Molecular Markers and Hypoxia in Renal Cell Carcinoma

de Alwis, Roger

2023

Document Version:
Publisher's PDF, also known as Version of record

Molecular Markers & Hypoxia in Renal Cell Carcinoma

Molecular Markers & Hypoxia in Renal Cell Carcinoma

Roger de Alwis

LUND
UNIVERSITY

DOCTORAL DISSERTATION

By due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at the Main Lecture Hall, Medicon Village, Lund on
Wednesday 24th of May, 09.00

Faculty opponent
Associate Professor Alan McIntyre
Faculty of Medicine and Health Sciences
University of Nottingham
England, United Kingdom

Organization: LUND UNIVERSITY

Document name: Doctoral thesis **Date of issue** 2023-05-24

Author(s): Roger de Alwis Sponsoring organization:

Title and subtitle: Molecular markers & Hypoxia in renal cell carcinoma

Abstract:

Renal cell carcinomas (RCCs) are a group of tumours that arise from the nephron within the kidney. Furthermore, the global average age of a RCC patient at diagnosis is 75 and they usually present with co-morbidities that may render invasive surgical/biopsy approaches risky. The most prevalent form of RCC is the clear cell RCC (ccRCC) subtype that displays pseudohypoxic activation of the two transcription factors HIF-1α and HIF-2α as the result of a non-functional pVHL protein.

In our **first paper** we aimed to decipher HIF-2α specific target genes operating in normal renal proximal tubule epithelial cells as well as ccRCC tumour cells. We pharmacologically emulated the loss of functional pVHL in renal proximal tubule cells whilst inhibiting HIF-2α transcriptional activity. Subsequent RNA-sequencing revealed potentially HIF-2α specific genes of which we selected *SEMA5B*, where its protein expression pattern matched our RNA-sequencing findings. We verified the HIF-2α regulatory specificity to *SEMA5B* in ccRCC cell lines, with other lines of evidence definitively demonstrating that HIF-2α but not HIF-1α specifically regulates the expression of *SEMA5B* in renal proximal tubule and ccRCC cells.

In **paper two**, we analysed transcription factor network and regulon activity in RCC subtypes using publicly available datasets. Using this analysis, we identified NFIA, a transcription factor that had similar regulon activity to HNF4A, a well characterised transcription factor in RCC. Based on this data, we examined the relationship between RNA expression of our selected transcription factor NFIA and TCGA-based RCC patient clinicopathological factors. We assessed the protein expression of NFIA and HNF4A in a large tissue-microarray consisting of ccRCC and papillary RCC (pRCC) tumours and found that NFIA expression can independently predict CSS in ccRCC patients.

In **paper three** we developed a workflow to enrich, detect and subtype RCC tumour cells from whole blood. We demonstrated that RCC tumour cells are suitable for this approach and showed that our methodology can isolate down to one spiked-in tumour cell from whole blood. Furthermore, through differential gene expression analyses between the three RCC subtypes, we identified transcriptomic markers that can be used to detect pRCC and ccRCC tumour cells from whole blood.

Key words: Kidney cancer, RCC, Liquid biopsy, CTC, Hypoxia

Classification system and/or index terms (if any) Supplementary bibliographical information

Language **ISSN and key title**: 1652-8220

ISBN: 978-91-8021-409-4

Recipient's notes **Number of pages**: 82

Price Security classification

Signature Date 2023-04-11

Molecular Markers & Hypoxia in Renal Cell Carcinoma

Roger de Alwis

LUND
UNIVERSITY

Coverphoto by Aimee de Alwis

All images created with BioRender.com

Faculty of Medicine

Department of Laboratory Medicine

ISSN 978-91-8021-409-4

ISBN 1652-8220

Printed in Sweden by Media-Tryck, Lund University, Lund 2023

An Ode to my Family

Table of Contents

List of Papers **11**

Abbreviations **13**

Popular Science Summary **15**

Abstract **17**

1 The Kidneys **19**

The Nephron 20

2 Hypoxia **23**

Oxygen Sensing and Hypoxia Response Machinery 23

Other Mechanisms of HIFα stabilisation/regulation 26

Differential regulation of gene expression by HIF-1α and HIF-2α 27

Tumour Hypoxia 28

HIF signalling in RCC: The HIFα Dichotomy 30

3 Kidney Cancer **33**

Renal Cell Carcinoma Subtypes 34

Clear Cell Renal Cell Carcinoma 34

Papillary renal cell carcinoma 36

Chromophobe renal cell carcinoma 37

Clinical Management of RCC 37

Disease manifestation and prognosis 37

Diagnosis and Imaging 40

The Current Landscape of RCC treatment 41

4 Liquid Biopsies in RCC **47**

Liquid biopsy fluids & analytes 47

Circulating Tumour Cells & Enrichment Approaches 48

CTC Applications in RCC 51

Circulating tumour DNA approaches and applications 52

The Present Investigation **55**
Overview and Aims 55
Paper I: SEMA5B is a HIF-2α specific target gene in renal proximal tubule cells 55
Paper II: Identification and validation of NFIA as a novel and prognostic marker in renal cell carcinoma 57
Paper III: Size-based isolation and detection of renal carcinoma cells from whole blood 60
Acknowledgments **63**
References **67**

List of Papers

Paper I

SEMA5B is a HIF-2α specific target gene in renal proximal tubule cells
Roger de Alwis, Sarah Schoch, David Lindgren, Mazharul Islam,
Christina Möller, Håkan Axelson.
Manuscript

Paper II

Identification and validation of NFIA as a novel prognostic marker in renal cell carcinoma
Roger de Alwis*, Sarah Schoch*, Mazharul Islam, Christina Möller,
Börje Ljungberg, Håkan Axelson.
Journal of Molecular Pathology: Clinical Research

Paper III

Size-based isolation and detection of renal carcinoma cells from whole blood
Roger de Alwis, Jennifer Hansson, David Lindgren, Sarah Schoch,
Alexander Tejera, Bianca Scholtz, Peter Elfving, Christina Möller, Helén Nilsson,
Martin Johansson, Håkan Axelson
Molecular and Clinical Oncology

*Denotes equal contribution

Abbreviations

ARE	antioxidant response element
ATP	Adenosine triphosphate
BAP1	BRCA1 associated protein-1
bHLH-PAS	basic helix-loop-helix/Per-ARNT-SIM
CAIX	Carbonic anhydrase IX
CBP	CREB-binding protein
ccRCC	clear cell renal cell carcinoma
CDKN2A	Cyclin dependent kinase inhibitor 2A
chRCC	chromophobe renal cell carcinoma
CSS	cancer-specific survival
CT	computed tomography
CTAD	C-terminal transactivation domain
CTC	circulating tumour cell
CTLA4	cytotoxic T-lymphocyte associated protein 4
CTM	circulating tumour microemboli
CXCR4	CXC motif chemokine receptor 4
EpCAM	Epithelial cell adhesion molecule
ETS1	ETS proto-oncogene 1
FFPE	Formalin fixed paraffin embedded
FOXI-1	Forkhead box I1
H&E	Haematoxylin and eosin
HIFa	Hypoxia inducible factor alpha
HNF4A	Hepatocyte nuclear factor 4A
HRE	Hypoxia response element
HSP90	Heat shock protein 90
ICI	Immune checkpoint inhibitor
ISUP	International society of Urological Pathology

Km	Michaelis-Menten constant
LDHA	Lactate dehydrogenase A
MET	tyrosine-protein kinase met
MRI	Magnetic resonance imaging
NFIA	Nuclear factor I A
NFIB	Nuclear factor I B
NRF2	NF-E2-related transcription factor
NTAD	N-termina transactivation domain
OS	Overall survival
p1RCC	type 1 papillary RCC
p2RCC	type 2 papillary RCC
PAX8	Paired box 8
PBRM1	polybromo 1
PD1	Programmed death 1
PET	positron emission tomography
PFKFB3	6-phosphofructo-2-kinase/fructose-2,6-biphosphotase
PFS	Progression-free survival
PHD	Prolyl hydroxylase domain isoforms
PKM	Pyruvate Kinase Muscle isozyme
pRCC	papillary renal cell carcinoma
PTEN	Phosphatase and tensin homolog
RACK1	Receptor for activated C Kinase 1
RPTEC/TERT1	Renal proximal tubule cells / Telomerase Reverse Transcriptase 1
SCENIC	single-cell regulatory network inference and clustering
SETD2	SET domain-containing 2
SRC-1	Steroid receptor coactivator-1
TCGA	The Cancer Genome Atlas
TIF-1	transcriptional intermediary factor 1
TKI	Tyrosine kinase inhibitor
TMA	Tissue microarray
TP53	tumour suppressor protein 53
US	Ultrasound
VEGF	Vascular endothelial growth factor
VHL	von-Hippel Lindau tumour-suppressor protein

Popular Science Summary

Renal cell carcinoma (RCC) is a type of kidney cancer that typically grows slowly and does not show symptoms until it has progressed to an advanced stage. RCC is more common in older individuals and is often associated with other health conditions, making invasive treatment or biopsy approaches risky. In the studies laid out in this thesis, we address the clinical challenges associated with RCC by investigating its biology. We identify and use molecular traits of RCC tissue to predict which patients may have worse survival. We also use these traits to detect RCC cells from whole blood.

The first paper describes how we identified a specific gene called *SEMA5B* that is regulated by a protein named HIF-2α in normal and cancerous kidney cells. SEMA5B may play a role in the development of blood vessels and the local cellular environment in which RCCs grow.

In the second paper, we analysed publicly available data to identify a protein called NFIA that could predict clinical outcomes in patients with the clear cell variety of RCC. We found that the presence and amount of NFIA could independently predict survival outcomes of RCC patients.

In the third paper, we developed a new method to isolate and detect RCC tumour cells in the blood. Tumours often shed their cells into the blood stream (termed circulating tumour cells, CTCs) and finding and identifying these CTCs can provide valuable information about a tumour, without having to surgically obtain tumour tissue. We used a size-based approach to isolate differentiative RCC tumour cells from other blood cells like immune cells, since the latter cell types are much smaller than the former. Furthermore, we identified molecular traits that could distinguish different varieties of RCC, and our methods could isolate tumour cells from as few as one tumour cell in 7.5ml of whole blood. This method may offer a new way to monitor RCC patients in scenarios where patients have a large or spread tumour.

Overall, these findings may help to improve the clinical management of RCC by identifying new targets for treatment, developing molecular markers for prognostication, and enabling non-invasive monitoring of the disease.

Abstract

Renal cell carcinomas (RCCs) are a group of tumours that arise from the nephron within the kidney. They are characterised by an indolent growth pattern and do not display overt clinical symptoms in patients with early to locally advanced tumours. Furthermore, the global average age of a RCC patient at diagnosis is 75 and they usually present with co-morbidities that may render invasive surgical/biopsy approaches risky. The most prevalent form of RCC is the clear cell RCC (ccRCC) subtype that displays pseudohypoxic activation of the two transcription factors HIF-1α and HIF-2α as the result of a non-functional pVHL protein. Although the general influence of these two transcription factors has been deciphered, the full extent of the tumour-promoting activities of HIF-2α via its target genes have not been elucidated. Given these limitations in the clinical management of RCC and understanding of its biology, we set out to address these issues through the papers included in this thesis.

In our **first paper** we aimed to decipher HIF-2α specific target genes operating in normal renal proximal tubule epithelial cells as well as ccRCC tumour cells. We pharmacologically emulated the loss of functional pVHL in renal proximal tubule cells whilst inhibiting HIF-2α transcriptional activity. Subsequent RNA-sequencing revealed potentially HIF-2α specific genes of which we selected *SEMA5B*, where its protein expression pattern matched our RNA-sequencing findings. We verified the HIF-2α regulatory specificity to *SEMA5B* in ccRCC cell lines, with other lines of evidence definitively demonstrating that HIF-2α but not HIF-1α specifically regulates the expression of *SEMA5B* in renal proximal tubule and ccRCC cells. Therefore, SEMA5B may have important role(s) in the context of ccRCC tumour vascularity and microenvironment.

With papers two and three, we sought to address issues in RCC prognostication and detection. In **paper two**, we analysed transcription factor network and regulon activity in RCC subtypes using publicly available datasets. Using this analysis, we identified NFIA, a transcription factor that had similar regulon activity to HNF4A, a well characterised transcription factor in RCC. Based on this data, we examined the relationship between RNA expression of our selected transcription factor NFIA and TCGA-based RCC patient clinicopathological factors such as grade, stage and

cancer-specific survival. We assessed the protein expression of NFIA and HNF4A in a large tissue-microarray consisting of ccRCC and papillary RCC (pRCC) tumours and found that NFIA expression can independently predict CSS in ccRCC patients. Molecular markers for the prognostication of RCC patients are not currently used in the clinic and we hope that our work can contribute towards their eventual implementation.

In **paper three** we developed a workflow to enrich, detect and subtype RCC tumour cells from whole blood. Given the poor performance of EpCAM based circulating tumour cell enrichment methods in RCC, we utilised a size-based isolation platform. We demonstrated that RCC tumour cells are suitable for this approach and showed that our methodology can isolate down to one spiked-in tumour cell from whole blood. Furthermore, through differential gene expression analyses between the three RCC subtypes, we identified transcriptomic markers that can be used to detect pRCC and ccRCC tumour cells from whole blood. This paper lays out a large extent of the methodology and fine-tuning required to isolate and detect RCC tumour cells from whole blood and may provide an additional way to monitor RCC patients in adjuvant and/or neoadjuvant settings.

1 The Kidneys

The kidneys are two bean shaped organs on either side of the vertebral column, sitting on the posterior wall of the abdomen within the retroperitoneal space of the body **(Figure 1).** They are responsible for the maintenance of a relatively constant fluid volume, electrolyte composition, excretion of metabolic waste and toxins, regulation of arterial pressure and regulation of acid-base balance amongst other key functions. An adult human kidney is roughly the size of a clenched fist and despite their unexceptional size or weight, they consume 25% of the ATP in a resting human, in order to carry out their functions [1]. Furthermore, renal tubular cells are packed with mitochondria to supply their energy demands [1].

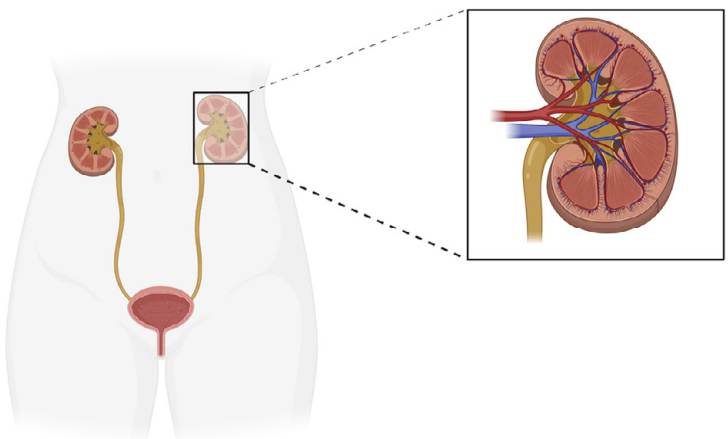

Figure 1 Anatomical location of the kidneys
The kidneys are two bean shaped organs located within the retroperitoneal abdominal space. Blood supply to the kidney enters via the renal artery (red), where it progressively branches out eventually forming glomerular capillaries. Drainage of blood occurs via the renal vein (blue).

In relation to their function, the kidneys release the enzyme renin from the juxtaglomerular apparatus. Renin functions by activating the vasoactive products angiotensin I and II. The angiotensins exert their functions by inducing vasoconstriction and the retention of salt and water, thereby increasing blood volume in circulation and consequently increasing blood pressure [1]. The kidneys are also responsible for the majority of the secreted erythropoietin, which is released in response to low oxygen levels (hypoxia) [1, 2]. Erythropoietin stimulates the production of erythrocytes that are capable of carrying oxygen to potentially hypoxic sites. The primary function of the kidney is to filter the blood, eliminating metabolic waste products and foreign toxins. This is apparent in the fact that the kidneys receive a large volume of blood via the renal arteries connected to the abdominal aorta, through which the kidney receives 22% of the cardiac output volume [3]. This supply of blood to the kidney far exceeds the kidneys' metabolic requirement and reflects its function in plasma filtration. Each day, an impressive 180 litres of primary filtrate is processed through the kidneys of which 99% is reabsorbed into the kidney, ultimately generating 1,5 litres of filtered waste in the form of urine [4]. This process is accomplished through the nephron, the smallest functional unit of the kidney.

The Nephron

Each human kidney possesses roughly one million nephrons, each capable of filtering plasma and forming urine. Each nephron (Figure 2) consists of glomerular capillaries named the glomerulus, contained within the Bowmans capsule. In the glomerulus, filtrate is collected and guided into a system of complex tubules. The capillaries in the glomerulus are impermeable to larger proteins, but smaller organic molecules, salts and ions pass freely through the capillary tuft. The first set of tubules encountered by the glomerular filtrate is the proximal convoluted tubule where almost all glucose, amino acids, 65% of water, potassium, sodium chloride and bicarbonate are reabsorbed. In addition to this reabsorption, proximal tubule cells also excrete metabolic waste products into the lumen. Apart from the convoluted nature of these tubules, they display a lumen facing brush border to increase surface area for reabsorption. Further down the tubules, the filtrate enters the loop of Henle that reaches into the medulla and has three distinct segments.

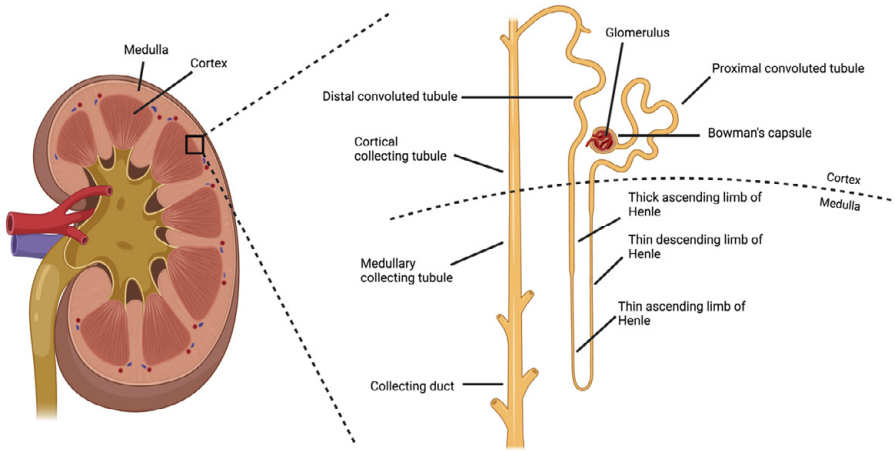

Figure 2 Nephron anatomy
Basic tubular segments of the nephron and their spatial arrangement within the kidney

Within the first of these, the thin descending limb, water is reabsorbed passively through the tubular layer, due to the hyper osmotic pressure of the medulla. As a result, filtrate reaching the bend of the loop consists of higher concentrations of salt and urea than blood plasma. The filtrate then returns towards the kidney cortex, through the thin ascending limb of the loop of Henle where sodium chloride diffuses out of the tubule into the surrounding tissue. In the final thick ascending limb segment of the loop of Henle, sodium chloride and potassium can be reabsorbed further even against a concentration gradient. This is achieved via active transport and thus cells in the ascending are packed with mitochondria to supply high ATP demand [1]. The ascending loop of Henle leads to the distal convoluted tube, where it possesses similar reabsorption properties to the segment preceding it. The distal convoluted tubule, with specialised epithelial cells called macula densa, is also in contact with the juxtamedullary apparatus, thereby allowing for negative feedback loops to regulate the glomerular filtration rate and renal blood flow. This in turn regulates blood pressure [1]. The distal tubules connect to the collecting duct via the consecutive tubular structures of the cortex connecting tubule, cortical collecting tubule, medullary collecting tubule and ultimately leading to the the collecting duct. Eight to ten cortical collecting ducts join to form a single collecting duct that runs downward into the medulla. These ducts merge to form progressively larger channels, where roughly 4000 nephrons can contribute to the volume of filtrate that passes through the collecting duct. Within the ducts, the filtrate closely resembles urine as opposed to plasma, which is then collected in the bladder [1, 3, 4].

2 Hypoxia

Oxygen plays a fundamental role in the physiology of humans and the vast majority of other species. Free molecular oxygen serves as a terminal electron acceptor in aerobic respiration that allows for the generation of cellular energy in the form of ATP. Other oxygen consuming biochemical reactions include fatty-acid desaturation and nucleic acid demethylation reactions which can change the configuration and function of DNA, RNA and histones [5]. The oxygen availability within cells is not always guaranteed and in fact, oxygen fluctuations and more specifically states of low oxygen (hypoxia) can occur. For example, hypoxia within tissues can occur due to several reasons including environmental conditions (e.g., high altitudes), increased rates of aerobic respiration (e.g., exercise) or as a result of damage to local vasculature that impedes delivery of oxygen to cells (e.g., wounds) [6]. Disease states, such as cancer, infection, inflammation and cardiovascular defects can also cause vascular insufficiency leading to hypoxia [7]. Under inadequate oxygen availability, aerobic organisms can undergo cellular dysfunction and eventual cell death [8]. Given the crucial role of oxygen in maintaining homeostasis and the various scenarios in which low-oxygen or hypoxia may be encountered, oxygen-sensing and subsequent hypoxia adaptation mechanisms are of utmost importance in maintaining biological function.

Oxygen Sensing and Hypoxia Response Machinery

Low oxygen sensing mechanisms and hypoxia response machinery are evolutionarily conserved in metazoans, signifying the crucial role of these signalling mechanisms in the survival and perpetuation of complex, multicellular organisms [6]. The conserved transcription factors hypoxia inducible factors (HIF) 1 and 2 play a pivotal role in the response to low oxygen levels and allow for an organism to improve the oxygen supply or undergo changes in gene expression that allows for adaptations to survive in hypoxic conditions. Other oxygen-sensitive regulatory systems operate in concert with the HIF transcription factors but is not within the scope of this thesis. The HIFs are heterodimer proteins and are members of the basic

helix-loop-helix/Per-Arnt-Sim homology (bHLH-PAS) protein family. They consist of two subunits: the alpha (α) subunit HIF-1α or HIF-2α (hereon collectively referred to as HIFα) and the beta unit, HIF-1β. While the expression of the latter is stable independent of oxygen levels, the alpha units are unstable under normoxic conditions. Apart from the bHLH-PAS domains, HIFα proteins also contain oxygen-dependent degradation domains (ODDD), N-terminal transactivation domains (NTAD) and C-terminal transactivation domains (CTAD) [9] (Figure 3).

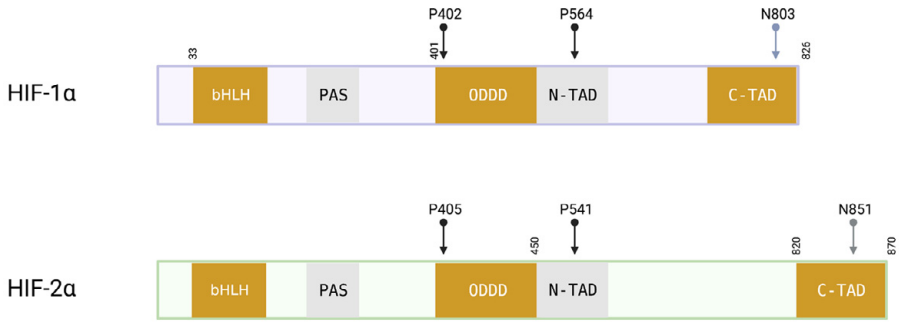

Figure 3: HIFα functional domains
Illustration depicting the functional domains present in HIF-1α and HIF-2α with amino acid numbers indicated. The two proteins share a high degree of protein structure similarity but regulate unique and shared target genes. HIF-1α and HIF-2α contain two proline residues (P) and one arginine residue (N) which can be recognised for hydroxylation by the enzymes PHD and FIH-1 respectively.

HIFα are only transcriptionally active upon heterodimerisation with HIF-1β and thus the stability of either HIF-1α and HIF-2α is crucial to the transcriptional response to hypoxia. There is a third HIFα isoform termed HIF-3α but this is less extensively studied and is understood to be primarily involved in the response to inflammation rather than hypoxia [10]. The stability of HIFα is strictly dependent upon the availability of oxygen – in normoxia HIFα has an exceptionally short half-life of less than 5 minutes [11, 12]. The rapid degradation of HIFα is mediated through the ODDD of the HIFα proteins, where this domain is subjected to hydroxylation on two conserved and highly specific proline residues by prolyl hydroxylase domain (PHD) enzymes, namely PHD1, PHD2 and PHD3 [13]. Oxygen is a key substrate for the PHD enzymes and all three have a Km for oxygen within the 230-250μM range, values above the oxygen concentration in aqueous solutions saturated by ambient air [14]. Intracellular oxygen levels are generally

lower than this Km range, and it is understood that this Km range for the PHDs ensure that they are highly sensitive to oxygen [15]. Prolyl hydroxylation by the PHD enzymes occur when the other necessary substrates of iron and α-ketoglutarate are sufficiently present [16]. These enzymes and their kinetics are the starting point for sensing oxygen levels or hypoxia in metazoans, and in turn their functions allow for a cell to mount an adaptation response to hypoxia. Hydroxylation of the HIFα ODDD by PHD enzymes generates a binding site for the ubiquitin-ligase complex containing the von Hippel-Lindau tumour-suppressor protein (pVHL) [17-19], a culprit in the tumorigenesis of renal cell carcinoma which will be explored in the next section. As a result of the pVHL containing ubiquitin ligase complex binding to the hydroxylated HIFα, the latter protein is polyubiquitylated and directed towards proteasomal degradation in normoxic conditions [19-21].

In hypoxic conditions, the necessary substrate of oxygen is insufficient for the PHDs to perform their function as hydroxylases on HIFα and thus HIFα is stabilised allowing it to translocate into the nucleus [22]. Once in the nucleus, HIFα heterodimerises with its ubiquitously available partner HIF1β to form the heterodimers HIF-1 or HIF-2. Here, these proteins bind to gene promoter regions, functioning as transcription factors [23]. It is within the nucleus that the HIFα heterodimers (i.e HIF-1/HIF-2, hereon referred to as HIFα) are able to exert their full potential as transcription factors by binding to DNA motifs on genes known as hypoxia response elements (HRE) [24]. One of the first genes found to have bound HIF was *erythropoietin(EPO)* [25] and later led to the definition of the consensus HRE sequence of ((A/G)CGTG), as other HIF bound genes were also found to contain this canonical sequence [26]. Once HIF is in contact with the HRE or in its proximity, other coactivators need to be recruited to form an intact initiation complex and this is mediated by the N- and CTAD. These domains recruit CBP/p300, SRC-1 and TIF-2 [27-29] although evidence for direct interaction has only been demonstrated between CBP/p300 and the CTAD [30]. The CBP/p300 complex possess histone acetyltransferase activity, allowing for chromatin modification that primes genes for transcription [31]. The mechanisms of oxygen dependent HIF signalling is summarised in Figure 4.

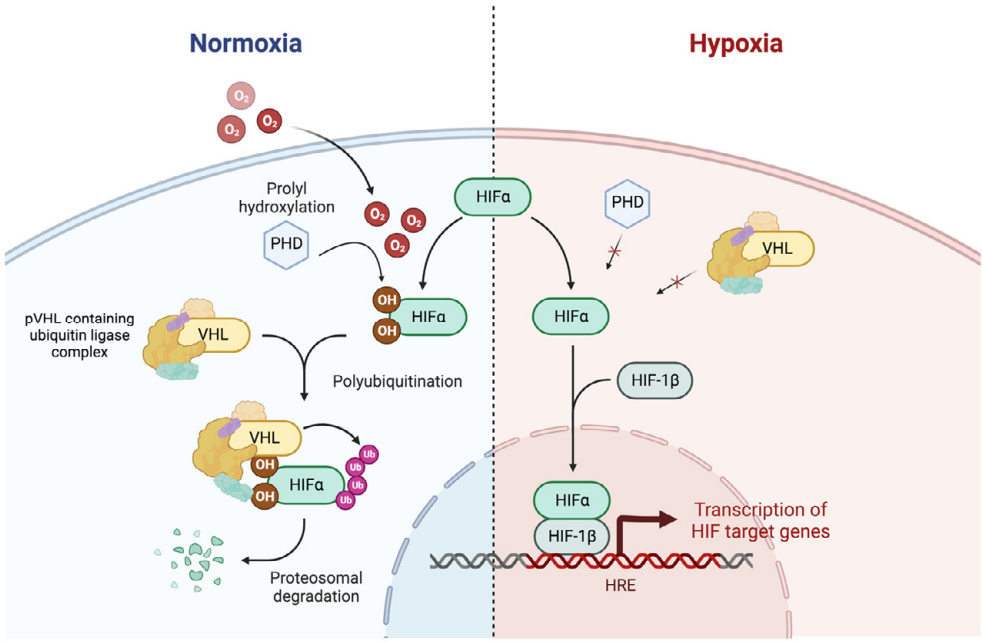

Figure 4: Hypoxia signalling
Simplified overview of HIF-1α and HIF-2α (HIFα) signalling in oxygenated and hypoxic conditions. In the presence of oxygen, HIFα is hydroxylated by the PHD enzymes, creating a site that is recognizable by the pVHL complex for polyubiquitination, destining HIFα for proteasomal degradation. In the absence of oxygen, PHD enzymes cannot perform their hydroxylation function and thus stabilised HIFα translocates into the nucleus and regulates the transcription of genes containing HREs.

Other Mechanisms of HIFα stabilisation/regulation

Apart from the described post-translational regulation of HIFα by the PHD enzymes, HIFα can also be regulated via two other proteins. The first of which is factor inhibiting HIF-1 (FIH-1). Within the CTAD of HIFα, a conserved asparaginyl residue is present and during normoxic conditions, it is hydroxylated by FIH-1 in a similar process to that of the PHD enzymes, where oxygen is an essential substrate [30]. However, unlike the degradation-based outcome of PHD dependent hydroxylation, hydroxylation by FIH-1 leads to steric hindrance at the dimerised HIF-1/2 CTAD site, preventing the localisation of necessary cofactors for the transactivation of transcription by HIF-1/2 within the nucleus (the function of HIF-1/2 as a transcription factor is described in more detail in the next paragraph) [30]. FIH-1 is functional at lower oxygen concentrations than the PHDs and may serve as a contingency measure to target and neutralise HIF-1α/2α escaping degradation at mild hypoxia [32, 33]. Furthermore, FIH-1 has a higher affinity towards HIF-1α

compared to HIF-2α and may contribute to differences in transcriptional activity of the two HIFα isoforms [34]. HIFα can also be regulated by oxygen-independent mechanisms and the most studied mechanisms involve HSP90, whereby this protein binds to HIFα and stabilises the latter protein. It has also been shown that upon dissociation of this HSP90/HIFα complex by HSP90 inhibitors, RACK1 binds to HIFα and recruits ubiquitin ligase protein complexes leading to the eventual degradation of HIFα [35].

Differential regulation of gene expression by HIF-1α and HIF-2α

HIF-1α and HIF-2α are involved in the coordinated response to low oxygen pressures through their effects on gene expression of hundreds of genes known to contain HREs [36]. Given that HIF-1α and HIF-2α display similar domain architecture, DNA binding and activation mechanisms they can be expected to target and regulate overlapping target genes. However, they also display preferential specificity towards target genes. This is in part governed by the differences in HIF stabilisation over time and varying oxygen availability [37]. These differences are further explained by the recruitment of differential transcription machinery and cofactors specific to either of the isoforms and can be dependent on the type of tissue. Current evidence suggests that induction of HIF-1/HIF-2 target genes is not solely dependent on either of the transcription factors but also influenced by other factors that are available and which HIFα isoform these entities can selectively interact with, thus enhancing the transcriptional power of one isoform over the other or neither [15]. For example, in the transcription of *Flk1* (VEGF-receptor 2), the endothelial cell transcription factor Ets1 cooperates with HIF-2α to drive *Flk1* expression in mouse endothelial cells [38]. Apparent responsiveness of target genes towards one HIFα isoform over the other can also be attributed to the level and/or activity of HIF-1α or HIF-2α within the examined cell type. For example, experiments in mouse embryonic fibroblasts revealed that endogenous HIF-2α bound to HREs but failed to elicit detectable target gene expression whilst HRE bound HIF-1α displayed target gene activity in the same setting. Upon overexpression of HIF-2α in these cells, the transcription of the examined target genes was effectively upregulated [39]. This suggests that the response to either HIF isoforms depends on expression and/or activity level of the HIFα isoforms present which could be a function of the cell type. In summary, HIF binding and gene

transactivation at promoter sites may be governed by the diversity of co-factors and other protein-protein interactions present [15].

Given the differences in the ability of either HIFα isoform to support the activation of target genes, it is not surprising that these two transcription factors under certain conditions can promote the expression of different target genes. HIF-1α is understood to be more closely associated with acute hypoxia and drives the expression of genes influencing the metabolic regulation of cells including the control of most glycolytic enzymes (e.g. PFKFB3, LDHA and PKM). Furthermore, the repertoire of genes upregulated by HIF-1 have also come to include genes involved in the maintenance of intracellular pH such as *CAIX* and *MCT4* [40, 41]. HIF-1α induction also leads to a multifaceted suppression of mitochondrial respiration that includes the induction of pyruvate dehydrogenase kinases. Broadly speaking, HIF-1 orchestrates expression of genes involved in metabolic re-wiring in order to divert cells from aerobic metabolism to anaerobic glycolysis when oxygen is limited, a process termed the 'Pasteur effect' [42].

HIF-2, on the other hand, is understood to be associated with chronic hypoxia and drives the expression of genes coding growth factors and regulates a hypoxic pro-growth, stem-cell like program in cells [43]. The outcome of this gene expression programme is aimed at overcoming hypoxia by improving oxygenation in the hypoxic cells or tissue for example, by improving vascular coverage [44]. The majority of the evidence available on the differential expression of target genes by HIF-1 and HIF-2 come from studies in cancer cells and will be explored in the following section.

Tumour Hypoxia

Whilst low oxygen sensing and response mechanisms by way of PHD enzymes and HIFα have evolutionarily evolved to maintain cellular function in fluctuating oxygen conditions, the same mechanisms are at play in tumours. The existence of hypoxia within human tumours was first reported by Thomlison and Gray in 1955 who correlated the presence of hypoxia with resistance to the anti-cancer therapies chemotherapy and radiation [45]. In general terms, solid tumours become hypoxic when the growth of the tumour mass outpaces the growth and coverage of local, functional vasculature that delivers oxygen to the tumour cells and microenvironment (Figure 5). Apart from the vasculature being inadequate in terms of coverage or quantity, these vessels are also malformed often containing blind ends, arteriovenous shunts and poor cellular junctions causing significant leakage [46].

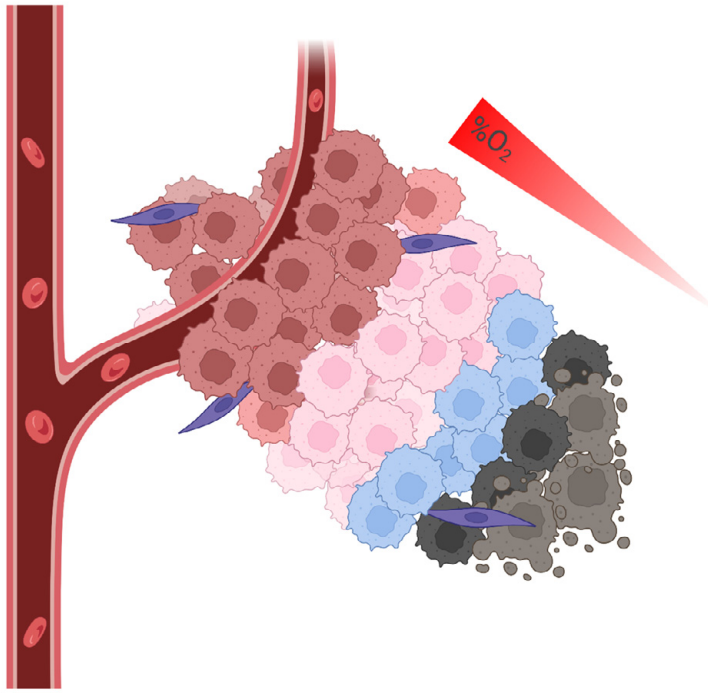

Figure 5: Tumour Hypoxia
Tumour cells in closest proximity to blood vessels are well oxygenated. Oxygen diffuses at a gradient through the tumour where tumour cells farthest from blood vessels have little to no access to oxygen, thus generating an area of hypoxic tumour with necrotic tissue.

Hypoxic tumours compound the inherent malignant properties of tumours that confer them aggressive and resistant to treatments such as chemotherapy and radiation. The presence of hypoxia in tumours is a poor prognostic factor in a majority of cancer types whereby the tumour phenotype displays a number of characteristics directly related to HIF signalling. For example, tumour hypoxia causes increased genetic instability, resistance to apoptosis, invasion, metastasis and malfunctional angiogenesis [46,47]. These mechanisms are a result of HIF regulated target genes operating in tumours and in most instances can be linked to the activity of one of the HIFα isoforms [48]. Although the isoform specific transcriptional effects of HIFα signalling have been largely elucidated within the context of physiological hypoxia, there are conflicting reports and less conclusive evidence in terms of isoform specific regulation of tumorigenesis and tumorigenicity. Clear cell renal cell carcinoma (ccRCC) as a disease model sheds light on this and is of most relevance within this thesis.

HIF signalling in RCC: The HIFα Dichotomy

A large majority (80%) of sporadic ccRCCs are characterised by constitutive, oxygen-independent hypoxia signalling due to loss of functional VHL [49]. An inactivating mutation in one of the *VHL* alleles and a heterozygous deletion of the chromosomal arm 3p results in the bi-allelic inactivation of *VHL* [49]. Additionally, in 3-10% of sporadic RCC cases, *VHL* is silenced via gene methylation [50]. The outcome of these genetic and epigenetic aberrations set the stage for dysregulated HIF signalling in ccRCC, whereby HIF-1α and HIF-2α are no longer subject to oxygen-dependent polyubiquitylation by the pVHL-ubuiqitin ligase complex and this phenomenon has been described as 'pseudohypoxia' (Figure 6). The stability of the HIFα transcription factors is significantly prolonged, allowing them to translocate, dimerise and exert their functions on downstream target genes. From an experimental and theoretical perspective, these features render ccRCC an ideal cancer type to model and study tumour hypoxia.

The contribution of HIF-1α and HIF-2α towards tumorigenicity has been an area of intense research over the last 20 years. Currently a majority of the collective evidence suggests a tumour-suppressive role for HIF-1α and a pro-tumorigenic role for HIF-2α in ccRCC. All VHL null, established ccRCC cell lines express HIF-2α whereas many do not express HIF-1α [51-53]. Secondly, tumours fail to grow in mouse xenograft experiments where HIF-2α knockout RCC cell lines are implanted [54, 55]. In line with this, overexpressing constitutively stabilised HIF-2α but not HIF-1α results in xenograft tumour growth despite pVHL activity [56-58]. Furthermore, multiple studies point to HIF-2α being necessary and sufficient for tumorigenesis when VHL has been tissue-specifically inactivated in genetically engineered mouse models [59-63].

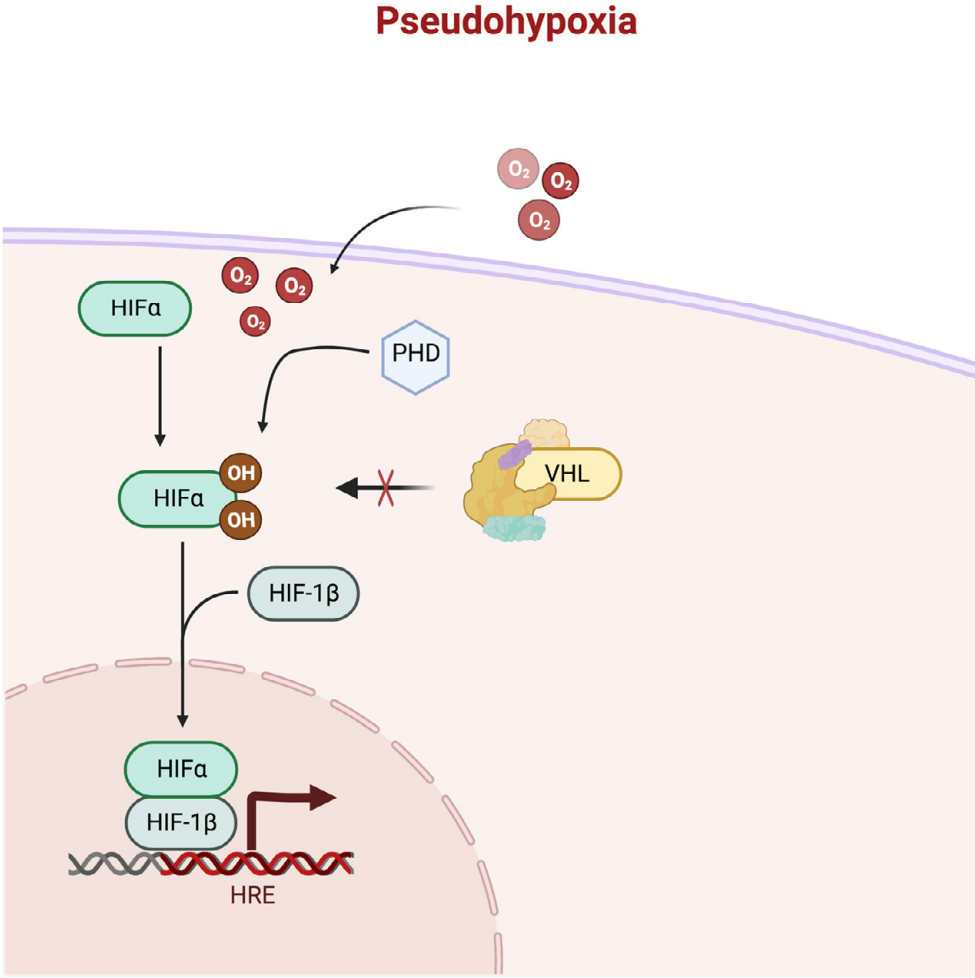

Figure 6: Pseudohypoxia
In ccRCC tumours VHL is inactivated in a large majority of cases causing the oxygen-independent stabilisation of the two HIFα isoforms HIF-1α and HIF-2α.

It has also been observed that the presence of HIF-2α in preneoplastic RCC lesions result in transformative processes leading to a neoplastic lesion [64]. Finally, HIF-2α single-nucleotide polymorphisms have been linked to an increased risk of kidney cancer within the general population [65]. Two of the prevailing explanations for these possibilities lie in the poor post-translational regulation of HIF-2α by FIH-1 and the downstream targets HIF-2α but not HIF-1α preferentially activate. HIF-1α is relatively more sensitive to FIH-1 as discussed previously and thus HIF-1α

transcriptional capacity may be dulled in comparison to that of HIF-2α. Either in conjunction or independently of this, HIF-2α may also have preference towards activating genes that are oncogenic in nature, such as genes required for a stem-cell like or dedifferentiated phenotype [66].

Multiple lines of evidence also suggest the tumour-suppressive effect of HIF-1α. The chromosomal arm 14q, which contains the *HIF1A* gene, is often deleted in clear cell RCC tumours and is associated with poor prognosis [51, 67]. Re-introduction of wild-type HIF-1α in ccRCC cell lines lacking endogenous HIF-1α suppresses their proliferative potential *in-vitro* and *in-vivo* [51, 56]. In line with this, short-hairpin RNA mediated knockdown of *HIF1α* in ccRCC cell lines already possessing endogenous HIF-1α promotes their proliferation at both *in-vitro* and *in-vivo* settings [51, 52]. Taking all of the experimental and observational evidence collectively, there is a strong case to be made that HIF-1α behaves as a tumour-suppressor whereas HIF-2α acts in an oncogenic manner in ccRCC tumours.

Even though some evidence exists that the HIF dichotomy may be the case in other truly hypoxic tumours, there is a tendency within the literature to extrapolate these ccRCC based findings onto other tumour types. It is likely that the roles HIF-1α and HIF-2α play within a tumour cannot be generalised and must be examined upon the context of oxygen availability, disease and cell type. It can be argued that convincing experimental evidence for the roles of HIF-1α and HIF-2α only exists within ccRCC and makes an undeniable argument for their polar and opposing roles in this disease.

3 Kidney Cancer

Kidney Cancer encompasses all types of cancers arising from the kidney. Even though the majority of epidemiological data refer to kidney cancer cases, a staggering majority (90%) of kidney cancers are histologically defined as renal cell carcinomas (RCC) [68]which arise from the tubular epithelium of the nephron. In 2020, there were an estimated 431,288 new cases of kidney cancer globally, making it the 14th most common cancer type [69]. Mortality from kidney cancer constituted 1.8% of global cancer deaths [70]. The incidence of kidney cancer is approximately twofold higher in men than in women, a pattern that holds true over time, geographic regions and age groups. The causes of this differential risk of developing kidney cancer between sexes are not yet understood [71, 72] but behavioural patterns may play a role. Geographically, kidney cancer incidence rates are higher in Europe and North America and incidence also appears to be associated with higher median incomes. This, however, is hypothesised to be due to higher rates of abdominal imaging (e.g CT & MRI) in high-income populations as small renal masses are found incidentally.

Kidney cancer but specifically RCC is an insidious disease, with an estimated average growth rate of 0.28cm a year [73]. The global mean age at diagnosis is approximately 75 years, although this can vary geographically: Sweden (67 years) [74] ,UK (74 years), India (67 years), China and Italy (82 years) [68]. The epidemiological causes of kidney cancer are poorly understood, but it is known that lifestyle and health factors such as smoking status [75], body weight [76], hypertension [77, 78], chronic kidney damage [79] and level of physical exercise [80] can contribute towards kidney cancer tumourigenesis. Exposure to environmental toxins have also been shown to be risk factors in the development of kidney cancer. Although the contribution of environmental toxins towards kidney cancer are challenging to quantify due to competing exposures and high variability in geographic risk factors, there is evidence that exposure to perflourinated chemicals and aristolochic acid can increase the risk of developing kidney cancer [81]. It could be argued that, given the blood filtering functions of the kidneys, tubular epithelial cells of the kidneys are exposed to a wide array of toxic chemicals compared to other organs such as the brain or prostate. Physiologically, there are safeguarding mechanisms against toxin exposure and subsequent kidney injury,

such as the kidney epitheliums' regenerative capacity. It may well be that the very same regenerative mechanisms become dysregulated upon repeated kidney injury, creating a favourable pro-growth environment for tumorigenesis. This is potentially evidenced by the increased likelihood of patients with chronic kidney damage to develop kidney tumours [82].

Renal Cell Carcinoma Subtypes

Renal cell carcinoma is the name given to a group of tumours arising from the renal epithelium and reportedly consists of at least 16 different subtypes. The classification is based upon histopathological and genetic characteristics [83]. The three primary and most common subtypes of RCC consist of clear cell RCC (ccRCC), papillary RCC (pRCC) and chromophobe RCC (chRCC)**.** They respectively represent 70-90%, 10-15% and 3-5% of all renal cell carcinomas [50]. Some of the other less common subtypes, with an incidence of less than 1% [84], will not be discussed in this thesis but include subtypes such as oncocytomas, collecting duct RCCs or unclassified RCCs.

Clear Cell Renal Cell Carcinoma

The classification of ccRCC was based on its histological appearance upon FFPE treatment and subsequent H&E staining. Under this processing, ccRCC cells appear to have a clear cytoplasm (Figure 7) due to the accumulation of glycogen and lipids which are displaced by the fixation and staining procedures. Coupled with an increased understanding of genetic aberrations and signalling mechanisms occurring in ccRCCs, the clear cell phenotype and to an extent the molecular features are understood to be a result of the metabolic rewiring that occurs in ccRCC due to constitutive HIF signalling [85, 86].

ccRCCs typically display a relatively low mutagenic load, however they are characterised by large chromosomal deletions and gains. Near universal deletion of chromosome arm 3p is observed in more than 90% of ccRCCs (Figure 8) as well as the loss of 14q (46%), gain of 5q (60%) and 7q (40%) [87-90]. In 90% of sporadic ccRCC cases, the second chromosome 3p arm is plagued by an inactivating mutation, homozygous deletion or silencing via methylation [87, 91, 92]. Together with the deletion of the other chromosomal arm 3p, the result is the bi-allelic inactivation of *VHL*. As a consequence of these genetic and epigenetic aberrations, pVHL loss of function can be interpreted as an obligate event in ccRCC and is often referred to as a hallmark of this subtype. This argument is further strengthened by

the observation that patients with VHL syndrome, a condition that is characterised by germline or mosaic VHL loss-of-function, have a high predisposition towards developing hereditary ccRCC (and other select tumour types) upon loss of the remaining wild-type *VHL* allele [93]. These observations are consistent with our understanding of the 'two-hit' *VHL* inactivation in sporadic ccRCC cases [93].

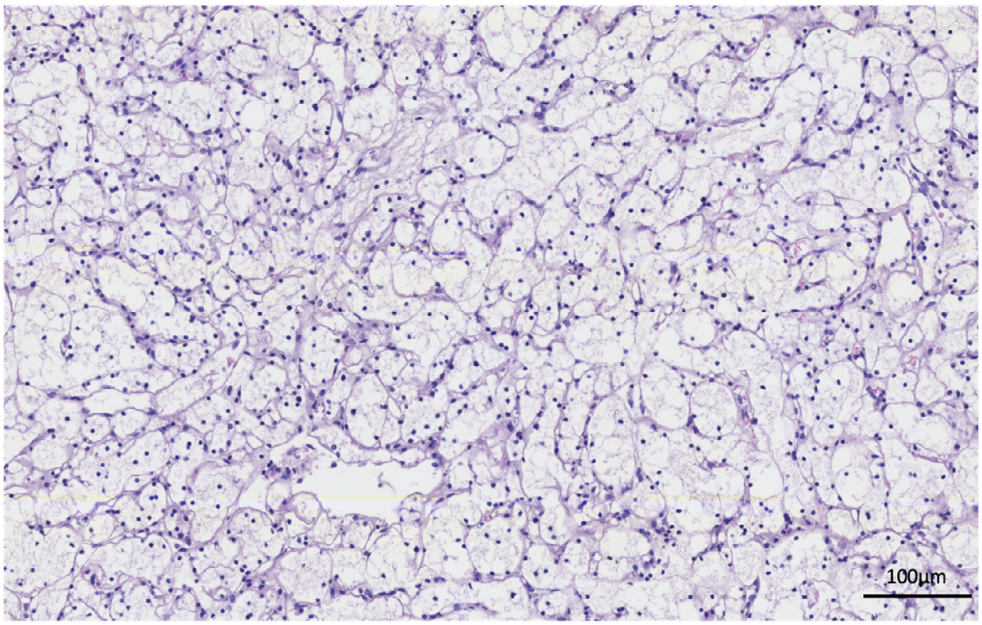

Figure 7: ccRCC histology section
H&E stained ccRCC tumour section displaying tumour cells with their distinctive 'clear' cytoplasm and high tumour vascularisation.

As mentioned previously, constitutive HIF signalling leads to the upregulation of a wide array of prototypical hypoxia target genes. Apart from the metabolic rewiring that enhances the Warburg effect [94] in ccRCC cells and contributes towards the clear cell phenotype, HIF directed overproduction of VEGF leads to highly vascularised tumours [88, 95]. There is great interest and potential clinical benefit in uncovering the exact tumour characteristics that arise as a result of aberrant HIF-1α or HIF-2α signalling.

Due to loss of 3p in ccRCC, three genes apart from *VHL* are also lost. Similarly to *VHL*, other genes on the remaining chromosome arm 3p have a propensity to be heterozygously inactivated. Namely these three genes include SET domain-containing 2 (*SETD2*), BRCA1-associated protein 1 (*BAP1*) and polybromo 1 (*PBRM1*) [87, 96]. *SETD2* is mutated in >40% of ccRCC tumours whilst *BAP1* and

PBRM1 are mutated in approximately 10-15% [97-99]. All three of these genes are involved in chromatin and histone remodelling and indicate that chromatin modification may allow a selective advantage in a VHL null setting [100-102]. In addition to these genetic aberrations, ccRCC (and pRCC) tumours are defined by a PAX8/ HNF-driven transcriptional programme that is present in normal proximal tubule renal epithelial cells [103, 104]. This makes a strong case for the fact that these cells may be the cell of origin in ccRCC [104]. Additionally, this transcription factor driven programme in ccRCC tumour can further distinguish this tumour subtype from other RCC subtypes such as chRCC.

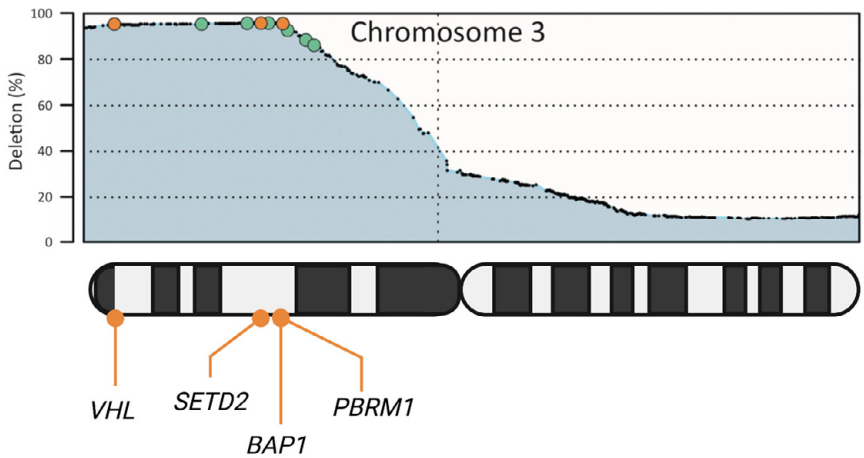

Figure 8: Chromosome 3p deletions in ccRCC
Frequency of Chromsome 3p deletions in ccRCC tumours and subsequently affected genes. Adapted from Lindgren at al. [85]

Papillary renal cell carcinoma

Papillary RCC is the second most common subtype of RCC but displays high heterogeneity. This is evidenced by the further subdivision of this tumour subtype into two histological types: type 1 and type 2. Type 1 is often multifocal and characterised by papillary structures that are lined with small tumour cells containing basophilic cytoplasm and small, uniform nuclei [105]. Type 2 is characterised by papillae covered with small cells containing eosinophilic cytoplasm and large irregular nuclei [106]. Type 1 pRCCs are clinically less aggressive and are linked to the overexpression and increased signalling of the MET receptor tyrosine kinase. *MET* alterations, through mutation, splice variation or gene

fusions have been found in 81% of all type 1 pRCCs and are thought to contribute towards tumorigenicity via increased cell proliferation [107]. Type 2 pRCC, the more aggressive of the two pRCC tumour types, display increased signalling in the NRF2-ARE pathway [108], a mechanism critical in the eradication of reactive oxygen species [109]. Furthermore, type 2 pRCC tumour can present with silencing of CDKN2A and mutations in the Hippo pathway tumour-suppressor NF2. Accumulated evidence suggests that there may be three or more distinct groups of tumours within the type2 pRCC subtype [110, 111] and thus pRCC as a whole represents a highly heterogenous RCC subtype.

Chromophobe renal cell carcinoma

Chromophobe RCC is the third most common subtype of RCC and is histologically characterised by perinuclear clearing and distinct cellular borders. It is the least aggressive subtype of RCC, potentially owing to the low number of somatic mutations found in these tumours. However, hypodiploidy is common and 80% of chRCC tumours have losses of chromosome 1, 2, 6, 10, 13, 17 and 21 [112, 113]. chRCCs are the most distinct subtype among the RCC subtypes, displaying a FOXI-1 driven transcriptional programme [104]. This transcriptional signature is understood to be remnant of chRCC pre-malignant cells and are thought to stem from the collecting duct [85]. In around half of chRCC tumours, aberrations in *TP53* and mTOR pathway genes are found. 10% of chRCC are observed to have rearrangements in the promoter region of the *TERT* gene, elevating its expression [112, 114]. Around 5-10% of chRCC patients develop metastases and these tumours are commonly found to have a combination of imbalanced chromosomal duplications and mutations in *TP53* and *PTEN* [112, 115].

Clinical Management of RCC

Disease manifestation and prognosis

Within the clinical realm of RCC patient care, it is a well-known and familiar occurrence that a large majority of RCC patients are diagnosed incidentally. Usually through abdominal imaging for non-related or non-specific causes, patients are unexpectedly plunged into a reality where they are dealing with a potential cancer diagnosis. In a recent study recruiting 608 RCC patients, a large majority (60%) of recruited patients had been diagnosed incidentally [116]. There are two potential

non-exclusive explanations for this; the first being that abdominal imaging has increased over the last decade, especially in the western hemisphere and thus the likelihood of discovering a renal mass is increased. The second reason for this are the elusive symptoms RCCs display. In early stages, the disease generally does not display any symptoms and at more advanced stages, the symptoms are usually bone pain, deterioration of performance status and persistent cough [117]. Aggressive disease classically presents with haematuria, flank pain and a palpable abdominal mass [118]. These symptoms, in both aggressive and advanced tumours are difficult to diagnose due to their non-specific nature and delayed onset once the tumour has metastasised. This is evidenced in a study where 36% of patients were diagnosed incidentally with stage III and IV RCC tumours respectively [116][119]. Due to the nature of how RCC clinically presents pre-diagnosis, early diagnosis is recognized as a key strategy to improve patient outcomes.

Prognostic factors in RCC can be classified into four categories: anatomical, histological, clinical and molecular. The former two categories are known to have more accumulated evidence in their support due to their historical use within the clinic. The classical anatomical prognostication score is based on the tumour, node and metastasis (TNM) classification (Table 1) which has been the most commonly used staging system in RCC [120]. The TNM staging system combines several well studied prognostic features of RCC tumours such as tumour size, invasion of the venous system, invasion into lymphatic system, extension into the adrenal gland, extension beyond the renal capsule or Gerota's fascia and number of distant sites with metastasis. In all RCC subtypes, prognosis is well evidenced to deteriorate with increasing T, N and M classification [121]. The TNM classification can be based upon imaging (iTNM) or upon pathology (pTNM) once the tumour has been surgically removed. pTNM usually allows for more definitive information on the local extension of the tumour and gives precise information on the location of the tumour border. TNM staging can provide critical information in planning initial treatment [122].

Histological prognostic factors aim to combine multiple tumour histological features and provide information on prognosis and tumour subtype. Although the four-tiered Fuhrman system was used up until 2012 to grade RCC tumours, it was viewed as suboptimal and thus has been replaced by the International Society of Urological Pathology (ISUP) grading system in 2012 [124]. Features taken into account in this current system include nuclear morphology, nucleolar prominence (eosinophilia), nuclear anaplasia, size of tumour cells, sarcomatoid and/or rhabdoid morphology [124, 125]. Although the ISUP grading system has been validated for use and provides prognostic insight into ccRCC and pRCC tumours, it has failed to show any link between ISUP grade and outcome for patients with chRCC tumours

[124]. Nevertheless, the ISUP grading system is applied to chRCC and other less common RCC subtypes but only for descriptive and diagnostic purposes and not for outcome prediction [124, 125].

Table 1: Tumour (T), Node (N), Metastasis (M) staging system used to define and describe the spread of RCC [110].

CATEGORY	DEFINITION	SUBDIVISION
TUMOUR		
TX	Primary tumour cannot be assessed	
T0	No evidence of primary tumour	
T1	Primary tumour is ≤7 cm in greatest dimension and confined within the renal capsule	1a: ≤4 cm in greatest dimension
		1b: Primary tumour is >4 but ≤7 cm in greatest dimension.
T2	Primary tumour is >7 cm in greatest dimension and confined within the renal capsule	2a: Primary tumour is >7 cm but ≤10 cm
		2b: Primary tumour is >10 cm
T3	Primary tumour extends into major veins or perinephric tissues but not into the ipsilateral adrenal gland and not beyond the perirenal (Gerota) fascia	3a: Primary tumour extends into the renal vein, renal sinus fat, and renal capsule but not beyond the perirenal (Gerota) fascia
		3b: Primary tumour invades the IVC below the diaphragm
		3c: Primary tumour invades the IVC above the diaphragm
T4	Primary tumour invades beyond the perirenal (Gerota) fascia or invades the ipsilateral adrenal gland	
REGIONAL LYMPH NODES		
NX	Lymph nodes cannot be assessed	
N0	No regional (retroperitoneal) lymph node metastasis	
N1	Regional (retroperitoneal) lymph node metastasis	
DISTANT METASTASIS		
M0	No distant metastasis	
M1	Distant lymph node or other metastasis, including noncontinuous adrenal involvement	

IVC=inferior vena cava

Clinical prognostic factors can relate to tumour 'external' characteristics such as patient performance status, presenting symptoms, paraneoplastic syndromes and peripheral blood-based values for calcium, albumin, haemoglobin and C-reactive protein [126, 127]. These are primarily used for routine risk-stratification and treatment decisions where the link to outcome has been best evidenced in the metastatic setting [126]. Molecular prognostic factors can be used for predicting treatment independent/dependent patient outcomes based on the molecule assessed.

For example, immunocytochemical analysis of CAIX, PTEN and CXCR4 as well as gene expression and methylation status profiling have been investigated [128] but Ljungberg and colleagues argue that these have not improved current prognostic systems and thus are not part of routine clinical care [124]. Notwithstanding, expression levels of *BAP1* and *PBRM1*, genes deleted in 90% of ccRCC can independently predict tumour recurrence [129]. Furthermore, a 16-gene signature has been shown to predict RCC relapse and has been validated in adjuvant trials [130]. However, this study seemingly fails to acknowledge which subtype of RCC this gene signature applies to; patients recruited had a 'majority clear cell compartment' but the authors do not definitively classify the tumour types enrolled in the trial. Prognostic information from cytokines and PD-L1 expression have also provided promising therapeutic results[131] but exploration of these molecules is not the routine in RCC clinical care [124] and will be explored later.

Diagnosis and Imaging

Most RCC masses are found incidentally during abdominal imaging for seemingly non-associated causes. However, upon finding a renal mass, the most common imaging modalities utilised to characterise them include computed tomography (CT), ultrasound (US) and magnetic resonance imaging (MRI) [132]. Contrast enhancement or restriction is the most critical characteristic when deeming a renal mass as malignant [133]. Positron emission tomography (PET) is increasingly being used for the characterisation of pRCCs however it is not advised in the use of ccRCC [124, 134, 135]. These imaging modalities are used to detect a renal mass and to obtain additional information on its size and spread into vessels, lymph nodes and/or distant organ sites. For renal masses, the Bosniak classification can provide risk of malignancy and guidance for management based on CT or MRI imaging [136]. Most cases of RCCs can be correctly assessed using one or more of these imaging modalities, however in instances where imaging information is insufficient, core-needle biopsies may be used. Tumour biopsies and subsequent histology aid in selecting patients for surveillance (in case of benign tumours/poor patient performance status) or prior to ablative and primary systemic treatment [137]. Unfortunately, there are many limitations of core-needle biopsies such as risk of infection, increased risk of procedure related tumour cell seeding and inability to capture tumour heterogeneity [138].

The Current Landscape of RCC treatment

Treatment of Localised Disease

Surgical removal of part (partial nephrectomy) or the entire (radical nephrectomy) kidney is still the only curative treatment for localised, low stage RCCs. Patients harbouring a T1 tumour benefit most from partial nephrectomy regardless of surgical approach based on renal function, oncological and quality of life outcomes [139]. This is based on multiple retrospective studies and one randomised controlled trial (RCT), where cancer-specific survival and other post-operative patient outcomes were assessed when partial or radical nephrectomy was employed. There are various techniques of surgery with regards to partial or radical nephrectomies. Surgeons have the option of performing open surgery, pure or robot-assisted laparoscopic surgery. Currently, there is no superiority between the techniques with regards to oncological outcomes however laparoscopic approaches have been shown to improve post-operative recovery given the relatively reduced loss of blood and use of analgesics [140]. The current surgical recommendation for small, localised tumours is one of the above-mentioned techniques depending on the equipment available to the surgeon as well as their level of expertise.

Treatment of Metastatic RCC (mRCC)

Within the context of metastatic RCC, surgery is most often not an option although patients may benefit from palliative cytoreductive procedures if the primary tumour site is impairing patient performance status or urological function [141]. RCCs most often metastasise to the lung, bone, liver, lymph nodes, adrenal gland and brain [142-144]. These tumours are notoriously poor responders to classical anti-cancer therapies such as radiotherapy and cytostatic drugs [145]. This observation could be explained by their relatively slow rate of growth (low proliferative rate) and (pseudo)hypoxic features. However, there is great promise in the new wave of anti-cancer therapies that rely upon unleashing the immune system, specifically immune checkpoint inhibitors (ICI). Although conceptually a familiar approach within mRCC where immunostimulatory therapies such as interleukin-2 and interferon-α were extensively used in the 1990s, these were limited by extensive systemic toxicities [127, 145, 146]. However, judged by tumour response to these therapies and accounts of spontaneous regression of RCC in the 80-ties and 90-ties [147, 148], RCC is touted to be an immunogenic tumour type that may benefit from the new-generation of immunotherapies. The treatment of mRCC can be classed into three eras: cytokine, tyrosine-kinase inhibitor (TKI) and the ICI eras (Figure 9**).**

Figure 9: Eras of mRCC therapies
Timeline of mRCC systemic therapies over the last three decades

Tyrosine-kinase inhibitor Era

The identification of VHL loss in RCC and the understanding of subsequent signalling that leads to the aberrant and overexpression of many hypoxia targets, including VEGF ligands and receptors, paved the way for the U.S Food and Drug Administration (FDA) approval of adjuvant TKI therapies targeting these molecules in 2006 [139, 149]. Sorafenib and sunitinib were the first anti-angiogenic TKIs approved by the FDA and was later followed by the approval of five other anti-angiogenic TKIs [149]. In addition to these, mTOR inhibitors and VEGF antibodies were also approved for use [150]. Given the highly vascular nature of ccRCCs, use of these therapies proved to regress but mostly stabilise disease and prolong survival of mRCC patients [43]. However, given the observed high adverse-event rates and poor tumour cytotoxicity in these patients [43], there was much to improve upon. With the emergence of ICIs, a transition would take place towards combination treatments with ICI and TKIs.

Immune checkpoint inhibitor era

Two of the most promising immune targets and hence investigated molecules are programmed death 1 (PD-1) and cytotoxic T-lymphocyte-associated antigen 4 (CTLA-4). CTLA-4 is an immunoinhibitory receptor present on T-cells and functions to dull the early activation of naïve and memory T-cells, hence termed an ‘immune checkpoint’. The ligands for CTLA-4 are present on antigen presenting cells or on tumour cell surfaces. In the tumour setting, tumour cells are able to escape immune destruction via the inhibition of these T-cells which can no longer target and kill these tumour cells [151]. Therefore, there has been immense interest in blocking this checkpoint interaction from occurring which has led the development of anti-CTLA-4 antibodies (ipilumumab) which can competitively bind to the CTLA-4 receptor, effectively blocking it from any activating signals thus circumventing the inhibition of the T-cell mediated immune response. Proof of

concept for this was first recorded in metastatic melanoma, where blockade of this immune checkpoint with ipilimumab led to the activation of the host immune system against tumour cells resulting in tumour shrinkage [152-154].

PD-1 is a transmembrane inhibitory receptor expressed on activated T-cells and has two known ligands, PD-L1 and PD-L2 that are normally expressed on antigen-presenting cells and pathologically on RCC cells as a means of immune evasion [155-157]. These ligands serve to bind and dampen the activity of T-cells. A retrospective analysis of 306 ccRCC patients who underwent nephrectomy revealed that patients with tissue expression of PD-L1 had significantly reduced 5-year CSS compared to those patients that did not display PD-L1 expression [158]. This data suggests that disrupting the inhibitory interaction between PD-1 and PD-L1/L2 (collectively referred to as PD-L) could serve as an attractive approach to restore T-cell capabilities against RCC tumours. RCC (along with 4 other tumour types) was one of the pioneer tumours to be included in testing this hypothesis, through an immunoglobulin that targeted PD-1, hence disrupting the PD-1/PD-L interaction. Bristol-Myers Squibb provided the first-in-human evidence that this approach could be efficacious in a pilot RCC patient, where they had a partial response [159]. Many patients and clinical trials later, the status quo of RCC immunotherapy guidelines sit on the recommendation that first-line treatment for metastatic ccRCC should be combination treatment with Pembrolixumab (anti-PD1 antibody) plus axitinib (VEGF TKI) [139]. This was borne out of the KEYNOTE-426 trial, where the combination treatment provided a median 5,6 month improvement in overall survival (OS) and 6% complete response rate improvement compared to treatment with only sunitinib (multiple tyrosine kinase inhibitor) [160, 161].

Presently, the standard of care has shifted from TKIs, mTOR inhibitors and VEGF antibodies towards first-line combination treatment with TKI and dual ICI via anti-CTLA-4 and/or anti-PD-1 [139]. Six phase 3 randomised controlled trials have shown that ICI combinations have superiority over sunitinib alone [162]. Higher numbers of ccRCC patients on ICI combinations have achieved durable remissions in this setting, but unfortunately convincing evidence is lacking for second- and third-line settings [139].Currently, the recommendation for these settings is VEGF targeted therapy (in VEGF TKI naïve patients) but clinical trials are ongoing to determine the best treatment strategies [139].

A Novel Therapeutic Target: HIF-2α

The non-functionality of the pVHL protein and subsequent oxygen-independent stabilisation of HIFα is a near-definitive observation in ccRCC. As previously discussed, there is a strong case to be made that HIF-2α drives and maintains many of the tumorigenic processes seen in ccRCC. Given this, HIF-2α would constitute

an attractive druggable target in an effort to minimise these pathological characteristics of ccRCC. In the absence of a known ligand-binding domain in HIF-2α and its structural similarity to HIF-1α, the latter has long been considered undruggable. However, biophysical studies revealed a small hydrophobic pocket that is able to bind small molecules in the PAS-B domain of HIF-2α [163, 164].

Although initial molecules showed promise in the possibility of antagonising HIF-2α, these showed poor cellular potency and poor physical properties. Further structural modifications and testing led to the development of the first-generation HIF-2α specific inhibitor PT2385, which displayed improved specificity and cellular potency in pre-clinical models [165].In a phase-I dose-escalation trial involving 51 patients with metastatic ccRCC who had not responded to any previous lines of therapy, the PT2385 treatment showed partial response in 21%, complete response in 2% and disease stability in 52%. Importantly, PT2385 treatment displayed minimal adverse events with major events being fatigue, anaemia and peripheral oedema which can be relatively well-managed [166].

Despite the clinically favourable profile of this first-generation inhibitor, its pharmacodynamics were highly variable within the *in-vivo* setting. This led to the development of PT2977 (Belzutifan), with other biophysical intermediaries tested along the way. Belzutifan has been trialled in VHL disease associated RCC as well as the metastatic ccRCC setting. In patients with VHL disease associated RCC, the objective response rate (ORR) was 49% with promising response rates in non-RCC tumours that VHL syndrome patients are known to develop [167]. Similarly to Belzutifans' predecessor, the most common toxicities patients experienced were anaemia and fatigue [93]. Based on this data, Belzutifan was approved by the FDA in 2021 for use in adults with VHL disease who require therapy for RCC and other VHL disease related tumours [168].

Within the sporadic, locally advanced or metastatic ccRCC setting, Belzutifan was trialled in the LITESPARK-001 study [169]. These patients had received at least one prior therapy, but Belzutifan monotherapy led to an ORR of 25% [169]. After a median follow up of more than 3 years, the median duration of response was not reached [169], indicating the drugs' ability to stabilise disease at the very least. Adverse event rates and types were similar to previous reports from clinical trials using Belzutifan for RCC [169]. Preliminary results from an on-going Phase II study assessing the combination of Belzutifan plus cabozanitib (multiple TKI) in patients that are treatment naïve or had received prior immunotherapy/ TKI showed at least 90% of patients showed tumour shrinkage [170]. This data demonstrates that combination of HIF-2α and TKI inhibition could offer more benefit than monotherapy with either drug. Additionally, these results make inlays towards establishing potential treatment avenues for pre-treated patients [171]. Although it

is unlikely that HIF-2α inhibitors will be the ‘magic bullet’ in the treatment of metastatic ccRCC tumour, combination therapy with other immunotherapies could be on the horizon and may re-define treatment paradigms for mRCC patients.

4 Liquid Biopsies in RCC

The term ‘liquid biopsy’ was coined 13 years ago by Klaus Pantel and Catherine Alix-Panabières and referred to the clinical utility in assessing circulating tumour cells (CTCs) to detect cancer [172]. Soon afterwards, the term evolved to encompass the use of any fluid analyte that can be obtained via minimally- or non-invasive approaches. Although the term liquid biopsy only came into use in 2010 and is perhaps strictly applicable to tumours, for centuries humans have been obsessed by the idea of gathering disease related information via minimally invasive approaches. This is evidenced by the most crude and antiquated application of this concept when the Oxford University physician Thomas Willis, in 1674, noted that the urine of some of his patients tasted sweet - an observation that would aid future researchers to isolate the cause and symptoms of diabetes mellitus [173]. Over the years, various other tests have been researched and implemented within the clinic ranging from blood counts to faecal occult blood tests. In modern oncology, the aims of these liquid biopsies are similar; to accurately find actionable tumour-related information via minimally invasive approaches. Given the characteristics of RCC onset, limitations in current diagnostics and therapy, the gulf between the standard of care and the standard that can be achieved with liquid biopsies is immense.

Liquid biopsy fluids & analytes

Theoretically, a liquid biopsy can be performed on any bodily fluid in which tumour-associated analytes are present. This can range from blood, urine, cerebrospinal fluid, saliva, cyst fluid, bone marrow and sputum. Peripheral blood provides a plethora of analytes that can be separated and isolated simultaneously and thus has the potential to provide a wealth of tumour related information. For example, blood can be used to isolate CTCs, circulating tumour DNA (ctDNA), circulating tumour RNA (ctRNA), tumour-derived exosomes (displaying tumour surface markers or containing nucleic acids), tumour proteins and tumour-educated platelets [174]. For the purposes of this thesis, the focus will remain on CTCs and ctDNA.

Circulating Tumour Cells & Enrichment Approaches

Circulating tumour cells are tumour cells that enter the systemic blood circulation after they have been shed by a tumour and were observed and reported in patients as early as the 1860s [175-177]. Entering the circulation is part of a multi-step process known as the metastatic cascade, most often taking place in aggressive tumours [178].

CTC analysis can be divided into three stages: enrichment, detection and characterisation. One of the most critical issues facing CTC enrichment methods are the low numbers of CTCs detected in peripheral circulation. Whilst in circulation, CTCs can face many hurdles such as anoikis, shearing forces and immune surveillance which may contribute towards their low numbers [179-181]. Furthermore, blood contains a vast number of cells and cell-types, and depending on the technology used for enrichment, red blood cell lysis and removal is a necessary step that may contribute to the loss of CTCs. The most common and efficient positive CTC enrichment strategies aim to exploit differences between tumour cells and blood cells. These strategies can be broadly classified as label dependent or label independent (Figure 10)**.** Label dependent strategies usually utilise an antibody-based approach whereas label independent approaches are typically based on differences in physical properties between CTCs and blood cells.

The most well-known and only FDA approved CTC enrichment technology is the CellSearch® system which relies on antibody tagged magnetic beads to pull down CTCs that express EpCAM [182]. The clinical utility of the CellSearch® system was demonstrated in the metastatic breast cancer setting [183], where the number of captured CTCs were shown to be associated with progression-free survival (PFS) and OS [184]. The CellSearch® system is also FDA-approved for use in other EpCAM expressing tumour types such as metastatic colorectal and prostate cancer [185, 186]. Unfortunately, the use of CellSearch® has been explored in RCC patients with somewhat disappointing results. Positive CTC detection (≥ 1 CTC) ranged from 16% - 46% across CellSearch® based RCC studies [187-189] and when compared to CTC detection rates in prostate cancer (90%) [190], RCC seems incompatible with the CellSearch® system.

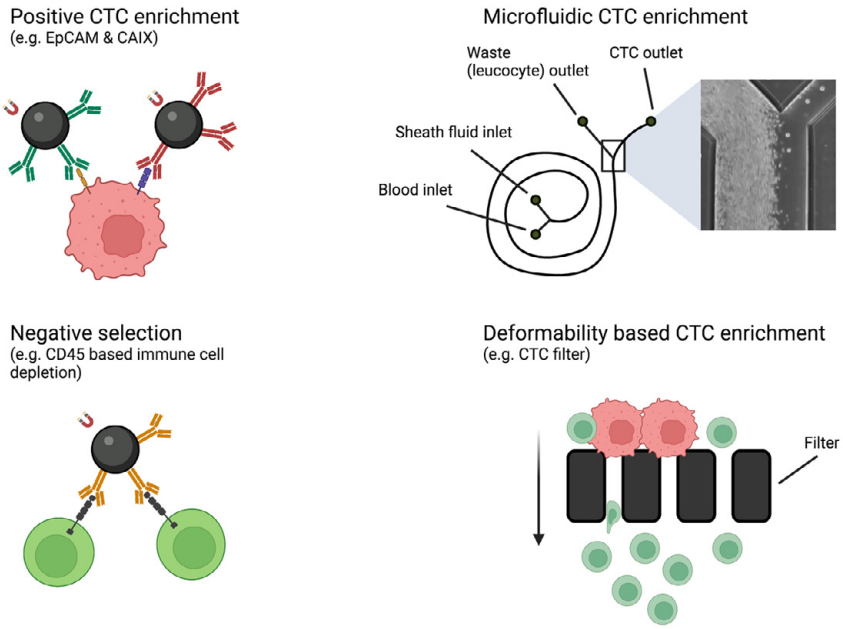

Figure 10: CTC enrichment approaches
Commonly employed CTC enrichment approaches with label-dependent methods (left) where antibodies against a target of interest (e.g. EpCAM) are attached to magnetic beads, which can subsequently be used to pull down tumour cells (positive selection) or used to remove leucocytes (negative selection). Label-independent methods (right) which exploit differences in physical properties between CTCs and leucocytes, such as their size or deformability. Microfluidic CTC enrichment based on ClearCell® FX platform [251].

Low and variable detection rates may be explained by the differences in EpCAM expression across RCC subtypes, where ccRCC is known to display the lowest expression out of the three primary subtypes [191, 192]. Another explanation for low detection rates may lie in the mixed-profile of RCC CTCs, where it has been shown that CTCs from RCC patients may have epithelial, mesenchymal, stem cell-like or mixed-cell profiles [193, 194]. In line with this, it was shown by one research group that RCC tumours may acquire EpCAM expression, demonstrated by differences in expression between primary and metastatic tumour sites in the same patient [195]. This may pose a niche opportunity for use of the CellSearch® system in ccRCC. Other approaches have included using other CTC targets in addition to EpCAM, such as the use of CAIX, a well-characterised HIF-1α target gene that is commonly expressed on ccRCC tumour cells [21, 196, 197]. Apart from CTC capture markers, exclusion markers can also be used to exclude blood cells such as leucoytes (negative selection), and these usually involve the cluster of differentiation markers such as CD45 [196]. It is generally accepted that methods

that use multiple targets in order to capture CTCs and exclude blood cells are superior to single target capture methods, since CTCs are heterogenous with regards to expression of cell surface molecules [198-200]. Multi-target capture strategies allow for a wider range of CTCs to be enriched whereas an EpCAM only approach will miss EpCAM negative CTCs.

Given the disappointing CTC detection rates with CellSearch® in RCC, a majority of studies exploring CTC enrichment have been intensely focused on label independent approaches that rely upon the biophysical properties of CTCs, such as size or deformability. Arguably, label independent approaches are not constrained by cell-surface marker expression or heterogeneity between CTCs [201, 202]. Size differences between CTCs and blood cells is a feature used to positively enrich CTC from whole blood samples. In general, tumour cells have been recorded to be larger than erythrocytes and leucocytes. In one study, dielectrophoretic field-flow fractionation was used to measure the diameter of tumour cells in the NCI-60 cell line panel in addition to blood cells, and it was found that tumour cell lines had a diameter in the range from 11.7-23.8μm and blood cells between 6.2-9.4μm [203]. Our group (Paper 3) as well as one other has shown that ccRCC tumour cell lines, including primary ccRCC cell lines, are on the higher end of the reported range, typically larger than 15μm [204, 205]. These findings classify ccRCC has a highly suitable tumour type for size-based enrichment approaches. CTCs can also be found in aggregates within the circulation, dubbed circulating tumour microemboli (CTM) and have been demonstrated to be present in the blood of RCC patients [206]. These are usually clusters of 2-50 CTCs together with leucocytes, cancer-associated fibroblasts, endothelial cells and platelets [207].These clusters, due to their increased size, may also be more suited for size and deformability-based isolation methods. Size-based isolation methods typically employ microfluidic chips, whereby blood cells and tumour cells are separated within a microfluidic capillary system. Deformability based separation techniques utilise a slit filtration system, relying on blood cells' smaller size and greater ability to 'deform' and pass through the slit filter whilst larger, less deformable tumour cells are captured. Although size-based isolation methods (in-comparison to label dependent approaches) have many advantages including improved processing speed and low cost of materials [208], they face issues with regards to microfluidic tube clogging, higher blood volume requirement and loss of smaller CTCs [209, 210].

Once the hurdle of capturing and isolating sufficient CTCs from whole blood has been overcome, these CTCs then need to be detected and characterised. The most commonly applied method for the detection and characterisation of CTCs has so far been immunocytochemical analysis, including immunofluorescence detection [198]. However, DNA alterations/mutations, RNA expression and protein

expression can also be queried via methods such as sequencing, digital droplet PCR (ddPCR) and fluorescence in-situ hybridisation (FISH). Furthermore, CTCs may be cultured *in-vitro* or xenografted *in-vivo*, where therapy sensitivity/resistance and related mechanisms can be assessed (Figure 11).

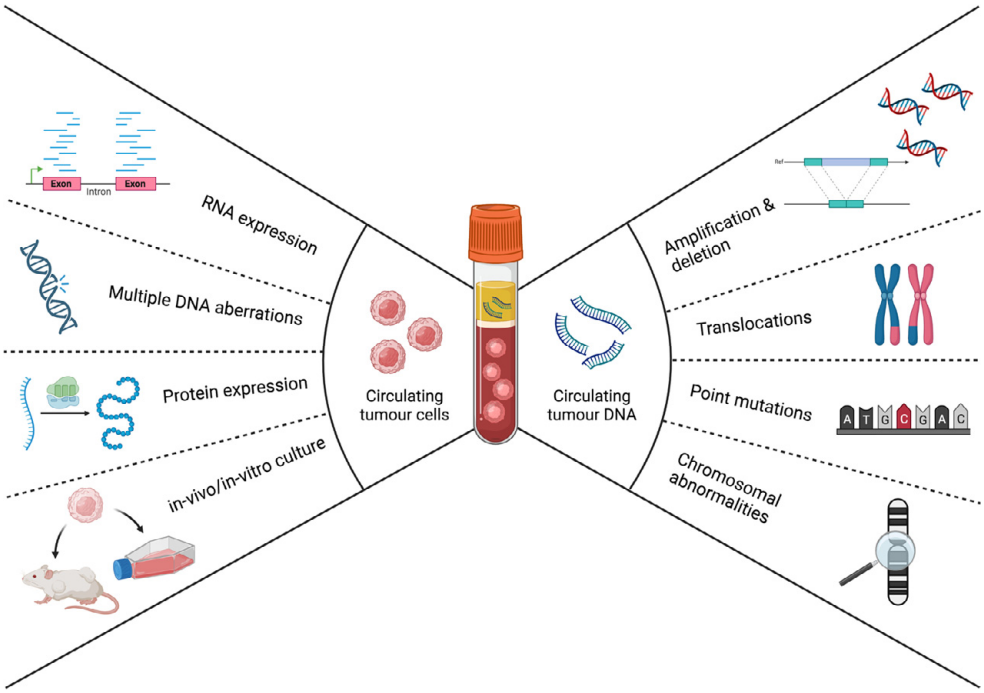

Figure 11: Possibilities with CTCs and ctDNA
Overview of query possibilities with circulating tumour cells and circulating tumour DNA. Adapted from gene-quantification.de

CTC Applications in RCC

Given the role of CTCs in the metastatic cascade, their applicability in RCC screening and diagnosis have been limited and will most likely require more sensitive techniques to be developed. However, the presence and number of CTCs, as well as their RNA expression or mutational status has been used to gain clinically relevant information such as patient prognosis and metastasis prediction, treatment selection and treatment monitoring [209]. Although much investigation has been carried out with regards to these clinical applications in other tumour types, efforts within RCC have had variable success.

Multiple studies have found that the presence of CTCs and/or CTMs in combination with tissue pathology were able to predict metastasis-free survival in

RCC patients [206, 211]. With regards to prognosis, it is a well-established fact that the presence of CTCs in the bloodstream correlates with poorer prognosis in many tumour types [184, 186, 200, 209]. With respect to RCC, multiple studies have also demonstrated that CTC counts, their characteristics and protein expression can significantly be associated with patient prognosis although there is a tendency for studies to correlate CTC counts to proxy predictive parameters such as tumour size [204]or Ki-67 expression [212], which in themselves are prognostic factors. CTCs in RCC have also been investigated to predict the effect of surgical technique on post-operative risk of metastasis and these studies conclude that open radical nephrectomies resulted in higher number of CTCs in the bloodstream in comparison to open partial nephrectomies or laparoscopic techniques [213, 214]. These findings make clear the importance of choosing the appropriate surgical techniques in an effort to minimise early post-operative metastasis thus improving patient prognosis.

CTCs have also been used for the selection and monitoring of treatment in RCC. The role of CTCs in predicting outcome of TKI therapy has been inconclusive, where one study demonstrated that CTC positive patients had a significantly poorer response to TKI therapy [215] but another study found no association between the two parameters [195]. Promising results, however, have been reported in relation to treatment response with ICIs, where PD-L1 expression in patient CTCs was positively associated with tumour response to PD-1/PD-L1 blockade [216]. Furthermore, for ccRCC patients undergoing anti-PD-1 therapy, the overall CTC count and number of CTCs expressing PD-L1 was dynamic where these changes correlated with improved disease outcome [217]. This data demonstrates the possibility of real-time immunotherapy response monitoring in ccRCC patients.

Circulating tumour DNA approaches and applications

Another analyte which may be used within the context of liquid biopsies is circulating tumour DNA. These are typically 50-150bp long fragments of DNA that are known to originate from tumour cells or CTCs during apoptosis [218-220] and can be found in blood plasma and urine [221]. ctDNA only represents a subset of cell-free DNA (cfDNA) in blood or urine, where the former constitutes $\geq$ 5–10% in late-stage tumours and $\leq$ 0.01–0.1% in early-stage tumours [222]. Thus, a key challenge in ctDNA analysis approaches is differentiating between ctDNA and cfDNA. Due to these reasons, somatic mutations are present at low frequencies in cfDNA (<3%) [223] and conventional next-generation sequencing technologies are not optimised to detect variants below allele frequencies of 5% [224]. However, with the advent of more sophisticated and sensitive sequencing methods, this is most

likely a hurdle that can be overcome. Genetic aberrations which may be queried with ctDNA analyses are summarised in Figure 11.

The last decade of research into ctDNA analysis in RCC has primarily focused on deciphering and optimising best approaches to detect ctDNA. Initially, studies used an approach where mutations found in the tumour guided detection of ctDNA in plasma [222, 225-228]. Although providing a high technical specificity, one drawback is that newly acquired mutations cannot be found within ctDNA with this approach. In one of these studies, the presence of ctDNA was investigated in 640 patients with various tumour types. ctDNA was detected in 40% of metastatic RCC patients, which classified RCC as a low-ctDNA tumour type [222]. This is surprising since the tumour guided approach strategy has potential to be one of the most technically sensitive methods [229]. However, two studies that employed a large number of tumour-specific DNA variants applied these patient tumour individualised mutation panels to detect ctDNA, resulting in improvement of ctDNA detection rates [225, 228]. Together, these studies as well as studies preceding these have demonstrated that targeted-sequencing approaches, allowing for deeper coverage, improved the rates of ctDNA detection in RCC patients [226, 228, 230]. This demonstrates that extremely sensitive methods are required for ctDNA detection and analysis in RCC, given its classification as a low ctDNA malignancy. Other notable approaches include global sequencing of plasma (cfDNA) [219] and targeted/global methylation analysis of plasma[231]. However these approaches yielded low sensitivity in RCC except for global methylation analyses [228, 232].

A key parameter that needs to be upheld when assessing ctDNA is the concordance between mutations in ctDNA and the tumour tissue [229]. This is important when applying ctDNA findings to prognosticate patients and monitor response to therapy. The concordance between tumour tissue and ctDNA was shown to be 77% in one study with mRCC patients [233], but further exploration of this in larger patient cohorts is required. Furthermore, ctDNA analysis has the potential to overcome spatial and temporal tumour heterogeneity and was evidenced in a study where 9/10 tumour region-specific mutations were detected in ctDNA [226]. Clinical applications of ctDNA show promise, where multiple studies have evidenced the correlation between ctDNA detection in various stages of RCC and higher risk of death [231], shorter PFS [234, 235], CSS [234] or OS [233, 235]. Together, this accumulated data warrants the further gathering of evidence on the applicability of ctDNA in prognosticating RCC patients and should be expanded to include larger and more diverse cohorts. With advances in technologies that are able to more sensitively and specifically detect ctDNA, the inclusion of these ctDNA based liquid biopsy in the clinical management of RCC patients is inevitable.

The Present Investigation

Overview and Aims

Renal cell carcinoma is a sophisticated and nuanced malignancy whereby its complexity is demonstrated by the ever-evolving number of subtypes, despite their shared organ of origin. These subtype dependent characteristics are relevant in the diagnosis and prognostication of RCC patients. Furthermore, the most prevalent subtype of RCC, ccRCC, is to a large extent defined by its constitutive hypoxia signalling. Therefore, this thesis aimed to explore the subtype dependent transcriptomics and protein expression, subsequently harnessing this knowledge to establish and build on current biopsy and prognostication approaches in RCC.

Specific aims of this thesis were:

I. To investigate and validate novel HIF-2α specific target genes in ccRCC

II. To explore novel transcription factors operating in ccRCC tumours that may inform tumour pathology and thus patient prognosis.

III. To develop a liquid biopsy workflow in order to isolate, differentiate and subtype tumour cells from whole blood

Paper I: SEMA5B is a HIF-2α specific target gene in renal proximal tubule cells

More than 80% of ccRCCs are characterised by the presence of the non-functional pVHL protein, leading to the constitutive activation of (pseudo)hypoxia signalling in these tumours [66]. Furthermore, over the last decade and a half, HIF-2α has been heavily implicated in the multi-faceted tumorigenic properties of ccRCC tumours, whilst HIF-1α is understood to contribute towards tumour-suppression [43, 51]. In this vein, we sought to explore novel, downstream HIF-2α specific target genes operating in normal kidney epithelial tubule cells and then sought to extend these findings to ccRCC.

With the goal of emulating the early events that take place in the tumorigenesis of renal proximal tubular epithelial cells, we characterised an inhibitor-cell line model. We utilised the renal proximal tubule epithelial/TERT1 (RPTEC/TERT1) cell line, one of the few renal proximal tubule epithelial cell lines that can be continuously cultured whilst retaining many of the *in-vivo* characteristics of the parent cells [236]. This cell line was treated with the highly selective pVHL inhibitor (VH298) that binds pVHL with high affinity, thereby diminishing its ability to polyubiquitinate HIFα, downstream of PHD hydroxylation [15]. In combination with pVHL inhibition, we also inhibited HIF-2α with PT2385, a predecessor of Belzutifan which is currently FDA approved for the treatment of hereditary ccRCC [167, 168]. Prior to exploratory analyses on this model, we ensured that single and combination treatment of RPTEC/TERT1 cells with these inhibitors elicited a predictable HIF-1α/ HIF-2α expression profile. Furthermore, we observed a marked reduction in HRE activity in 786-O (only HIF-2α positive) upon HIF-2α inhibition, warranting the use of this inhibitor cocktail in exploring HIF-2α target genes in renal/RCC cells.

RPTEC/TERT1 cells with pVHL and HIF-2α inhibition were subjected to mRNA-sequencing along with RPTEC/TERT1 cells with only pVHL inhibition. Differential gene expression (DGE) analysis was utilised to build a candidate list of potentially HIF-2α specific genes. Amongst the top ten most differentially expressed genes, a well described HIF-2α target gene *SLC7A5* was present, validating our approach for seeking out HIF-2α targets. Nonetheless, from this list we selected *SEMA5B* to further validate the regulatory links to HIF-2α. The other genes on this list were excluded due to unconfirmed protein expression pattern (upon inhibitor treatment), poorly performing antibodies or poor expression in ccRCC tumours (TCGA). We were able to definitively validate the HIF-2α regulatory link to *SEMA5B* in ccRCC cell lines as well via pharmacological and RNA interference-based methods. With equal importance, we ruled out HIF-1α regulation of *SEMA5B* with the same methods. The role of SEMA5B in ccRCC tumour is not known due to conflicting evidence. However, our data linking SEMA5B to HIF-2α may provide some hints towards its functional capacity within tumours given the established understanding of HIF-2α and its oncogenic role in ccRCC.

Further lines of evidence, such as HIF-2α expression correlation analysis with SEMA5B as well as the identification of HIF-2 binding sites on the HRE of SEMA5B reinforce our conclusions. We also noticed in our data that knocking down HIF-1α had the tendency to increase SEMA5B expression in RCC cell lines, possibly hinting at an opposing regulatory role for this transcription factor. This

divergent regulatory influence of the two transcription factors HIF-1α and HIF-2α would not be surprising given their polar classification within ccRCC [66]. Furthermore, it is well understood that the target gene specificity of HIF-2α (and HIF-1α) is highly cell-type and local transcription machinery dependent [15]. Given this, the conserved regulation between HIF-2α and SEMA5B between normal renal proximal tubule cells and in transformed ccRCC tumour cells could suggest that this pathway is important for tumorigenesis and the maintenance of tumorigenicity. Specifically, SEMA5B is known to be cleaved from the cell surface and functions in a paracrine manner as a repulsive cue for neuronal axons [237, 238]. Other data suggests that SEMA5B negatively regulates endothelial cell tube-formation and proliferation [239]. Extrapolating from these findings, it could be suggested that SEMA5B elicits an anti-angiogenic function within the tumour microenvironment, providing an inhibitory cue in order to balance VEGF driven angiogenesis. This would be especially relevant in ccRCC tumours where exaggerated levels of VEGF signalling are present irrespective of oxygen availability.

In conclusion, we present multiple lines of evidence definitively linking the expression of HIF-2α to SEMA5B in normal and transformed renal proximal tubule cells, suggesting that HIF-2α regulates SEMA5B expression. Along these lines, we contribute to the known pool of HIF-2α targets in RCC and uncover some of the pseudohypoxia biology in these tumours. Further work is warranted to decipher the exact function(s) of SEMA5B in ccRCC. Depending on these results, the pharmacological inhibition of SEMA5B and other HIF-2α targets may be a necessary reality in the event that HIF-2α inhibition does not pass clinical trials. Targeting HIF-2α specific genes, downstream of HIF-2α, is likely to reduce the rate of adverse events in a patient, due to its targeted and specific nature as opposed to inhibiting an entire signalling cascade that may be required for non-tumour, physiological processes. This could especially be relevant and useful in combination therapy regimens where there is an increased likelihood of unpleasant side-effects for the patient.

Paper II: Identification and validation of NFIA as a novel and prognostic marker in renal cell carcinoma

Patient prognostication tools are an essential component of the clinical management of RCC patients and often guide therapy selection and disease monitoring frequency. Unfortunately, the use of molecular markers such as CAIX, PTEN or CXCR4 to prognosticate RCC patients is not the current standard of care, due to

their apparent prognostic inferiority compared to anatomical and clinical prognostic factors [139]. Given this, we aimed to explore novel molecular prognostic markers whilst using the previously characterised expression profile of the transcription factor HNF4A in the nephron [104] to guide our exploration. Furthermore, we strived to add to the currently available prognostic information of HNF4A in RCC patients.

A bioinformatic analysis (SCENIC) was conducted on RNA expression profiles of ccRCC, pRCC and chRCC tumours within the TCGA, where transcription factor regulon activity was elucidated. Based on this analysis, *NFIA* conformed best to the subtype specificity and regulon activity levels to that of *HNF4A*, qualifying *NFIA* for further exploration. We observed that *NFIA* expression was elevated in ccRCC and pRCC but not chRCC tumours within the TCGA, confirming our observation from our SCENIC analysis. This subtype specific expression may in part be related to the cell of origin from which ccRCC and pRCC tumours arise from, as single-cell RNA data revealed that renal proximal tubule cells display elevated levels of *NFIA*. We explored the relationship between *NFIA* expression in the TCGA RCC tumours and CSS, concluding that high *NFIA* expression favoured CSS. This data encouraged us to explore the relationship between NFIA protein expression and RCC patient outcomes in a tumour-microarray (TMA) consisting of 313 ccRCC and 50 pRCC patients. We did not explore protein expression in chRCC patients, as initial exploration of this in the TCGA cohort proved that expression of NFIA was lacking in chRCC tumours and did not correlate with any clinical parameters.

In the TMA we found that NFIA and HNF4A expression negatively correlated with ccRCC stage and grade. Furthermore, we consolidated our TCGA-based findings that increased *NFIA* and *HNF4A* expression yielded favourable CSS. We extended these investigations into the pRCC subtype, and although favourable CSS was still associated with higher NFIA expression, we could not see any convincing stage or grade dependent tendencies. Expression of HNF4A protein completely lacked any association to stage, grade or CSS in the pRCC subtype. This is the first time that investigation of stage, grade and CSS in conjunction with HNF4A protein expression has been carried out in a large cohort of RCC patients that includes pRCC tumours.

We also tested whether NFIA could be used as an independent prognostic factor in ccRCC patients. Multivariable Cox-regression analysis was performed to determine if NFIA could be used to prognosticate ccRCC patients and showed that prognostication was achieved independently of tumour grade, tumour stage and four other clinicopathological parameters.

The biological relevance of NFIA expression in ccRCC tumours is not yet understood. However, multiple reports describe *NFIA* as a tumour-suppressor gene

in brain tumours [240], melanomas [241] leukaemia [242]and lymphomas [243] whilst other reports suggest an oncogenic role for NFIB (a NFIA family member with high DNA-binding sequence homology[244]) in breast cancer [245, 246]. Furthermore, it is known that NFIA is required for embryonic development of the renal system. Mice that are NFIA heterozygous or homozygous null display developmental renal abnormalities [247]. In development of other organ systems, it is understood that NFIA contributes towards cell differentiation and proliferation [248]. Taken together, it could be postulated that NFIA is important for the correct development of the kidney via proliferative and differentiation processes.

The above observations could explain the relationship between NFIA and ccRCC tumour aggressivity. NFIA expression might be retained in the adult kidneys, in cells such as renal progenitor cells which function to repopulate damaged kidney tubule cells [249]. It may well be that upon kidney injury (a known risk factor for RCC tumorigenesis [250]), the NFIA driven differentiation and proliferative programme 'kicks in' but in an unknown way becomes aberrant, giving rise to preneoplastic or early neoplastic lesions. This hypothesis would be in line with the high NFIA expression we describe in early stage and grade RCC tumours. In grade four ccRCC tumours, where we demonstrate a significantly lowered expression of NFIA, dedifferentiation of tumour cells is extensive. Without further investigation, it is challenging to postulate whether increased dedifferentiation is a result of lowered NFIA levels or if reduced NFIA expression is a result of other events, such as VHL inactivation in early ccRCC tumours and the subsequent HIF-2α expression in advanced stage/grade ccRCC tumours [66]. HIF-2α expression in ccRCC, especially in the absence of HIF-1α, in high grade tumours may provide the necessary influence over the transcriptome to maintain a dedifferentiated, stem-cell like phenotype [43] ruling out tumour dependence on an aberrant NFIA driven programme. Given these complexities, it is problematic to categorise *NFIA* as tumour-promoting or suppressing in ccRCC without extensive further work, but as it stands, our findings would hint at the latter.

Finally, further work is required to validate the clinical utility of NFIA expression via the interrogation of larger, more diverse patient cohorts. Ideally, other staining approaches, antibodies and quantification approaches also need to be validated to test if the relationship between NFIA and CSS remains. If the correlations and relationship between NFIA and the clinical parameters examined withstands workflow variations and large patient cohorts, NFIA may serve to be a clinically useful prognostic biomarker.

Paper III: Size-based isolation and detection of renal carcinoma cells from whole blood

Despite advances and clinical implementation of CTC enrichment methods, RCC patients have failed to benefit from these. The advent and implementation of CellSearch® in the clinic for breast cancer and prostate cancer has been successful, however due to the variable but generally poor expression of EpCAM on ccRCC and pRCC tumours, high enrichment rates have not been achieved with these tumours [185, 188, 195]. Instead of applying a label-dependent approach to isolate RCC CTCs, we validated and developed a workflow that utilises a label-independent enrichment platform, ClearCell® FX, to isolate RCC CTCs [251]. We also analysed publicly available RCC tumour subtype-specific transcriptomic data to identify gene expression profiles which could be applied to characterise enriched RCC CTCs. Given the low, yet possibly interfering background of leukocytes in our CTC enriched sample, identified gene expression profiles were confirmed to be negative in healthy whole blood.

RCC CTC enrichment was achieved via the ClearCell® FX platform that utilises a microfluidic chip to separate leucocytes from tumour cells. This platform had not been previously tested in the RCC setting, thus we investigated cell diameters of common RCC cell lines and primary RCC cells. We demonstrated that RCC tumour cells lines are generally much larger than leucocytes, suggesting the ClearCell® FX platform was suitable for RCC CTC enrichment. Through spike-in experiments, we showed that this platform could be used to enrich ccRCC tumour cells from whole blood with high sensitivity, where one ccRCC tumour cell in whole blood could be isolated and detected. In general, the ClearCell® FX platform demonstrated a recovery efficiency greater than 50% and 60% for RCC cell lines and primary RCC cells respectively.

Although CTC isolation and analyses are not currently at the forefront of the RCC liquid biopsy realm, I do believe that their full potential has not yet been uncovered. Recently, a large amount of research has focused on ctDNA analysis, and great advancement has been made in isolating and deciphering tumour related information from these nucleic acid fragments. However, it could be argued that the potential for CTCs to inform clinical decisions is much greater than that of ctDNA in the post-metastatic setting. For instance, a wealth of information can be obtained from isolated CTCs, ranging from insights into therapy resistance/sensitivity (if isolated and cultured), analysis of tumour DNA, analysis of tumour RNA expression and insights into protein expression. Furthermore, the cross talk between CTCs and other blood cells is being investigated, where the biological significance of these within the metastatic cascade is of clinical interest [252, 253]. Limitations of ctDNA

are especially highlighted in RCC patients, due to their low ctDNA detection rates [222]. The CTC field is perhaps suffering from a lag in technological advancement, and though the analyses of CTCs in RCC may have been exhausted for the time being, I would not rule CTCs out of clinical contention.

Acknowledgments

The work in this thesis has taken me a long five years and I am wholeheartedly grateful for everyone that has contributed towards my time in Lund and development as a scientist. **Håkan**, I am certain you have contributed the most to this development. Thank you for giving me an opportunity in your lab, despite my shoddy student internet connection disconnecting our interview video call multiple times. I was sure I was doomed at the time, but you had the patience to wait for me to return and still listened to what I had to say. Once I started in your lab as a fresh-faced PhD student, I soon learned that in fact, you had A LOT of patience. Thank you for always encouraging my professional growth, having a nurturing approach towards your students and genuinely having an 'open-door' policy, as opposed to that being something supervisors just say. I am especially grateful for the freedom you grant for unhindered scientific and technical exploration and I count myself lucky that I had this opportunity to work with you!

Sarah and Mazhar, I am grateful for all of your input towards this thesis and contributions towards our publications. I appreciate you sharing all of your knowledge and expertise on experimental techniques and always being willing to help. **Christina**, thank you for teaching me the basics of FFPE, cell culture and for running countless patient TMA staining optimisation cycles. I would've been so lost without your histology expertise! **Sara A**, thanks for being cheery in labs and providing candy whenever I looked stressed. **David L**, I appreciate all of your bioinformatic and theoretical contributions to the work in this thesis and your attention to detail! I especially appreciate your constructive criticism, as without it, growth is not possible. **Jennifer**, I appreciate you giving me an introduction to most of the molecular methods in this thesis. It was fun to kickstart the CTC project as a duo. **Siran**, thank you for all of your western blot troubleshooting expertise and advice.

Gjendine, Jessica and Anson, you have been great office mates and friends. Thank you for making sure there were snacks available at all times and spicing up the workday with venting sessions. **Matteo, Katarzyna, Lexi, Rebecca, Renee** I'm grateful for the lunch banter and work escapism. **Elinn** thank you for sharing all of your HIF wisdom and my pain in HIF western blotting!

I am extremely grateful to the lab managers **Margareta, Sara L** and **Maria** without whom TCR would be a very unproductive place. **Margareta**, thank you for your ever helpful and positive attitude! Thank you **Johanna** for all of your behind the scenes accounting and rallying efforts in organising TCR social events.

Sophie L, thank you for taking the time to help us wrap our heads around independent prognostic factors and being an allround cheerful soul. **Gottfrid**, thanks for your help with the slide scanner and scanning programmes! **Bianca**, **Peter, Martin and Helen N** I appreciate your input in the CTC study and provision of patient material. Thank you **Börje** for providing us with the TMA and all of your clinical input. Massive thank you to the staff at the Lund blood donation centre for your cooperation with the research community. Thank you to **Håkan Rundgren** and other ClearCell FX technicians in helping with troubleshooting and doing your utmost to get the platform back in action.

Thank you **Gunilla** from MediaTryck for being so understanding and for all of your help in thesis printing matters.

I am grateful for the **CanFaster** programme that pushed us outside of our comfort zone and I am especially thankful to **Sara Ek** and **Jana** for exemplifying the ethos of the programme and your efforts in shaping 'non-typical' PhD students.

I would also like to thank my half-time opponents, **Sofie and Kristina**, for providing valuable feedback as well as encouraging and guiding me for the next chapter.

I would also like to thank all of the **cancer patients** and healthy blood donors that made the work in this thesis possible- you are the past, present and future of translational cancer research. Additionally, I am grateful for all of the funding agencies that have made this work possible.

Thank you to **everyone at TCR** for creating a giving and collaborative environment and to anyone I have borrowed reagents/protocols/cells/time from!

James and David, this PhD journey together has been infinitely better with you two in it. James, you have followed me around to two offices like a bad smell, but I wouldn't have had it any other way. Thank you for being a great friend that listened when things were rough and celebrated with me when things went well. **David**, you were my first friend in Lund, and I am glad that hasn't changed. Thank you for offering unsolicited project advice, forcing me to be more social, sharing my love for the sun and messy dinner parties. **Claudio**, thanks for keeping David in check, for Italian cheeses and being the friendly Kämnärs neighbour I could borrow from in emergencies. **Chris and Xavier**, I appreciate your affinity for a good time and the impressive dinner party spreads.

To the people and puppy this thesis is dedicated to:

Amma & Thatha, this thesis is a testament of your abilities. I wouldn't be the curious-minded scientist I am without all of your encouragement and guidance from a young age. From allowing me to dismantle and 'fix' (expensive) electronics around the house to providing me with science encyclopaedias, you always nurtured my curiosity about the world. You give the best advice and have always taught me the value of integrity. I hate the fact that I had to spend so much time away from you in order to pursue my passions, but you never complained and always supported me and I am eternally grateful for that. **Kevin**, thank you for showing me that life doesn't need to be so difficult if you don't want it to be and for being a little bro I don't need to worry about.

Rosie, Kiera, Mya and **Joe**, thanks for always referring to me as the 'smart' one and making me feel good about doing a PhD, even if it was not the case at times.

My wonderful Wife, there are no number of words I can write here that will do justice to how much you have supported me through these years. From taking a leap of faith and moving to an alien country with me 5 years ago to building our lives and getting married in Sweden, it feels like we have been through five lifetimes worth of experiences. Although you didn't see first-hand how many experiments failed in the lab, you definitely heard about it at the end of each day and I am grateful that you always knew what to say. You are also the reason why I went back and tried again and again. I am forever grateful that you decided to join me on this tortuous journey and this accomplishment is every bit yours as it is mine (especially the stellar cover art!).

Finally, this thesis wouldn't be complete without giving credit to my little four-legged PhD trauma therapist, **Pepper**.

References

1.AC Guyton JH. Textbook of medical physiology. 11 ed. Philadelphia: Elsevier Saunders; 2006.

2.Jelkmann W. Regulation of erythropoietin production. J Physiol. 2011;589(Pt 6):1251-8.

3.A McDonough ST, B Brebber, F Rector. The Kidney Physiology and Pathophysiology Philadelphia: Elsevier 2011.

4.Lote C. Principles of renal physiology. 5 ed. New York: Springer; 2012.

5.Lee P, Chandel NS, Simon MC. Cellular adaptation to hypoxia through hypoxia inducible factors and beyond. Nat Rev Mol Cell Biol. 2020;21(5):268-83.

6.Hammarlund EU, Flashman E, Mohlin S, et al. Oxygen-sensing mechanisms across eukaryotic kingdoms and their roles in complex multicellularity. Science. 2020;370(6515).

7.Chen PS, Chiu WT, Hsu PL, et al. Pathophysiological implications of hypoxia in human diseases. J Biomed Sci. 2020;27(1):63.

8.Kaelin WG, Jr., Ratcliffe PJ. Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. Mol Cell. 2008;30(4):393-402.

9.Bracken CP, Whitelaw ML, Peet DJ. The hypoxia-inducible factors: key transcriptional regulators of hypoxic responses. Cell Mol Life Sci. 2003;60(7):1376-93.

10.Cuomo F, Dell'Aversana C, Chioccarelli T, et al. HIF3A Inhibition Triggers Browning of White Adipocytes via Metabolic Rewiring. Front Cell Dev Biol. 2021;9:740203.

11.Yu AY, Frid MG, Shimoda LA, et al. Temporal, spatial, and oxygen-regulated expression of hypoxia-inducible factor-1 in the lung. Am J Physiol. 1998;275(4):L818-26.

12.Huang LE, Arany Z, Livingston DM, et al. Activation of hypoxia-inducible transcription factor depends primarily upon redox-sensitive stabilization of its alpha subunit. J Biol Chem. 1996;271(50):32253-9.

13.Fong GH, Takeda K. Role and regulation of prolyl hydroxylase domain proteins. Cell Death Differ. 2008;15(4):635-41.

14.Hirsila M, Koivunen P, Gunzler V, et al. Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor. J Biol Chem. 2003;278(33):30772-80.

15.Ivan M, Kaelin WG, Jr. The EGLN-HIF O(2)-Sensing System: Multiple Inputs and Feedbacks. Mol Cell. 2017;66(6):772-9.

16.Markolovic S, Wilkins SE, Schofield CJ. Protein Hydroxylation Catalyzed by 2-Oxoglutarate-dependent Oxygenases. J Biol Chem. 2015;290(34):20712-22.

17.Masson N, Willam C, Maxwell PH, et al. Independent function of two destruction domains in hypoxia-inducible factor-alpha chains activated by prolyl hydroxylation. EMBO J. 2001;20(18):5197-206.

18.Jaakkola P, Mole DR, Tian YM, et al. Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. Science. 2001;292(5516):468-72.

19.Ivan M, Kondo K, Yang H, et al. HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. Science. 2001;292(5516):464-8.

20.Duan DR, Pause A, Burgess WH, et al. Inhibition of transcription elongation by the VHL tumor suppressor protein. Science. 1995;269(5229):1402-6.

21.Schodel J, Grampp S, Maher ER, et al. Hypoxia, Hypoxia-inducible Transcription Factors, and Renal Cancer. Eur Urol. 2016;69(4):646-57.

22.Semenza GL. The Genomics and Genetics of Oxygen Homeostasis. Annu Rev Genomics Hum Genet. 2020;21:183-204.

23.Wenger RH, Stiehl DP, Camenisch G. Integration of oxygen signaling at the consensus HRE. Sci STKE. 2005;2005(306):re12.

24.Semenza GL. Hypoxia-inducible factors in physiology and medicine. Cell. 2012;148(3):399-408.

25.Semenza GL, Wang GL. A nuclear factor induced by hypoxia via de novo protein synthesis binds to the human erythropoietin gene enhancer at a site required for transcriptional activation. Mol Cell Biol. 1992;12(12):5447-54.

26.Elvidge GP, Glenny L, Appelhoff RJ, et al. Concordant regulation of gene expression by hypoxia and 2-oxoglutarate-dependent dioxygenase inhibition: the role of HIF-1alpha, HIF-2alpha, and other pathways. J Biol Chem. 2006;281(22):15215-26.

27.Arany Z, Huang LE, Eckner R, et al. An essential role for p300/CBP in the cellular response to hypoxia. Proc Natl Acad Sci U S A. 1996;93(23):12969-73.

28.Ema M, Hirota K, Mimura J, et al. Molecular mechanisms of transcription activation by HLF and HIF1alpha in response to hypoxia: their stabilization and redox signal-induced interaction with CBP/p300. EMBO J. 1999;18(7):1905-14.

29.Carrero P, Okamoto K, Coumailleau P, et al. Redox-regulated recruitment of the transcriptional coactivators CREB-binding protein and SRC-1 to hypoxia-inducible factor 1alpha. Mol Cell Biol. 2000;20(1):402-15.

30.Lando D, Peet DJ, Whelan DA, et al. Asparagine hydroxylation of the HIF transactivation domain a hypoxic switch. Science. 2002;295(5556):858-61.

31.Beischlag TV, Taylor RT, Rose DW, et al. Recruitment of thyroid hormone receptor/retinoblastoma-interacting protein 230 by the aryl hydrocarbon receptor nuclear translocator is required for the transcriptional response to both dioxin and hypoxia. J Biol Chem. 2004;279(52):54620-8.

32.Dayan F, Roux D, Brahimi-Horn MC, et al. The oxygen sensor factor-inhibiting hypoxia-inducible factor-1 controls expression of distinct genes through the bifunctional transcriptional character of hypoxia-inducible factor-1alpha. Cancer Res. 2006;66(7):3688-98.

33.Koivunen P, Hirsila M, Gunzler V, et al. Catalytic properties of the asparaginyl hydroxylase (FIH) in the oxygen sensing pathway are distinct from those of its prolyl 4-hydroxylases. J Biol Chem. 2004;279(11):9899-904.

34.Bracken CP, Fedele AO, Linke S, et al. Cell-specific regulation of hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha stabilization and transactivation in a graded oxygen environment. J Biol Chem. 2006;281(32):22575-85.

35.Liu YV, Semenza GL. RACK1 vs. HSP90: competition for HIF-1 alpha degradation vs. stabilization. Cell Cycle. 2007;6(6):656-9.

36.Semenza GL. A compendium of proteins that interact with HIF-1alpha. Exp Cell Res. 2017;356(2):128-35.

37.Holmquist-Mengelbier L, Fredlund E, Lofstedt T, et al. Recruitment of HIF-1alpha and HIF-2alpha to common target genes is differentially regulated in neuroblastoma: HIF-2alpha promotes an aggressive phenotype. Cancer Cell. 2006;10(5):413-23.

38.Elvert G, Kappel A, Heidenreich R, et al. Cooperative interaction of hypoxia-inducible factor-2alpha (HIF-2alpha) and Ets-1 in the transcriptional activation of vascular endothelial growth factor receptor-2 (Flk-1). J Biol Chem. 2003;278(9):7520-30.

39.Hu CJ, Iyer S, Sataur A, et al. Differential regulation of the transcriptional activities of hypoxia-inducible factor 1 alpha (HIF-1alpha) and HIF-2alpha in stem cells. Mol Cell Biol. 2006;26(9):3514-26.

40.McIntyre A, Harris AL. The Role of pH Regulation in Cancer Progression. Recent Results Cancer Res. 2016;207:93-134.

41.Kaluz S, Kaluzova M, Liao SY, et al. Transcriptional control of the tumor- and hypoxia-marker carbonic anhydrase 9: A one transcription factor (HIF-1) show? Biochim Biophys Acta. 2009;1795(2):162-72.

42.Seagroves TN, Ryan HE, Lu H, et al. Transcription factor HIF-1 is a necessary mediator of the pasteur effect in mammalian cells. Mol Cell Biol. 2001;21(10):3436-44.

43.Choueiri TK, Kaelin WG, Jr. Targeting the HIF2-VEGF axis in renal cell carcinoma. Nat Med. 2020;26(10):1519-30.

44.Befani C, Liakos P. The role of hypoxia-inducible factor-2 alpha in angiogenesis. J Cell Physiol. 2018;233(12):9087-98.

45.Thomlinson RH, Gray LH. The histological structure of some human lung cancers and the possible implications for radiotherapy. Br J Cancer. 1955;9(4):539-49.

46.McIntyre A, Harris AL. Metabolic and hypoxic adaptation to anti-angiogenic therapy: a target for induced essentiality. EMBO Mol Med. 2015;7(4):368-79.

47.Singleton DC, Rouhi P, Zois CE, et al. Hypoxic regulation of RIOK3 is a major mechanism for cancer cell invasion and metastasis. Oncogene. 2015;34(36):4713-22.

48.Walsh JC, Lebedev A, Aten E, et al. The clinical importance of assessing tumor hypoxia: relationship of tumor hypoxia to prognosis and therapeutic opportunities. Antioxid Redox Signal. 2014;21(10):1516-54.

49.Rini BI, Campbell SC, Escudier B. Renal cell carcinoma. Lancet. 2009;373(9669):1119-32.

50.Hsieh JJ, Purdue MP, Signoretti S, et al. Renal cell carcinoma. Nat Rev Dis Primers. 2017;3:17009.

51.Shen C, Beroukhim R, Schumacher SE, et al. Genetic and functional studies implicate HIF1alpha as a 14q kidney cancer suppressor gene. Cancer Discov. 2011;1(3):222-35.

52.Gordan JD, Lal P, Dondeti VR, et al. HIF-alpha effects on c-Myc distinguish two subtypes of sporadic VHL-deficient clear cell renal carcinoma. Cancer Cell. 2008;14(6):435-46.

53.Maxwell PH, Wiesener MS, Chang GW, et al. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. Nature. 1999;399(6733):271-5.

54.Kondo K, Kim WY, Lechpammer M, et al. Inhibition of HIF2alpha is sufficient to suppress pVHL-defective tumor growth. PLoS Biol. 2003;1(3):E83.

55.Zimmer M, Doucette D, Siddiqui N, et al. Inhibition of hypoxia-inducible factor is sufficient for growth suppression of VHL-/- tumors. Mol Cancer Res. 2004;2(2):89-95.

56.Raval RR, Lau KW, Tran MG, et al. Contrasting properties of hypoxia-inducible factor 1 (HIF-1) and HIF-2 in von Hippel-Lindau-associated renal cell carcinoma. Mol Cell Biol. 2005;25(13):5675-86.

57.Murakami M, Muranishi S. [Insulin suppository in the treatment of diabetic patients]. Nihon Rinsho. 1990;48 Suppl:1000-5.

58.Maranchie JK, Vasselli JR, Riss J, et al. The contribution of VHL substrate binding and HIF1-alpha to the phenotype of VHL loss in renal cell carcinoma. Cancer Cell. 2002;1(3):247-55.

59.Kim WY, Safran M, Buckley MR, et al. Failure to prolyl hydroxylate hypoxia-inducible factor alpha phenocopies VHL inactivation in vivo. EMBO J. 2006;25(19):4650-62.

60.Rankin EB, Higgins DF, Walisser JA, et al. Inactivation of the arylhydrocarbon receptor nuclear translocator (Arnt) suppresses von Hippel-Lindau disease-associated vascular tumors in mice. Mol Cell Biol. 2005;25(8):3163-72.

61.Rankin EB, Rha J, Selak MA, et al. Hypoxia-inducible factor 2 regulates hepatic lipid metabolism. Mol Cell Biol. 2009;29(16):4527-38.

62.Rankin EB, Rha J, Unger TL, et al. Hypoxia-inducible factor-2 regulates vascular tumorigenesis in mice. Oncogene. 2008;27(40):5354-8.

63.Rankin EB, Biju MP, Liu Q, et al. Hypoxia-inducible factor-2 (HIF-2) regulates hepatic erythropoietin in vivo. J Clin Invest. 2007;117(4):1068-77.

64.Mandriota SJ, Turner KJ, Davies DR, et al. HIF activation identifies early lesions in VHL kidneys: evidence for site-specific tumor suppressor function in the nephron. Cancer Cell. 2002;1(5):459-68.

65.Purdue MP, Johansson M, Zelenika D, et al. Genome-wide association study of renal cell carcinoma identifies two susceptibility loci on 2p21 and 11q13.3. Nat Genet. 2011;43(1):60-5.

66.Shen C, Kaelin WG, Jr. The VHL/HIF axis in clear cell renal carcinoma. Semin Cancer Biol. 2013;23(1):18-25.

67.Klatte T, Rao PN, de Martino M, et al. Cytogenetic profile predicts prognosis of patients with clear cell renal cell carcinoma. J Clin Oncol. 2009;27(5):746-53.

68.Bukavina L, Bensalah K, Bray F, et al. Epidemiology of Renal Cell Carcinoma: 2022 Update. Eur Urol. 2022;82(5):529-42.

69.Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209-49.

70.Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.

71.Scelo G, Larose TL. Epidemiology and Risk Factors for Kidney Cancer. J Clin Oncol. 2018;36(36):JCO2018791905.

72.Scelo G, Li P, Chanudet E, et al. Variability of Sex Disparities in Cancer Incidence over 30 Years: The Striking Case of Kidney Cancer. Eur Urol Focus. 2018;4(4):586-90.

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

74.Hemminki K, Forsti A, Hemminki A, et al. Progress in survival in renal cell carcinoma through 50 years evaluated in Finland and Sweden. PLoS One. 2021;16(6):e0253236.

75.Liu X, Peveri G, Bosetti C, et al. Dose-response relationships between cigarette smoking and kidney cancer: A systematic review and meta-analysis. Crit Rev Oncol Hematol. 2019;142:86-93.

76.Lauby-Secretan B, Scoccianti C, Loomis D, et al. Body Fatness and Cancer--Viewpoint of the IARC Working Group. N Engl J Med. 2016;375(8):794-8.

77.Haggstrom C, Rapp K, Stocks T, et al. Metabolic factors associated with risk of renal cell carcinoma. PLoS One. 2013;8(2):e57475.

78.Macleod LC, Hotaling JM, Wright JL, et al. Risk factors for renal cell carcinoma in the VITAL study. J Urol. 2013;190(5):1657-61.

79.Weikert S, Boeing H, Pischon T, et al. Blood pressure and risk of renal cell carcinoma in the European prospective investigation into cancer and nutrition. Am J Epidemiol. 2008;167(4):438-46.

80.Moore SC, Lee IM, Weiderpass E, et al. Association of Leisure-Time Physical Activity With Risk of 26 Types of Cancer in 1.44 Million Adults. JAMA Intern Med. 2016;176(6):816-25.

81.Stanifer JW, Stapleton HM, Souma T, et al. Perfluorinated Chemicals as Emerging Environmental Threats to Kidney Health: A Scoping Review. Clin J Am Soc Nephrol. 2018;13(10):1479-92.

82.Stengel B. Chronic kidney disease and cancer: a troubling connection. J Nephrol. 2010;23(3):253-62.

83.Moch H, Cubilla AL, Humphrey PA, et al. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. Eur Urol. 2016;70(1):93-105.

84.Ng KL, Rajandram R, Morais C, et al. Differentiation of oncocytoma from chromophobe renal cell carcinoma (RCC): can novel molecular biomarkers help solve an old problem? J Clin Pathol. 2014;67(2):97-104.

85.Lindgren D, Sjolund J, Axelson H. Tracing Renal Cell Carcinomas back to the Nephron. Trends Cancer. 2018;4(7):472-84.

86.Chen F, Zhang Y, Senbabaoglu Y, et al. Multilevel Genomics-Based Taxonomy of Renal Cell Carcinoma. Cell Rep. 2016;14(10):2476-89.

87.Cancer Genome Atlas Research N. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature. 2013;499(7456):43-9.

88.Sato Y, Yoshizato T, Shiraishi Y, et al. Integrated molecular analysis of clear-cell renal cell carcinoma. Nat Genet. 2013;45(8):860-7.

89.Beroukhim R, Brunet JP, Di Napoli A, et al. Patterns of gene expression and copy-number alterations in von-hippel lindau disease-associated and sporadic clear cell carcinoma of the kidney. Cancer Res. 2009;69(11):4674-81.

90.Moore LE, Jaeger E, Nickerson ML, et al. Genomic copy number alterations in clear cell renal carcinoma: associations with case characteristics and mechanisms of VHL gene inactivation. Oncogenesis. 2012;1(6):e14.

91.Nickerson ML, Jaeger E, Shi Y, et al. Improved identification of von Hippel-Lindau gene alterations in clear cell renal tumors. Clin Cancer Res. 2008;14(15):4726-34.

92.Young AC, Craven RA, Cohen D, et al. Analysis of VHL Gene Alterations and their Relationship to Clinical Parameters in Sporadic Conventional Renal Cell Carcinoma. Clin Cancer Res. 2009;15(24):7582-92.

93.Kaelin WG, Jr. Von Hippel-Lindau disease: insights into oxygen sensing, protein degradation, and cancer. J Clin Invest. 2022;132(18).

94.Ward PS, Thompson CB. Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. Cancer Cell. 2012;21(3):297-308.

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

96.Dalgliesh GL, Furge K, Greenman C, et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature. 2010;463(7279):360-3.

97.Varela I, Tarpey P, Raine K, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. Nature. 2011;469(7331):539-42.

98.Guo G, Gui Y, Gao S, et al. Frequent mutations of genes encoding ubiquitin-mediated proteolysis pathway components in clear cell renal cell carcinoma. Nat Genet. 2011;44(1):17-9.

99.Hakimi AA, Chen YB, Wren J, et al. Clinical and pathologic impact of select chromatin-modulating tumor suppressors in clear cell renal cell carcinoma. Eur Urol. 2013;63(5):848-54.

100.Ho TH, Park IY, Zhao H, et al. High-resolution profiling of histone h3 lysine 36 trimethylation in metastatic renal cell carcinoma. Oncogene. 2016;35(12):1565-74.

101.Simon JM, Hacker KE, Singh D, et al. Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. Genome Res. 2014;24(2):241-50.

102.Kanu N, Gronroos E, Martinez P, et al. SETD2 loss-of-function promotes renal cancer branched evolution through replication stress and impaired DNA repair. Oncogene. 2015;34(46):5699-708.

103.Patel SA, Hirosue S, Rodrigues P, et al. The renal lineage factor PAX8 controls oncogenic signalling in kidney cancer. Nature. 2022;606(7916):999-1006.

104.Lindgren D, Eriksson P, Krawczyk K, et al. Cell-Type-Specific Gene Programs of the Normal Human Nephron Define Kidney Cancer Subtypes. Cell Rep. 2017;20(6):1476-89.

105.Delahunt B, Eble JN. Papillary renal cell carcinoma: a clinicopathologic and immunohistochemical study of 105 tumors. Mod Pathol. 1997;10(6):537-44.

106.Zbar B, Tory K, Merino M, et al. Hereditary papillary renal cell carcinoma. J Urol. 1994;151(3):561-6.

107.Cancer Genome Atlas Research N, Linehan WM, Spellman PT, et al. Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma. N Engl J Med. 2016;374(2):135-45.

108.Ooi A, Dykema K, Ansari A, et al. CUL3 and NRF2 mutations confer an NRF2 activation phenotype in a sporadic form of papillary renal cell carcinoma. Cancer Res. 2013;73(7):2044-51.

109.Ooi A, Wong JC, Petillo D, et al. An antioxidant response phenotype shared between hereditary and sporadic type 2 papillary renal cell carcinoma. Cancer Cell. 2011;20(4):511-23.

110.Murugan P, Jia L, Dinatale RG, et al. Papillary renal cell carcinoma: a single institutional study of 199 cases addressing classification, clinicopathologic and molecular features, and treatment outcome. Mod Pathol. 2022;35(6):825-35.

111.Lobo J, Ohashi R, Helmchen BM, et al. The Morphological Spectrum of Papillary Renal Cell Carcinoma and Prevalence of Provisional/Emerging Renal Tumor Entities with Papillary Growth. Biomedicines. 2021;9(10).

112.Davis CF, Ricketts CJ, Wang M, et al. The somatic genomic landscape of chromophobe renal cell carcinoma. Cancer Cell. 2014;26(3):319-30.

113.Speicher MR, Schoell B, du Manoir S, et al. Specific loss of chromosomes 1, 2, 6, 10, 13, 17, and 21 in chromophobe renal cell carcinomas revealed by comparative genomic hybridization. Am J Pathol. 1994;145(2):356-64.

114.Durinck S, Stawiski EW, Pavia-Jimenez A, et al. Spectrum of diverse genomic alterations define non-clear cell renal carcinoma subtypes. Nat Genet. 2015;47(1):13-21.

115.Casuscelli J, Weinhold N, Gundem G, et al. Genomic landscape and evolution of metastatic chromophobe renal cell carcinoma. JCI Insight. 2017;2(12).

116.Vasudev NS, Wilson M, Stewart GD, et al. Challenges of early renal cancer detection: symptom patterns and incidental diagnosis rate in a multicentre prospective UK cohort of patients presenting with suspected renal cancer. BMJ Open. 2020;10(5):e035938.

117.Sun R, Breau RH, Mallick R, et al. Prognostic impact of paraneoplastic syndromes on patients with non-metastatic renal cell carcinoma undergoing surgery: Results from Canadian Kidney Cancer information system. Can Urol Assoc J. 2021;15(4):132-7.

118.Thorstenson A, Bergman M, Scherman-Plogell AH, et al. Tumour characteristics and surgical treatment of renal cell carcinoma in Sweden 2005-2010: a population-based study from the national Swedish kidney cancer register. Scand J Urol. 2014;48(3):231-8.

119.The National Cancer Registration and Analysis Service [Available from: http://ncin.org.uk/publications/survival_by_stage.

120.Sun M, Shariat SF, Cheng C, et al. Prognostic factors and predictive models in renal cell carcinoma: a contemporary review. Eur Urol. 2011;60(4):644-61.

121.Keegan KA, Schupp CW, Chamie K, et al. Histopathology of surgically treated renal cell carcinoma: survival differences by subtype and stage. J Urol. 2012;188(2):391-7.

122.Parosanu A, Stanciu IM, Pirlog C, et al. Prognostic Models for Renal Cell Carcinoma in the Era of Immune Checkpoint Therapy. Cureus. 2022;14(10):e30821.

123.Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin. 2017;67(2):93-9.

124.Ljungberg B, Albiges L, Abu-Ghanem Y, et al. European Association of Urology Guidelines on Renal Cell Carcinoma: The 2019 Update. Eur Urol. 2019;75(5):799-810.

125.Delahunt B, Eble JN, Egevad L, et al. Grading of renal cell carcinoma. Histopathology. 2019;74(1):4-17.

126.Heng DY, Xie W, Regan MM, et al. Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multicenter study. J Clin Oncol. 2009;27(34):5794-9.

127.Motzer RJ, Bacik J, Murphy BA, et al. Interferon-alfa as a comparative treatment for clinical trials of new therapies against advanced renal cell carcinoma. J Clin Oncol. 2002;20(1):289-96.

128.Inamura K. Renal Cell Tumors: Understanding Their Molecular Pathological Epidemiology and the 2016 WHO Classification. Int J Mol Sci. 2017;18(10).

129.Joseph RW, Kapur P, Serie DJ, et al. Clear Cell Renal Cell Carcinoma Subtypes Identified by BAP1 and PBRM1 Expression. J Urol. 2016;195(1):180-7.

130.Rini BI, Escudier B, Martini JF, et al. Validation of the 16-Gene Recurrence Score in Patients with Locoregional, High-Risk Renal Cell Carcinoma from a Phase III Trial of Adjuvant Sunitinib. Clin Cancer Res. 2018;24(18):4407-15.

131.Lopez-Beltran A, Henriques V, Cimadamore A, et al. The Identification of Immunological Biomarkers in Kidney Cancers. Front Oncol. 2018;8:456.

132.Vogel C, Ziegelmuller B, Ljungberg B, et al. Imaging in Suspected Renal-Cell Carcinoma: Systematic Review. Clin Genitourin Cancer. 2019;17(2):e345-e55.

133.Rossi SH, Prezzi D, Kelly-Morland C, et al. Imaging for the diagnosis and response assessment of renal tumours. World J Urol. 2018;36(12):1927-42.

134.Ma H, Shen G, Liu B, et al. Diagnostic performance of 18F-FDG PET or PET/CT in restaging renal cell carcinoma: a systematic review and meta-analysis. Nucl Med Commun. 2017;38(2):156-63.

135.Jena R, Narain TA, Singh UP, et al. Role of positron emission tomography/computed tomography in the evaluation of renal cell carcinoma. Indian J Urol. 2021;37(2):125-32.

136.Silverman SG, Pedrosa I, Ellis JH, et al. Bosniak Classification of Cystic Renal Masses, Version 2019: An Update Proposal and Needs Assessment. Radiology. 2019;292(2):475-88.

137.Marconi L, Dabestani S, Lam TB, et al. Systematic Review and Meta-analysis of Diagnostic Accuracy of Percutaneous Renal Tumour Biopsy. Eur Urol. 2016;69(4):660-73.

138.Quencer KB. Renal Mass Biopsy. Tech Vasc Interv Radiol. 2021;24(4):100774.

139.Ljungberg B, Albiges L, Abu-Ghanem Y, et al. European Association of Urology Guidelines on Renal Cell Carcinoma: The 2022 Update. Eur Urol. 2022;82(4):399-410.

140.MacLennan S, Imamura M, Lapitan MC, et al. Systematic review of oncological outcomes following surgical management of localised renal cancer. Eur Urol. 2012;61(5):972-93.

141.Culp SH, Tannir NM, Abel EJ, et al. Can we better select patients with metastatic renal cell carcinoma for cytoreductive nephrectomy? Cancer. 2010;116(14):3378-88.

142.Liu R, Qiu K, Wu J, et al. Cost-effectiveness analysis of nivolumab plus cabozantinib versus sunitinib as first-line therapy in advanced renal cell carcinoma. Immunotherapy. 2022;14(11):859-69.

143.Lee CH, DiNatale RG, Chowell D, et al. High Response Rate and Durability Driven by HLA Genetic Diversity in Patients with Kidney Cancer Treated with Lenvatinib and Pembrolizumab. Mol Cancer Res. 2021;19(9):1510-21.

144.DeVita VT Jr. LT, Rosenberg SA. DeVita, Hellman, and Rosenberg's Cancer: Principles and Practice of Oncology,. 8 ed. Philadelphia: Lippincott Williams & Wilkins; 2008.

145.Motzer RJ, Jonasch E, Agarwal N, et al. Kidney Cancer, Version 3.2022, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2022;20(1):71-90.

146.Negrier S, Escudier B, Lasset C, et al. Recombinant human interleukin-2, recombinant human interferon alfa-2a, or both in metastatic renal-cell carcinoma. Groupe Francais d'Immunotherapie. N Engl J Med. 1998;338(18):1272-8.

147.Fairlamb DJ. Spontaneous regression of metastases of renal cancer: A report of two cases including the first recorded regression following irradiation of a dominant metastasis and review of the world literature. Cancer. 1981;47(8):2102-6.

148.Vogelzang NJ, Priest ER, Borden L. Spontaneous regression of histologically proved pulmonary metastases from renal cell carcinoma: a case with 5-year followup. J Urol. 1992;148(4):1247-8.

149.Larroquette M, Peyraud F, Domblides C, et al. Adjuvant therapy in renal cell carcinoma: Current knowledges and future perspectives. Cancer Treat Rev. 2021;97:102207.

150.Choueiri TK, Motzer RJ. Systemic Therapy for Metastatic Renal-Cell Carcinoma. N Engl J Med. 2017;376(4):354-66.

151.Topalian SL, Drake CG, Pardoll DM. Targeting the PD-1/B7-H1(PD-L1) pathway to activate anti-tumor immunity. Curr Opin Immunol. 2012;24(2):207-12.

152.Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010;363(8):711-23.

153.Phan GQ, Yang JC, Sherry RM, et al. Cancer regression and autoimmunity induced by cytotoxic T lymphocyte-associated antigen 4 blockade in patients with metastatic melanoma. Proc Natl Acad Sci U S A. 2003;100(14):8372-7.

154.Robert C, Thomas L, Bondarenko I, et al. Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. N Engl J Med. 2011;364(26):2517-26.

155.Keir ME, Butte MJ, Freeman GJ, et al. PD-1 and its ligands in tolerance and immunity. Annu Rev Immunol. 2008;26:677-704.

156.Sharpe AH, Freeman GJ. The B7-CD28 superfamily. Nat Rev Immunol. 2002;2(2):116-26.

157.Tykodi SS. PD-1 as an emerging therapeutic target in renal cell carcinoma: current evidence. Onco Targets Ther. 2014;7:1349-59.

158.Thompson RH, Kuntz SM, Leibovich BC, et al. Tumor B7-H1 is associated with poor prognosis in renal cell carcinoma patients with long-term follow-up. Cancer Res. 2006;66(7):3381-5.

159.Brahmer JR, Tykodi SS, Chow LQ, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. N Engl J Med. 2012;366(26):2455-65.

160.Rini BI, Plimack ER, Stus V, et al. Pembrolizumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. N Engl J Med. 2019;380(12):1116-27.

161.Powles T, Plimack ER, Soulieres D, et al. Pembrolizumab plus axitinib versus sunitinib monotherapy as first-line treatment of advanced renal cell carcinoma (KEYNOTE-426): extended follow-up from a randomised, open-label, phase 3 trial. Lancet Oncol. 2020;21(12):1563-73.

162.Bedke J, Albiges L, Capitanio U, et al. The 2021 Updated European Association of Urology Guidelines on Renal Cell Carcinoma: Immune Checkpoint Inhibitor-based Combination Therapies for Treatment-naive Metastatic Clear-cell Renal Cell Carcinoma Are Standard of Care. Eur Urol. 2021;80(4):393-7.

163.Rogers JL, Bayeh L, Scheuermann TH, et al. Development of inhibitors of the PAS-B domain of the HIF-2alpha transcription factor. J Med Chem. 2013;56(4):1739-47.

164.Scheuermann TH, Li Q, Ma HW, et al. Allosteric inhibition of hypoxia inducible factor-2 with small molecules. Nat Chem Biol. 2013;9(4):271-6.

165.Wallace EM, Rizzi JP, Han G, et al. A Small-Molecule Antagonist of HIF2alpha Is Efficacious in Preclinical Models of Renal Cell Carcinoma. Cancer Res. 2016;76(18):5491-500.

166.Courtney KD, Infante JR, Lam ET, et al. Phase I Dose-Escalation Trial of PT2385, a First-in-Class Hypoxia-Inducible Factor-2alpha Antagonist in Patients With Previously Treated Advanced Clear Cell Renal Cell Carcinoma. J Clin Oncol. 2018;36(9):867-74.

167.Jonasch E, Donskov F, Iliopoulos O, et al. Belzutifan for Renal Cell Carcinoma in von Hippel-Lindau Disease. N Engl J Med. 2021;385(22):2036-46.

168.FDA Approves Belzutifan for Cancers Associated with Von Hippel-Lindau Disease [press release]. 2021.

169.Jonasch E, Bauer TM, Papadopoulos KP, et al. Phase 1 LITESPARK-001 (MK-6482-001) study of belzutifan in advanced solid tumors: Update of the clear cell renal cell carcinoma (ccRCC) cohort with more than 3 years of total follow-up. Journal of Clinical Oncology. 2022;40(16_suppl):4509-.

170.Choueiri TK, Bauer TM, McDermott DF, et al. Phase 2 study of the oral hypoxia-inducible factor 2α (HIF-2α) inhibitor MK-6482 in combination with cabozantinib in patients with advanced clear cell renal cell carcinoma (ccRCC). Journal of Clinical Oncology. 2021;39(6_suppl):272-.

171.Choueiri TK, McDermott DF, Merchan J, et al. Belzutifan plus cabozantinib for patients with advanced clear cell renal cell carcinoma previously treated with immunotherapy: an open-label, single-arm, phase 2 study. Lancet Oncol. 2023.

172.Pantel K, Alix-Panabieres C. Circulating tumour cells in cancer patients: challenges and perspectives. Trends Mol Med. 2010;16(9):398-406.

173.Eknoyan G, Nagy J. A history of diabetes mellitus or how a disease of the kidneys evolved into a kidney disease. Adv Chronic Kidney Dis. 2005;12(2):223-9.

174.Poulet G, Massias J, Taly V. Liquid Biopsy: General Concepts. Acta Cytol. 2019;63(6):449-55.

175.Parkinson DR, Dracopoli N, Petty BG, et al. Considerations in the development of circulating tumor cell technology for clinical use. J Transl Med. 2012;10:138.

176.Lone SN, Nisar S, Masoodi T, et al. Liquid biopsy: a step closer to transform diagnosis, prognosis and future of cancer treatments. Mol Cancer. 2022;21(1):79.

177.Ko J, Baldassano SN, Loh PL, et al. Machine learning to detect signatures of disease in liquid biopsies - a user's guide. Lab Chip. 2018;18(3):395-405.

178.Eslami SZ, Cortes-Hernandez LE, Alix-Panabieres C. The Metastatic Cascade as the Basis for Liquid Biopsy Development. Front Oncol. 2020;10:1055.

179.Ganesh K, Massague J. Targeting metastatic cancer. Nat Med. 2021;27(1):34-44.

180.Luo YT, Cheng J, Feng X, et al. The viable circulating tumor cells with cancer stem cells feature, where is the way out? J Exp Clin Cancer Res. 2018;37(1):38.

181.Anvari S, Osei E, Maftoon N. Interactions of platelets with circulating tumor cells contribute to cancer metastasis. Sci Rep. 2021;11(1):15477.

182.Riethdorf S, O'Flaherty L, Hille C, et al. Clinical applications of the CellSearch platform in cancer patients. Adv Drug Deliv Rev. 2018;125:102-21.

183.Cristofanilli M, Budd GT, Ellis MJ, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med. 2004;351(8):781-91.

184.Bidard FC, Peeters DJ, Fehm T, et al. Clinical validity of circulating tumour cells in patients with metastatic breast cancer: a pooled analysis of individual patient data. Lancet Oncol. 2014;15(4):406-14.

185.Resel Folkersma L, Olivier Gomez C, San Jose Manso L, et al. Immunomagnetic quantification of circulating tumoral cells in patients with prostate cancer: clinical and pathological correlation. Arch Esp Urol. 2010;63(1):23-31.

186.Negin BP, Cohen SJ. Circulating tumor cells in colorectal cancer: past, present, and future challenges. Curr Treat Options Oncol. 2010;11(1-2):1-13.

187.Gradilone A, Iacovelli R, Cortesi E, et al. Circulating tumor cells and "suspicious objects" evaluated through CellSearch(R) in metastatic renal cell carcinoma. Anticancer Res. 2011;31(12):4219-21.

188.Allard WJ, Matera J, Miller MC, et al. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. Clin Cancer Res. 2004;10(20):6897-904.

189.Bai M, Zou B, Wang Z, et al. Comparison of two detection systems for circulating tumor cells among patients with renal cell carcinoma. Int Urol Nephrol. 2018;50(10):1801-9.

190.de Bono JS, Scher HI, Montgomery RB, et al. Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. Clin Cancer Res. 2008;14(19):6302-9.

191.Seligson DB, Pantuck AJ, Liu X, et al. Epithelial cell adhesion molecule (KSA) expression: pathobiology and its role as an independent predictor of survival in renal cell carcinoma. Clin Cancer Res. 2004;10(8):2659-69.

192.Eichelberg C, Chun FK, Bedke J, et al. Epithelial cell adhesion molecule is an independent prognostic marker in clear cell renal carcinoma. Int J Cancer. 2013;132(12):2948-55.

193.Werner S, Stenzl A, Pantel K, et al. Expression of Epithelial Mesenchymal Transition and Cancer Stem Cell Markers in Circulating Tumor Cells. Adv Exp Med Biol. 2017;994:205-28.

194.Nel I, Gauler TC, Bublitz K, et al. Circulating Tumor Cell Composition in Renal Cell Carcinoma. PLoS One. 2016;11(4):e0153018.

195.Basso U, Facchinetti A, Rossi E, et al. Prognostic Role of Circulating Tumor Cells in Metastatic Renal Cell Carcinoma: A Large, Multicenter, Prospective Trial. Oncologist. 2021;26(9):740-50.

196.Bade RM, Schehr JL, Emamekhoo H, et al. Development and initial clinical testing of a multiplexed circulating tumor cell assay in patients with clear cell renal cell carcinoma. Mol Oncol. 2021;15(9):2330-44.

197.Dorai T, Sawczuk I, Pastorek J, et al. Role of carbonic anhydrases in the progression of renal cell carcinoma subtypes: proposal of a unified hypothesis. Cancer Invest. 2006;24(8):754-79.

198.Li M, Li L, Zheng J, et al. Liquid biopsy at the frontier in renal cell carcinoma: recent analysis of techniques and clinical application. Mol Cancer. 2023;22(1):37.

199.Broncy L, Paterlini-Brechot P. Circulating Tumor Cells for the Management of Renal Cell Carcinoma. Diagnostics (Basel). 2018;8(3).

200.Paterlini-Brechot P. Circulating Tumor Cells: Who is the Killer? Cancer Microenviron. 2014;7(3):161-76.

201.Harouaka RA, Nisic M, Zheng SY. Circulating tumor cell enrichment based on physical properties. J Lab Autom. 2013;18(6):455-68.

202.Koch C, Joosse SA, Schneegans S, et al. Pre-Analytical and Analytical Variables of Label-Independent Enrichment and Automated Detection of Circulating Tumor Cells in Cancer Patients. Cancers (Basel). 2020;12(2).

203.Gascoyne PR, Shim S, Noshari J, et al. Correlations between the dielectric properties and exterior morphology of cells revealed by dielectrophoretic field-flow fractionation. Electrophoresis. 2013;34(7):1042-50.

204.Klezl P, Pospisilova E, Kolostova K, et al. Detection of Circulating Tumor Cells in Renal Cell Carcinoma: Disease Stage Correlation and Molecular Characterization. J Clin Med. 2020;9(5).

205.De Alwis R, Hansson J, Lindgren D, et al. Size-based isolation and detection of renal carcinoma cells from whole blood. Mol Clin Oncol. 2022;16(5):101.

206.Guan Y, Xu F, Tian J, et al. The prognostic value of circulating tumour cells (CTCs) and CTC white blood cell clusters in patients with renal cell carcinoma. BMC Cancer. 2021;21(1):826.

207.Santoni M, Cimadamore A, Cheng L, et al. Circulating Tumor Cells in Renal Cell Carcinoma: Recent Findings and Future Challenges. Front Oncol. 2019;9:228.

208.Rushton AJ, Nteliopoulos G, Shaw JA, et al. A Review of Circulating Tumour Cell Enrichment Technologies. Cancers (Basel). 2021;13(5).

209.Habli Z, AlChamaa W, Saab R, et al. Circulating Tumor Cell Detection Technologies and Clinical Utility: Challenges and Opportunities. Cancers (Basel). 2020;12(7).

210.Coumans FA, van Dalum G, Beck M, et al. Filtration parameters influencing circulating tumor cell enrichment from whole blood. PLoS One. 2013;8(4):e61774.

211.Guan Y, Xu F, Tian J, et al. Prognostic value of circulating tumor cells and immune-inflammatory cells in patients with renal cell carcinoma. Urol Oncol. 2022;40(4):167 e21- e32.

212.Song J, Yu Z, Dong B, et al. Clinical significance of circulating tumour cells and Ki-67 in renal cell carcinoma. World J Surg Oncol. 2021;19(1):156.

213.Haga N, Onagi A, Koguchi T, et al. Perioperative Detection of Circulating Tumor Cells in Radical or Partial Nephrectomy for Renal Cell Carcinoma. Ann Surg Oncol. 2020;27(4):1272-81.

214.Wang Z, Zhang P, Chong Y, et al. Perioperative Circulating Tumor Cells (CTCs), MCTCs, and CTC-White Blood Cells Detected by a Size-Based Platform Predict Prognosis in Renal Cell Carcinoma. Dis Markers. 2021;2021:9956142.

215.Nayak B, Panaiyadiyan S, Singh P, et al. Role of circulating tumor cells in patients with metastatic clear-cell renal cell carcinoma. Urol Oncol. 2021;39(2):135 e9- e15.

216.Tan Z, Yue C, Ji S, et al. Assessment of PD-L1 Expression on Circulating Tumor Cells for Predicting Clinical Outcomes in Patients with Cancer Receiving PD-1/PD-L1 Blockade Therapies. Oncologist. 2021;26(12):e2227-e38.

217.Chen BH, Kao CC, Xu T, et al. Determining programmed cell death ligand 1 expression in circulating tumor cells of patients with clear cell renal cell carcinoma and its correlation with response to programmed cell death protein 1 inhibitors. Int J Urol. 2022;29(9):947-54.

218.Underhill HR, Kitzman JO, Hellwig S, et al. Fragment Length of Circulating Tumor DNA. PLoS Genet. 2016;12(7):e1006162.

219.Mouliere F, Chandrananda D, Piskorz AM, et al. Enhanced detection of circulating tumor DNA by fragment size analysis. Sci Transl Med. 2018;10(466).

220.Markus H, Chandrananda D, Moore E, et al. Refined characterization of circulating tumor DNA through biological feature integration. Sci Rep. 2022;12(1):1928.

221.Tivey A, Church M, Rothwell D, et al. Circulating tumour DNA - looking beyond the blood. Nat Rev Clin Oncol. 2022;19(9):600-12.

222.Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med. 2014;6(224):224ra24.

223.Shin HT, Choi YL, Yun JW, et al. Prevalence and detection of low-allele-fraction variants in clinical cancer samples. Nat Commun. 2017;8(1):1377.

224.Spencer DH, Tyagi M, Vallania F, et al. Performance of common analysis methods for detecting low-frequency single nucleotide variants in targeted next-generation sequence data. J Mol Diagn. 2014;16(1):75-88.

225.Corro C, Hejhal T, Poyet C, et al. Detecting circulating tumor DNA in renal cancer: An open challenge. Exp Mol Pathol. 2017;102(2):255-61.

226.Smith CG, Moser T, Mouliere F, et al. Comprehensive characterization of cell-free tumor DNA in plasma and urine of patients with renal tumors. Genome Med. 2020;12(1):23.

227.Wan JCM, Heider K, Gale D, et al. ctDNA monitoring using patient-specific sequencing and integration of variant reads. Sci Transl Med. 2020;12(548).

228.Lasseter K, Nassar AH, Hamieh L, et al. Plasma cell-free DNA variant analysis compared with methylated DNA analysis in renal cell carcinoma. Genet Med. 2020;22(8):1366-73.

229.Geertsen L, Koldby KM, Thomassen M, et al. Circulating Tumor DNA in Patients with Renal Cell Carcinoma. A Systematic Review of the Literature. Eur Urol Open Sci. 2022;37:27-35.

230.Hahn AW, Gill DM, Maughan B, et al. Correlation of genomic alterations assessed by next-generation sequencing (NGS) of tumor tissue DNA and circulating tumor DNA (ctDNA) in metastatic renal cell carcinoma (mRCC): potential clinical implications. Oncotarget. 2017;8(20):33614-20.

231.Jung M, Ellinger J, Gevensleben H, et al. Cell-Free SHOX2 DNA Methylation in Blood as a Molecular Staging Parameter for Risk Stratification in Renal Cell Carcinoma Patients: A Prospective Observational Cohort Study. Clin Chem. 2019;65(4):559-68.

232.Nuzzo PV, Berchuck JE, Korthauer K, et al. Detection of renal cell carcinoma using plasma and urine cell-free DNA methylomes. Nat Med. 2020;26(7):1041-3.

233.Bacon JVW, Annala M, Soleimani M, et al. Plasma Circulating Tumor DNA and Clonal Hematopoiesis in Metastatic Renal Cell Carcinoma. Clin Genitourin Cancer. 2020;18(4):322-31 e2.

234.Yamamoto Y, Uemura M, Fujita M, et al. Clinical significance of the mutational landscape and fragmentation of circulating tumor DNA in renal cell carcinoma. Cancer Sci. 2019;110(2):617-28.

235.Lin YL, Wang YP, Li HZ, et al. Aberrant Promoter Methylation of PCDH17 (Protocadherin 17) in Serum and its Clinical Significance in Renal Cell Carcinoma. Med Sci Monit. 2017;23:3318-23.

236.Simon BR, Wilson MJ, Wickliffe JK. The RPTEC/TERT1 cell line models key renal cell responses to the environmental toxicants, benzo[a]pyrene and cadmium. Toxicol Rep. 2014;1:231-42.

237.O'Connor TP, Cockburn K, Wang W, et al. Semaphorin 5B mediates synapse elimination in hippocampal neurons. Neural Dev. 2009;4:18.

238.Browne K, Wang W, Liu RQ, et al. Transmembrane semaphorin5B is proteolytically processed into a repulsive neural guidance cue. J Neurochem. 2012;123(1):135-46.

239.Grundmann S, Lindmayer C, Hans FP, et al. FoxP1 stimulates angiogenesis by repressing the inhibitory guidance protein semaphorin 5B in endothelial cells. PLoS One. 2013;8(9):e70873.

240.Song HR, Gonzalez-Gomez I, Suh GS, et al. Nuclear factor IA is expressed in astrocytomas and is associated with improved survival. Neuro Oncol. 2010;12(2):122-32.

241.Perna D, Karreth FA, Rust AG, et al. BRAF inhibitor resistance mediated by the AKT pathway in an oncogenic BRAF mouse melanoma model. Proc Natl Acad Sci U S A. 2015;112(6):E536-45.

242.Micci F, Thorsen J, Panagopoulos I, et al. High-throughput sequencing identifies an NFIA/CBFA2T3 fusion gene in acute erythroid leukemia with t(1;16)(p31;q24). Leukemia. 2013;27(4):980-2.

243.Mao X, Chaplin T, Young BD. Integrated genomic analysis of sezary syndrome. Genet Res Int. 2011;2011:980150.

244.Chen KS, Lim JWC, Richards LJ, et al. The convergent roles of the nuclear factor I transcription factors in development and cancer. Cancer Lett. 2017;410:124-38.

245.Kreike B, van Kouwenhove M, Horlings H, et al. Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas. Breast Cancer Res. 2007;9(5):R65.

246.Moon HG, Hwang KT, Kim JA, et al. NFIB is a potential target for estrogen receptor-negative breast cancers. Mol Oncol. 2011;5(6):538-44.

247.Lu W, Quintero-Rivera F, Fan Y, et al. NFIA haploinsufficiency is associated with a CNS malformation syndrome and urinary tract defects. PLoS Genet. 2007;3(5):e80.

248.Gronostajski RM. Roles of the NFI/CTF gene family in transcription and development. Gene. 2000;249(1-2):31-45.

249.Huang J, Kong Y, Xie C, et al. Stem/progenitor cell in kidney: characteristics, homing, coordination, and maintenance. Stem Cell Res Ther. 2021;12(1):197.

250.Nagy A, Walter E, Zubakov D, et al. High risk of development of renal cell tumor in end-stage kidney disease: the role of microenvironment. Tumour Biol. 2016;37(7):9511-9.

251.Lee Y, Guan G, Bhagat AA. ClearCell(R) FX, a label-free microfluidics technology for enrichment of viable circulating tumor cells. Cytometry A. 2018;93(12):1251-4.

252.Pereira-Veiga T, Schneegans S, Pantel K, et al. Circulating tumor cell-blood cell crosstalk: Biology and clinical relevance. Cell Rep. 2022;40(9):111298.

253.Pantel K, Alix-Panabieres C. Crucial roles of circulating tumor cells in the metastatic cascade and tumor immune escape: biology and clinical translation. J Immunother Cancer. 2022;10(12).

Paper II

The Journal of Pathology: Clinical Research
J Pathol Clin Res 2023
Published online 22 March 2023 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cjp2.316

ORIGINAL ARTICLE

Identification and validation of NFIA as a novel prognostic marker in renal cell carcinoma

Roger de Alwis[1†], Sarah Schoch[1†], Mazharul Islam[1], Christina Möller[1], Börje Ljungberg[2] and Håkan Axelson[1*]

[1]*Division of Translational Cancer Research, Department of Laboratory Medicine Lund University, Lund, Sweden*
[2]*Department of Surgical and Perioperative Sciences, Urology and Andrology, Umeå University, Umeå, Sweden*

**Correspondence to: Håkan Axelson, Division of Translational Cancer Research, Department of Laboratory Medicine, Medicon Village, Building 404 A3, Scheelevägen 8, 404A3, 223 63 Lund, Sweden. E-mail: hakan.axelson@med.lu.se*

[†]*These authors contributed equally to this work.*

Abstract

Prognostic tools are an essential component of the clinical management of patients with renal cell carcinoma (RCC). Although tumour stage and grade can provide important information, they fail to consider patient- and tumour-specific biology. In this study, we set out to find a novel molecular marker of RCC by using hepatocyte nuclear factor 4A (HNF4A), a transcription factor implicated in RCC progression and malignancy, as a blueprint. Through transcriptomic analyses, we show that the nuclear factor I A (NFIA)-driven transcription network is active in primary RCC and that higher levels of NFIA confer a survival benefit. We validate our findings using immunohistochemical staining and analysis of a 363-patient tissue microarray (TMA), showing for the first time that NFIA can independently predict poor cancer-specific survival in clear cell RCC (ccRCC) patients (hazard ratio = 0.46, 95% CI = 0.24–0.85, *p* value = 0.014). Furthermore, we confirm the association of HNF4A with higher grades and stages in ccRCC in our TMA cohort. We present novel data that show HNF4A protein expression does not confer favourable prognosis in papillary RCC, confirming our survival analysis with publicly available *HNF4A* RNA expression data. Further work is required to elucidate the functional role of NFIA in RCC as well as the testing of these markers on patient material from diverse multi-centre cohorts, to establish their value for the prognostication of RCC.

Keywords: renal cell carcinoma; biomarker; immunohistochemistry; cancer-specific survival

Received 3 October 2022; Revised 10 January 2023; Accepted 24 February 2023

No conflicts of interest were declared.

Introduction

Renal cell carcinoma (RCC) comprises a range of tumour types arising from the epithelial cells of the nephron, the functional unit of the kidney [1]. The nephron is a tubular structure covered with a single epithelial layer, with highly specialised cell types executing discrete and highly specific uptake and excretion of substances in a section-specific and strictly regulated fashion [2,3]. According to the WHO classification, there are more than 50 established and provisional distinct subtypes of RCC, each with a defined morphology and genetic characteristic [4]. These characteristics can be partly traced back to the cell type in the nephron from which the tumour originated. With the advent of single-cell sequencing, the transcriptional wiring in different sections of the nephron is currently being delineated, providing a detailed map of the different cell types along the nephron [5].

The three major RCC types are clear cell RCC (ccRCC), papillary RCC (pRCC), and chromophobe RCC (chRCC), together representing 95% of all RCCs. ccRCC and pRCC represent 75 and 15% of all RCCs, respectively, and it is believed that they stem from the proximal cells of the nephron [6]. Indeed, by comparison with normal proximal tubule cells, ccRCC and pRCC maintain a range of proximal characteristics, including key transcription factors (TFs). In contrast, chRCC shows strong transcriptional similarities to intercalated distal nephron cells [7,8]. All major

RCC subtypes have been extensively characterised with respect to genetic aberrations, and it is likely that the genetic landscape of each subtype is selected for based on the constitution of the transcriptional wiring of the cell of origin. For example, ccRCC is initiated by a functional loss of the von Hippel–Lindau (VHL) tumour suppressor gene leading to a constitutive activation of the hypoxia-inducible factors (HIF1A and HIF2A) and their respective transcriptional programme [9,10]. Recent data indeed indicate that the oncogenic effect of HIF activation in proximal tubule cells is dependent on transcriptional cooperation with tissue-specific transcriptional networks, including PAX8 and HNF1B [11].

Another transcriptional regulator implicated in RCC biology is the hepatocyte nuclear factor 4A (HNF4A), a TF belonging to the nuclear receptor family, primarily expressed in liver, gut, kidney, and pancreatic beta-cells, where it acts as a master transcriptional regulator [12]. In the kidney, HNF4A is specifically expressed in the proximal tubule cells, regulating key target genes vital to the designated function of this part of the nephron [13–15]. We have shown that the HNF4A transcriptional activity is maintained and characterises ccRCC and pRCC, which are derived thereof. Furthermore, ccRCC tumours with increased malignancy display lower levels of HNF4A expression and related target genes, indicating that de-differentiation and loss of lineage fidelity play a major role in RCC tumorigenesis [8]. *HNF4A* expression has been extensively studied at a transcriptional level, but validation at the protein level in RCC patients is limited to one study with only 30 ccRCC patients [16]. In our current study, we consolidate these findings utilising a larger patient cohort whilst including pRCC, the second most common RCC subtype.

Based on the concept that loss of lineage fidelity is associated with RCC progression, we investigated other TFs that mimic *HNF4A* expression in non-tumour kidney tissue (herein referred to as normal kidney) and RCC. One standout gene was nuclear factor 1A (*NFIA*), expressed in the proximal tubules of the nephron. NFIA transcriptionally defines and regulates the development of multiple organs [17] including liver [18], brain [19], kidney [20], and lung [21]. *NFIA* has been studied in brain tumours where it acts as a tumour suppressor gene [22,23] and in oesophageal cancer; *NFIA* expression provides prognostic information [24]. To the best of our knowledge, neither the protein expression nor the role of NFIA in RCC has been previously studied.

In the current study, we confirm the grade and subtype-specific expression of HNF4A in RCCs and explore its utility as a prognostic marker. Second, we show that NFIA, a proximal tubule-associated TF, is a novel prognostic marker for ccRCC and pRCC, with similar association to tumour aggressivity as HNF4A.

Materials and methods

Patient characteristics and tissue collection

The study comprised 376 patients with histopathologically verified RCC. Subtype-specific patient characteristics are summarised in Table 1. All patients were diagnosed at the Department of Urology, Umeå University Hospital, Sweden between 1990 and 2010. At metastatic disease, most patients were treated with interferon or hormones, whilst a minority had palliative treatment only. All patients were subject to yearly follow-up, screened in the medical records and screened for being alive in the Swedish National Population Register. The last follow-up was done in December 2020. Cancer-specific survival (CSS) time was defined as the time from diagnosis to the date of death from RCC or alive at the end of December 2020. Histopathological classification of RCC type was performed according to the Heidelberg classification [25]. Nuclear grading was performed according to Fuhrman *et al* [26]. The updated TNM classification 2017 was used for tumour stage grouping [27]. In the stage grouping, patients with Nx were joined with N0, and patients with Mx joined with M0. The distribution of Fuhrman grades, tumour, node, and metastases status across grouped stages in ccRCC and pRCC cohorts is presented in supplementary material,

Table 1. TMA patient cohort characteristics

Characteristics	ccRCC ($n = 313$)	pRCC ($n = 50$)
Age range	30–90	20–90
Median age	67	67
Sex		
Female	138	16
Male	175	34
Grouped staging		
I	112	19
II	48	13
III	71	9
IV	81	8
Grade		
G1	30	7
G2	95	20
G3	130	16
G4	58	5
Median follow-up (years)	4.8	6.8
Dead by disease	153	22

Figures S1 and S2. Tumour size, defined as the largest tumour diameter, was measured primarily on the computed tomography (CT) or magnetic resonance imaging (MRI) scans. The median (range) tumour size was 7 (1–25) cm. The venous invasion was defined as tumour invasion in major renal veins, verified microscopically in the renal hilum. All patients were followed with clinical and radiological examinations.

Tissue microarray construction

For tissue microarray (TMA) construction, tumour sections were screened by a pathologist and representative tumour cores (0.6 mm in diameter) were arranged in recipient TMA paraffin TMA blocks, as described previously [28]. Most tumours were represented by 2–4 valid tissue cores. Eight ccRCC and two pRCC patients were excluded from analysis due to loss of material during processing. All samples from included patients were histopathologically re-evaluated and TMAs, containing duplicate 1.00 mm cores, were constructed from both non-tumour and tumour kidney tissue using a tissue array machine (Beecher Instruments, Microarray Technology, MD, USA).

Immunohistochemistry staining

The TMA blocks were sliced into 4 μm sections and treated according to standard procedures including de-paraffinisation and rehydration. Immunostaining was performed using Dako's Autostainerplus with the EnVisionFlex High pH-kit (Dako, CA, USA) and the following antibodies: HNF4A (HPA004712, Atlas Antibodies, Bromma, Sweden) and NFIA (HPA006111, Atlas Antibodies, Bromma, Sweden) both at 1:200 dilution.

Image acquisition and data analysis

Digital images of slides were extracted using a Zeiss Axion scanner (Jenna, Germany) at ×200 magnification. Image analysis was processed in QuPath v.0.3.2 (Queen's University, Belfast, Northern Ireland). TMA slides were de-arrayed and pre-processed as previously described [29]. Cells were detected using cell detection and cell classifiers were created and trained for each tumour type to separate tumour cells from stromal and immune cells. To evaluate the staining intensity, 3,3'-Diaminobenzidine (DAB) staining intensity classification was performed, and the Allred score was calculated by QuPath for each core. For further statistical analysis, the average of the Allred score of all cores from each patient was calculated.

Ethical approval

The patients had provided informed consent, orally before year 2000, and informed and written consent from year 2000. The study was reviewed and approved by the Ethical Review Board (Dnr: 2015-146-31M and Dnr: 2018-296-32M) and the Ethical Board of Sweden (Dnr: 2019-02579). All procedures regarding patients were in accordance with the Helsinki Declaration. The data used were anonymised and throughout the project, all data were treated under the regulations of the General Data Protection Regulation Act.

TCGA analysis

The Cancer Genome Atlas (TCGA) data were downloaded from the GDC data portal (https://portal.gdc.cancer.gov/) using the R package TCGAbiolinks [30]. Additional clinical data regarding grade of the KIRC cohort were downloaded separately from firebrowse.org. For the data analysis, transcripts per million unstranded data of tumour samples were extracted for each RCC type, followed by a log2 transformation after adding an offset of 1. TF network activity was analysed using the R package single-cell regulatory network inference and clustering (SCENIC) according to Aibar *et al* [31]. In the first step, potential TF targets were identified by GENIE3 followed by a TF-motif enrichment analysis and identification of direct targets (regulons) using RcisTarget. The last step consisted of scoring the activity of regulons performed by AUCell. Finally, the network activity was binarised.

Single-cell RNA-sequencing analysis

Single-cell RNA-seq data of ccRCC patients were obtained from Obradovic *et al* [32]. Raw data were downloaded from https://data.mendeley.com/datasets/nc9bc8dn4m/1. For the re-analysis, the data were subset into normal adjacent tissue and further split into a CD45 positive and a CD45 negative subset. For the analysis in this paper, only the subset of CD45 negative of adjacent normal tissue was used. Re-processing of the data was performed in R using Seurat (v.4.1.1) [33] and harmony (v.0.1.0) [34] was used for data integration. Initial Quality control comprised the removal of cells with fewer than 200 and more than 5,000 features detected as well as a mitochondrial content >25%. Furthermore, features needed to be expressed in a minimum of 50 cells to be included in the data.

Before data integration, the data were normalised using Seurat's function NormalizeData followed by a

feature selection to find the 2,000 most variable genes. In addition, the data were scaled using ScaleData and a principal component analysis was performed. Finally, RunHarmony was used for data integration to adjust for patient-to-patient variation. For visualisation, uniform manifold approximation and projection was performed on the first 20 dimensions of the harmony reductions. The Louvain algorithm implemented in Seurat was used for cluster analysis with a resolution parameter of 0.5. Seurat's function FindAllMarkers was applied for differential gene expression analysis to annotate clusters. After initial annotation, the data were again subset to include only cells of the nephron. This subset was re-processed once again using the same workflow as stated above.

Statistical analysis

Data were compiled, analysed, and visualised using R (v.4.2.1). Survival analysis was performed by using the log-rank test and the Kaplan–Meier method, where patients who died from non-RCC causes and/or had poor or missing cores were censored. As a threshold for the Kaplan–Meier analysis, the median was used for the TCGA data and an Allred score of 4 for the TMA analysis. Pearson correlation was used for the correlation analysis between the HNF4A and NFIA staining. Multivariable testing was performed on all clinical parameters that were found to be significant with univariate analysis using the Cox regression model. $P < 0.05$ was considered statistically significant.

Results

Analyses of TFs in RCC cohorts

Our previous work investigating transcriptional programs in the normal nephron and RCC revealed *HNF4A* as a tumour-defining transcriptional marker in ccRCC and pRCC [8]. Experimental studies have corroborated the role and significance of *HNF4A* in RCC tumorigenesis and lineage fidelity [13,14,35]. In our current study, we set out to identify other novel TFs, apart from HNF4A, that may be operating within RCC tumours to build upon the molecular definitions of the disease as well as understand clinical implications. We identified expression signatures of the histological RCC subgroups that can be intimately associated with specific TF networks. We used the SCENIC package that is designed to infer TFs and gene regulatory networks or 'regulons'. Using this approach, we could corroborate the signifying *HNF4A* activity in ccRCC (KIRC) and pRCC (KIRP) within the TCGA data sets (Figure 1A). In contrast, and as previously reported, the *FOXI1* signature is highly enriched in chRCC [8]. Thus, our SCENIC analysis defined key TFs seem to be transcriptionally active and defined the different molecular subtypes of RCC. Using our analysis, we identified *NFIA*, a TF not previously studied within the context of RCC. The expression of *NFIA* and its target genes are highly enriched in the ccRCC and pRCC cohorts, whilst the same is undetected or extremely low in chRCC which is known to arise from the distal segments of the nephron. The network enrichment signature of *NFIA* in ccRCC and pRCC subtype cohorts mimics the enrichment signature to that of *HNF4A*. These observations were further evidenced through our correlation analysis of *HNF4A* and *NFIA* RNA in the TCGA cohort, where a significantly positive correlation was seen in ccRCC ($R = 0.5$) and pRCC (0.45) tumours (Figure 1B) (validation in supplementary material, Figure S3). In line with this, our analysis of *HNF4A* and *NFIA* in normal kidney single-cell data (Figure 1C) show an enrichment of these TFs in the proximal tubular cells S2 and S3 (Figure 1D).

Pan-TCGA *HNF4A*/*NFIA* expression

We next sought to dissect the expression pattern of *NFIA* in TCGA-based RCC cohorts, relate this expression to *HNF4A* and finally validate the observations from our SCENIC analysis. To broadly understand the expression of *NFIA* in cancer and more specifically in RCC, we analysed its expression at the transcriptional level in the TCGA cohort which included 33 cancer types (Figure 2A). In line with previous reports on the role of NFIA in development and cancer, it is broadly expressed across many tumour types [22]. This includes ccRCC (KIRC) and pRCC (KIRP) whilst having relatively low expression in chRCC which is in line with our SCENIC analysis. Given that *NFIA* expression is primarily enriched within the proximal tubular segments of the nephron and to an extent is conserved in ccRCC and pRCC tumours, this data further validate our understanding of the proximal origin of these RCC subtypes [36,37]. Based on this reasoning, it is unsurprising that chRCCs relatively lack *NFIA* expression as they are thought to arise from the intercalated cells of the collecting duct. We also explored the expression of *HNF4a* within the same tumour panel and as expected, there was a greater disparity in expression level across tumours, reflecting the highly tissue-specific expression of this TF. As previously reported, expression of *HNF4A*

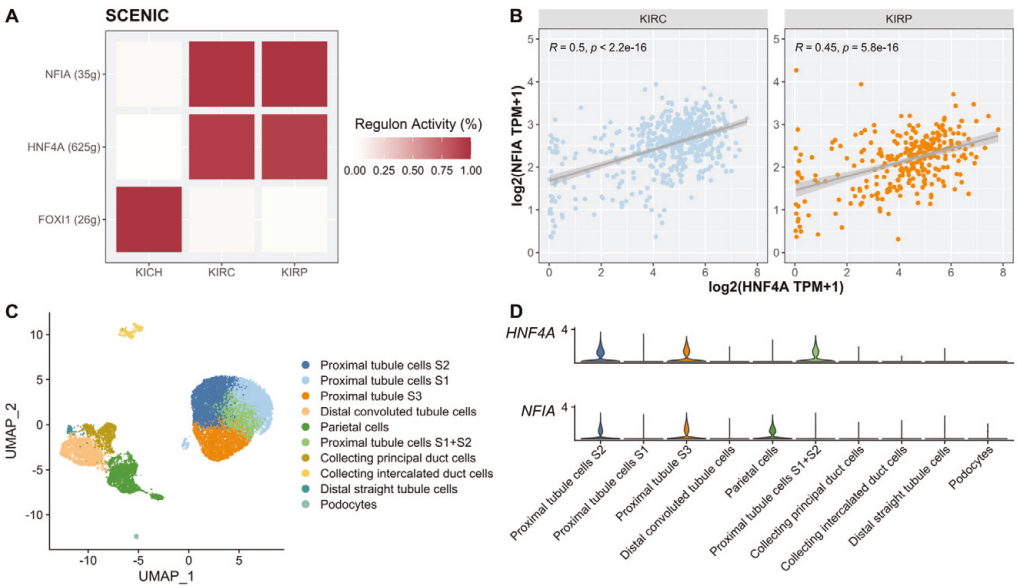

Figure 1. Analysis of TCGA RCC cohorts and kidney single-cell RNA-seq data. (A) Regulon activity of target genes by selected transcription factors within RCC subtypes in the TCGA cohort, with number of target genes indicated in brackets. (B) Correlation of *NFIA* and *HNF4* RNA expression in ccRCC ($n = 540$) and pRCC ($n = 290$) samples within the TCGA patient cohort. R = Pearson's correlation coefficient. Line is line of best fit; shaded area represents 95% CI. (C) Re-analysis dependent clustering of kidney single-cell RNA-seq data. (D) RNA expression of *NFIA* and *HNF4A* in normal kidney cell types based on re-analysis.

within the RCC cohorts is largely restricted to ccRCC and pRCC whilst remaining relatively low in chRCC reflecting the cell of origin characteristics of the respective tumour subtypes [8] (Figure 2B). As a result of the low levels of *HNF4A* and *NFIA* expression in the distal segments of the nephron and chRCC, we excluded this subtype from subsequent analyses.

HNF4A and *NFIA* expression correlates with favourable survival

To better understand the role HNF4A and NFIA play in RCC, we analysed TCGA-based RNA expression of *NFIA* and *HNF4A* in relation to grade (Figure 3A) and stage in ccRCC and only stage in pRCC (TCGA grade data is not available for pRCC) (Figure 3B,D). Higher grades displayed lower levels of *NFIA* or *HNF4A* expression in ccRCC ($p < 0.0001$ and $p < 0.01$, respectively). Furthermore, we dichotomised tumours based on expression level for *HNF4A* or *NFIA* and examined patient overall survival between these two assigned groups (Figure 3C,E). Within the TCGA ccRCC and pRCC tumours, there is a tendency for higher stage tumours to display relatively lower levels of *NFIA* expression in comparison to low-stage tumours. In line with these observations, when RCC tumours were grouped by subtype, Kaplan–Meier survival plots show that high *NFIA* expression confers better CSS probability in ccRCC ($p < 0.0001$) and pRCC ($p < 0.041$). High *HNF4A* expression correlated with favourable CSS probability in ccRCC ($p > 0.001$) but not in pRCC.

HNF4A and NFIA expression confers better CSS probability in the validation TMA cohort

To validate our TCGA data-derived findings, we opted to immunohistochemically stain a TMA of RCC tumours consisting of ccRCC (Figure 4A) and pRCC (Figure 4B) for HNF4A and NFIA proteins. We related the Allred score (a combination indicator of positively stained cell percentage and intensity) of these TMA cores to tumour grade, stage, and patient CSS (Figures 5 and 6). Lower levels of NFIA were observed with higher ccRCC tumour stages ($p < 0.05$) and grades ($p < 0.0001$) (Figure 5A). A similar association was observed for HNF4A in relation to tumour grade ($p < 0.001$) but not

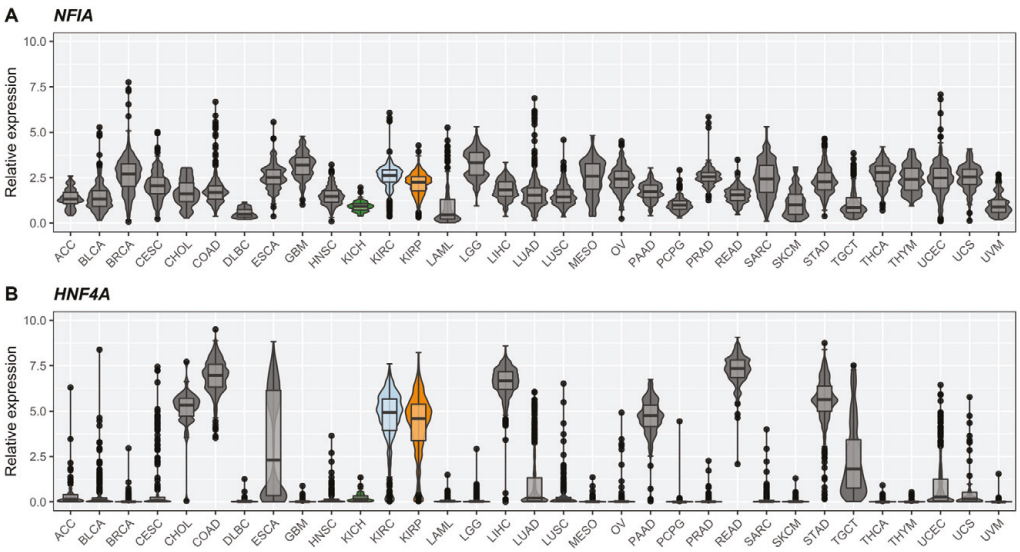

Figure 2. Pan-TCGA expression of *NFIA* and *HNF4A*. Relative RNA expression (transcripts per million) of (A) *NFIA* and (B) *HNF4A* in 33 cancer types from TCGA data including ccRCC (KIRC, $n = 540$), pRCC (KIRP, $n = 290$), and chRCC (KICH, $n = 65$) RCC subtypes.

stage (Figure 5C). These results corroborate the mRNA data, showing that the loss of NFIA and HNF4A protein is associated with increased tumour malignancy in ccRCC. In concordance, we were able to confirm that both high NFIA ($p < 0.0001$) and HNF4A ($p < 0.0001$) protein expression confers favourable survival in ccRCC patients (Figure 5B,D).

Similarly to the TCGA data analysis in pRCC, lower NFIA levels were only observed in stage III compared to stage I ($p < 0.05$) and no trend with relation to grade was observed (Figure 6A). Furthermore, a stage or grade-dependent HNF4A expression trend was not observed (Figure 6C). In line with the TCGA data, high HNF4A protein expression did not result in a survival benefit for patients with pRCC tumours (Figure 6D). For pRCC patients, high NFIA expression conferred a survival benefit ($p = 0.02$) (Figure 6B). As the pRCC cohort encompasses a limited number of patients ($n = 50$), analysis on a larger pRCC cohort is warranted to further clarify the role of HNF4A and NFIA in this RCC subtype.

Prognostic significance of NFIA and HNF4A in ccRCC

To investigate whether NFIA expression was an indicator of CSS in ccRCC patients in the TMA cohort, we performed univariable Cox regression analyses. High NFIA expression in ccRCC patients indicated a favourable effect on CSS (hazard ratio [HR] = 0.44, 95% CI = 0.32–0.62, p value < 0.001) (Table 2). Other clinicopathological parameters were also assessed via univariable Cox regression analysis to identify parameters that would influence CSS in ccRCC patients. When univariably tested, six out of nine parameters indicated a statistically significant effect on CSS. Subsequently, multivariable Cox regression analysis was performed to ascertain whether NFIA expression level could serve as a prognostic factor independently of these six clinicopathological features. We found that high NFIA expression could independently indicate better CSS in ccRCC patients (HR = 0.46, 95% CI = 0.24–0.85, p value = 0.014). A similar univariable and multivariable analysis was performed for HNF4A and although HNF4A univariably had a significant effect on CSS, we failed to detect such an effect on CSS when adjusted for other clinicopathological features (supplementary material, Table S1), suggesting that HNF4A protein expression is not an independent prognostic marker in ccRCC.

Discussion

Although the clinical outlook for RCC patients has improved in the last decade, mainly due to the use of

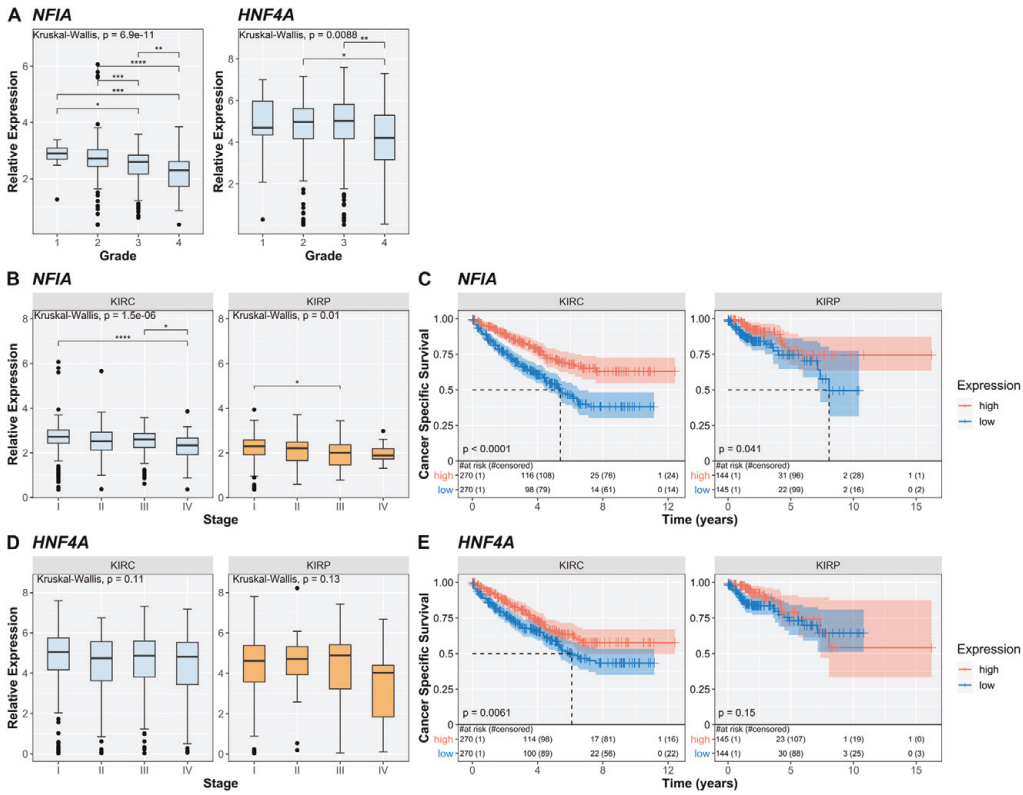

Figure 3. *HNF4A* and *NFIA* expression confers favourable survival in TCGA RCC cohorts. (A) Expression of *NFIA* and *HNF4A* in ccRCC patients within TCGA, split by Fuhrman grade. $n = 14$ (G1), 236 (G2), 206 (G3), and 76 (G4). (B) Expression of *NFIA* in ccRCC and pRCC tumours within TCGA, split by stage. ccRCC, $n = 272$ (stage I), 59 (stage II), 123 (stage III), and 83 (stage IV). pRCC, $n = 172$ (stage I), 21 (stage II), 52 (stage III), and 15 (stage IV). (C) Kaplan–Meier survival analysis of ccRCC and pRCC patients within TCGA based on high (>median) or low (<median) expression of *NFIA*. ccRCC, $n = 270$ (high), 270 (low). pRCC, $n = 145$ (high), $n = 145$ (low). (D) *HNF4A* expression in ccRCC and pRCC tumours within TCGA, split by stage. (E) Kaplan–Meier survival analysis of ccRCC and pRCC patients within TCGA based on high (>median) or low (<median) expression of *HNF4A*.

molecular therapies such as kinase and checkpoint inhibitors, RCC remains a difficult to treat disease entity. Up to 30% of RCC patients develop tumour recurrence after being considered disease-free at primary diagnosis [38–40]. This has highlighted the importance of not only gaining further information about RCC biology but also identifying patients that may have poorer prognoses. TNM classification and histological grade are currently the standard for determining RCC patient prognosis; however, its accuracy remains suboptimal as it is likely to disregard patient-related factors [41]. In the current study, we have identified a novel molecular marker, NFIA, that can independently prognosticate RCC patients. We have validated our findings through comparisons to the well-established RCC and proximal tubule TF, HNF4A, and presented novel data on its expression in pRCC with relation to clinical parameters.

HNF4A is one of the most extensively studied tissue-specific TF regulatory networks [42]. In line with this, we based our search for a novel TF that is operating within RCC in a similar manner to HNF4A. Our SCENIC analysis identified a gene coding for TF NFIA, which operates within the same RCC subtypes as HNF4A. On this basis, we founded our subsequent analyses of NFIA in comparison to HNF4A, using the

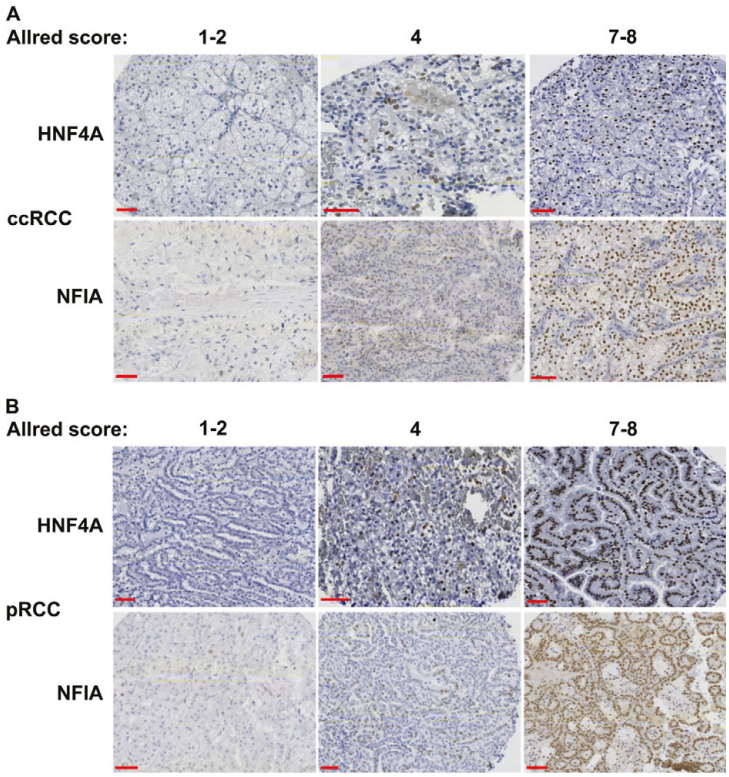

Figure 4. HNF4A and NFIA immunohistochemistry (IHC) staining on TMA tumours. (A) Representative IHC staining of HNF4A and NFIA on ccRCC tumours within the TMA, with Allred score range indicated. Scale bar represents 50 μm. (B) Representative IHC staining of HNF4A and NFIA on pRCC tumours within the TMA, with Allred score indicated. Scale bar indicates 50 μm.

former as a blueprint for how an influential TF within RCC may be uncovered. Our analysis of *HNF4A* and *NFIA* in the single-cell kidney transcriptome and pan-TCGA cohorts further likened the expression of *NFIA* to *HNF4A*. Both TFs are primarily expressed in the proximal tubule of the kidney, more specifically the S2 and S3 populations. HNF4A is an important lineage defining TF in proximal tubule development [13,14] and maintenance [43]. In fact, studies also show that NFIA is a key factor in development where haploinsufficiency leads to a range of perinatal abnormalities including renal defects [20]. Given the key role of these two TFs in proximal tubule development and homeostasis, it is no surprise that there is a large overlap in expression within the subpopulations of proximal tubular cells. Our pan-TCGA expression analysis of the two TFs shows that they are widely expressed across multiple tumour types, but more importantly, within RCC, the tumour subtype-specific expression pattern is conserved between the two TFs. This likely indicates the retention of some cell of origin gene expression characteristics even after transformation which warrants further investigation into the potential role of NFIA in RCC.

We also looked at the correlation of these two TFs in RCC subtype cohorts within the TCGA, which was subsequently validated by correlation analysis using the TMA cohort (supplementary material, Figure S3). In both analyses, NFIA expression positively correlated significantly with HNF4A expression. This could indicate that *NFIA* may be under the regulation of HNF4A or they are regulated by overlapping signalling pathways. Alternatively, they operate independently but display similar associations to disease

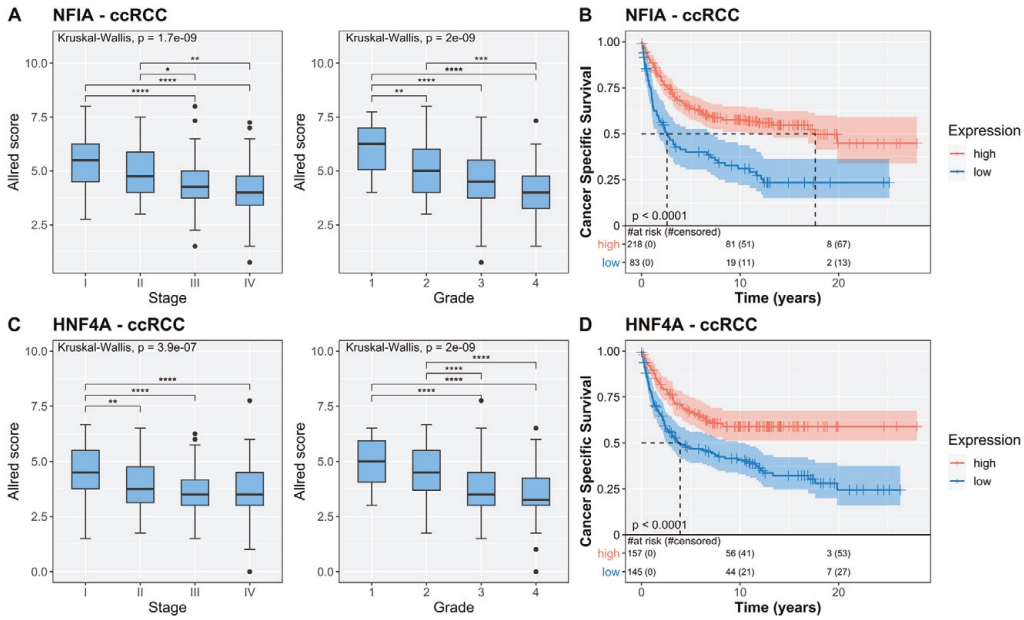

Figure 5. HNF4A and NFIA expression confers favourable CSS in the ccRCC patient TMA. (A) Allred score of ccRCC tumours stained with NFIA, split by stage and grade. Stages I–IV, $n = 108$, 47, 69, and 80. Grades I–IV, $n = 26$, 94, 130, and 55. (B) Kaplan–Meier survival analysis of ccRCC tumours based on high (≥4, $n = 219$) and low (<4, $n = 83$) NFIA Allred score. (C) Allred score of ccRCC tumours stained with HNF4A, split by stage and grade. Sample numbers same (A). (D) Kaplan–Meier survival analysis of ccRCC tumours split by high (≥4, $n = 158$) and low (<4, $n = 145$) HNF4A Allred score. Post hoc log-rank test $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$.

progression in ccRCC. Further work is required to elucidate the mechanisms pertaining to their expression, e.g. through the analysis of binding sites on both genes. Regardless of the regulatory mechanisms in place, the positive correlation provides further relevance to NFIA in RCC by associating it with HNF4A, a well-described TF behaving as a tumour suppressor gene in ccRCC [16].

To gain further clinically relevant information pertaining to NFIA, we compared its expression to clinical variables such as tumour grade, stage, and CSS, and compared this data to that of HNF4A. We performed these analyses on the TCGA RCC patient cohort and were able to validate our findings with analyses of the TMA. In both TCGA and TMA cohorts, we observed that the expression of NFIA significantly decreased with increasing tumour malignancy. Interestingly, this trend was not replicated with HNF4A, where our TMA data showed a significant decrease in HNF4A expression with increasing grade and stage, observations that were not replicated in the TCGA data set. Additionally, high tumour expression of NFIA or HNF4A conferred longer CSS compared to low-expressing ccRCC tumours in both TCGA and TMA cohorts. Given the discrepancy in HNF4A expression in relation to tumour stage and grade between the TCGA and TMA cohorts, we believe that our TMA data portray the biologically relevant setting, as it is based on protein data. Post-translational regulation is likely to play a role in HNF4A abundance in tumours, which could serve as an explanation for why differences in HNF4A abundance between stages and grades (in ccRCC) or stages (in pRCC) are not detected in RNA data from the TCGA. Furthermore, our TMA data are able to validate previously published TMA data from ccRCC by Gao *et al* [16]. We are the first to validate that HNF4A does not correlate with stage or grade and its expression level does not confer any CSS advantage in pRCC. Despite pRCCs and ccRCCs both arising from proximal tubule segments of the nephron, pRCCs are genetically and transcriptionally less uniform than ccRCC tumours [44,45], where the latter can be defined by a near-universal loss of VHL function [46]. This could explain the inability of HNF4A expression to segregate patients based on stage, grade or CSS in pRCC. However,

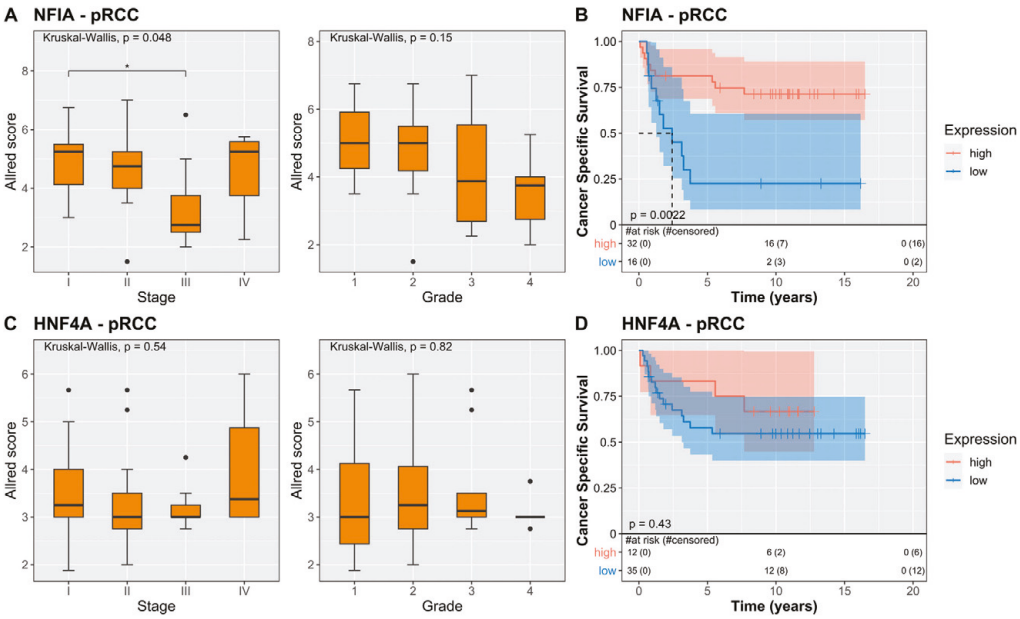

Figure 6. NFIA but not HNF4A expression confers longer CSS in pRCC TMA. (A) Allred score of pRCC tumours stained with NFIA, split by stage and grade. Stages I–IV, $n = 19$, 13, 9, and 7. Grades I–IV, $n = 7$, 20, 16, and 5. (B) Kaplan–Meier survival analysis of pRCC tumours based on high (≥4, $n = 32$) and low (<4, $n = 16$) NFIA Allred score. (C) Allred score of pRCC tumours stained with HNF4a, split by stage and grade. Same sample numbers as (A). (D) Kaplan–Meier survival analysis of pRCC tumours split by high (≥4, $n = 35$) and low (<4, $n = 12$) HNF4a Allred score. Post hoc log-rank test $^*p < 0.05$.

Table 2. Evaluation of the prognostic role of NFIA on CSS in univariable and multivariable analyses

		Univariable			Multivariable		
Covariate	Units	Hazard ratio	95% CI	*P* value	Hazard ratio	95% CI	*P* value
NFIA	<4 Allred Score	ref.					
	≥4 Allred Score	0.44	0.32–0.62	<0.001	0.46	0.24–0.85	0.014
Grade	III–IV	ref.					
	I–II	0.36	0.25–0.53	<0.001	0.30	0.15–0.61	0.001
Stage	III–IV	ref.					
	I–II	0.14	0.10–0.21	<0.001	0.14	0.06–0.33	<0.001
Age		0.99	0.98–1	0.16			
Sex		0.91	0.66–1.3	0.59			
C-reactive protein		1	1–1	<0.001	1.01	1.01–1.02	<0.001
Weight		0.99	0.97–1	0.037	1.01	0.99–1.03	0.951
Venous invasion	No venous growth	ref.					
	Sinus vein	2.01	1.04–3.89	0.038	2.45	0.85–7.03	0.096
	Renal vein	2.37	1.57–3.59	<0.001	0.95	0.40–2.26	0.916
	Vena cava	2.61	1.68–4.06	<0.001	0.68	0.31–1.46	0.312
Tumour diameter		1.01	1.01–1.02	<0.001	1.00	0.99–1.01	0.842
Bilateralness	No						
	Primary	1.19	0.49–2.9	0.703			
	Secondary	1.07	0.34–3.36	0.906			

ref., reference category.

our TMA-based NFIA data suggest that the expression of this TF in pRCC can segregate patients based on stage and CSS but not grade, despite the high biological variability displayed by pRCC tumours.

Founded on our results that NFIA and HNF4A can segregate patients based on stage, grade and CSS in the ccRCC setting, we explored the potential use of NFIA and HNF4A as independent prognostic marker in ccRCC. HNF4A was unable to independently predict CSS, but we found that NFIA is able to predict poor CSS in ccRCC patients independently of clinicopathological parameters such as tumour grade and stage. Although HNF4A is previously described in RCC literature, we show that NFIA may be a better prognostication marker. Such evaluations are critical to refine the pool of potential prognostication markers within the management of RCC. With studies of this nature, there are challenges with regard to the applicability of immunohistochemistry-based methods pertaining to reagents and quantification. Additionally, observations need to be validated with other independent cohorts. Nevertheless, robust prognostication tools are essential for the effective clinical management of RCC patients and the initiation and subsequent translation of these biomarker investigations are a key step in improving patient survival.

Acknowledgements

This research was supported by the CanFaster Grant EU-H2020-MSCA-COFUND-2016-754299, the Swedish Cancer Society (Grant no. CAN2018/1153), and Regional ALF Funds (Grant no. 2018-176).

Author contributions statement

HA and RdA conceived the study. RdA, MI and CM performed the experiments and collected data. RdA, MI and SS carried out data analyses. SS adapted software and performed bioinformatic analyses. RdA and SS prepared figures. BL collected clinical samples and provided clinical data. HA obtained funding. RdA drafted the manuscript. All authors revised the manuscript and approved the submitted version.

Ethics approval and consent to participate

The study was reviewed and approved by the Ethical Review Board (Dnr: 2015-146-31M and Dnr: 2018-296-32M) and the Ethical Board of Sweden (Dnr: 2019-02579). All procedures regarding patients were in accordance with the Helsinki Declaration. The patients had informed consent, orally before year 2000, and informed and written consent from year 2000.

Data availability statement

Data are available upon request to the corresponding author.

References

1. Hsieh JJ, Purdue MP, Signoretti S, *et al*. Renal cell carcinoma. *Nat Rev Dis Primers* 2017; **3**: 17009.
2. Cheval L, Pierrat F, Rajerison R, *et al.* Of mice and men: divergence of gene expression patterns in kidney. *PLoS One* 2012; **7**: e46876.
3. Lee JW, Chou CL, Knepper MA. Deep sequencing in microdissected renal tubules identifies nephron segment-specific transcriptomes. *J Am Soc Nephrol* 2015; **26**: 2669–2677.
4. Cimadamore A, Cheng L, Scarpelli M, *et al*. Towards a new WHO classification of renal cell tumor: what the clinician needs to know – a narrative review. *Transl Androl Urol* 2021; **10**: 1506–1520.
5. Liao J, Yu Z, Chen Y, *et al*. Single-cell RNA sequencing of human kidney. *Sci Data* 2020; **7**: 4.
6. Cohen HT, McGovern FJ. Renal-cell carcinoma. *N Engl J Med* 2005; **353**: 2477–2490.
7. Lindgren D, Sjolund J, Axelson H. Tracing renal cell carcinomas back to the nephron. *Trends Cancer* 2018; **4**: 472–484.
8. Lindgren D, Eriksson P, Krawczyk K, *et al*. Cell-type-specific gene programs of the normal human nephron define kidney cancer subtypes. *Cell Rep* 2017; **20**: 1476–1489.
9. Nickerson ML, Jaeger E, Shi Y, *et al*. Improved identification of von Hippel-Lindau gene alterations in clear cell renal tumors. *Clin Cancer Res* 2008; **14**: 4726–4734.
10. Shen C, Kaelin WG Jr. The VHL/HIF axis in clear cell renal carcinoma. *Semin Cancer Biol* 2013; **23**: 18–25.
11. Patel SA, Hirosue S, Rodrigues P, *et al*. The renal lineage factor PAX8 controls oncogenic signalling in kidney cancer. *Nature* 2022; **606**: 999–1006.
12. Lazarevich NL, Fleishman DI. Tissue-specific transcription factors in progression of epithelial tumors. *Biochemistry* 2008; **73**: 573–591.
13. Marable SS, Chung E, Park JS. Hnf4a is required for the development of Cdh6-expressing progenitors into proximal tubules in the mouse kidney. *J Am Soc Nephrol* 2020; **31**: 2543–2558.
14. Marable SS, Chung E, Adam M, *et al*. Hnf4a deletion in the mouse kidney phenocopies Fanconi renotubular syndrome. *JCI Insight* 2018; **3**: e97497.
15. Thiagarajan RD, Georgas KM, Rumballe BA, *et al*. Identification of anchor genes during kidney development defines ontological relationships, molecular subcompartments and regulatory pathways. *PLoS One* 2011; **6**: e17286.

16. Gao Y, Yan Y, Guo J, *et al*. HNF-4α downregulation promotes tumor migration and invasion by regulating E-cadherin in renal cell carcinoma. *Oncol Rep* 2019; **42**: 1066–1074.
17. Gronostajski RM. Roles of the NFI/CTF gene family in transcription and development. *Gene* 2000; **249**: 31–45.
18. Cardinaux JR, Chapel S, Wahli W. Complex organization of CTF/NF-I, C/EBP, and HNF3 binding sites within the promoter of the liver-specific vitellogenin gene. *J Biol Chem* 1994; **269**: 32947–32956.
19. das Neves L, Duchala CS, Tolentino-Silva F, *et al*. Disruption of the murine nuclear factor I-A gene (Nfia) results in perinatal lethality, hydrocephalus, and agenesis of the corpus callosum. *Proc Natl Acad Sci U S A* 1999; **96**: 11946–11951.
20. Lu W, Quintero-Rivera F, Fan Y, *et al*. NFIA haploinsufficiency is associated with a CNS malformation syndrome and urinary tract defects. *PLoS Genet* 2007; **3**: e80.
21. Bachurski CJ, Kelly SE, Glasser SW, *et al*. Nuclear factor I family members regulate the transcription of surfactant protein-C. *J Biol Chem* 1997; **272**: 32759–32766.
22. Chen KS, Lim JWC, Richards LJ, *et al*. The convergent roles of the nuclear factor I transcription factors in development and cancer. *Cancer Lett* 2017; **410**: 124–138.
23. Chen KS, Lynton Z, Lim JWC, *et al*. NFIA and NFIB function as tumour suppressors in high-grade glioma in mice. *Carcinogenesis* 2021; **42**: 357–368.
24. Yang B, Zhou ZH, Chen L, *et al*. Prognostic significance of NFIA and NFIB in esophageal squamous carcinoma and esophagogastric junction adenocarcinoma. *Cancer Med* 2018; **7**: 1756–1765.
25. Kovacs G, Akhtar M, Beckwith BJ, *et al*. The Heidelberg classification of renal cell tumours. *J Pathol* 1997; **183**: 131–133.
26. Fuhrman SA, Lasky LC, Limas C. Prognostic significance of morphologic parameters in renal cell carcinoma. *Am J Surg Pathol* 1982; **6**: 655–663.
27. Brierley JD, Gospodarowicz MK, Wittekind C. *TNM Classification of Malignant Tumours* (8th edn). Wiley-Blackwell: Oxford, 2017.
28. Hedberg Y, Ljungberg B, Roos G, *et al*. Expression of cyclin D1, D3, E, and p27 in human renal cell carcinoma analysed by tissue microarray. *Br J Cancer* 2003; **88**: 1417–1423.
29. Bankhead P, Loughrey MB, Fernández JA, *et al*. QuPath: open source software for digital pathology image analysis. *Sci Rep* 2017; **7**: 16878.
30. Colaprico A, Silva TC, Olsen C, *et al*. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* 2016; **44**: e71.
31. Aibar S, González-Blas CB, Moerman T, *et al*. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 2017; **14**: 1083–1086.
32. Obradovic A, Chowdhury N, Haake SM, *et al*. Single-cell protein activity analysis identifies recurrence-associated renal tumor macrophages. *Cell* 2021; **184**: 2988–3005.e16.
33. Hao Y, Hao S, Andersen-Nissen E, *et al*. Integrated analysis of multimodal single-cell data. *Cell* 2021; **184**: 3573–3587.e29.
34. Korsunsky I, Millard N, Fan J, *et al*. Fast, sensitive and accurate integration of single-cell data with harmony. *Nat Methods* 2019; **16**: 1289–1296.
35. Lu J, Chen Z, Zhao H, *et al*. ABAT and ALDH6A1, regulated by transcription factor HNF4A, suppress tumorigenic capability in clear cell renal cell carcinoma. *J Transl Med* 2020; **18**: 101.
36. Zhang Y, Narayanan SP, Mannan R, *et al*. Single-cell analyses of renal cell cancers reveal insights into tumor microenvironment, cell of origin, and therapy response. *Proc Natl Acad Sci U S A* 2021; **118**: e2103240118.
37. Harlander S, Schönenberger D, Toussaint NC, *et al*. Combined mutation in Vhl, Trp53 and Rb1 causes clear cell renal cell carcinoma in mice. *Nat Med* 2017; **23**: 869–877.
38. Lucca I, de Martino M, Hofbauer SL, *et al*. Comparison of the prognostic value of pretreatment measurements of systemic inflammatory response in patients undergoing curative resection of clear cell renal cell carcinoma. *World J Urol* 2015; **33**: 2045–2052.
39. Haas NB, Manola J, Uzzo RG, *et al*. Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): a double-blind, placebo-controlled, randomised, phase 3 trial. *Lancet* 2016; **387**: 2008–2016.
40. Lam JS, Shvarts O, Leppert JT, *et al*. Postoperative surveillance protocol for patients with localized and locally advanced renal cell carcinoma based on a validated prognostic nomogram and risk group stratification system. *J Urol* 2005; **174**: 466–472 discussion 72; quiz 801.
41. Zhou L, Cai X, Liu Q, *et al*. Prognostic role of C-reactive protein in urological cancers: a meta-analysis. *Sci Rep* 2015; **5**: 12733.
42. Lazarevich NL, Shavochkina DA, Fleishman DI, *et al*. Deregulation of hepatocyte nuclear factor 4 (HNF4) as a marker of epithelial tumors progression. *Exp Oncol* 2010; **32**: 167–171.
43. Sasaki S, Hara A, Sakaguchi M, *et al*. Hepatocyte nuclear factor 4alpha regulates megalin expression in proximal tubular cells. *Biochem Biophys Rep* 2019; **17**: 87–92.
44. Cancer Genome Atlas Research Network, Linehan WM, Spellman PT, *et al*. Comprehensive molecular characterization of papillary renal-cell carcinoma. *N Engl J Med* 2016; **374**: 135–145.
45. Chen F, Zhang Y, Şenbabaoğlu Y, *et al*. Multilevel genomics-based taxonomy of renal cell carcinoma. *Cell Rep* 2016; **14**: 2476–2489.
46. Rathmell WK, Chen S. VHL inactivation in renal cell carcinoma: implications for diagnosis, prognosis and treatment. *Expert Rev Anticancer Ther* 2008; **8**: 63–73.

SUPPLEMENTARY MATERIAL ONLINE

Figure S1. Correlation analysis of HNF4A and NFIA expression in the TMA cohort

Figure S2. Distribution of Fuhrman grades across T-grouped stages within ccRCC and pRCC patient cohorts

Figure S3. Distribution of T, N, and M status' within T-grouped stages in ccRCC and pRCC patients

Table S1. Evaluation of the prognostic value of HNF4A on CSS in univariable and multivariable analyses

Identification and validation of NFIA as a novel prognostic marker in RCC

Roger de Alwis, Sarah Schoch *et al*, *J Pathol Clin Res*, https://doi.org/10.1002/cjp2.316

Supplementary Material

Supplementary Figures S1 – S3
Supplementary Table S1

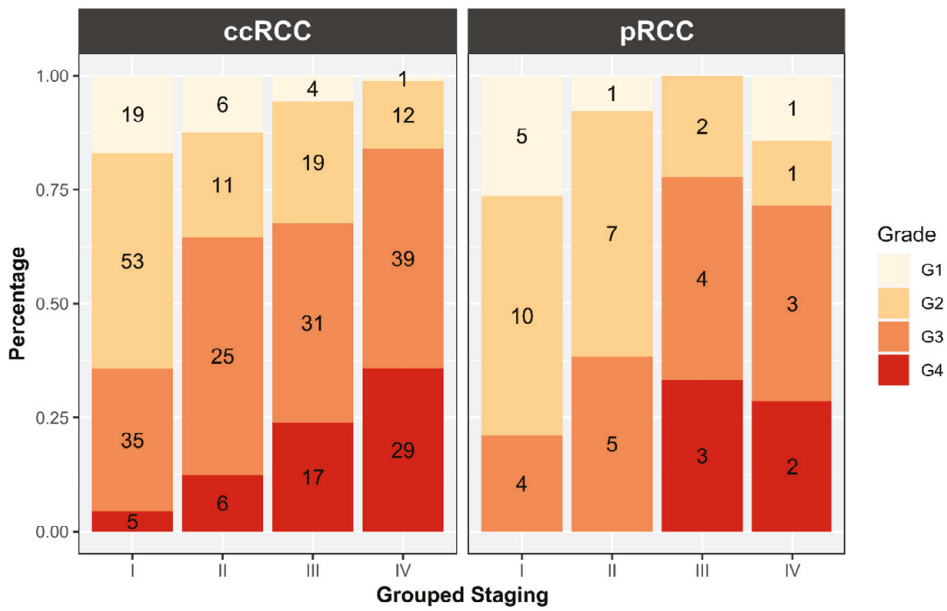

Figure S1. Distribution of Furhman grades across T grouped stages within ccRCC and pRCC patient cohorts.

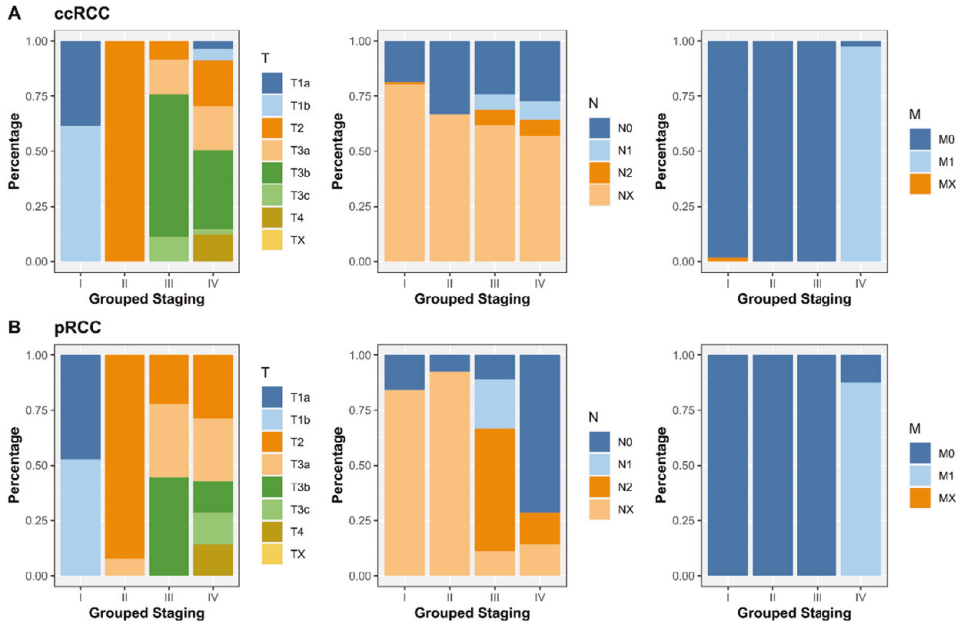

Figure S2. (A) Distribution of T, N and M status' within T grouped stages in ccRCC patients. (B) Distribution of T, N and M status' within T grouped stages in pRCC patients.

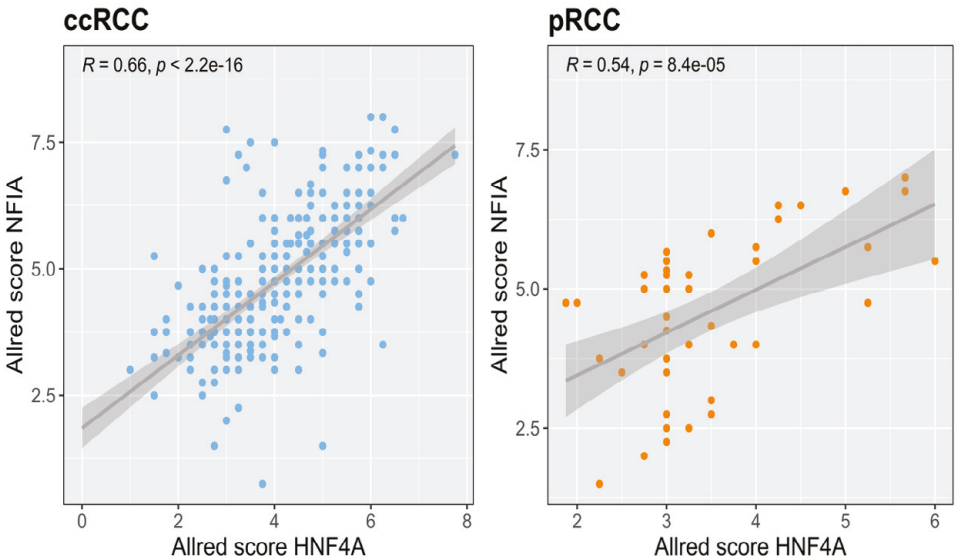

Figure S3. Correlation analysis of HNF4A and NFIA expression in the TMA cohort.

Table S1. Evaluation of the prognostic value of HNF4A on CSS in univariable and multivariable analyses.

		Univariable			Multivariable		
Covariate	**Units**	**Hazard Ratio**	**95% CI**	**p value**	**Hazard Ratio**	**95% CI**	**p value**
HNF4A	<4 Allred Score	ref.*					
	≥4 Allred Score	0.50	0.36-0.69	<0.001	1.19	0.66-2.13	0.566
Grade	III-IV	ref.*					
	I-II	0.36	0.25-0.53	<0.001	0.31	0.15-0.63	0.001
Stage	III-IV	ref.*					
	I-II	0.14	0.10-0.21	<0.001	0.10	0.04-0.23	<0.001
Age		0.99	0.98-1.00	0.162			
Sex		0.91	0.66-1.27	0.591			
C-reactive protein		1.01	1.01-1.01	<0.001	1.01	1.01-1.02	<0.001
Weight		0.99	0.97-1.00	0.037			
Venous Invasion	No venous growth	ref.*					
	Sinus vein	2.01	1.04-3.89	0.038			
	Renal vein	2.37	1.57-3.59	<0.001	0.63	0.28-1.42	0.264
	Vena cava	2.37	1.57-3.59	p<0.001	0.59	0.27-1.27	0.175
Tumour Diameter		1.01	1.01-1.02	<0.001	1.00	0.99-1.01	0.664
Bilateralness	None						
	Primary	1.19	0.49-2.90	0.703			
	Secondary	1.07	0.34-3.36	0.906			

*ref: reference category

Paper III

Size-based isolation and detection of renal carcinoma cells from whole blood

ROGER DE ALWIS[1], JENNIFER HANSSON[1,2], DAVID LINDGREN[1,3], SARAH SCHOCH[1], ALEXANDER TEJERA[1], BIANCA SCHOLTZ[4], PETER ELFVING[4], CHRISTINA MÖLLER[1], HELÉN NILSSON[5], MARTIN JOHANSSON[5] and HÅKAN AXELSON[1]

[1]Division of Translational Cancer Research, Department of Laboratory Medicine, Lund University, 223 81 Lund; [2]Kyowa Kirin AB, 164 40 Kista; [3]Division of Clinical Genetics, Lund University, 22184 Lund; [4]Department of Urology, [5]Center for Molecular Pathology, Department of Translational Medicine, Skåne University Hospital, 214 28 Malmö, Sweden

Received November 23, 2021; Accepted February 18, 2022

DOI: 10.3892/ mco.2022.2534

Abstract. Renal cell carcinoma (RCC) is a tumour type with an indolent growth pattern and rather vague symptoms. The present study developed a platform for liquid biopsy of RCC based upon the isolation of circulating tumour cells (CTCs). Founded on the observation that RCC tumour cells are considerably larger than leucocytes, the present study employed a microfluidics-based system for isolation of RCC CTCs from whole blood. Using this system, it was revealed that 66% of spiked-in RCC tumour cells could be retrieved using this approach. Furthermore, it was demonstrated that these cells could be molecularly detected with digital PCR using RCC-specific genes down to one tumour cell, whilst avoiding detection in samples lacking tumour cells. Finally, subtype specific transcripts were identified to distinguish the different subtypes of RCC, which were then validated in patient tumours. The present study established a novel workflow for the isolation of RCC CTCs from whole blood, with the potential to detect these cells irrespective of subtype.

Introduction

Recently, liquid biopsies have garnered attention for their applications in personalised medicine. A liquid biopsy, as opposed to a tissue biopsy, requires that the analytical sample is derived from the bodily fluid of a patient by a minimally invasive method. This includes but is not limited to peripheral-vein blood, saliva, urine and cerebrospinal fluid. When considering blood as a bio-source there are potential components that can be considered such as circulating tumour cells (CTCs), platelets or plasma as a source of circulating tumour (ct) nucleic acids (1). The choice of component used in a liquid biopsy assay will depend on the disease being queried, the potential for gathering tumour derived information and downstream applications.

Circulating tumour cells shed into systemic circulation by the primary tumour or metastases (2). Tumour cells enter the circulation via intravasation and typically an epithelial-mesenchymal transition as part of the metastatic cascade. However, CTCs may also be found in clusters in the circulation (3). CTCs and/or circulating tumour cell clusters can be isolated from blood in systemic circulation at which point their DNA, RNA, protein expression or functionality can be studied (4). If efficiently isolated, CTCs are rich sources for multi-omic analyses that may not only aid in assessment of disease or disease progression but also in understanding aspects of tumour processes.

The majority (>90%) of solid lesions found in the kidney are renal cell carcinomas (5) and patients with the disease may benefit from the advancement of liquid biopsy based molecular detection (6). More than 60% of RCC patients are asymptomatic, and often radiologic diagnoses are made incidentally (5,7). 20% of these cases are metastatic at diagnosis (8,9). For local disease, surgical resection is performed, however, 30% of these patients will experience recurrence within 5 years after surgery (9,10). Despite these circumstances, currently there are no cost-effective or practical approaches to detecting early metastases or monitoring recurrence (11).

The three most common RCC subtypes are clear cell RCC (ccRCC), papillary RCC (p1RCC & p2RCC) and chromophobe RCC (chRCC) accounting for 75-80%, 15% and 5% of cases, respectively (12,13). These tumour entities arise along the nephron with ccRCC and pRCC originating from cells of the proximal tubules and chRCC from the collecting duct. As a consequence, they are transcriptionally distinct, with ccRCCs and pRCCs retaining much of the HNF-driven transcriptional program found in the proximal tubules and chRCC retaining the FOXI1-driven programme defining collecting duct cells (12). Additionally, almost all ccRCC tumours lack a functional VHL protein, rendering them pseudo-hypoxic due to the

Correspondence to: Professor Håkan Axelson, Division of Translational Cancer Research, Department of Laboratory Medicine, Lund University, Building 404 A3, Medicon Village, Scheelevägen 2, 223 81 Lund, Sweden
E-mail: hakan.axelson@med.lu.se

Key words: circulating tumour cells, kidney cancer, liquid biopsy, molecular diagnostics

accumulation of hypoxia inducible factor (HIF) proteins (14). These transcriptional differences between the subtypes set the stage for liquid biopsy approaches that could allow for subtype specific monitoring of RCC tumours.

With regards to blood based liquid biopsies, it has been observed that RCC is a poor shedder of circulating tumour DNA (ctDNA) and hence may be better suited for CTC analyses (15). Importantly, RCC cells are not suitable for the commonly used and FDA approved EpCAM based isolation strategies due to their poor expression of this surface marker, despite being of epithelial origin (16,17). Antigen-dependent enrichment methods, such as the use of the surface marker CA9 in ccRCC, has been previously employed to successfully detect CTCs (18,19). However, this type of enrichment is limited to one subtype or population of RCC CTCs. In order to broadly enrich CTCs from a wider range of RCC subtypes, an antigen-independent approach is desirable such as cell size-based enrichment of CTCs (16,20). We have addressed this need by employing the ClearCell FX platform which uses a size-based, microfluidic enrichment approach. This platform enriches fully intact and viable CTCs directly from whole blood by exploiting the biophysical disparities between CTCs and other cells found in the blood through the Dean Flow Fractionation principle (21). Cells travelling through a spiral microfluidic channel experience counter rotating flow vortices (collectively called a Dean vortex), which channel larger cells towards the inner wall of the channel and smaller cells towards the outer wall. In addition to this, there are cell diameter dependent wall-induced and shear-induced lift forces which further focus the cells along opposing inner wall of the channel. Direct enrichment in buffer is beneficial, compared to the Parsortix system for example, where cells are enriched inside a cassette and washing out these cells may pose a risk of cell loss (22). Enriched, intact CTCs on a limited background of leucocytes within this buffer can then be processed downstream for applications such as gene expression querying. We have evaluated the performance of this platform in relation to enriching RCC cells and then explored and optimized methods for specifically detecting these cells through the use of subtype specific biomarkers.

Materials and methods

Primary patient tissue, primary cells, cell culture and healthy blood controls. A total of 16 primary RCC patient tumour samples (Table I) and 5 primary RCC cell were kindly provided by Dr Helén Nilsson, Department of Translational Medicine, Lund University for validation of selected biomarkers. Patient tumour samples were retrieved from the material collection established after Lund University ethical committee approval (approval no. LU680-08), where informed written consent was obtained before archival. Human renal tissue dissociation and culturing was performed as described in Appendix S1 and previously described in Hansson *et al* (23). Tumour tissues obtained consisted of ccRCC, p1RCC, p2RCC and chRCC tissues and cells that were predetermined by a trained pathologist. ccRCC cell lines SNU-349 (Korean Cell Line Bank) and 786-O (ATCC) were cultured in RPMI-1640 (Corning, Manassas, USA) and DMEM (Corning, Manassas, USA) respectively, both medium supplemented with 10% foetal bovine serum (Gibco, MA, USA). All cell lines were incubated in a humidified chamber in 5% CO_2 at 37˚C, were STR authenticated, had a passage number of not more than 20 and were mycoplasma free. Healthy blood for controls were obtained from the Blood donation centre in Lund with informed written consent from patients and ethical permission obtained by the regional health care provider (Region Skåne, Lund, Sweden; approval no. 2018:19).

Identification of RCC subtype specific marker genes. *CA9* and *SLC6A3* were selected as ccRCC subtype specific markers based on a literature search and previously published data from Hansson *et al* (23). The subtype specific expression of these markers was confirmed in the TCGA data set as described below. The same approach was also used to identify subtype specific markers for the other two subtypes, pRCC and chRCC. Level 3 RNA-sequencing data processed using the 'UNC V2 RNA-Seq' workflow was downloaded from The Cancer Genome Atlas (TCGA) data portal (https://portal.gdc.cancer.gov/), as described in Lindgren *et al* (12). Statistical analyses were performed using the R software (http://www.r-project.org). Differential gene expression between clear cell, papillary and chromophobe RCC (KIRC, KIRP and KICH, respectively) tumours was determined using the Limma Bioconductor package (24). For the analysis of taxonomy groups, samples from all 3 TCGA kidney cancer projects [chromophobe RCC (KICH), clear cell (KIRC), papillary kidney carcinoma (KIRP)] were merged into one single dataset (25). Molecular RCC taxonomy groups as well as clinical and histopathologic parameters were used as presented in the article by Chen *et al* (26). The selection criteria for sub-type specific expression were minimal expression in human blood (based on the blood datasets found in the Human Protein Atlas, available from https://www.proteinatlas.org) and distinct RCC sub-type specific expression. For the two types of papillary RCC defined by Chen *et al* (26) (p.e1 and p.e2) optimal markers for the more common p.e1 subtype (but also displaying elevated expression in p.e2 tumours) were selected (Table SI). These selection criteria resulted in a panel of 7 genes (Table II).

Verification of subtype specific gene expression. cDNA synthesis for cell lines and primary tumours was performed using the High-Capacity RNA-to-cDNA kit. Quantitative PCR (qPCR) was performed on a QuantStudio 7 Flex Real-Time PCR system using TaqMan Gene Expression Master Mix and TaqMan Probes (Table III) (all from Applied Biosystems, Vilnius, Lithuania). Relative gene expression from qPCR data was calculated using the double delta Cq method and normalised to *β-actin* levels (27).

Cell size measurements. Cell diameter of RCC cell lines and primary cell lines were measured on a Nucleocounter NC-3000 (ChemoMetec, Allerod, Denmark) using NC-Slide A8 with the Cell Viability and Cell Count Assay, according to manufacturer's instructions.

Establishment of ClearCell FX system performance for RCC cells. Whole blood (7.5 ml) was processed for each run on the ClearCell FX system (Biolidics, Singapore). Firstly, red blood cells (RBCs) were lysed by a 10-minute

Table I. Patient tumour characteristics for biomarker panel validation.

Patient ID	RCC subtype	Furhman grade	Stage
R294T	ccRCC	F2	pT1a
R320T	ccRCC	F1	pT1b
R375T	ccRCC	F3(I)	pT3a
R363T	ccRCC	F3(I)	pT3a
R256T	chRCC	F3	pT3a
R308T	chRCC	-	pT1a
R377T	chRCC	-	pT1a
R275T	chRCC	F3(I)	F3 (I)
R221T	p1RCC	F3	pT2b
R163T	p1RCC	F2	pT1b
R376T	p1RCC	F2(I)	pT1b
R188T	p1RCC	F1	pT3a
R290T	p2RCC	F4	pT3b
R303T	p2RCC	F2(I)	pT1b
R321T	p2RCC	F2(I)	pT1b
R199T	p2RCC	F4	pT3a

ccRCC, clear cell renal cell carcinoma; chRCC, chromophobe RCC; p1RCC, papillary type 1 RCC; p2RCC, papillary type 2 RCC. (I) denotes tumour grading according to 2016 WHO/ISUP system (39).

Table II. Biomarkers selected to distinguish RCC subtypes.

RCC subtype	Selected markers
Clear Cell RCC	*SLC6A3, CA9*
Papillary RCC	*SOSTDC1, SLC34A2, LRRN4*
Chromophobe RCC	*SLC26A7, ATP6V0A4*

Table III. TaqMan probes used for subtype specific markers.

Gene name	TaqMan probe
CA9	Hs00154208_m1
SLC6A3	Hs00997374_m1
SOSTDC1	Hs00383602_m1
SLC34A2	Hs00197519_m1
LRRN4	Hs00379905_m1
SLC26A7	Hs01104163_m1
ATP6V0A4	Hs00220986_m1

incubation with a RBC lysis buffer (G-Bioscience, St. Louis, MO, USA) and discarded via centrifugation and removal of supernatant. The resulting cell pellet containing nucleated cells was resuspended in 4 ml of ClearCell FX re-suspension buffer prior to being loaded onto the ClearCell FX system and running 'Protocol 1'. Renal cancer cell line SNU-349 and primary RCC cells were used to establish the recovery efficiency of the ClearCell FX tumour cell isolation system for RCC cells. Cells were labelled by incubating in serum free media (Corning, Manassas, USA) with a final concentration of 25 μM CellTracker Green CMFDA (Invitrogen, Carlsbad, USA) dye for 45 min. Cells were then washed with DPBS and diluted to varying numbers before being counted and spiked into 7.5 ml of healthy donor blood. The spiked blood sample was then processed and run through the ClearCell FX system according to manufacturer's instructions. The isolated output cells were then pelleted, resuspended in a lower volume and aliquoted into a 96-well plate. Labelled and isolated cells were counted using an inverted immunofluorescence microscope to calculate the recovery efficiency.

RNA extraction, global-preamplification and cDNA clean-up. RNA extraction for ClearCell FX output samples were performed according to manufacturer's instructions using the Norgen Biotek Single Cell RNA Purification kit (Norgen Biotek Corp., Thorold, Canada) and RNA was eluted in 14 μl of PCR clean water. Global pre-amplification of RNA and cDNA clean-up was performed on RNA extracts from ClearCell FX system outputs using the SMART-Seq v4 Ultra Low Input RNA kit (Takara Bio Inc., Shiga, Japan) according to manufacturer instructions. Pre-amplified cDNA was purified using AMPure XP magnetic bead solution (Beckman Coulter, USA) and eluted in 20 μl TE buffer. RNA concentration, RNA integrity number and pre-amplified cDNA concentration from ClearCell FX system outputs was assessed on the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, USA). Primary patient tumour tissue (<30 mg) was processed by disruption in a TissueLyser using Trizol (Qiagen, Hilden, Germany). Homogenization was achieved by flushing the disrupted sample through a QIAshredder column. Subsequent RNA extraction was performed using the Qiagen RNeasy Mini (Qiagen, Hilden, Germany) kit according to manufacturer's instructions. RNA concentration and quality from cultured cells and primary tumour material was assessed on a Nanodrop 2000 (Thermo Fisher Scientific, Waltham, USA).

Real-time monitoring of global pre-amplification. Global pre-amplification was monitored in real time to obtain the optimal cycle number by adding 0.1X SYBR Green (Sigma Aldrich, Darmstadt, Germany) and allowing the reaction to run for 40 cycles (28).

Statistical analyses. Mann-Whitney U, mean and SEM calculations were performed on GraphPad Prism software (Graph Pad Software, San Diego, USA). Kruskal-Wallis with Dunn's comparison calculations were performed on R-software (Vienna, Austria) and GraphPad Prism.

Results

Tumour cell isolation strategy. In this study we set out to develop a method for efficient isolation of CTCs from RCC patient blood. Previous studies have indicated that label-free, size-based enrichment is particularly well suited for RCC, which led us to employ the size-based ClearCell FX system for our study (16,29).

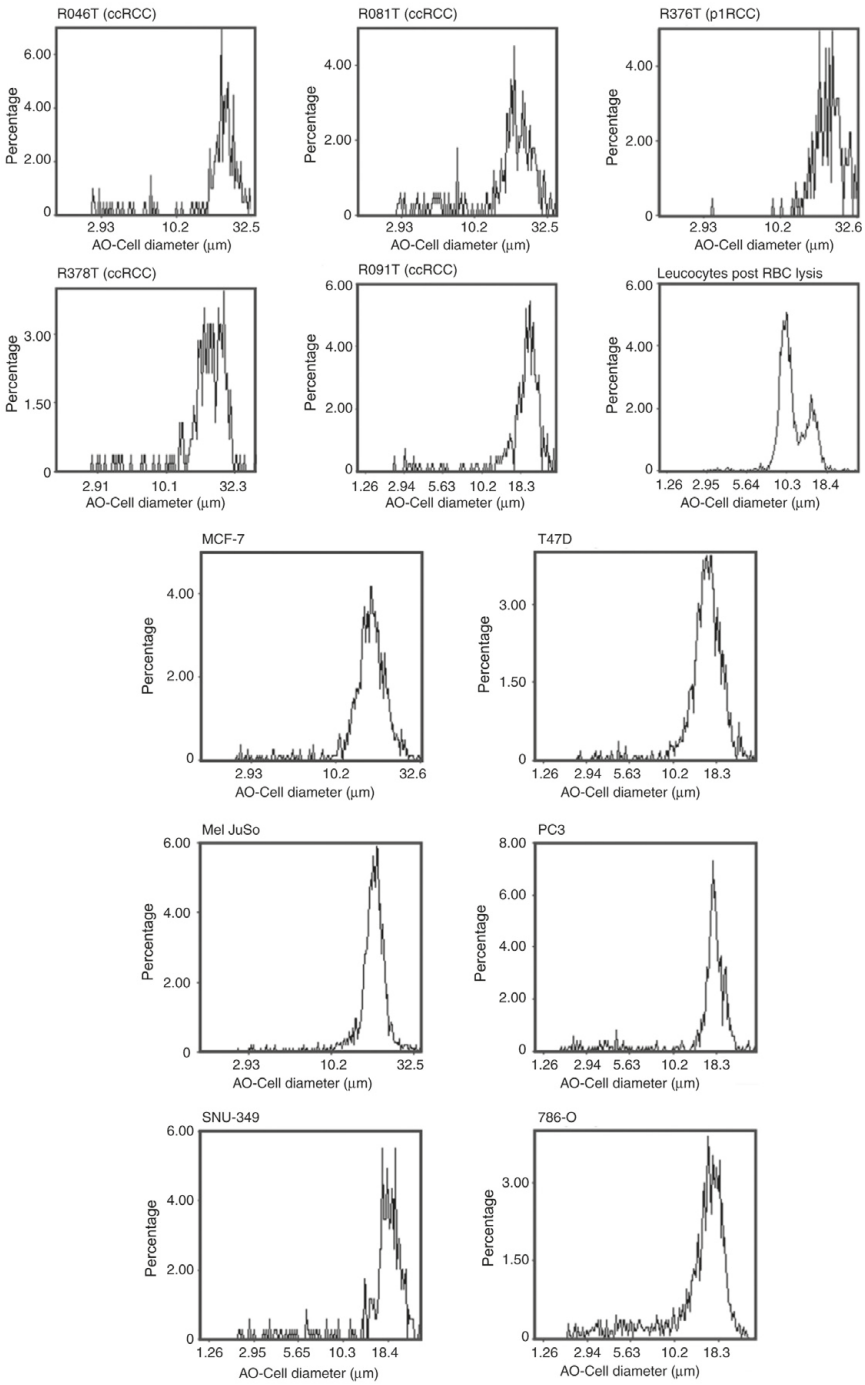

Figure 1. Cell diameters of primary RCC, cultured RCC and non-RCC cells. AO, Acridine Orange; ccRCC, clear cell renal cell carcinoma; p1RCC, papillary type 1 RCC; RBC, red blood cell lysis.

In order to confirm that RCC tumour cells were large enough to be successfully enriched in a size-dependent system, we measured the cell diameters of two ccRCC cell lines along with 5 primary cell lines (ccRCC and p1RCC). This data showed that cultured RCC tumour cells SNU-349 and 786-O are approximately 18 μm in diameter whilst cultured primary tumour cells are larger and measure between 20-30 μm (Fig. 1). These values are greater than the lower ClearCell FX isolation size threshold of 14 μm suggesting that our choice of enrichment method is likely suitable for isolation of CTCs from RCC patients. A small fraction of leucocytes from whole blood are larger than 14 μm, these are also presumably enriched and isolated together with the tumour cells. Furthermore, we measured the cell diameter of four non-RCC tumour cell lines (MCF-7, T47D, Mel JuSo, PC3) to confirm that tumour cells in general, and RCC cells in particular, are larger than most leucocytes.

ClearCell FX system performance on RCC cells. To establish the recovery efficiency of the ClearCell FX system with RCC cells, we performed cell spike-in experiments. For these experiments we employed the ccRCC cell line SNU-349 and primary RCC cell lines. The cells were labelled with a fluorescent tracker and the number of fluorescent cells were counted (range of cells spiked-in 10-232) before mixing them with 7.5 ml of healthy donor blood. After lysis of the reticulocytes, the blood sample was subjected to size-based isolation using the ClearCell FX system. The output, containing a background of approximately 10,000 leukocytes and the enriched labelled SNU-349 cells were thereafter analysed using a fluorescent microscope. These experiments show that on average, the system is able to recover 50% of SNU-349 cells and 66% of primary RCC cells that are spiked into whole blood (Fig. 2A). Spiked-in and recovered cell numbers are reported in Table SII. We also performed immunofluorescence staining on ClearCell FX enriched samples originating from whole blood spiked with SNU-349 cells. To distinguish RCC cells from leucocytes we stained cells for CD45 (leukocytes) and CA9 (a well-established marker for ccRCC cells). We were able to clearly distinguish CA9 positive, CD45 negative SNU-349 cells on a background of only CD45 positive leucocytes (Fig. 2B)

Assessment of RNA quality, pre-amplification and assay reproducibility. Next, we wanted to establish and validate a method to detect RCC-specific RNA transcripts from the ClearCell FX enriched CTC fraction. First, RNA was extracted and subjected to quality and quantity assessment (Fig. S1). Once the quality was deemed acceptable (RIN >5), we performed global pre-amplification of reverse transcribed cDNA to yield sufficient sample material and to increase the relatively low transcript copy numbers from the enriched SNU-349 cells present on a background of leukocytes in the sample. Prior to global pre-amplification of cDNA from the CTC enriched samples, the pre-amplification process was monitored with quantitative PCR using SNU-349 cDNA, in order to determine the optimal number of cycles required (Fig. S1B). A successful pre-amplification reaction should yield an increase in fragments in the 400-10,000 base-pair range, which was observed in our samples via fragment analysis. (Fig. S1C).

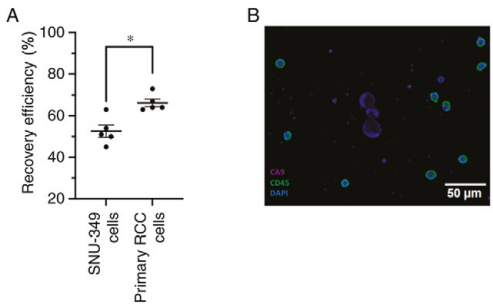

Figure 2. ClearCell FX recovery efficiency and immunofluorescence-based detection of RCC cells. (A) SNU-349 and primary RCC cell line spike-in and recover efficiencies (%). Horizontal lines represent average recovery efficiency and error bars are standard error of mean (SEM). Mann-Whitney U Test, *P<0.05 (two-tailed). (B) Immunofluorescence-based detection of spiked SNU-349 cells enriched from whole-blood using the ClearCell FX system. The output material from 7.5 ml of whole blood spiked with SNU349 cells were immobilized by cytospin centrifugation and labelled with antibodies against CD45 and CA9 for detection of leukocytes and ccRCC cells, respectively. Fluorophore signal is merged. Purple, CA9; Green, CD45; Blue, DAPI. RCC, renal cell carcinoma.

Digital PCR detection of ccRCC cells after size-based enrichment. In order to identify ccRCC specific marker genes suitable for assessing the presence of ccRCC cells in our spiked-in samples, we analysed data from the publicly available TCGA (The Cancer Genome Atlas) database and the subsequent sub-type classification by Chen *et al* (26), as defined when pooling the three RCC datasets. Using Limma analyses, we extracted a gene list with the most differentially expressed genes, when comparing ccRCC to pRCC and chRCC (Table SI). *CA9* and *SLC6A3* were amongst the most differentially expressed genes in ccRCC and were selected for further exploration (Fig. 3A). *CA9* is a well-documented hypoxia-driven gene and *SLC6A3* displays a highly ccRCC specific expression pattern, as described by us and others (23,30). We tested the reproducibility of the digital PCR assays of these marker genes on pre-amplified cDNA (Fig. 3B). Increasing the input cDNA yielded higher copies/μl in a linear fashion. Next, we assessed the limit of detection of our assays within the context of enriched tumour cell samples. We performed a titration of tumour cells (SNU-349) spiked into 7.5 ml of whole blood resulting in samples with defined numbers of tumour cells after enrichment. We reliably detected transcripts of *CA9* at the one cell level with more transcripts being detected with higher cell numbers. Similarly, *SLC6A3* transcripts could also be detected from enriched samples containing ≥1 to 121 cells (Fig. 3C). However, transcripts were less reliably detected at the one cell level with *SLC6A3*. Enriched samples from healthy blood controls (4 samples per marker) were also tested, consistently giving readouts of less than 1 copy/μl.

RCC subtype-specific biomarker panel. With the aim of broadening the applicability of our method to include pRCC and chRCC, we identified genes that are differentially and specifically expressed within these subtypes (Table SI). When selecting markers for p1RCC and p2RCC, markers were

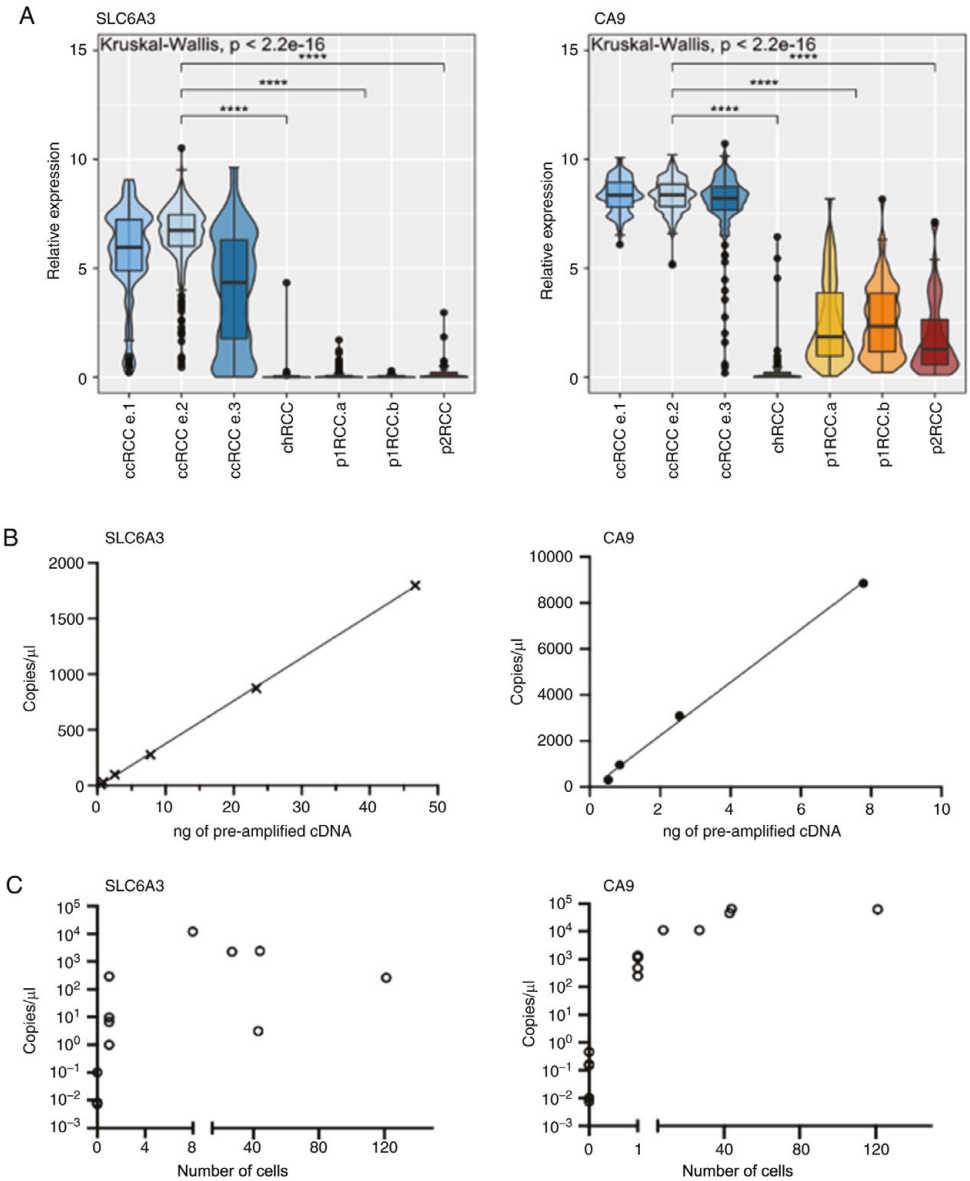

Figure 3. Specificity, reproducibility and detection of ccRCC cells with *SLC6A3* and *CA9*. (A) Specificity of *SLC6A3* and *CA9* expression in RCC tumours within TCGA database. (n=103 ccRCC.e1, 255 ccRCC e.2, 137 ccRCC e.3, 77 chRCC, 134 p1RCCa, 71 p1RCCb, 52 p2RCC). ****P<0.0001. (B) Digital PCR assay efficiency of *SLC6A3* and *CA9* over increasing pre-amplified cDNA inputs with line of best fit. (C) Digital PCR based copies/µl measurement of SNU-349 spiked blood samples after ClearCell FX enrichment and downstream processing. Copies/µl also shown for 4 enriched blood samples without spiked-in SNU-349 cells. ccRCC, clear cell renal cell carcinoma; chRCC, chromophobe RCC; p1RCC, papillary type 1 RCC; p2RCC, papillary type 2 RCC; cDNA, complementary DNA.

chosen so they indicated a papillary subtype and not based on a papillary type 1 or 2 subtype. The selection criteria for a marker included minimal or absent expression in human blood, based on analyses of blood datasets within the Human Protein Atlas (31,32) and high RCC subtype specific expression based on Limma analyses of TCGA expression data, which is based

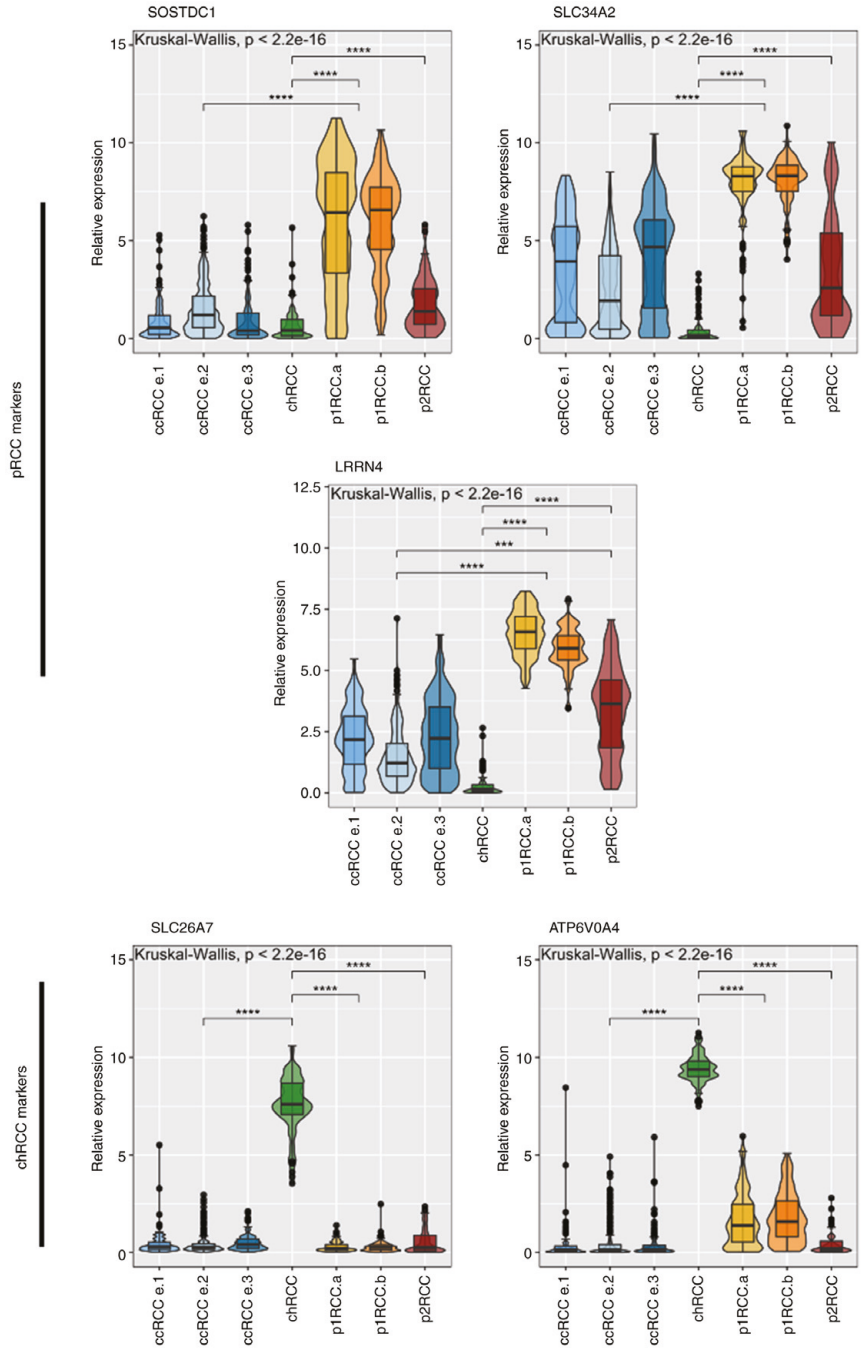

Figure 4. Relative expression of non-ccRCC markers derived from Limma analysis of TCGA expression data. Relevant subtypes are shown according to Chen *et al* (26) taxonomy. ccRCC cohorts and p1RCC cohorts grouped for Dunn's comparison test (n=103 ccRCC.e1, 255 ccRCC e.2, 136 ccRCC e.3, 77 chRCC, 134 p1RCCa, 71 p1RCCb, 52 p2RCC). ***P<0.001, ****P<0.0001. ccRCC, clear cell renal cell carcinoma; chRCC, chromophobe RCC; p1RCC, papillary type 1 RCC; p2RCC, papillary type 2 RCC.

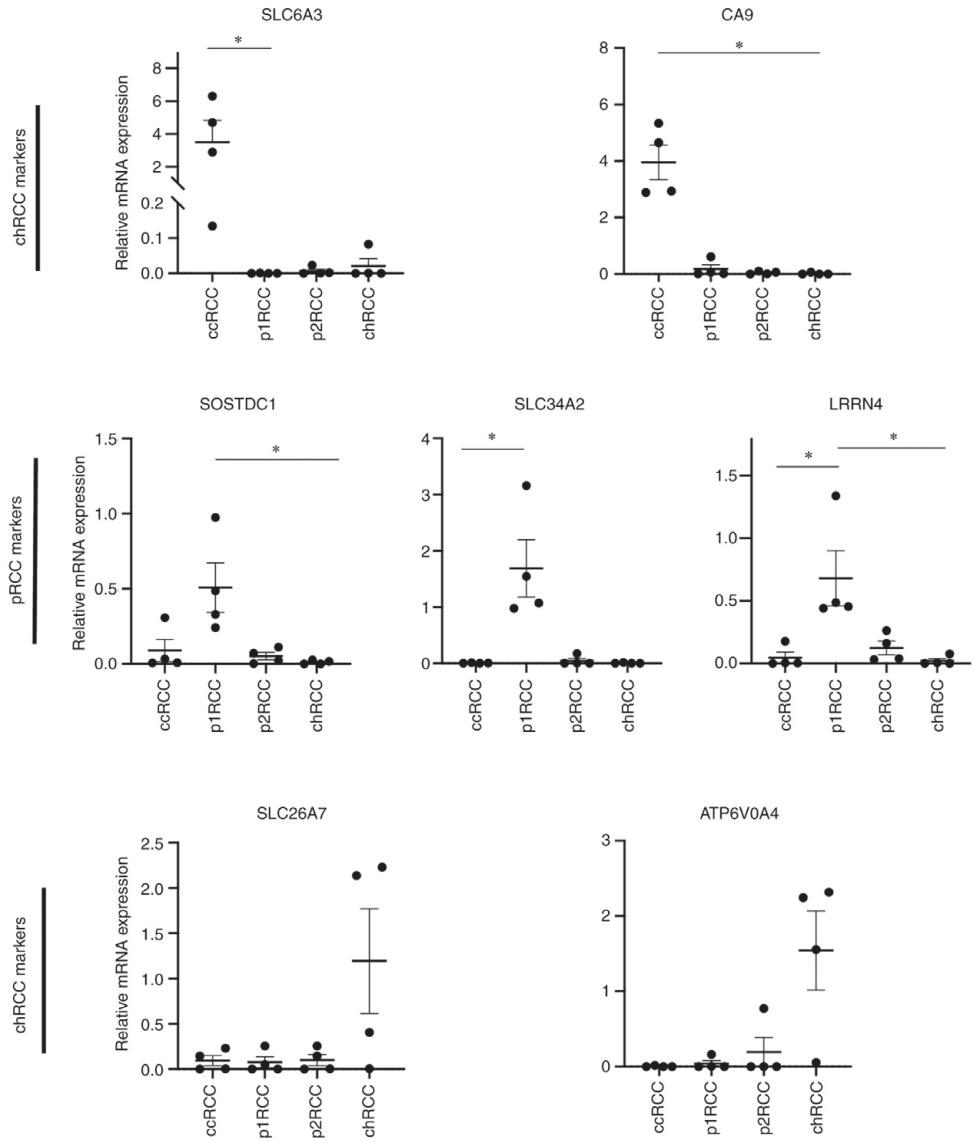

Figure 5. qPCR based mRNA expression relative to beta-actin for all selected markers in patient primary RCC tumour tissue. Horizontal lines represent average mRNA expression and error bars represent SEM (n=4 for each subtype). Data were analysed using Kruskal-Wallis with post-hoc Dunn's test. *P<0.05. Comparisons not reaching significance not indicated. ccRCC, clear cell renal cell carcinoma; chRCC, chromophobe RCC; p1RCC, papillary type 1 RCC; p2RCC, papillary type 2 RCC.

on the Chen *et al* (26) classification of RCC tumours. These selection criteria resulted in a panel of 5 genes in addition to the ccRCC markers *CA9* and *SLC6A3* (Fig. 4).

To validate the sub-type specific expression of the selected markers we analysed RNA extracted from primary tumour samples (Fig. 5). Only RCC subtypes that were confirmed to be one of the three primary subtypes were included in this patient cohort (Table II). Our expression data was in-line with the pattern seen in the TCGA data set. The two markers for the ccRCC subtype (*SLC6A3, CA9*) clearly distinguished this subtype from the rest in the tumour samples tested. Using a joint set of markers to distinguish the papillary RCC subtype

also allowed for clear separation of this subtype from the other two. Finally, the markers for the chRCC subtype distinguished these tumours extremely well, showing elevated RNA expression in a subtype specific manner. We additionally tested all markers on 6 healthy-volunteer blood samples and found them to be negative for our RCC specific markers (Fig. S2).

Discussion

Despite the recent surge in liquid biopsy development and clinical implementation in other cancer types, RCC seems to have largely missed this wave. Potential explanations range from the poor compatibility of RCC cells with EpCAM based isolation methods to the use of cell-free DNA, where RCC tumours have shown to be poor shedders of ctDNA (15,16). Additionally, CTC based liquid biopsy approaches have their inherent pre-analytical challenges (33). The first of these is the sparsity of CTCs in the blood of patients, hence the enrichment procedure requires high recovery efficiency. Secondly, the isolation procedure is often a trade-off between efficiency and specificity; high CTC numbers are desirable but are difficult to obtain with minimal contamination of leucocytes. Finally, sample material is often limited, and methods are required to generate sufficient analytical material for multiple downstream analyses, such as in the case of querying the expression of multiple disease specific transcripts.

Here we develop a method that overcomes the hurdles associated with a CTC based liquid biopsy approach in RCC. The use of the ClearCell FX system successfully isolated spiked-in RCC tumour cells from whole blood with a recovery rate of 66% for primary cells, which is higher than previously reported recovery rates with other isolation systems such as CellSearch®, RosetteSep® or Parsortix® (16,29,34). These observations are in line with the characteristics of RCC tumour cells as our data and others' show that tumour cells including RCC cells are typically larger than 15 μm, suiting them for isolation with a size-based approach that enriches cells larger than 14 μm (17,35).

After enrichment, our data show that we are able to successfully detect ccRCC tumour cells present in the enriched sample via dPCR and immunofluorescence. With one of the ccRCC markers, *CA9*, we were able to consistently and reliably detect down to one ccRCC tumour cell via dPCR in the ClearCell FX enriched sample. The other ccRCC marker *SLC6A3* was marginally less consistent at low cell numbers. This is likely due to the differences in transcript numbers of these two genes present in ccRCC cells, where the absolute transcript levels in ccRCCs are higher for *CA9* than for *SLC6A3*, both in cultured ccRCC cells and within the ccRCC cohort of the TCGA. We still reliably detected *SLC6A3* transcripts in enriched samples that contained 8 or more ccRCC cells, demonstrating the relevance of the described workflow in relation to the reported number of CTCs isolated from ccRCC patients using a size-based approach, that may range from 1 to >100 cells per 10 ml whole blood (17,36). Importantly, the pre-amplification step introduced in our workflow generates an excessive amount of cDNA for multiple analyses.

Finally, we curated a 7-marker gene panel to distinguish the three major RCC subtypes. This is facilitated by the fact that RCC subtypes arise from differing cells of origin and have further defining features based on their subtype specific oncogenetic alterations (37). For example, ccRCC and papillary RCC arise from the proximal segment of the nephron whereas chRCCs arise from the collecting duct. These anatomical distances translate into transcriptomic and functional differences since cells perform distinct functions along segments of the nephron (12,38). As a consequence, chRCC specific markers are unchallenging to define and show a pronounced subtype specific pattern. In stark contrast, ccRCCs and pRCCs arise from the same proximal segment of the nephron, making them far more difficult to molecularly define. However, the virtually universal loss of VHL and the resulting pseudo-hypoxic drive can be leveraged to distinguish ccRCCs from pRCCs, while pRCC specific markers are less precise. This is evident in our primary tumour tissue analysis where overlapping expression of pRCC specific markers with ccRCC markers is observed. Due to this, identification of further pRCC specific markers is warranted. The lack of available cell lines as well as the rarity of these RCC subtypes posed a challenge in validating these markers within this workflow. Ultimately, these markers will require enriched CTCs from pRCC or chRCC patient blood for validation.

Although a few primary ccRCC tumours show raised expression for non-ccRCC specific markers, we predict it would be straightforward to assign them as ccRCC due to the relatively high expression of their respective markers *CA9* and *SLC6A3*. Vice-versa, the expression of the combination of the papillary markers can be leveraged against the low expression of non-papillary markers to designate a papillary subtype. Thus, overlapping expression of subtype specific markers is unlikely to affect the overall sub-type designation of a CTC isolate for these reasons. Regardless, it is warranted and required to explore the use of these markers on CTCs from RCC patient blood.

There were certain limitations to the present study. Firstly, size-based CTC isolation platforms can miss smaller CTCs, as CTC sizes in patients may vary more than observed in this study with cultured and primary cells. Furthermore, the scarce availability of CTCs in patient blood can add to the recovery sufficiency of CTCs.

In conclusion, we established a clear workflow for the isolation and detection of RCC tumour cells from whole blood with obvious implications for use with CTCs in RCC patients. Further work is required to experimentally validate our novel 7-marker gene panel on a larger cohort of patient tumours and to demonstrate its ability in classifying tumour subtype. This should be complemented with a larger cohort of healthy controls as well as RCC patients with other confounding conditions, such as inflammatory or kidney diseases.

Acknowledgements

Not applicable.

Funding

This research was funded by The European Union (grant no. EU-H2020-MSCA-COFUND-2016-754299), The Swedish Cancer Foundation (grant no. CAN2018/1153),

Cancera Stiftelesen and Region Skåne ALF funding (grant no. 2018-176).

Availability of data and materials

For limma analyses, level 3 RNA-seq data of chromophobe renal cell carcinoma (KICH), clear cell kidney carcinoma (KIRC), papillary kidney carcinoma (KIRP) were downloaded from The Cancer Genome Atlas (TCGA) data portal (https://portal.gdc.cancer.gov) as described in (12). The other datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

HA, JH and RDA conceived and designed the study. HA, JH, DL and RDA developed and designed the methods. DL implemented computational software for TCGA analysis. RDA verified the reproducibility of the results. HA, RDA and SS synthesised and analysed data. RDA, AT and CM performed experiments and collected data. BS and PE performed surgery, provided clinical advice and contributed to study design. HN and MJ provided primary patient tissue and established primary cell lines. HA and RDA wrote and prepared the original draft; all authors reviewed and edited the manuscript. HA, DL, SS and RDA prepared figures. HA provided supervision, project administration and funding acquisition. HA and RDA confirm the authenticity of all the raw data. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Regional Healthcare provider (Region Skåne) and Lund University ethical committee (approval nos. LU680-08 and 2018:19). Informed written consent was obtained from all subjects involved in the study.

Patient consent for publication

Written informed consent was obtained from patients for publication of results.

Competing interests

The authors declare that they have no competing interests.

References

1. Mattox AK, Bettegowda C, Zhou S, Papadopoulos N, Kinzler KW and Vogelstein B: Applications of liquid biopsies for cancer. Sci Transl Med 11: eaay1984, 2019.
2. Mader S and Pantel K: Liquid biopsy: Current status and future perspectives. Oncol Res Treat 40: 404-408, 2017.
3. Vanharanta S and Massague J: Origins of metastatic traits. Cancer Cell 24: 410-421, 2013.
4. Rossi E and Zamarchi R: Single-cell analysis of circulating tumor cells: How far have we come in the-omics era? Front Genet 10: 958, 2019.
5. Ljungberg B, Albiges L, Abu-Ghanem Y, Bensalah K, Dabestani S, Fernández-Pello S, Giles RH, Hofmann F, Hora M, Kuczyk MA, *et al*: European association of urology guidelines on renal cell carcinoma: The 2019 update. Eur Urol 75: 799-810, 2019.
6. Hernandez-Yanez M, Heymach JV and Zurita AJ: Circulating biomarkers in advanced renal cell carcinoma: Clinical applications. Curr Oncol Rep 14: 221-229, 2012.
7. Curti BD: Renal cell carcinoma. JAMA 292: 97-100, 2004.
8. Belldegrun AS, Klatte T, Shuch B, LaRochelle JC, Miller DC, Said JW, Riggs SB, Zomorodian N, Kabbinavar FF, Dekernion JB and Pantuck AJ: Cancer-specific survival outcomes among patients treated during the cytokine era of kidney cancer (1989-2005): A benchmark for emerging targeted cancer therapies. Cancer 113: 2457-2463, 2008.
9. Dabestani S, Thorstenson A, Lindblad P, Harmenberg U, Ljungberg B and Lundstam S: Renal cell carcinoma recurrences and metastases in primary non-metastatic patients: A population-based study. World J Urol 34: 1081-1086, 2016.
10. Klatte T, Rossi SH and Stewart GD: Prognostic factors and prognostic models for renal cell carcinoma: A literature review. World J Urol 36: 1943-1952, 2018.
11. Finley DS, Pantuck AJ and Belldegrun AS: Tumor biology and prognostic factors in renal cell carcinoma. Oncologist 16 (Suppl 2): S4-S13, 2011.
12. Lindgren D, Eriksson P, Krawczyk K, Nilsson H, Hansson J, Veerla S, Sjölund J, Höglund M, Johansson ME and Axelson H: Cell-type-specific gene programs of the normal human nephron define kidney cancer subtypes. Cell Rep 20: 1476-1489, 2017.
13. Cohen HT and McGovern FJ: Renal-cell carcinoma. N Engl J Med 353: 2477-2490, 2005.
14. Rathmell WK and Chen S: VHL inactivation in renal cell carcinoma: Implications for diagnosis, prognosis and treatment. Expert Rev Anticancer Ther 8: 63-73, 2008.
15. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM, *et al*: Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 6: 224ra24, 2014.
16. Maertens Y, Humberg V, Erlmeier F, Steffens S, Steinestel J, Bögemann M, Schrader AJ and Bernemann C: Comparison of isolation platforms for detection of circulating renal cell carcinoma cells. Oncotarget 8: 87710-87717, 2017.
17. Klezl P, Pospisilova E, Kolostova K, Sonsky J, Maly O, Grill R, Pawlak I and Bobek V: Detection of circulating tumor cells in renal cell carcinoma: Disease Stage Correlation and Molecular Characterization. J Clin Med 9: 1372, 2020.
18. Liu S, Tian Z, Zhang L, Hou S, Hu S, Wu J, Jing Y, Sun H, Yu F, Zhao L, *et al*: Combined cell surface carbonic anhydrase 9 and CD147 antigens enable high-efficiency capture of circulating tumor cells in clear cell renal cell carcinoma patients. Oncotarget 7: 59877-59891, 2016.
19. Chudasama D, Katopodis P, Stone N, Haskell J, Sheridan H, Gardner B, Urnovitz H, Schuetz E, Beck J, Hall M, *et al*: Liquid biopsies in lung cancer: Four emerging technologies and potential clinical applications. Cancers (Basel) 11: 331, 2019.
20. Rizzo MI, Ralli M, Nicolazzo C, Gradilone A, Carletti R, Di Gioia C, De Vincentiis M and Greco A: Detection of circulating tumor cells in patients with laryngeal cancer using ScreenCell: Comparative pre- and post-operative analysis and association with prognosis. Oncol Lett 19: 4183-4188, 2020.
21. Lee Y, Guan G and Bhagat AA: ClearCell(R) FX, a label-free microfluidics technology for enrichment of viable circulating tumor cells. Cytometry A 93: 1251-1254, 2018.
22. Xu L, Mao X, Imrali A, Syed F, Mutsvangwa K, Berney D, Cathcart P, Hines J, Shamash J and Lu YJ: Optimization and evaluation of a novel size based circulating tumor cell isolation system. PLoS One 10: e0138032, 2015.
23. Hansson J, Lindgren D, Nilsson H, Johansson E, Johansson M, Gustavsson L and Axelson H: Overexpression of functional SLC6A3 in clear cell renal cell carcinoma. Clin Cancer Res 23: 2105-2115, 2017.
24. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK: Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43: e47, 2015.
25. Nacer DF, Liljedahl H, Karlsson A, Lindgren D and Staaf J: Pan-cancer application of a lung-adenocarcinoma-derived gene-expression-based prognostic predictor. Brief Bioinform 22: bbab154, 2021.

26. Chen F, Zhang Y, Senbabaoglu Y, Ciriello G, Yang L, Reznik E, Shuch B, Micevic G, De Velasco G, Shinbrot E, *et al*: Multilevel Genomics-Based taxonomy of renal cell carcinoma. Cell Rep 14: 2476-2489, 2016.
27. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408, 2001.
28. Kroneis T, Jonasson E, Andersson D, Dolatabadi S and Stahlberg A: Global preamplification simplifies targeted mRNA quantification. Sci Rep 7: 45219, 2017.
29. Cappelletti V, Verzoni E, Ratta R, Vismara M, Silvestri M, Montone R, Miodini P, Reduzzi C, Claps M, Sepe P, *et al*: Analysis of single circulating tumor cells in renal cell carcinoma reveals phenotypic heterogeneity and genomic alterations related to progression. Int J Mol Sci 21: 1475, 2020.
30. Pastorekova S and Gillies RJ: The role of carbonic anhydrase IX in cancer development: Links to hypoxia, acidosis, and beyond. Cancer Metastasis Rev 38: 65-77, 2019.
31. Thul PJ, Akesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, Alm T, Asplund A, Björk L, Breckels LM, *et al*: A subcellular map of the human proteome. Science 356: eaal3321, 2017.
32. Chen J, Cheung F, Shi R, Zhou H and Lu W; CHI Consortium: PBMC fixation and processing for Chromium single-cell RNA sequencing. J Transl Med 16: 198, 2018.
33. Neumann MHD, Bender S, Krahn T and Schlange T: ctDNA and CTCs in Liquid Biopsy-current status and where we need to progress. Comput Struct Biotechnol J 16: 190-195, 2018.
34. Gradilone A, Iacovelli R, Cortesi E, Raimondi C, Gianni W, Nicolazzo C, Petracca A, Palazzo A, Longo F, Frati L and Gazzaniga P: Circulating tumor cells and 'suspicious objects' evaluated through CellSearch® in metastatic renal cell carcinoma. Anticancer Res 31: 4219-4221, 2011.
35. Vona G, Sabile A, Louha M, Sitruk V, Romana S, Schütze K, Capron F, Franco D, Pazzagli M, Vekemans M, *et al*: Isolation by size of epithelial tumor cells: A new method for the immunomorphological and molecular characterization of circulatingtumor cells. Am J Pathol 156: 57-63, 2000.
36. El-Heliebi A, Kroneis T, Zohrer E, Haybaeck J, Fischereder K, Kampel-Kettner K, Zigeuner R, Pock H, Riedl R, Stauber R, *et al*: Are morphological criteria sufficient for the identification of circulating tumor cells in renal cancer? J Transl Med 11: 214, 2013.
37. Lindgren D, Sjolund J and Axelson H: Tracing renal cell carcinomas back to the nephron. Trends Cancer 4: 472-484, 2018.
38. Knepper M and Burg M: Organization of nephron function. Am J Physiol 244: F579-F589, 1983.
39. Moch H, Cubilla AL, Humphrey PA, Reuter VE and 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 70: 93-105, 2016.

Appendix S1

Dissociation of human renal tissue. The cortical tissue farthest from the tumor was selected and put in ice-cold Dulbecco's modified Eagle's medium (Invitrogen, Carlsbad, CA) supplemented with 10% fetal calf serum and 1% penicillin/streptomycin. Samples were rinsed, minced, and subjected to overnight collagenase treatment at 37°C in a processing medium consisting of Ham's F-12/Dulbecco's modified Eagle's medium [1:1 (v/v); Invitrogen], supplemented with 5% fetal calf serum, 1% penicillin/streptomycin, collagenase IV at a final concentration of 300 U/ml (Invitrogen), and deoxyribonuclease I type II at a final concentration of 200 U/ml (Sigma-Aldrich, St. Louis, MO). After trituration by slow repeated pipetting through a 10-ml pipette, the resulting tissue suspension was serially passed through tissue strainers with mesh sizes of 100 and 70 μm, respectively, thereby excluding glomeruli from the preparation. The suspension was treated with 1X trypsin–EDTA for 5 minutes and then was passed through a 20-μm strainer, which resulted in single cells.

Figure S1. Optimisation of downstream processing of ClearCell FX output. (A) Representative image of Total Eukaryote RNA Pico Chip electropherogram and RIN values for RNA extracted from ClearCell FX output containing 27 and 8 SNU-349 cells amongst background of leucocytes. (B) qPCR monitoring of pre-amplification cycles using SYBR green for 500pg and 50pg total RNA input from SNU-349 cells. Vertical red line represents selected number of cycles for pre-amplification of future pre-amplification reactions. (C) cDNA fragment analysis of pre-amplified cDNA compared to non-pre-amplified cDNA from SNU-349 cells. nt, nucleotides; rRNA, ribosomal RNA.

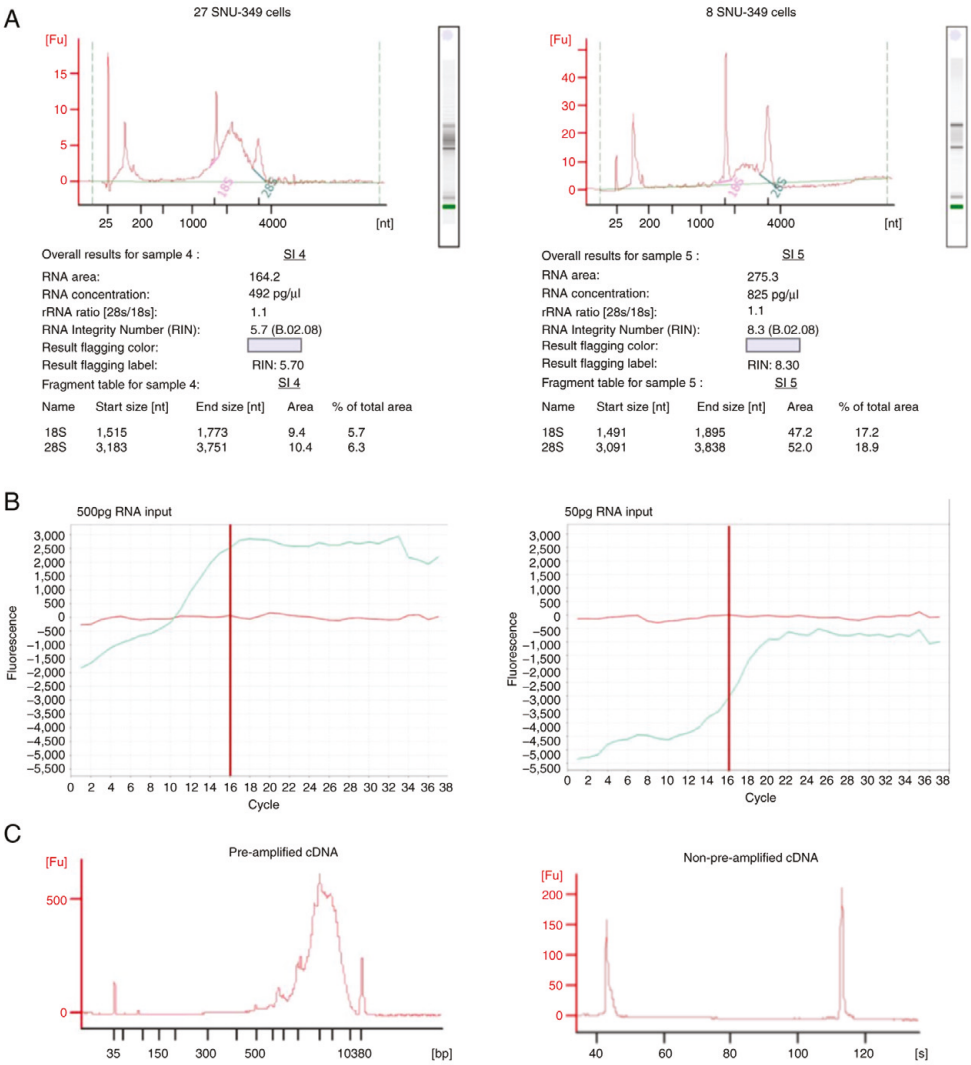

Figure S2. Expression of all subtype specific markers in 6 healthy blood samples enriched for CTCs.

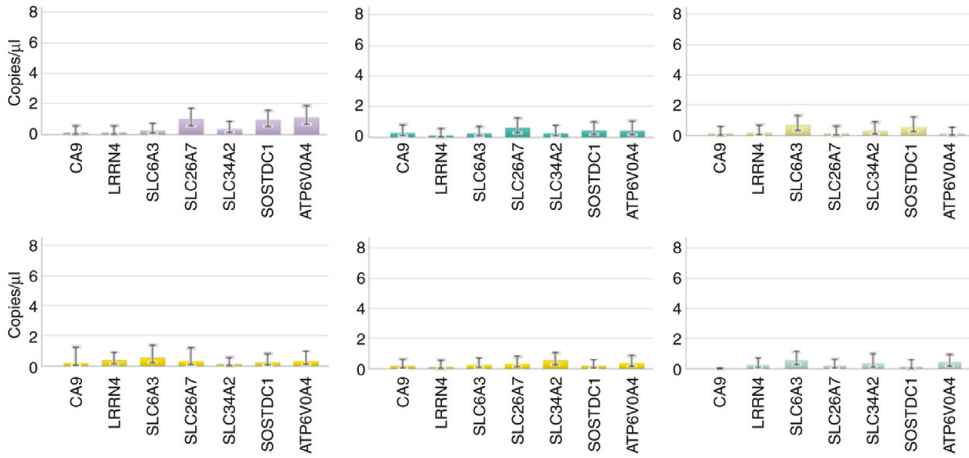

Table SI A. Limma analyses for differentially expressed transcripts.

ENSG	ENSGv	gene_ version	SYMBOL	gene_biotype	chr	start	end	strand	ENTREZID	ccRCC… chRCC	ccRCC… pRCC_e1	ccRCC… pRCC_e2	F	P.Value	adj.P.Val
ENSG00000102755	ENSG00000102755.9	11	FLT1	protein_coding	13	28300344	28495145	-	2321	2,248	4,635	3,537	1442,73687	0	0
ENSG00000107159	ENSG00000107159.11	12	CA9	protein_coding	9	35673856	35681159	+	768	7,916	5,785	6,422	1711,63179	0	0
ENSG00000112715	ENSG00000112715.19	21	VEGFA	protein_coding	6	43770184	43786487	+	7422	2,458	5,108	3,765	1787,64862	0	0
ENSG00000113739	ENSG00000113739.9	10	STC2	protein_coding	5	173314713	173329503	-	8614	3,295	4,803	3,300	1469,88963	0	0
ENSG00000129521	ENSG00000129521.12	13	EGLN3	protein_coding	14	33924231	34462774	-	112399	5,084	3,628	4,351	1513,10883	0	0
ENSG00000185633	ENSG00000185633.9	10	NDUFA4L2	protein_coding	12	57234903	57240715	-	56901	5,749	8,005	6,851	3214,68276	0	0
ENSG00000167772	ENSG00000167772.10	11	ANGPTL4	protein_coding	19	8363289	8374373	+	51129	7,043	5,856	5,332	1397,9088	3.95005e-320	1.31730993920891e-317
ENSG00000154269	ENSG00000154269.13	14	ENPP3	protein_coding	6	131628442	131747418	+	5169	7,017	3,851	3,899	1291,03896	0,00E+00	5,19E-306
ENSG00000171208	ENSG00000171208.8	9	NETO2	protein_coding	16	47077703	47143997	-	81831	2,035	3,697	2,765	1251,09834	6,81E-304	1,86E-301
ENSG00000117266	ENSG00000117266.14	15	CDK18	protein_coding	1	205504595	205532793	+	5129	3,785	2,598	2,282	1236,30429	3,59E-302	9,43E-300
ENSG00000184661	ENSG00000184661.12	13	CDCA2	protein_coding	8	25458997	25507920	+	157313	2,993	3,315	2,835	1233,44823	7,76E-302	2,01E-299
ENSG00000164283	ENSG00000164283.11	12	ESM1	protein_coding	5	54977864	55022671	-	11082	2,479	5,329	3,415	1172,40349	1,56E-294	3,80E-292
ENSG00000091879	ENSG00000091879.12	13	ANGPT2	protein_coding	8	6499651	6563409	-	285	2,315	4,351	3,440	1145,61081	3,15E-291	7,28E-289
ENSG00000140945	ENSG00000140945.14	16	CDH13	protein_coding	16	82626803	83800640	+	1012	2,108	3,631	2,899	1065,14028	6,45E-281	1,26E-278
ENSG00000261371	ENSG00000261371.4	5	PECAM1	protein_coding	17	64319415	64413776	-	5175	2,035	3,596	2,744	1053,24473	2,44E-279	4,61E-277
ENSG00000143248	ENSG00000143248.11	12	RGS5	protein_coding	1	163111121	163321791	-	8490	2,417	5,347	3,974	1012,76183	7,26E-274	1,29E-271
ENSG00000159167	ENSG00000159167.10	11	STC1	protein_coding	8	23841915	23854807	-	6781	2,797	4,514	4,043	1006,3332	5,58E-273	9,62E-271
ENSG00000135245	ENSG00000135245.9	9	HILPDA	protein_coding	7	128455849	128458418	+	29923	4,572	3,872	4,166	994,419478	2,50E-271	4,25E-269
ENSG00000186918	ENSG00000186918.12	13	ZNF395	protein_coding	8	28345585	28402701	-	55893	2,413	2,391	2,614	983,572046	8,25E-270	1,38E-267
ENSG00000115112	ENSG00000115112.7	7	TFCP2L1	protein_coding	2	121216587	121285207	-	29842	-6,329	-2,416	-2,175	976,629358	7,86E-269	1,26E-266
ENSG00000187730	ENSG00000187730.7	8	GABRD	protein_coding	1	2019298	2030758	+	2563	2,217	3,642	3,282	974,545007	1,55E-268	2,42E-266
ENSG00000162654	ENSG00000162654.8	8	GBP4	protein_coding	1	89181148	89198932	-	115361	2,476	3,536	2,762	952,004175	2,57E-265	3,83E-263
ENSG00000148346	ENSG00000148346.10	11	LCN2	protein_coding	9	128149071	128153455	+	3934	-3,316	-6,751	-3,500	939,283317	1,79E-263	2,60E-261
ENSG00000164342	ENSG00000164342.11	12	TLR3	protein_coding	4	186069152	186088069	+	7098	3,194	2,132	2,359	930,298252	3,70E-262	5,16E-260
ENSG00000134716	ENSG00000134716.8	9	CYP2J2	protein_coding	1	59893308	59926790	-	1573	6,110	5,038	4,721	916,629308	3,85E-260	5,19E-258
ENSG00000161267	ENSG00000161267.10	11	BDH1	protein_coding	3	197509783	197573323	-	622	-3,405	-2,173	-2,396	873,922671	1,10E-253	1,32E-251
ENSG00000134954	ENSG00000134954.13	14	ETS1	protein_coding	11	128458761	128587558	-	2113	2,402	2,727	2,424	860,976414	1,12E-251	1,26E-249
ENSG00000082684	ENSG00000082684.13	14	SEMA5B	protein_coding	3	122909193	123028605	-	54437	4,588	2,998	3,685	858,601472	2,62E-251	2,91E-249
ENSG00000140479	ENSG00000140479.15	16	PCSK6	protein_coding	15	101297142	101525202	-	5046	3,286	3,652	3,255	841,567125	1,25E-248	1,32E-246
ENSG00000050767	ENSG00000050767.14	15	COL23A1	protein_coding	5	178237618	178590555	-	91522	6,427	2,441	2,951	827,598689	2,11E-246	2,12E-244
ENSG00000133574	ENSG00000133574.8	9	GIMAP4	protein_coding	7	150567277	150573955	+	55303	2,106	2,826	2,215	827,435451	2,25E-246	2,24E-244
ENSG00000086205	ENSG00000086205.15	16	FOLH1	protein_coding	11	49146635	49208670	-	2346|219595	2,325	3,404	2,450	812,224062	6,46E-244	6,11E-242
ENSG00000185585	ENSG00000185585.18	19	OLFML2A	protein_coding	9	124777158	124814885	+	169611	2,252	3,768	3,555	807,780958	3,43E-243	3,17E-241
ENSG00000181577	ENSG00000181577.14	15	C6orf223	protein_coding	6	44000580	44005958	+	221416	2,240	4,478	4,088	800,365073	5,64E-242	5,07E-240
ENSG00000112214	ENSG00000112214.9	10	FHL5	protein_coding	6	96562548	96616636	+	9457	2,384	3,396	2,793	792,942013	9,48E-241	8,41E-239
ENSG00000142319	ENSG00000142319.17	17	SLC6A3	protein_coding	5	1392790	1445430	-	6531	5,700	5,705	5,558	766,163239	2,96E-236	2,38E-234
ENSG00000171992	ENSG00000171992.11	12	SYNPO	protein_coding	5	150601080	150659220	+	11346	2,189	3,402	2,367	755,029001	2,37E-234	1,81E-232
ENSG00000113532	ENSG00000113532.11	12	ST8SIA4	protein_coding	5	100806935	100903266	-	7903	2,659	2,720	2,356	716,802269	1,17E-227	8,04E-226
ENSG00000181104	ENSG00000181104.6	6	F2R	protein_coding	5	76716043	76735781	+	2149	2,297	3,042	2,274	714,820592	2,65E-227	1,79E-225
ENSG00000161638	ENSG00000161638.9	10	ITGA5	protein_coding	12	54395261	54419460	-	3678	2,115	2,653	2,244	703,280507	3,16E-225	2,04E-223
ENSG00000176387	ENSG00000176387.6	6	HSD11B2	protein_coding	16	67430652	67437553	+	3291	-5,531	2,379	2,089	701,674087	6,17E-225	3,95E-223
ENSG00000175538	ENSG00000175538.9	10	KCNE3	protein_coding	11	74454841	74467729	-	10008	2,617	2,472	2,196	699,110124	1,80E-224	1,14E-222
ENSG00000163083	ENSG00000163083.5	5	INHBB	protein_coding	2	120346143	120351808	+	3625	2,280	3,736	3,120	668,917812	6,76E-219	3,90E-217
ENSG00000135218	ENSG00000135218.16	17	CD36	protein_coding	7	80369575	80679277	+	948	2,001	3,835	2,927	665,197389	3,38E-218	1,94E-216

ENSG00000100024	ENSG00000100024.13	14	UPB1	protein_coding	22	24494107	24528390	+	51733	3,202	2,891	2,218	661,097834	2,01E-217	1,11E-215
ENSG00000144476	ENSG00000144476.5	5	ACKR3	protein_coding	2	236567787	236582358	+	57007	2,577	2,913	2,822	645,324478	2,05E-214	1,07E-212
ENSG00000163909	ENSG00000163909.7	7	HEYL	protein_coding	1	39624153	39639945	-	26508	2,031	3,338	2,432	635,555459	1,59E-212	8,04E-211
ENSG00000117228	ENSG00000117228.9	9	GBP1	protein_coding	1	89052319	89065360	-	2633	3,157	2,551	2,034	634,057046	3,12E-212	1,56E-210
ENSG00000088899	ENSG00000088899.13	14	LZTS3	protein_coding	20	3162617	3173592	-	9762	-2,369	-2,275	-2,058	630,583239	1,49E-211	7,32E-210
ENSG00000183722	ENSG00000183722.7	7	LHFP	protein_coding	13	39342892	39603528	-	10186	2,101	2,917	2,047	605,816655	1,21E-206	5,40E-205

Table description: Multi-group limma analysis between the main RCC Chen taxonomy entities. Clear cell samples within the CC-e1, e-2 and e-3 were merged into one group and the P.e-1a and P.e-1b subgroups were merged into one group. The between group mean log2 expression level difference is supplied for each individual gene together with adjusted p-values and Limma F statistics.

Table SI B.

ENSG	ENSGv	gene_ version	SYMBOL	gene_biotype	chr	start	end	strand	ENTREZID	pRCC_e1… ccRCC	pRCC_e1… chRCC	pRCC_e1… pRCC_e2	F	P.Value	adj.P.Val
ENSG00000148346	ENSG00000148346.10	11	LCN2	protein_coding	9	128149071	128153455	+	3934	6,751	3,435	3,250	939,2833173	1,79E-263	2,60E-261
ENSG00000125872	ENSG00000125872.7	7	LRRN4	protein_coding	20	6040778	6054049	-	164312	4,451	5,991	2,978	780,498526	1,13E-238	9,50E-237
ENSG00000167779	ENSG00000167779.6	7	IGFBP6	protein_coding	12	53097436	53102345	+	3489	3,432	5,986	2,351	646,1023263	1,45E-214	7,67E-213
ENSG00000136155	ENSG00000136155.15	16	SCEL	protein_coding	13	77535674	77645263	+	8796	3,711	3,792	2,470	570,8964251	1,77E-199	6,83E-198
ENSG00000165215	ENSG00000165215.6	6	CLDN3	protein_coding	7	73768997	73770270	-	1365	3,931	5,245	2,180	522,8151273	4,13E-189	1,30E-187
ENSG00000132821	ENSG00000132821.10	11	VSTM2L	protein_coding	20	37903104	37945350	+	128434	4,312	4,036	3,661	439,0380866	1,90E-169	4,28E-168
ENSG00000204385	ENSG00000204385.9	10	SLC44A4	protein_coding	6	31863192	31879046	-	80736	3,970	5,469	4,424	407,4749017	1,97E-161	3,78E-160
ENSG00000144214	ENSG00000144214.8	9	LYG1	protein_coding	2	99284238	99304742	-	129530	2,118	6,262	2,608	406,3530595	3,85E-161	7,37E-160
ENSG00000158747	ENSG00000158747.12	13	NBL1	protein_coding	1	19640554	19658456	+	4681	2,553	4,915	2,696	383,8087932	3,56E-155	6,13E-154
ENSG00000058085	ENSG00000058085.13	14	LAMC2	protein_coding	1	183186238	183244900	+	3918	3,172	3,801	2,417	373,3938501	2,39E-152	3,88E-151
ENSG00000171243	ENSG00000171243.7	7	SOSTDC1	protein_coding	7	16461481	16530580	-	25928	4,673	5,175	4,121	348,82601	1,70E-145	2,48E-144
ENSG00000131037	ENSG00000131037.13	14	EPS8L1	protein_coding	19	55072020	55087923	+	54869	2,899	2,031	2,805	327,4980028	2,52E-139	3,31E-138
ENSG00000168077	ENSG00000168077.12	13	SCARA3	protein_coding	8	27633868	27676776	+	51435	2,595	4,041	2,488	320,9251115	2,23E-137	2,81E-136
ENSG00000168703	ENSG00000168703.5	5	WFDC12	protein_coding	20	45123425	45124465	-	128488	2,802	2,988	2,420	317,6963684	2,05E-136	2,53E-135
ENSG00000175315	ENSG00000175315.2	2	CST6	protein_coding	11	66011841	66013505	+	1474	2,727	2,269	2,589	310,8809271	2,31E-134	2,73E-133
ENSG00000114638	ENSG00000114638.6	7	UPK1B	protein_coding	3	119173517	119205153	+	7348	3,556	4,732	3,481	307,7385288	2,08E-133	2,40E-132
ENSG00000157765	ENSG00000157765.10	11	SLC34A2	protein_coding	4	25655301	25678748	+	10568	4,710	7,499	4,463	306,6733524	4,40E-133	5,02E-132
ENSG00000189056	ENSG00000189056.12	13	RELN	protein_coding	7	103471784	103989516	-	5649	2,717	2,658	2,160	292,3053853	1,21E-128	1,27E-127
ENSG00000164825	ENSG00000164825.3	3	DEFB1	protein_coding	8	6870575	6878022	-	1672	2,797	-2,832	2,616	282,8065835	1,20E-125	1,18E-124
ENSG00000152527	ENSG00000152527.12	13	PLEKHH2	protein_coding	2	43637273	43767987	+	130271	2,246	3,392	2,734	263,6353006	1,94E-119	1,71E-118
ENSG00000176842	ENSG00000176842.13	14	IRX5	protein_coding	16	54930862	54934485	+	10265	2,099	4,094	2,373	257,7960204	1,67E-117	1,43E-116
ENSG00000124875	ENSG00000124875.8	9	CXCL6	protein_coding	4	73836497	73849064	+	6372	3,290	4,722	3,174	248,5496162	2,15E-114	1,74E-113
ENSG00000105825	ENSG00000105825.10	11	TFPI2	protein_coding	7	93885397	93890991	-	7980	3,660	5,260	2,203	245,2688732	2,80E-113	2,23E-112
ENSG00000184292	ENSG00000184292.6	6	TACSTD2	protein_coding	1	58575423	58577773	-	4070	3,750	3,165	4,326	230,8502269	2,75E-108	2,00E-107
ENSG00000088992	ENSG00000088992.16	17	TESC	protein_coding	12	117038923	117099479	-	54997	2,387	3,404	2,065	207,9420212	4,68E-100	2,93E-99
ENSG00000175121	ENSG00000175121.10	11	WFDC5	protein_coding	20	45109452	45115172	-	149708	2,592	3,030	2,409	189,2573417	4,80E-93	2,67E-92
ENSG00000124107	ENSG00000124107.5	5	SLPI	protein_coding	20	45252239	45254564	-	6590	4,941	5,307	4,185	180,1992345	1,52E-89	7,95E-89
ENSG00000124134	ENSG00000124134.7	8	KCNS1	protein_coding	20	45092310	45101112	-	3787	2,042	2,477	2,235	174,3610578	2,99E-87	1,50E-86
ENSG00000007062	ENSG00000007062.10	11	PROM1	protein_coding	4	15963076	16084378	-	8842	3,093	4,400	3,021	157,6359558	1,65E-80	7,32E-80
ENSG00000171345	ENSG00000171345.12	13	KRT19	protein_coding	17	41523617	41528308	-	3880	3,125	5,998	4,881	152,2045331	2,90E-78	1,23E-77
ENSG00000175793	ENSG00000175793.11	11	SFN	protein_coding	1	26863138	26864457	+	2810	2,754	3,787	2,373	149,6142797	3,50E-77	1,45E-76
ENSG00000138207	ENSG00000138207.11	13	RBP4	protein_coding	10	93591687	93601744	-	5950	3,628	7,331	4,207	149,0665467	5,93E-77	2,45E-76
ENSG00000142102	ENSG00000142102.14	15	PGGHG	protein_coding	11	289135	296107	+	80162	2,054	4,281	3,407	139,7478584	5,28E-73	2,03E-72
ENSG00000159212	ENSG00000159212.11	12	CLIC6	protein_coding	21	34669389	34718227	+	54102	2,729	2,944	3,516	114,84785	5,37E-62	1,71E-61
ENSG00000102854	ENSG00000102854.13	15	MSLN	protein_coding	16	760762	768865	+	10232	2,715	3,938	3,194	108,537253	4,29E-59	1,29E-58
ENSG00000001626	ENSG00000001626.13	14	CFTR	protein_coding	7	117465784	117715971	+	1080	0,195	-6,247	0,271	3215,350829	0	0
ENSG00000004939	ENSG00000004939.12	13	SLC4A1	protein_coding	17	44248385	44268141	-	6521	-0,366	-8,105	-0,272	1720,627111	0	0
ENSG00000035720	ENSG00000035720.6	7	STAP1	protein_coding	4	67558728	67607337	+	26228	-0,195	-7,623	0,079	3349,998773	0	0
ENSG00000065320	ENSG00000065320.7	8	NTN1	protein_coding	17	9021542	9244000	+	9423	-0,273	-5,332	-0,037	2890,289665	0	0
ENSG00000065882	ENSG00000065882.14	15	TBC1D1	protein_coding	4	37891087	38139175	+	23216	-0,117	-3,149	-0,192	1663,055542	0	0
ENSG00000072954	ENSG00000072954.5	6	TMEM38A	protein_coding	19	16661127	16690029	+	79041	1,224	-4,248	-0,131	1837,197478	0	0
ENSG00000099822	ENSG00000099822.2	2	HCN2	protein_coding	19	589893	617159	+	610	0,363	-5,241	-0,498	1461,806946	0	0
ENSG00000100362	ENSG00000100362.11	12	PVALB	protein_coding	22	36800684	36819479	-	5816	-0,042	-9,959	-1,984	1716,68979	0	0
ENSG00000102678	ENSG00000102678.6	6	FGF9	protein_coding	13	21671383	21704498	+	2254	0,623	-5,082	0,469	2367,433302	0	0
ENSG00000102755	ENSG00000102755.9	11	FLT1	protein_coding	13	28300344	28495145	-	2321	-4,635	-2,388	-1,098	1442,736875	0	0
ENSG00000105929	ENSG00000105929.14	15	ATP6V0A4	protein_coding	7	138706295	138799560	-	50617	1,327	-7,672	1,239	2240,767691	0	0
ENSG00000107159	ENSG00000107159.11	12	CA9	protein_coding	9	35673856	35681159	+	768	-5,785	2,131	0,636	1711,63179	0	0
ENSG00000107954	ENSG00000107954.9	10	NEURL1	protein_coding	10	103493979	103592552	+	9148	-0,277	-3,803	-0,300	1454,391136	0	0
ENSG00000109684	ENSG00000109684.13	14	CLNK	protein_coding	4	10486395	10684865	-	116449	-0,199	-6,400	0,021	5678,837823	0	0
ENSG00000112530	ENSG00000112530.10	11	PACRG	protein_coding	6	162727132	163315492	+	135138	0,962	-4,348	0,510	1683,738541	0	0

Table SI C.

ENSG	ENSGv	gene_ version	SYMBOL	gene_biotype	chr	start	end	strand	ENTREZID	chRCC... ccRCC	chRCC... pRCC_e1	chRCC... pRCC_e2	AveExpr	F	P.Value	adj.P.Val
ENSG00000167749	ENSG00000167749.10	11	KLK4	protein_coding	19	50906352	50910738	-	9622	2,641	-3,237	2,308	2,94E-17	646,080701	1,47E-214	7,73E-213
ENSG00000243955	ENSG00000243955.4	5	GSTA1	protein_coding	6	52791664	52803910	-	2938	-3,495	3,492	2,326	-2,95E-16	559,206403	5,17E-197	1,90E-195
ENSG00000177508	ENSG00000177508.11	11	IRX3	protein_coding	16	54283304	54286763	-	79191	-2,050	2,278	-2,701	-3,69E-16	521,722426	7,22E-189	2,25E-187
ENSG00000148795	ENSG00000148795.5	6	CYP17A1	protein_coding	10	102830531	102837533	-	1586	3,880	2,590	3,051	-1,74E-17	204,38054	9,67E-99	5,91E-98
ENSG00000158022	ENSG00000158022.6	6	TRIM63	protein_coding	1	26051304	26068436	-	84676	3,379	3,528	2,758	2,62E-17	177,564644	1,63E-88	8,37E-88
ENSG00000004846	ENSG00000004846.15	16	ABCB5	protein_coding	7	20615207	20777038	+	340273	2,197	2,268	2,189	8,17E-18	148,52615	9,99E-77	4,11E-76
ENSG00000002726	ENSG00000002726.18	20	AOC1	protein_coding	7	150824627	150861504	+	26	-3,379	2,274	-2,668	8,83E-17	148,439828	1,09E-76	4,47E-76
ENSG00000104722	ENSG00000104722.12	13	NEFM	protein_coding	8	24913012	24919098	+	4741	3,228	3,763	2,070	6,16E-17	145,645463	1,63E-75	6,57E-75
ENSG00000186198	ENSG00000186198.3	3	SLC51B	protein_coding	15	65045370	65053396	+	123264	3,238	2,762	2,627	-1,25E-16	129,509664	1,47E-68	5,22E-68
ENSG00000103154	ENSG00000103154.8	9	NECAB2	protein_coding	16	83968632	84002776	+	54550	2,298	3,516	2,623	3,00E-17	126,156718	4,43E-67	1,54E-66
ENSG00000211452	ENSG00000211452.9	10	DIO1	protein_coding	1	53891239	53911086	+	1733	2,907	3,430	2,120	-3,65E-17	118,499208	1,18E-63	3,87E-63
ENSG00000157315	ENSG00000157315.4	4	TMED6	protein_coding	16	69343248	69351809	-	146456	2,449	3,740	2,619	-2,78E-17	114,206351	1,05E-61	3,33E-61
ENSG00000198074	ENSG00000198074.8	9	AKR1B10	protein_coding	7	134527592	134541408	+	57016	5,372	6,158	3,662	1,28E-16	114,190583	1,07E-61	3,39E-61
ENSG00000102575	ENSG00000102575.9	10	ACP5	protein_coding	19	11574660	11579008	-	54	2,637	4,491	2,224	-3,72E-16	105,741541	8,57E-58	2,52E-57
ENSG00000171759	ENSG00000171759.7	9	PAH	protein_coding	12	102836885	102958410	-	5053	3,075	2,811	4,187	5,18E-17	80,272681	1,75E-45	4,16E-45
ENSG00000197901	ENSG00000197901.10	11	SLC22A6	protein_coding	11	62936385	62984983	-	9356	2,691	5,635	4,235	-2,37E-16	78,4004721	1,52E-44	3,56E-44
ENSG00000103485	ENSG00000103485.16	17	QPRT	protein_coding	16	29679008	29698699	+	23475|105	2,929	3,801	3,129	2,24E-16	76,5448377	1,31E-43	3,02E-43
ENSG00000158296	ENSG00000158296.12	13	SLC13A3	protein_coding	20	46557823	46684467	-	64849	3,342	3,210	2,891	2,62E-17	69,309587	6,51E-40	1,41E-39
ENSG00000185499	ENSG00000185499.15	16	MUC1	protein_coding	1	155185824	155192916	-	4582	-2,150	-3,701	-2,717	4,26E-16	68,3854202	1,96E-39	4,21E-39
ENSG00000148702	ENSG00000148702.13	14	HABP2	protein_coding	10	113550837	113589602	+	3026	2,817	5,075	2,976	2,19E-16	55,294478	1,58E-32	3,01E-32
ENSG00000001626	ENSG00000001626.13	14	CFTR	protein_coding	7	117465784	117715971	+	1080	-0,076	-6,518	-0,271	-6,16E-17	3215,35083	0	0
ENSG00000004939	ENSG00000004939.12	13	SLC4A1	protein_coding	17	44248385	44268141	-	6521	-0,094	-7,834	0,272	-4,71E-17	1720,62711	0	0
ENSG00000035720	ENSG00000035720.6	7	STAP1	protein_coding	4	67558728	67607337	+	26228	-0,274	-7,702	-0,079	1,15E-16	3349,99877	0	0
ENSG00000065320	ENSG00000065320.7	8	NTN1	protein_coding	17	9021542	9244000	+	9423	-0,236	-5,295	0,037	2,64E-17	2890,28966	0	0
ENSG00000065882	ENSG00000065882.14	15	TBC1D1	protein_coding	4	37891087	38139175	+	23216	0,074	-2,957	0,192	-2,01E-16	1663,05554	0	0
ENSG00000072954	ENSG00000072954.5	6	TMEM38A	protein_coding	19	16661127	16690029	+	79041	1,355	-4,117	0,131	2,20E-16	1837,19748	0	0
ENSG00000099822	ENSG00000099822.2	2	HCN2	protein_coding	19	589893	617159	+	610	0,860	-4,743	0,498	-9,37E-17	1461,80695	0	0
ENSG00000100362	ENSG00000100362.11	12	PVALB	protein_coding	22	36800684	36819479	-	5816	1,942	-7,974	1,984	-7,63E-17	1716,68979	0	0
ENSG00000102678	ENSG00000102678.6	6	FGF9	protein_coding	13	21671383	21704498	+	2254	0,154	-5,551	-0,469	-8,72E-18	2367,4333	0	0
ENSG00000102755	ENSG00000102755.9	11	FLT1	protein_coding	13	28300344	28495145	-	2321	-3,537	-1,290	1,098	1,01E-16	1442,73687	0	0
ENSG00000105929	ENSG00000105929.14	15	ATP6V0A4	protein_coding	7	138706295	138799560	-	50617	0,088	-8,911	-1,239	1,96E-17	2240,76769	0	0
ENSG00000107159	ENSG00000107159.11	12	CA9	protein_coding	9	35673856	35681159	+	768	-6,422	1,494	-0,636	-4,14E-16	1711,63179	0	0
ENSG00000107954	ENSG00000107954.9	10	NEURL1	protein_coding	10	103493979	103592552	+	9148	0,023	-3,502	0,300	-1,20E-17	1454,39114	0	0
ENSG00000109684	ENSG00000109684.13	14	CLNK	protein_coding	4	10486395	10684865	-	116449	-0,220	-6,421	-0,021	-2,75E-17	5678,83782	0	0
ENSG00000112530	ENSG00000112530.10	11	PACRG	protein_coding	6	162727132	163315492	+	135138	0,452	-4,859	-0,510	-6,81E-18	1683,73854	0	0
ENSG00000112715	ENSG00000112715.19	21	VEGFA	protein_coding	6	43770184	43786487	+	7422	-3,765	-1,308	1,343	2,89E-16	1787,64862	0	0
ENSG00000113739	ENSG00000113739.9	10	STC2	protein_coding	5	173314713	173329503	-	8614	-3,300	-0,005	1,502	-6,16E-17	1469,88963	0	0
ENSG00000116039	ENSG00000116039.10	11	ATP6V1B1	protein_coding	2	70935882	70965406	+	525	0,407	-7,958	-0,176	1,25E-16	1636,78261	0	0
ENSG00000129521	ENSG00000129521.12	13	EGLN3	protein_coding	14	33924231	34462774	-	112399	-4,351	0,732	-0,723	2,72E-17	1513,10883	0	0
ENSG00000131910	ENSG00000131910.4	4	NR0B2	protein_coding	1	26911489	26913966	-	8431	1,303	-5,089	1,267	3,53E-17	1503,6986	0	0
ENSG00000132677	ENSG00000132677.11	12	RHBG	protein_coding	1	156369212	156385219	+	57127	0,477	-6,228	0,285	6,51E-17	3441,54946	0	0
ENSG00000140519	ENSG00000140519.11	12	RHCG	protein_coding	15	89471398	89496613	-	51458	1,153	-9,115	1,225	2,18E-18	3366,23437	0	0
ENSG00000142449	ENSG00000142449.11	12	FBN3	protein_coding	19	8065402	8149846	-	84467	0,078	-4,417	0,067	-1,20E-17	2517,95741	0	0

ENSG00000142611	ENSG00000142611.15	16	PRDM16	protein_coding	1	3069168	3438621	+	63976	-0,365	-3,253	0,137	-2,23E-17	1778,0422	0	0
ENSG00000143001	ENSG00000143001.4	4	TMEM61	protein_coding	1	54980792	54992293	+	199964	0,778	-4,733	-0,737	-7,25E-17	1521,94729	0	0
ENSG00000146755	ENSG00000146755.9	10	TRIM50	protein_coding	7	73312539	73328082	-	135892	0,497	-5,938	-0,612	-4,36E-18	1660,82716	0	0
ENSG00000147606	ENSG00000147606.7	8	SLC26A7	protein_coding	8	91209494	91398152	+	115111	0,142	-6,978	0,279	-8,72E-17	3322,63571	0	0
ENSG00000147614	ENSG00000147614.3	3	ATP6V0D2	protein_coding	8	85987323	86154228	+	245972	0,396	-8,966	-0,082	-9,26E-17	2252,98915	0	0
ENSG00000151418	ENSG00000151418.10	11	ATP6V1G3	protein_coding	1	198523222	198540945	-	127124	-0,054	-8,116	0,041	-1,58E-17	5728,12703	0	0
ENSG00000153822	ENSG00000153822.12	13	KCNJ16	protein_coding	17	70053429	70135608	+	3773	0,485	5,904	-0,832	2,46E-16	1463,51327	0	0
ENSG00000125872	ENSG00000125872.7	7	LRRN4	protein_coding	20	6040778	6054049	-	164312	1,473	3,012	-2,978	-4,20E-17	780,498526	1,13E-238	9,50E-237

Table SI D.

ENSG	ENSGv	gene_ version	SYMBOL	gene_biotype	chr	start	end	strand	ENTREZID	chRCC... ccRCC	chRCC... pRCC_e1	chRCC... pRCC_e2	AveExpr	F	P.Value	adj.P.Val
ENSG00000001626	ENSG00000001626.13	14	CFTR	protein_coding	7	117465784	117715971	+	1080	6,442	6,247	6,518	-6,16E-17	3215,35083	0	0
ENSG00000004939	ENSG00000004939.12	13	SLC4A1	protein_coding	17	44248385	44268141	-	6521	7,740	8,105	7,834	-4,71E-17	1720,62711	0	0
ENSG00000035720	ENSG00000035720.6	7	STAP1	protein_coding	4	67558728	67607337	+	26228	7,428	7,623	7,702	1,15E-16	3349,99877	0	0
ENSG00000065320	ENSG00000065320.7	8	NTN1	protein_coding	17	9021542	9244000	+	9423	5,059	5,332	5,295	2,64E-17	2890,28966	0	0
ENSG00000065882	ENSG00000065882.14	15	TBC1D1	protein_coding	4	37891087	38139175	+	23216	3,032	3,149	2,957	-2,01E-16	1663,05554	0	0
ENSG00000072954	ENSG00000072954.5	6	TMEM38A	protein_coding	19	16661127	16690029	+	79041	5,472	4,248	4,117	2,20E-16	1837,19748	0	0
ENSG00000099822	ENSG00000099822.2	2	HCN2	protein_coding	19	589893	617159	+	610	5,603	5,241	4,743	-9,37E-17	1461,80695	0	0
ENSG00000100362	ENSG00000100362.11	12	PVALB	protein_coding	22	36800684	36819479	-	5816	9,916	9,959	7,974	-7,63E-17	1716,68979	0	0
ENSG00000102678	ENSG00000102678.6	6	FGF9	protein_coding	13	21671383	21704498	+	2254	5,705	5,082	5,551	-8,72E-18	2367,4333	0	0
ENSG00000105929	ENSG00000105929.14	15	ATP6V0A4	protein_coding	7	138706295	138799560	-	50617	8,999	7,672	8,911	1,96E-17	2240,76769	0	0
ENSG00000107954	ENSG00000107954.9	10	NEURL1	protein_coding	10	103493979	103592552	+	9148	3,526	3,803	3,502	-1,20E-17	1454,39114	0	0
ENSG00000109684	ENSG00000109684.13	14	CLNK	protein_coding	4	10486395	10684865	-	116449	6,200	6,400	6,421	-2,75E-17	5678,83782	0	0
ENSG00000112530	ENSG00000112530.10	11	PACRG	protein_coding	6	162727132	163315492	+	135138	5,310	4,348	4,859	-6,81E-18	1683,73854	0	0
ENSG00000116039	ENSG00000116039.10	11	ATP6V1B1	protein_coding	2	70935882	70965406	+	525	8,365	7,782	7,958	1,25E-16	1636,78261	0	0
ENSG00000131910	ENSG00000131910.4	4	NR0B2	protein_coding	1	26911489	26913966	-	8431	6,392	6,355	5,089	3,53E-17	1503,6986	0	0
ENSG00000132677	ENSG00000132677.11	12	RHBG	protein_coding	1	156369212	156385219	+	57127	6,705	6,514	6,228	6,51E-17	3441,54946	0	0
ENSG00000140519	ENSG00000140519.11	12	RHCG	protein_coding	15	89471398	89496613	-	51458	10,268	10,339	9,115	2,18E-18	3366,23437	0	0
ENSG00000142449	ENSG00000142449.11	12	FBN3	protein_coding	19	8065402	8149846	-	84467	4,495	4,484	4,417	-1,20E-17	2517,95741	0	0
ENSG00000142611	ENSG00000142611.15	16	PRDM16	protein_coding	1	3069168	3438621	+	63976	2,888	3,390	3,253	-2,23E-17	1778,0422	0	0
ENSG00000143001	ENSG00000143001.4	4	TMEM61	protein_coding	1	54980792	54992293	+	199964	5,511	3,996	4,733	-7,25E-17	1521,94729	0	0
ENSG00000146755	ENSG00000146755.9	10	TRIM50	protein_coding	7	73312539	73328082	-	135892	6,435	5,327	5,938	-4,36E-18	1660,82716	0	0
ENSG00000147606	ENSG00000147606.7	8	SLC26A7	protein_coding	8	91209494	91398152	+	115111	7,120	7,257	6,978	-8,72E-17	3322,63571	0	0
ENSG00000147614	ENSG00000147614.3	3	ATP6V0D2	protein_coding	8	85987323	86154228	+	245972	9,362	8,884	8,966	-9,26E-17	2252,98915	0	0
ENSG00000151418	ENSG00000151418.10	11	ATP6V1G3	protein_coding	1	198523222	198540945	-	127124	8,062	8,157	8,116	-1,58E-17	5728,12703	0	0
ENSG00000153822	ENSG00000153822.12	13	KCNJ16	protein_coding	17	70053429	70135608	+	3773	-5,419	-6,735	-5,904	2,46E-16	1463,51327	0	0
ENSG00000156284	ENSG00000156284.5	5	CLDN8	protein_coding	21	30214006	30216073	-	9073	7,646	7,698	7,710	-1,91E-18	1925,21746	0	0
ENSG00000158089	ENSG00000158089.13	14	GALNT14	protein_coding	2	30910467	31155202	-	79623	-6,042	-5,969	-5,401	2,16E-16	1609,46221	0	0
ENSG00000162399	ENSG00000162399.6	6	BSND	protein_coding	1	54998933	55010883	+	7809	6,563	6,610	6,483	-7,14E-17	9957,70823	0	0
ENSG00000165621	ENSG00000165621.7	8	OXGR1	protein_coding	13	96985719	96994730	-	27199	4,307	4,429	4,359	3,35E-17	2502,86033	0	0
ENSG00000167748	ENSG00000167748.9	10	KLK1	protein_coding	19	50819148	50823787	-	3816	8,750	8,687	7,270	-5,45E-18	1667,05233	0	0
ENSG00000168269	ENSG00000168269.8	8	FOXI1	protein_coding	5	170105897	170109725	+	2299	9,336	9,428	9,381	-1,12E-16	11645,4956	0	0
ENSG00000168785	ENSG00000168785.6	7	TSPAN5	protein_coding	4	98470367	98658629	-	10098	3,847	5,652	5,285	-1,10E-16	1514,09782	0	0
ENSG00000168878	ENSG00000168878.15	16	SFTPB	protein_coding	2	85657314	85668741	-	6439	5,143	5,098	5,173	5,27E-17	1738,86893	0	0
ENSG00000168907	ENSG00000168907.12	13	PLA2G4F	protein_coding	15	42139034	42156636	-	255189	6,468	6,373	6,194	-5,99E-18	4923,2739	0	0
ENSG00000173253	ENSG00000173253.13	15	DMRT2	protein_coding	9	1049858	1057552	+	10655	7,961	8,052	7,996	5,38E-17	9454,13854	0	0
ENSG00000174562	ENSG00000174562.12	13	KLK15	protein_coding	19	50825289	50837213	-	55554	5,199	5,212	4,610	-2,04E-17	1798,84601	0	0
ENSG00000184012	ENSG00000184012.10	11	TMPRSS2	protein_coding	21	41464551	41531116	-	7113	6,517	5,700	5,613	4,14E-17	1718,9789	0	0
ENSG00000184908	ENSG00000184908.16	17	CLCNKB	protein_coding	1	16043736	16057308	+	1188	8,388	6,791	7,961	-1,23E-16	1657,0891	0	0
ENSG00000185133	ENSG00000185133.12	13	INPP5J	protein_coding	22	31122731	31134696	+	27124	7,256	6,218	6,507	1,18E-16	3227,77655	0	0
ENSG00000185352	ENSG00000185352.8	8	HS6ST3	protein_coding	13	96090839	96839562	+	266722	4,448	4,656	4,367	4,56E-17	1747,84025	0	0
ENSG00000186766	ENSG00000186766.7	7	FOXI2	protein_coding	10	127737235	127741186	+	399823	5,395	5,656	5,653	-3,27E-18	1503,88512	0	0
ENSG00000187715	ENSG00000187715.12	13	KBTBD12	protein_coding	3	127915232	127987671	+	166348	3,310	3,272	3,310	8,85E-18	1953,35988	0	0
ENSG00000188175	ENSG00000188175.8	9	HEPACAM2	protein_coding	7	93188586	93226524	-	253012	7,968	8,017	8,003	3,81E-18	7538,16223	0	0
ENSG00000197943	ENSG00000197943.8	9	PLCG2	protein_coding	16	81739097	81962693	+	5336	4,662	5,781	5,047	-1,90E-16	2050,69584	0	0
ENSG00000214128	ENSG00000214128.9	10	TMEM213	protein_coding	7	138797952	138838101	+	155006	8,839	7,942	8,871	6,05E-17	2491,59109	0	0
ENSG00000241233	ENSG00000241233.2	3	KRTAP5-8	protein_coding	11	71538025	71539207	+	57830	4,801	4,882	4,891	1,53E-17	2409,36213	0	0
ENSG00000263155	ENSG00000263155.4	5	MYZAP	protein_coding	15	57591941	57685364	+	100820829	4,848	5,098	5,010	-2,19E-17	2908,13569	0	0
ENSG00000185274	ENSG00000185274.10	11	WBSCR17	protein_coding	7	71132169	71713600	+	64409	6,564	6,815	6,615	1,96E-17	1424,85168	5.92878775009496e-323	2.12102381759647e-320
ENSG00000146021	ENSG00000146021.13	14	KLHL3	protein_coding	5	137617500	137736090	-	26249	3,820	3,976	4,068	-6,27E-18	1420,33436	1.77863632502849e-322	6.24943635424593e-320
ENSG00000101638	ENSG00000101638.12	13	ST8SIA5	protein_coding	18	46667821	46759257	-	29906	2,878	2,965	2,695	9,13E-18	1401,15656	1.79395236004957e-320	6.19259904234835e-318
ENSG00000157703	ENSG00000157703.14	15	SVOPL	protein_coding	7	138594285	138701352	-	136306	2,266	2,191	2,234	9,40E-18	1397,98237	3.87989751679131e-320	1.31622200665677e-317

Table SII. Spiked-in and recovered cell numbers.

Cell type	Cells spiked in	Cells recovered	Recovery efficiency
Cell line			
SNU-349	120	54	45%
SNU-349	10	5	50%
SNU-349	232	118	51%
SNU-349	229	123	54%
SNU-349	93	59	63%
Primary cells			
R046T	161	103	64%
R081T	218	145	67%
R376T	166	105	63%
R378T	183	118	64%
R091T	176	129	73%