	0.036	-0.6
K21_1	ND	5.4	5.9	21.6	0.016	0.017	0.062	-0.4
K23_1	5.1	0.0	0.0	5.3	0.006	0.011	NA	-0.4
K27_1	4.9	8.1	8.3	11.5	0.025	NA	0.037	1.0
K39_1	14.2	17.2	15.5	23.1	0.090	0.078	0.147	3.3
K40_1	ND	4.0	4.5	5.6	0.016	0.016	NA	1.0
K42_1	3.4	0.0	0.0	2.4	0.007	0.007	NA	1.5
K44_1	ND	1.2	3.3	6.8	0.011	0.018	NA	0.1

ID	Mutation detected*	SCNA detected	tMAD >0.015	mFAST-SeqS z-score >3
K02_1	FALSE	FALSE	TRUE	FALSE
K05_1	TRUE	TRUE	TRUE	FALSE
K08_1	FALSE	TRUE	FALSE	FALSE
K11_1	FALSE	TRUE	FALSE	FALSE
K12_1	TRUE	TRUE	TRUE	TRUE
K13_1	FALSE	TRUE	FALSE	FALSE
K18_1	TRUE	TRUE	TRUE	FALSE
K19_1	FALSE	TRUE	FALSE	FALSE
K20_1	TRUE	TRUE	TRUE	FALSE
K21_1	FALSE	TRUE	TRUE	FALSE
K23_1	TRUE	TRUE	FALSE	FALSE
K27_1	TRUE	TRUE	TRUE	FALSE
K39_1	TRUE	TRUE	TRUE	TRUE
K40_1	FALSE	TRUE	TRUE	FALSE
K42_1	TRUE	FALSE	FALSE	FALSE
K44_1	FALSE	TRUE	FALSE	FALSE

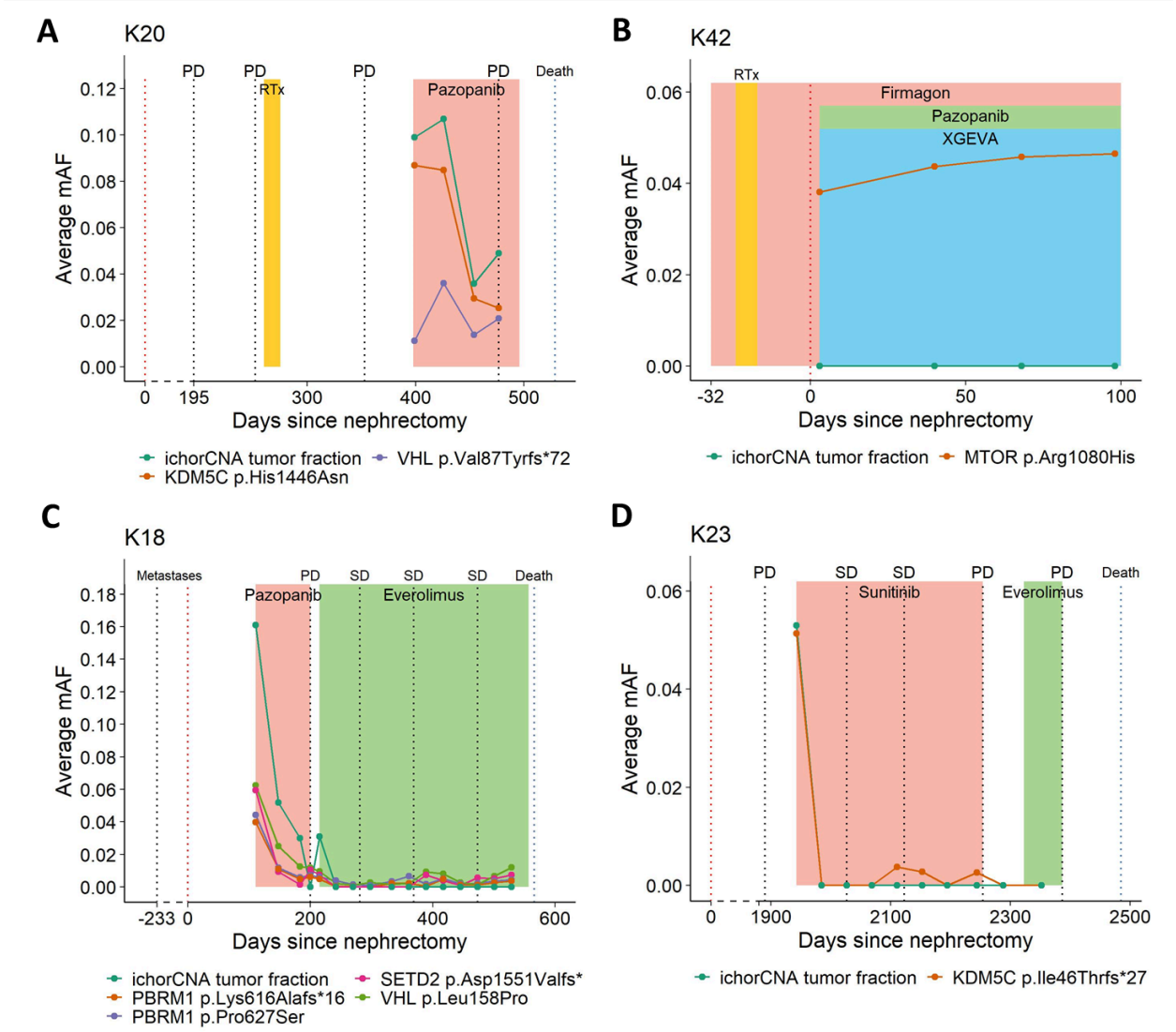
* Mutation detected using a custom QIASeq panel covering 10 frequently mutated genes

Supplementary figures



Supplementary Figure 1: Sensitivity assessment of the QIAseq mutation analysis kit.

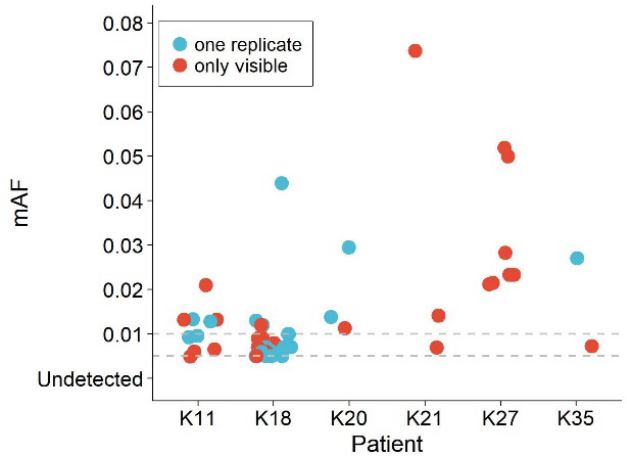
The detection rates of all mutations (n=10) covered by the panel in the Seraseq ctDNA Reference Material v2 at defined mAFs are shown. Parallel analysis of two replicates revealed similar findings. Due to the detection of variants (3/10) down to a mAF of 5×10^{-3}, the maximum achievable sensitivity of this assay was set to 5×10^{-3}. The wildtype (WT) sample clearly showed in both replicates a negative result. The detection rates of the covered mutations (y-axis) at various mAFs (x-axis) and the number of mutations called by the assay are displayed. Detected and missed mutations are indicated by green and grey, respectively.



Supplementary Figure 2: Longitudinal monitoring of ctDNA mAF dynamics.

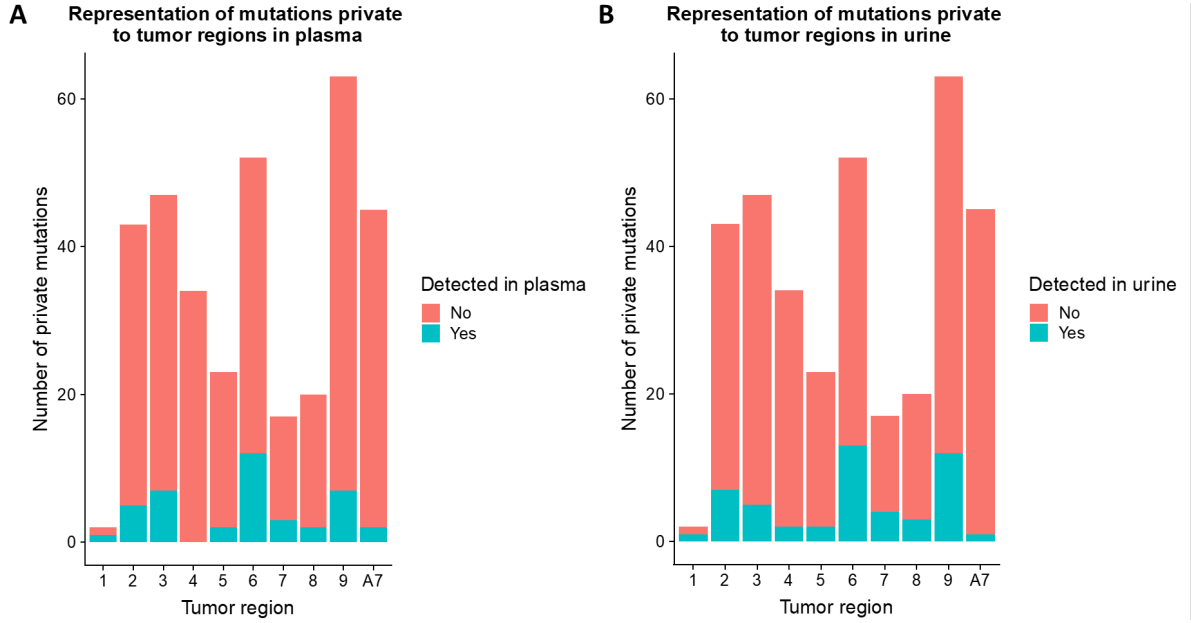
[Figure and legend adapted from Genome Medicine (Smith & Moser et al., 2020)]

(A-D) Plots demonstrating multiple measurements of ctDNA over time. In each panel, the time of diagnosis (red) and the time of patients' death (blue) are indicated as vertical dotted lines. SD denotes stable disease, PD progressive disease, and PR partial response. The ichorCNA tumor fraction is shown in green, and the respective somatic mutations detected by the QIAseq panel are displayed in various colors. Therapeutic interventions are shown in multiple colors (RTx= radiotherapy). Plots for patients K11, K21, K27, K35, and K39 are provided in Figure 20.



Supplementary Figure 3: Accurate detection of response-specific ctDNA mAF dynamics is affected by the performance of the variant caller.

Summary of mutations even with highly confident variant calls (≥ 0.01 mAF; dotted line) which were missed by the variant mutation caller. Many of the mutations were only present in one replicate (displayed in blue) and were therefore filtered, and in some follow-up plasma samples, mutations were clearly detected by scrutiny of the sequencing data but not called by the recommended variant caller (depicted in red). Horizontal dotted line at 0.005 displays the analytical sensitivity of the QIAseq panel.



Supplementary Figure 4: Mutations private to individual tumor regions and their representation in liquid biopsy fluids.

[Figure and legend originally published in Genome Medicine (Smith & Moser et al., 2020)]

(A) Representation of mutations private to different tumor regions in plasma. All but one region had at least one private mutation called in this compartment.

(B) As for plot A but for urine. All regions had at least one private mutation called in urine.