UNIVERSITY OF GENOA, ITALY

Department of Internal Medicine, School of Medicine

XXXI cycle Ph.D. course:

Translational Medicine in Oncology and Hematology

CELL-FREE METHYLATED DNA IMMUNOPRECIPITATION AND HIGH THROUGHPUT SEQUENCING TECHNOLOGY: DIAGNOSTIC VALUE IN PATIENTS WITH RENAL CELL CARCINOMA.

Tutor:

Francesco Boccardo, MD
Full Professor of Medicine
University of Genova, Italy
AOU San Martino - IST, Genoa, Italy

Ph.D student:

Pier Vitale Nuzzo, MD

Supervisors:

Matthew L. Freedman, M.D.
Associate Professor of Medicine
Harvard Medical School
Dana-Farber Cancer Institute, Boston,MA, USA

Toni K. Choueiri, MD
Full Professor of Medicine
Harvard Medical School
Dana-Farber Cancer Institute, Boston,MA, USA

2017-2018

Declaration

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

Individuals and organizations that have contributed to the research for this dissertation:

- Academic Unit of Medical Oncology, IRCCS San Martino University Hospital - IST National Cancer Research Institute, University of Genoa, Genoa, Italy.
- Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA, USA.
- Dana Farber Cancer Institute, Boston, MA; Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA
- Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA;
- The Eli and Edythe L. Broad Institute, Cambridge, MA 02142, USA.
- Princess Margaret Cancer Centre, University Health Network, Toronto, ON.

Table of Contents

Acknowledgements

I would like to thank Prof. Matthew Freedman for giving me the opportunity to be a part of his research team and for letting me conduct my PhD under his supervision. Thank you for all of your scientific advice, discussions, corrections and for giving me the opportunity to experience different research cancer areas at high level.

I would like to express my sincere gratitude to Prof. Toni Choueiri, for providing me this opportunity, for his patient guidance, enthusiastic encouragement and useful critiques of this research work. Thanks for letting me be a part of your clinical research family and the precious networking.

I also 1acknowledge De Carvalho Daniel's laboratory from the *Princess Margaret Cancer Centre, Toronto, ON* for their scientific and precious collaboration in this research study. My sincere thanks also go to Ankur Shen Roxana for all the computational assistance and bioinformatics analysis.

I would also like to express my deep thanks to all members and lovely friends of Freedman's laboratory. I am proud to work with a team full of ambitious people who continue to deliver excellent results. Thanks for the support and friendship. You're the most amazing coworkers.

In particular, would like to express my sincere gratitude to Sandor Spisak for the continuous support of my PhD study, for his patience, motivation, and immense knowledge. Your guidance helped me to increase my scientific abilities in the research area. Thanks for sharing your scientific knowledge and tricks with me and for teaching me a lot of scientific technologies and procedures.

I gratefully acknowledge the patients for participating in this study and my loving family for their support and encouragement through my entire life.

Last but least, I express my gratitude to Prof. Francesco Boccardo and Prof.ssa Alessandra Rubagotti for all the things they taught me during my residency and for the opportunity to be in Boston. I am grateful to Prof. Boccardo, who raised me with a love of science and supported me in all my pursuits and has been my guide of science, endurance and patience. I'm grateful for all of his help, continued support and the invaluable role in my life.

1. INTRODUCTION

Renal cell carcinoma (RCC) is the third most common urological cancer of men and represents approximately 3% of all malignant disease[1]. In 2012, 338,000 new cases were newly diagnosed and cancer-related deaths were 143,000 worldwide[2].

Nowadays RCC diagnosis is commonly incidental, based on radiologic procedures performed for other indications[3,4]. At diagnosis, nearly 70% of patients presents with localized disease, whereas 15% has regional disease and 15% distant metastases[5]. For those with localized RCC and selected patients with advanced RCC, radical or partial nephrectomy (RN) can be curative[6,7]. However, around one third of patients who underwent resection of localized disease eventually recurs distant[8]. Known prognostic classifications based on clinic-pathological features, such as TNM stage and Fuhrman grade, showed limited ability to estimate risk of recurrence for these patients[9,10].

Therefore, there is an urgent need to identify simple and reliable markers to enhance diagnostic accuracy and better determine the risk of relapse following nephrectomy.

In the past few years, as the concordance of genetic alterations between cell-free DNA (cfDNA) and matched tumor biopsies was validated[11], there has been increasing enthusiasm over the use of cfDNA as a potential blood cancer biomarker. In fact, cfDNA could serve as a "liquid biopsy" enabling comprehensive tumor genomic profiling from blood at various time points, without the shortcomings of invasive tissue biopsies[11,12].

Obstacles to analyzing traditional circulating tumor DNA (ctDNA) relate to sensitivity given the limited number of recurrent mutations available in order to properly distinguish between tumor and normal circulating cfDNA in a cost-effective manner and the technical artefacts as a result of sequencing. On the other hand, the use of cell-free methylated DNA (cfmeDNA) fragments from cfDNA would overcome these issues.

The advantages of developing cfmeDNA as a biomarker are three-fold: 1) methylation status differs between normal and malignant tissues[13]; 2) methylation status is tissue-specific, which permits inference on tissue origin[14,15]; and 3) the 'target size' for methylation is larger than individually identifying a handful of somatically acquired genetic alterations, thereby improving dramatically its sensitivity and dynamic range.

The group of Daniel De Carvalho *(Princess Margaret Cancer Centre, University Health Network, Toronto, ON)*, developed a novel, optimized technology, namely the cfMeDIP-seq (cell-free methylated DNA immunoprecipitation and high-throughput sequencing) for genome-wide bisulfite-free plasma DNA methylation profiling. This methodology permits an enrichment for CpG rich, thereby overcoming the fragmented nature of plasma cfDNA, and thus enhancing cost-effectiveness[16]. Their first approach was the optimization of the gold standard low-input MeDIP-seq protocol, reducing the 100 ng input of DNA to only 1-10ng input DNA[17]

In this thesis, we will evaluate the sensitivity and specificity of cfMeDIP-seq technology to early detect RCC in a plasma samples of RCC patients at stage I, II and III, collected before RN.

2. RENAL CELL CARCINOMA

2.1 Epidemiology

Globally, the incidence of RCC varies widely from region to region, with the highest rates observed in the Czech Republic and North America[18]. In the United States, there are approximately 65,000 new cases and almost 15,000 deaths from RCC each year[1]. In the European Union, there were approximately 84,000 cases of RCC and 35,000 deaths due to kidney cancer in 2012[19].

RCC is approximately 50 percent more common in men compared with women occurs predominantly in the sixth to eighth decade of life with median age at diagnosis around 64 years of age[1,20].

Within the United States, Asian Americans or Pacific Islanders have the lowest incidence of renal cancers compared with American Indians/Alaska natives, Hispanic/Latinos, Whites, or African Americans[1].

Nowadays RCC diagnosis is commonly incidental, based on radiologic procedures performed for other indications [3,4]. At diagnosis, nearly 70% of patients presents with localized disease, whereas 15% has regional disease and 15% distant metastases[5]. For those with localized RCC and selected patients with advanced RCC, RN can be curative[6,7]. However, around one third of patients who underwent resection of localized disease eventually recurs distantly[8].

The five-year survival rate of patients with kidney cancer has doubled over the last 50 years, from 34 percent in 1954 to 62 percent in 1996, and to 73 percent from 2005 to 2011[21]. The incidence of RCC has risen threefold higher than the mortality rate, mostly due to earlier detection of these tumors at smaller sizes and curative surgical treatment[22].

2.2 Risk factors

Smoking is a well-established risk factor for RCC. The relative risks for RCC for all smokers, current smokers, and former smokers were 1.31, 1.36, and 1.16, respectively[23]. Another known risk is the hypertension, which seems to be independent of anti-hypertensive medications or obesity[24].

Obesity is another risk factor for RCC in both men and women[25]. The relative risk (RR) of RCC increased progressively with baseline body mass index (BMI). For patients with newly diagnosed RCC, excess body weight is associated with a lower stage and lower grade disease[26]. Furthermore, in patients with metastatic disease, RCC is associated with a longer overall survival for those with excess body weight compared with those with normal or below normal body weight[27]. Acquired cystic disease of the kidney is a definitive risk factor for RCC. It develops in approximately 35 to 50 percent of chronic dialysis patients, approximately 6 percent who eventually develop RCC[28].

Occupational exposure to toxic compounds, such as cadmium, asbestos, and petroleum by-products, has been associated with an increased risk of RCC, although studies of occupational exposures are often limited by the lack of specific exposure details[29].

The other risk factor that have been linked to RCC is the use of analgesics. Epidemiologic studies have demonstrated an increased risk for RCC with heavy use of aspirin, non-steroidal anti-inflammatory drugs (NSAIDS), and acetaminophen, although the risk may vary depending upon the agent[30].

Approximately 2-3% of cases are familial and several hereditary RCC syndromes have been described including von Hippel-Lindau syndrome, hereditary papillary RCC, hereditary leiomyomatosis RCC, Birt-Hogg-Dubé, and tuberous sclerosis[31,32]. Among these, the von Hippel-Lindau syndrome is most notable.

Additional clinical factors may increase the risk of developing RCC, such as diabetes mellitus[33], polycystic kidney disease[34], and dietary factors such as the intake of nitrite from processed meat sources[35], reproductive factors (eg, increasing number of pregnancies), and prior radiation therapy (RT)[36]. Conversely, alcohol intake is associated with a protective effect on the risk of RCC in both men and women[37] and the use of oral contraceptives in women may reduce risk[38].

2.3 Pathology

Histopathological examination of tumor tissue is necessary to confirm the diagnosis of RCC. Then, RCC is classified histologically into different subtypes.

The classification reflects the morphology, growth pattern, cell of origin, histochemical, and molecular basis of the different types of adenocarcinomas[39,40].

Clear cell carcinoma is the most common subtype and accounts for 75% of the cases[41]. Clear cell refers to the high lipid content in cytoplasm that is dissolved during histological preparation, resulting in a lucent or clear cytoplasm. In the remaining cases, papillary carcinoma, chromophobe carcinoma and collecting duct carcinoma are described, which represent 10-15%, 5-10% and 1% of the cases, respectively[39,40].

Because clear cell carcinoma is the most common subtype of RCC, much work has been done to classify these cancers based on genetic alterations.

The common genetic abnormalities in RCC was found in von-Hippel Lindau (VHL) gene[42]. The VHL gene is found on chromosome 3 (3p25 to 26) and plays a pivotal role in the development of clear cell RCC in patients with VHL disease. In addition, VHL gene alterations appear to be important in the pathogenesis of sporadic RCC[43,44]. In one report of 187 patients with sporadic RCC, somatic mutations or promoter hypermethylation in the VHL gene was observed in 58 percent of cases[42]. Other reports using high throughput methodologies have demonstrated improved identification of VHL alterations; up to 91 percent of patients with clear cell RCC harbor a VHL gene alteration through genetic or epigenetic mechanisms[44].

2.4 Clinical manifestation

Patients with RCC can present with a range of symptoms; unfortunately, many patients are asymptomatic until the disease is advanced. At presentation, approximately 25 percent of individuals either have distant metastases or advanced locoregional disease[45].

Patients with localized disease can present with a wide array of symptoms and/or laboratory abnormalities, or they may be diagnosed incidentally. In the past, the most common presenting symptoms were hematuria, abdominal mass, pain, and weight loss[45]. In contemporary series, there is an increased frequency of incidental diagnosis due to radiologic procedures performed for other indications. This shift in pattern of presentation along with improvements in therapy may have contributed to better outcomes in RCC[46].

Patients with RCC can also present with or subsequently develop systemic symptoms or paraneoplastic syndromes[47]. In some instances, these may be due to ectopic production of various hormones (eg, erythropoietin, parathyroid hormone-related protein [PTHrP], gonadotropins, human chorionic somatomammotropin, an adrenocorticotropic hormone [ACTH]-like substance, renin, glucagon, insulin)[47].

2.5 Diagnostic evaluation

The usual first diagnostic test is abdominal computed tomography (CT) or, occasionally, abdominal ultrasound[48]. Although ultrasonography is less sensitive than CT in detecting a renal mass, it is useful to distinguish a simple benign cyst from a more complex cyst or a solid tumor.

If the ultrasonography is not sufficient, the patient should undergo CT before and after injection of iodinated contrast. On CT, a simple cyst has a smooth appearance without a clearly delineated wall, has no enhancement with intravascular contrast, and is the density of water. CT urography allows imaging of both the renal parenchyma and the collecting system[49].

Magnetic resonance imaging (MRI) may be useful when ultrasonography and/or CT are inconclusive or if iodinated contrast cannot be administered because of allergy or poor renal function. MRI is particularly helpful in cases where a neoplasm is diagnosed as it evaluates for tumor growth into the collecting system or the vessels better than the other modalities[50].

2.6 Tissue diagnosis

RN is used in most cases to obtain tissue for diagnosis of RCC prior to treatment, although the diagnosis of RCC is occasionally established by a biopsy of a metastasis. After the presumptive diagnosis has been made based upon imaging studies, the patient must be evaluated for the extent of local involvement and the presence of metastatic disease prior to surgery.

The role of percutaneous biopsy is more limited, although it may be used for a small renal mass if there is a high index of suspicion for a metastatic lesion to the kidney, lymphoma, or a focal kidney infection[51]. A biopsy can also be used to confirm a diagnosis of RCC in patients who are not surgical candidates prior to initiating appropriate medical treatment, although biopsy of a metastatic lesion is often preferable[52].

2.7 Staging imaging studies

The extent of local and regional involvement is determined primarily by abdominal CT, which is extremely accurate in staging RCC. Other procedures that may be useful for assessing for distant metastases include bone scan, CT of the chest, MRI, and PET/CT.

Bone scan is indicated only in patients with bone pain and/or an elevated serum alkaline phosphatase[53]. CT of the chest is useful to evaluate for evidence of pulmonary or mediastinal lymph node metastases[54]. MRI scanning with gadolinium is superior to CT for evaluation of the inferior vena cava and right atrium when tumor involvement is suspected[55]. PET scanning has high sensitivity and specificity for the primary lesion. Although PET or PET/CT may be more sensitive than radionuclide scanning for the detection of bone metastases, it is expensive and has limited use for routine staging[56].

2.8 TNM Staging System

The eighth (2017) Tumor, Node, Metastasis (TNM) staging system is used for staging all histologic variants of RCC. This system is supported by both the American Joint Committee on Cancer (AJCC) and the International Union for Cancer Control (UICC)[49]. These TNM criteria use the anatomic extent of disease to define prognostic stage groups.

The TNM system is shown in the table (table 1). In this system, tumors limited to the kidney are classified as T1 or T2 based upon size. T3 tumors extend into the renal vein or perinephric tissues but not beyond the Gerota fascia, while T4 tumors extend beyond the Gerota fascia, including direct extension into the ipsilateral adrenal gland. Nodal and distant metastases are simply classified as absent or present.

3. CIRCULATING CELL-FREE DNA

3.1 Origin and characteristics of cfDNA

cfDNA are single- or double-stranded extracellular fragments of deoxyribonucleic acids with a molecular weight ranging from 0.16 kb to 21kb that circulate in plasma, serum, and other bodily fluids outside of cells[11].

The presence of cfDNA in human blood was initially reported in 1948 by Mandel and Métais in healthy individuals, pregnant women, and sick individuals[57]. However, this did not arise much interest until 1977 when Leon et al. suggested the possibility of exploiting cfDNA as tumor biomarker for the first time[58]. Only 10 years later, Stroun et al. demonstrated the presence of neoplastic characteristics in the cfDNA of cancer patients[59]. Subsequently, several studies reported the possible use of cfDNA as a biomarker for cancer patients and a variety of alterations in cfDNA including point mutations, deletions, DNA hypermethylation, copy number variation (CNV), microsatellite instabilities (MI), and losses of heterozygosity (LOH) have been investigated[11].

CfDNA can be present in the bloodstream as "naked" DNA, free from any binding to other molecules or surfaces[60,61], wrapped around histone proteins in nucleosome structures[62], or internalized in vesicles[63] which protect it from nucleases and prevent its exposure to the immune system, or in virtosome structures, bound to particular structures, such as apoptotic bodies, serum carrier proteins, or anti-DNA antibodies[64]. Furthermore, cfDNA can be observed attached to the exterior of the cell membrane or adsorbed on the surface of blood cells from which it can be freed into the circulatory system[65,66].

The biological mechanisms which underlie cfDNA shedding into the bloodstream remain to be fully elucidated. In healthy individuals, cfDNA could result from cell apoptosis or necrosis or

living cells actively releasing DNA fragments[67,68]. The analysis of cfDNA fragmentation revealed that most of the cfDNA fragments measure between 160bp and 200bp indicating that the majority is produced by apoptosis. However, DNA fragments of higher molecular length, similar to necrotic cell DNA, were also detected in several samples[11].

CtDNA is a fragment of DNA that originates from tumor cells and thus harbors cancer-related mutations[69]. This characteristic differentiates it from cfDNA which more broadly defines any DNA freely circulating in the bloodstream. ctDNA can vary between 0.01% and more than 90% of total cfDNA depending on different clinical cancer histories[69,70]. CtDNA was detectable at lower concentration in patients with localized cancers and at high concentration in the circulation of most patients with metastatic cancer, reaching more than 75% in metastatic pancreatic, ovarian, colorectal, bladder, breast, melanoma, hepatocellular and head and neck cancers[71]. However in RCC, ctDNA was detectable in less than 50% of patients with metastatic disease and in 20% of patients of primary disease[71].

The mechanisms hypothesized underpining the release of ctDNA comprise active secretion of tumor cells, including primary tumor, lysis of tumor cells that circulate in the blood or micrometastases shed by tumor, passive release from apoptotic and necrotic tumor cells, or other physiological events induced by microenvironmental stress and treatment pressure[68,72,73].

Currently, there are no data available on the kinetics of cfDNA in RCC. Some reports in other tumors suggested that the clearance of cfDNA occurs rapidly, ranging from 15 minutes to several hours[70 29]. The elimination process is characterized by a rapid phase with a mean half-life of approximately 1 hour, of which liver, spleen, and kidney may be responsible, and a slower phase with a mean half-life of approximately 13 hours, probably due to the activity of plasma nucleases[60,69,74].

Although cancer patients have higher levels of cfDNA compared to healthy individuals, the cfDNA concentrations alone does not seem to be useful to distinguish healthy individual from patients with benign and malignant disease. In cancer patients, the concentration of cfDNA in the plasma or in the serum varies up to 50 times the normal level[71], based on tumor burden, stage, cellular turnover, response to therapy, and the presence of DNAse inhibitors[71]. However increase of cfDNA is also observed in several other pathologies, such as multiple sclerosis[75], stroke[76],

hemodialysis[77], pancreatitis[78], trauma[76], aging[79], and after transplantation[80], systemic lupus erythematosus (LES)[81], and cardiovascular diseases[82,83].

3.2 Technical aspects

The numerous discrepancies in the results of the studies of cfDNA reported in the literature so far are mainly due to the lack of standardization of the various protocols for sample handling and processing and to the different techniques used for cfDNA detection and analysis. These issues constitute a major bias when comparing data from different studies and are one of the greatest obstacles in translating cfDNA to clinical practice.

3.2.1 Blood sampling and processing

Various blood sampling and processing factors can affect concentrations and integrity of cfDNA and they include the type of serum preparation (with or without a coagulation accelerator)[84], the use of different anticoagulants for plasma collection[84,85], the time between collection and centrifugation[86], the storage temperature of blood before centrifugation[86,87], centrifugation forces[70,88,89] time and temperature-dependent cryopreservation of samples[90]. A reliable protocol should be developed following careful consideration of the above factors based on the literature data.

Despite serum samples having a concentration of cfDNA 3-24 fold higher than plasma, the latter is a better source for cfDNA[91-94]. Much of the cfDNA in the serum is generated in the original collection tube during the process of clothing. Therefore, the tube contains a higher proportion of cfDNA deriving from white blood cells, which makes detecting ctDNA more difficult[91-94].

Although there is no universal protocol, current knowledge suggests to isolate cfDNA using 5-10ml of blood which is collected in tubes handled with EDTA or cfDNA collection tubes which prevent blood cell lysis[85,95]. Blood must be processed within 4 hours following the blood draw to avoid significant increase in plasma concentration and DNA integrity[90]. Moreover, centrifugation ensures the absence of cells in the plasma and, in this regard, a second high-speed cycle is highly recommended. The number of freeze-thaw cycles does not affect the DNA yield, but it leads to

fragmentation of DNA[90,96]. Therefore, plasma must be aliquoted and stored at −80°C or −20°C for up to three months for cfDNA concentration and fragmentation analysis or nine months for specific sequence detection. In fact, prolonged storage of whole plasma or serum DNA leads to substantial DNA degradation, with an annual degradation rate of approximately 30%[90,96,97].

3.2.2 Detection of cfDNA

The efficiency and quality of the cfDNA extraction process are influenced by the protocol chosen to isolate cfDNA. That is also partly responsible for the high variability of cfDNA data observed among studies. The main methods used for extracting cfDNA can be divided in "commercial methods", a ready-to-use extraction kits based on anion-exchange method, silica-membrane technology, or magnetic-particle technology, and "non-commercial methods", such as phenol-chloroform extraction, guanidine-resin method, alcohol precipitation, and salting-out method.

Although many of "non-commercial methods" achieve high yields and enable extraction of more small-sized fragments, they include use of toxic solvents, take longer time, and, due to the high coefficient of variability in results, they are not appropriate for clinical analyses. Those are the reasons why most of the studies were performed by "commercial methods". They have good repeatability and reproducibility and thus enable standardization for clinical studies.

Once cfDNA was isolated, the two fundamental approaches to analyze cfDNA are based on targeted approaches which allow detection of specific known mutations, and untargeted approaches that allow identification of events without a priori.

PCR assays had been a widely used method for quantifying DNA as the concentration of cfDNA is within the range of nanograms for accurate measurement. Quantitative PCR (qPCR) measures signals from labeled probes during amplification of different target genes (i.e., hTERT, b-globin) or total cfDNA, or quantification of tumor specific mutations. It is a very sensitive and reproducible method with high sensitivity[98,99]. The disadvantage of qPCR is the high risk of false positivity and the limitation of the single assay, for which a high number of separate assays are needed to cover the increasing number of relevant mutations[98,99].

Although PCR can be considered the current standard for detection of multiple mutations (multiplexing) from the same sample, with 0.01%-0.1% sensitivity and high specificity, it has

limitations in terms of small input of DNA applicable for individual assays[100]. Instead the BEAMing (beads, emulsion, amplification, and magnetic), which detects DNA segments by means of amplification using primers with known tag sequences and covalently bound to magnetic beads, nevertheless, has a complex workflow, it is time-consuming, and has a relatively high cost per sample, which make its implementation in a clinical study setting less feasible.

In contrast with these approaches that target hotspot mutations new technologies based on untarget approaches have been developed, such as next generation sequencing (NGS). It is based on the analysis of millions of short sequences from DNA molecules and the subsequent comparison to a reference sequence without knowing a priori the genotype of the tumor[101,102]. Despite the clear advantage of enabling detection of multiple somatic alterations simultaneously with high sensitivity and specificity, NGS is an expensive and time consuming technique[101,103]. A capture based NGS method of ctDNA detection termed CAPP (cancer personalized profiling by deep sequencing) recently showed a significantly higher sensitivity with a lower cost[104]. Finally, the whole exome sequencing (WES) and whole genome sequencing (WGS) techniques not only allow screening of mutations but also of rearrangements and of CNV, providing a comprehensive genomic profiling of ctDNA[105]. However, they are still very expensive to be routinely applied for clinical diagnosis.

Due to its high fragmentation, contamination by non tumoral cfDNA, low amount, and fast clearance, detection of ctDNA remains challenging and requires ultrasensitive analytical assays.

4. CIRCULATING CELL-FREE DNA AND RENAL CELL CARCINOMA

The use of cfDNA could help determining early signs of disease and distinguishing benign from malignant renal lesions, based on its concentration, integrity and genetic and epigenetic abnormalities (table 2).

4.1 CfDNA levels

In this regard, three studies with more than 300 RCC patients and 141 control subjects overall examined cfDNA levels in patients with RCC[106-108]. Although the interpretation of these studies as a whole is hampered by the technical differences between them (serum versus plasma, different

cfDNA isolation methods, specific DNA elements used to quantify cfDNA, etc.) as well as the diverse disease stages for cancer patients, they showed that cfDNA levels were higher in patients with RCC compared with healthy controls.

In the older of these study the preoperative plasma cfDNA content of RCC patients was eight times higher than that of healthy controls (26.4±48.3ng/ml versus controls 3.2±1.5ng/ml)[107]. The statistical significance of this result was maintained even when sex, age, histology, tumor size, grading, and pathologic TNM stratifications were applied. Nevertheless, the area under the curve (AUC) obtained with receiver operating characteristic (ROC) analysis demonstrated a moderate careful discriminating power[107].

Using real-time PCR, Wan et al. also reported increased cfDNA levels in RCC patients compared with localized RCC or controls[108]. However, in this case, there was a significant difference in plasma cfDNA levels among RCC patients with different Fuhrman grade, TNM stage, and tumor size suggesting an association with the tumor aggressivity[108].

Finally, the largest of the 3 reports analyzed cfDNA levels using hydrolysis probes for the ring finger protein 185 gene (RNF185). In patients with RCC, cfDNA levels were significantly higher than in patients with benign tumors, indicating that cfDNA may aid in the differential diagnosis of solid renal masses[106]. Furthermore, total cfDNA levels were higher in patients with metastatic RCC and necrotic RCC[106]. Overall, these studies suggest that elevated cfDNA levels in RCC patients are associated with high degree necrosis, apoptosis, or success of therapy. However, as shown by the ROC analysis of the report by Perego et al, data are not sufficiently robust to distinguish between cancer and non-cancer. This is in keep with the fact that cfDNA levels are influenced by concomitant diseases and other clinical factors. Therefore, currently cfDNA blood concentrations alone cannot be considered a reliable diagnostic tool.

4.2 CfDNA integrity

Postulating that cfDNA fragments have different lengths in cancer vs. non malignant diseases, two studies aimed to analyze the size differences of the cfDNA fragments[109,110]. Using quantitative real-time PCR, Hauser et al. analyzed the serum cfDNA integrity defined as a ratio of the longer fragment actine-beta gene 384 (ACTB384) and the shorter fragment actine-beta gene 106

(ACTB106)[110]. The levels of both DNA fragments were increased in the preoperative serum of RCC patients (n=35) compared to the healthy individuals (n=54)[110]. The significantly higher presence and increased integrity of the long fragment in cancer patients compared to the healthy subjects supported the thesis of necrotic origins and indicated that cfDNA is fragmented to a higher degree in cancer patients[110]. Similarly to Hauser's analysis, also Gang et al. found higher concentration of long fragments of a housekeeping gene glyceraldehyde-3-phosphate dehydrogenase in the preoperative serum of RCC patients (n=78) compared to those in the healthy control group (n=42), suggesting that necrosis is a more frequent event in patients with tumor[109]. Besides, cfDNA integrity was correlated also with tumor stage and size in this cohort[109]. Although the results of the two studies supported the idea that ctDNA is more likely released by necrotic cells and thus it is composed of larger fragments compared to non-tumor-derived cfDNA, likely released by apoptotic cells, the limited number of samples analyzed, the lack of standardization in measuring long and short fragments of cfDNA and the modest sensitivity, specificity and AUC values for short and long fragment detection are not enough to suggest that cfDNA size should be considered as a reliable diagnostic biomarker.

4.3 CfDNA genetic and epigenetic abnormalities

As opposed to cfDNA levels and sizes, the detection of genetic (i.e. MI) and epigenetic (i.e., CpG island hypermethylation) alterations within ctDNA may be useful to increase the rate of early and accurate diagnoses. Eight studies searched for cancer-associated ctDNA alterations common across all RCC[106,107,111-116] (table 3 and 4).

Goessl et al., applying only 4 markers for MI in chromosome 3p, was the first to detect plasma ctDNA genetic alterations in RCC patients[116]. While 63% of patients (n=40) had LOH in at least one locus and 35% in more than one locus, microsatellite instability was found only in one patient, indicating that DNA mismatch repair genes are unlikely to play a major role in renal carcinogenesis[116]. However, those alterations were not associated with tumor stage and only a portion was found in the corresponding primary tumor[116]. Surprisingly, plasma cfDNA alterations were not found in the healthy controls (n=10)[116]. A following report confirmed similar frequency of microsatellite alterations in RCC patients serum[111]. In this case, 60% of patients with malignant renal tumor (n=25) had at least one alteration in the 28 microsatellite markers from 20

chromosome regions studied[111]. Interestingly, also in this study none of the controls (n=16) showed serum cfDNA and no association between cfDNA alteration and tumor stage was found[111]. Conversely, a later study found microsatellite alterations in 74% of RCC patients using 9 markers from 3 chromosomal regions and in 87% of RCC using 20 markers for MI[112]. Most of these alterations were observed on chromosomes 3p and 5q and showed a strong association with nuclear grading and a weak correlation with advanced tumor stages[112]. This method showed 3 false positive alterations in 20 healthy controls, resulting in 85% specificity[112]. In this regard, it should be noted that the microsatellite alteration analysis was possible only in 9 of the 54 RCC patients of Perego's cohort[107]. In this study, 55.6% of patients harbored at least one of the 5 microsatellite markers located on chromosome 3p and these alterations were present also in the corresponding primary tumor[107].

Methylation of knows or putative genes occurs frequently in the early stages of RCC. Particularly, methylation of Ras association domain family member 1A (RASSF1A) was detected frequently in RCC patients. In this regard, a recent analysis showed that RASSF1A is methylated in 22.9% of RCC patients[115], while a previous study by De Martino et al. demonstrated methylation of RASSF1A in 45.9% of patients[106]. Several other studies reported the methylation of this gene with rates ranging from 11% to 62.9% of serum samples of patients with RCC[113,114]. These differences in the RASSF1 gene methylation levels can be explained by the use of different CpG islands for analyses and of a bisulfate-based assay in lieu of the methylation-sensitive restriction enzyme assay which may be more sensitive.

Although Ellinger et al. demonstrated a 100% correlation between DNA hypermethylation of the RASSF1A promoter and papillary RCC[117], De Martino et al. found no association of RASSF1A methylation with the papillary RCC analysed[106].

In addition, also the methylation of other genes with a key-role in renal carcinogenesis seem to have a diagnostic role. In Hoque's cohort, 67% of serum samples of RCC patients had methylated at least one of the 9 gene promoters[113]. Besides RASSF1A, methylation was detected frequently in TIMP3 and CDH1 genes[113]. Interestingly, the 6 control patients who displayed serum methylation were smokers[113]. Moreover, VHL methylation was detected more frequently in patients with clear cell RCC than in other subtypes, with an ROC value of 0.694[106].

Hauser et al. found that 85% of the investigated patients has at least one of the 8 genes studied methylated and six of the 8 genes were significantly iper-methylated in RCC patients compared with healthy individuals[115]. Despite the ROC analysis showing a high specificity ranging between 85.2% and 100% for serum cfDNA methylation, sensitivity was low in single-gene analyses[115].

5. CIRCULATING CELL-FREE METHYLATED DNA AND RENAL CELL CARCINOMA

5.1 DNA methylation and cancer

DNA methylation is a kind of epigenetic modifications, which adds a methyl group to the 5th carbon of cytosine (5-methylcytosine, 5 mC) by DNA methyltransferases (DNMTs) and prefers to occur within CpG dinucleotides[118]. mC can be oxidized to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) by ten-eleven translocation (TET) proteins. The latter two are excised by thymine DNA glycosylase (TDG) coupled with base excision repair (BER), which results in unmodified cytosine[119,120]. This process is known as active DNA demethylation (Figure 1). The dynamic regulation of DNA methylation is a very vital process in cell fate determination and development [121]. In somatic cells, 5mC is primarily restricted to palindromic CpG dinucleotides, and about 60%–80% of CpGs are methylated[122]. Generally, methylated DNA is mainly enriched at promoters and is correlated with inhibition of transcription initiation, so DNA methylation has been known as a repressive marker in genome[118]. Aberrant DNA methylation is a noticeable feature of cancer cells [123,124]. The cancer genome is globally hypomethylated , while hypermethylation of tumor suppressor genes is an early event in many tumors, which promotes cancer progression[118,125]. Growing evidence suggests that impairment of active DNA demethylation may contribute to cancer initiation[126]. Active DNA demethylation is TET-mediated. TET1 was initially identified owing to its fusion to mixed-lineage leukemia (MLL) in patients with acute myeloid leukemia (AML)[127,128].Around the time when TET-mediated oxidation of 5 mC was discovered, multiple studies reported TET2-inactivating mutations in myeloid disorders[128]. In addition to TET2 mutations, TET1 and TET3 mutations were observed in haematopoietic malignancies[129]. Mutations of TET proteins are also observed in various solid tumors, which causes aberrant active DNA demethylation, as shown by the reduction

of 5hmC [126,128]. Therefore, global decrease of 5hmC may be broadly used as a diagnostic biomarker for cancers[119,130]. As to another enzyme in demethylation, TDG has also been implicated in various cancers, although it may be due to TDG's role in mismatch repair[130].

5.2 CfmeDNA as a potential biomarker in cancer.

Although cfDNA represents a promising biomarker capable of providing diagnostic, prognostic and predictive information in many tumor types, clinical routine practice has been slow to adopt liquid biopsy, primarily because scientists have struggled in identifying the ideal technical approach that would yield robust and reproducible results[131].One of the primary challenges of cfDNA is that cancer mutation profiles can be highly variable between patients and across tumor types. This diversity of mutations represents a significant complication for the development and clinical application of cancer tests using cfDNA [132]. This may affect sensitivity for many tumor types, which is a critical issue in using cfDNA. On the other hand, CpG island methylation is characterized as stable[133]. The stability of this DNA methylation can be leveraged by analyzing cfmeDNA in patient plasma[133].

When tumor DNA is shed into the bloodstream, these patterns also become detectable in plasma and serum [133]. Furthermore, methylation changes are a common feature of different cancer types, and occur early in cancer development, typically repressing the expression of tumor suppressor genes[134]. Thus, DNA methylation may offer a more consistent and broadly applicable biomarker of tumor DNA in the blood than mutations. These blood-based methylated cfDNA of tumors represent the basis of our proposed biomarker development.

Another obstacles to analyzing traditional ctDNA relate to sensitivity given the limited number of recurrent mutations available in order to properly distinguish between tumor and normal circulating cfDNA in a cost-effective manner and the technical artefacts as a result of sequencing. On the other hand, the use of methylated DNA fragments from cfDNA would overcome these issues.

The advantages of developing cfmeDNA as a biomarker are three-fold: 1) methylation status differs between normal and malignant tissues[13]; 2) methylation status is tissue-specific, which permits inference on tissue origin[14,15]; and 3) the ‘target size’ for methylation is larger than

individually identifying a handful of somatically acquired genetic alterations, thereby improving dramatically its sensitivity and dynamic range.

cfmeDNA has the potential to be excellent blood cancer biomarkers. Many proof-of-principle studies have indicated that cfmeDNA represent informative and worthwhile blood biomarkers in various tumor types and settings[135-139] Thus, by varying numbers of differentially methylated regions, coverage, and ctDNA abundance, sensitivity is improved , with lower sequencing depth, thereby reducing the cost of detection, classification, and monitoring of cancer. In reality this practice is labor intensive, and challenging given the low-abundance and fragmented nature of plasma cfDNA.[140] Indeed most of the previous plasma methylation profiling has been limited to locus-specific PCR-based assays.[141-143] Alternative approach of cfDNA sequencing have been attempted[144,145], but were esteemed inefficient[146], costly, and with limited information recovery.

5.3 Cell-free methylated DNA immunoprecipitation and high-throughput sequencing methodology

To overcome these technical issues, Daniel De Carvalho developed a novel, optimized technology, namely the **cfMeDIP-seq (cell-free methylated DNA immunoprecipitation and high-throughput sequencing)** for genome-wide bisulfite-free plasma DNA methylation profiling[16]. This methodology permits an enrichment for CpG rich, thereby overcoming the fragmented nature of plasma cfDNA, and thus enhancing cost-effectiveness. Their first approach was the optimization of the gold standard low-input MeDIP-seq protocol, reducing the 100 ng input of DNA to only 1-10ng input DNA[17], and compared low-input cfMeDIP-seq vs. the original standard MeDIP-seq using colorectal cancer samples, which showed robust CpG enrichment and inter-replicate correlation. Subsequently, they compared cfMeDIP-seq to ultra-deep unique molecular identifiers based, hybrid capture mutation sequencing across a serial dilution of colorectal cancer DNA into multiple myeloma. Near-perfect linear associations between the observed and expected number of differentially methylated regions were detected. Furthermore, they found that cfMEDIP-seq had the ability to enrich ctDNA through biased sequencing of CpG-rich sequences, which are often hypermethylated in cancer compared to normal tissues[147]. cfMeDIP-seq of plasma cfDNA had the ability to detect tumor-derived DNA methylation events in ctDNA, where 45,173 differentially methylated CpGs were found between early-stage

pancreatic tumor and normal tissues, as opposed to 14,716 differentially methylated regions between cases and controls cfDNA. Based on the discovery that tumor-specific differentially methylated regions in the plasma of early-stage pancreatic cancer are detectable with cfMEDIP-seq, they then sought to assess whether it could also classify various cancer types from healthy controls. cfMeDIP-seq was then performed on a discovery cohort of 189 plasma samples obtained from seven tumor types, including RCC (early-stage pancreatic cancer, colorectal cancer, breast cancer, lung cancer, bladder cancer, acute myeloid leukemia, RCC) and healthy controls. Subsequently the performance of cfMeDIP profiles in cancer detection and classification was validated in an external ensemble of 199-sample cohort. cfMeDIP-seq was shown to be highly accurate in distinguishing tumor and normal tissues. Averaging the class probabilities output by E100 for each sample yielded high receiver operating characteristics for distinguishing between acute myeloid leukemia vs. other cancers (0.980), pancreatic cancer vs. other (0.918), lung cancer vs. others (0.971), normal vs. others (0.969). Taken together, the De Carvalho laboratory developed a robust and sensitive, bisulfite-free methodology for immunoprecipitation-based profiling of methylation patterns in cfDNA, underlying a highly potent utility of cfDNA methylation profiles as the basis for non-invasive and cost-effective early tumor detection for cancer interception.[16]

5.4 Preliminary results

Our research team at the Dana-Farber Cancer Institute, in collaboration De Carvalho laboratory, have produced two pilot data using early stage/metastatic RCC samples by applying the MeDIP-seq technology, and compared them to 50 RCC samples and 24-non RCC control samples obtained from the De Carvalho group. Our findings showed that the metastatic renal cancer samples from Boston (in blue) clustered well with the renal samples from Toronto (in green), and that both clusters were obviously distinct from the normal clusters (in red, Figure 2A). Subsequently we performed the same analysis using six non-metastatic, early-stage RCC samples (4 patients had stage I disease, 2 patients had stage II disease). Our findings showed that all the renal cancer samples from Boston (in blue) once again clustered well with the ones from Toronto (in green), clearly distinct from the non-cancer control samples (in red), evidence that the methylation profile of cfDNA is also detected in early-stage RCC (Figure 2B).

6. RATIONALE AND AIMS

6.1 Rationale and Hypothesis

The aforementioned findings are proof-of-concept and preliminary studies that suggest that the imaging methods including CT scans and ultrasonography, are currently used for initial diagnosis of RCC, which is subsequently confirmed by histologic analysis. These modalities have limitations for distinguishing the various types of RCC and are sometimes slow and labor intensive. In addition, asymptomatic RCC could be a leading cause of failed early detection of kidney cancer. Despite significant progress in the medical treatment of metastatic RCC, RN remains the only effective treatment for localized RCC. However, no clinically relevant screening assay is currently available to detect asymptomatic RCC and there is, therefore, an urgent need for validated markers of RCC. The use of cfDNA, in particular his methylation pattern, could help determining early signs of disease and distinguishing benign from malignant renal lesions.

Because the 'target size' for methylation is larger than individually identifying a handful of somatically acquired genetic alterations, we hypothesize that cfmeDNA is a more sensitive biomarker than targeting plasma DNA variants to diagnose patients with RCC.

6.2 Aim of the study

Objective: To determine the feasibility of using cfMeDIP-seq technology to detect clear cell RCC (ccRCC) cfmeDNA in a plasma samples collected before RN at different pathologic stages – pT1, pT2, pT3 and metastatic disease.

Aim 1: To evaluate if cfMeDIP-seq technology is able to distinguish ccRCC methylation pattern from healthy patients pattern.

Aim 2: To evaluate the sensitivity and specificity of cfmeDNA to detect ccRCC methylation pattern in a plasma samples collected before RN at different pathologic stages – pT1, pT2, pT3 and metastatic disease.

7. MATERIAL AND METHODS

7.1 Patient selection

We selected a cohort of 46 patients (test cohort) who had histologically confirmed diagnosis of ccRCC between February 2005 and May 2015, and subsequently were followed up at DFCI. These cohort was selected based on the availability of a corresponding plasma sample collected before the time of surgery and cryopreserved up to processing. Patients demography is summarized in table 5. Healthy patients consisted of 10 plasma samples collected between November 2017 and December 2018 at DFCI. The training cohorts consisted 24 normal samples and 165 plasma samples from cancer patients, of which 20 were ccRCC described in Shen et al paper[16].

7.2 Ethical remarks

This research project was governed under a Dana Farber Cancer Institute *Institutional Review Board* (Protocol IRB# 18-515, Role of plasma cell-free methylated DNA immunoprecipitation and high-throughput sequencing).

7.3 Sample acquisition

All patients provided written informed consent, and all samples were obtained upon approval of the IRB, according to all relevant ethical regulations. Renal cancer plasma samples were obtained from the Dana Farber Cancer Institute – GELB Center Biobank.

7.4 Specimen processing for patient cfDNA

Plasma samples were collected using EDTA and acid citrate dextrose tubes and were kept frozen until use. cfDNA was extracted from 1 ml of plasma using the QIAamp Circulating Nucleic Acid Kit (Qiagen) and quantified through Qubit before use.

7.5 cfMeDIP–seq protocol

A schematic representation of the cfMeDIP–seq protocol is shown in figure 3. Before cfMeDIP, the samples were subjected to library preparation using Kapa HyperPrep Kit (Kapa Biosystems), following the manufacturer's protocol with minor modifications. In brief, after end-repair and A-tailing, samples were ligated to 0.181 µM of NEBNext adaptor (NEBNext Multiplex Oligos for Illumina kit, New England BioLabs) by incubating at 20 °C for 20 min and purified with AMPure XP beads (Beckman Coulter). The eluted library was digested using the USER enzyme (New England BioLabs) followed by purification with Qiagen MinElute PCR Purification Kit (MinElute columns) before MeDIP.

The prepared libraries were combined with the filler λ DNA (to ensure the total amount of DNA (cfDNA + filler) was 100 ng) and subjected to MeDIP with Diagenode MagMeDIP kit (C02010021). The filler DNA consists of a mixture of unmethylated and in vitro methylated λ amplicons of different CpG densities (Supplementary Table 6), similar in size to adaptor-ligated cfDNA libraries. Its addition ensures a constant ratio of antibody to input DNA and helps to maintain similar immunoprecipitation efficiency across samples regardless of available cfDNA, while minimizing non-specific binding by the antibody and DNA loss due to binding to plasticware. For MeDIP, the prepared library/filler DNA mixture was combined with 0.3 ng of control methylated and 0.3 ng of the control unmethylated Arabidopsis thaliana DNA provided in the kit, and the buffers. The mixture was heated to 95 °C for 10 min, then immediately placed into an ice water bath for 10 min. Each sample was partitioned into two 0.2 ml PCR tubes: one for the 10% input control (7.9 µl) and the other for the sample to be subjected to immunoprecipitation (79 µl). The included 5-mC monoclonal antibody 33D3 (C15200081) from the MagMeDIP kit was diluted 1:15 before generating the diluted antibody mix and was added to the sample. Washed magnetic beads (following the manufacturer's instructions) were also added before incubation at 4 °C for 17 h. The samples were purified using the Diagenode iPure Kit v2 (C03010015) and eluted in 50 µl of buffer C. The success of the reaction (QC1) was validated by qPCR to detect recovery of the spiked-in methylated and unmethylated A. thaliana DNA. The percentage recovery of unmethylated spiked-in DNA should be <1% (relative to input control, adjusted for input control being 10% of the overall sample) and the percentage specificity of the reaction should be >99% (as calculated by (1 – [recovery of spiked-in unmethylated control DNA over recovery of spiked-in methylated control DNA]) × 100), before proceeding to the next step. The optimal number of cycles to amplify each library was determined by qPCR, after which the

samples were amplified using Kapa HiFi Hotstart Mastermix and NEBNext multiplex oligos, added to a final concentration of 0.3 μM. The final libraries were amplified as follows: activation at 95 °C for 3 min, followed by predetermined cycles of 98 °C for 20 s, 65 °C for 15 s and 72 °C for 30 s and a final extension of 72 °C for 1 min. The amplified libraries were purified using Beads purification. All the final libraries were submitted for BioAnalyzer analysis before sequencing at the Novogene Corporation on an Illumina HiSeq 2500, SBS V4 chemistry, single read 50 bp, multiplexed as twelve samples per lane. After sequencing, the sequenced reads were aligned to λ and hg19 using Bowtie20 with the default settings. On the basis of virtually no alignment to the λ genome, the filler DNA does not interfere with the generation of sequencing data. The generated SAM files from hg19 alignment were converted to BAM format, ensuring the removal of duplicate reads, and the reads were then sorted and indexed using SAMtools21 before subsequent analysis with the R package MEDIPS22. The CpG enrichment score, as a quality control measure for the immunoprecipitation reaction, was calculated as part of the MEDIPS package.

7.6 Concentration of cfDNA, calculation and visualization of differentially methylated regions from cfDNA of patients with renal cell carcinoma and healthy donors

The t-test was applied to compare mean values of cfDNA isolated in ccRCC with those isolated in healthy patients. A significance level of p=0.05 was used and the 95% Confidence Intervals (CIs) were also calculated. All P values were two-tailed. The IBM software Statistical Package for Social Sciences (SPSS) version 19.0 for Windows (SPSS Inc. Chicago, Illinois, USA) was used for data analysis. DMRs between cfDNA samples from patients with ccRCC and healthy donors were calculated using MEDIPS and DESeq2 R packages22,34. For each sample, we computed counts per 300 bp non-overlapping windows, filtered out windows with less than 10 counts across all samples and fit a negative binomial model to call DMRs at FDR < 0.1 (Wald test). z-scores of DMR RPKM values with Euclidean distance and Ward clustering were used for visualization.

7.7 Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized. Plasma samples were blinded during the sample preparation and sequencing. Data analysis was performed unblinded.

7.8 Bioinformatic simulation of tumour-specific features and probability of detection by sequencing depth

cfMeDIP-seq paired end data were concatenated into single fastqs and were subsequently processed using the pipeline described in Shen et al16, to yield RPKMs. These were then transformed into log2 (counts per million) estimates and the data were reduced to the 505,000 300-bp windows of the genome that map to CpG islands, shores, shelves and FANTOM5 enhancers.

In order to assess classification performance, we used the ensemble of elastic net models previously trained in Shen et al study[16]. Specifically, we used the classifiers designed to distinguish between normal samples and seven types of cancer (bladder cancer, renal cancer, lung cancer, breast cancer,pancreatic cancer,colorectal cancer and AML). This model was trained on 24 normal samples and 165 plasma samples from cancer patients, of which 20 were RCC.

To build a robust validation cohort to assess performance, we combined the 46 RCC samples generated in our lab with 10 normal samples processed in our lab and 62 normal samples previously generated for validation in Shen et al study[16].

Class probabilities were generated for our validation cohort, and then ROC curves were constructed to estimate performance both across the dataset, and separately for pathological stage I, II, III and metastatic samples. We also examined the distributions of class probabilities for both the normal samples and the tumours in our dataset.

8. RESULTS

Concentration of cfDNA in plasma samples of patients with renal cancer and healthy donors.

Plasma samples in RCC patients were collected from 1 to 109 days before nephrectomy (mean 26.5±28.8 days).

The results of the SYBR Green I fluorescence measurements showed that the concentrations of cfDNA in patients with RCC range from 1.95 to 260 ng/µL (mean 19.8±39.8 ng/µL). The range of cfDNA concentration in healthy donors was from 3.3 to 17.8 ng/mL (mean 7.9±3.4 ng/µL).

The amount of cfDNA was significantly higher in RCC patients compared to healthy donors (p=0.03).

Quality control of cfMeDIP-seq from cfDNA from patients with RCC and healthy controls.

Specificity of reaction was calculated using methylated and unmethylated spiked-in *A. thaliana DNA* for each sample. As we expected, the recovery of unmethylated spiked in *A. thaliana DNA* was less than 1% and the specificity of each reaction was more than 99% before the sequencing process.

cfMeDIP-seq analysis.

The elastic net models designed to detect cancer in contrast to normal samples, previously published in Shen paper[16] showed high performance in discriminating between cancers and normal samples across our validation cohort (AUROC = 0.89) (figure 4a). The ROC analysis comparing the samples of each pathological stage versus normal samples obtained similar AUROC values among the different pathologic stages (figure 4b). Indeed, as represented in the bar graphs in figure 5 and 6, all tumors, and only a fraction of normal samples, had class probabilities of close to 1 for being called malignant.

While applying a naïve probability threshold of 0.5 yields poor sensitivity at high specificity (Specificity = 1, sensitivity = 0.31) across the dataset, the high ROC values suggest that remarkably higher sensitivity should be possible with an optimised threshold, albeit with concomitant drops in specificity.

9. Discussion

The use of liquid biopsies for cancer detection and management is rapidly gaining prominence. In particular, the analysis of ctDNA has numerous potential clinical applications. However, certain settings, such as cancer screening and the detection of minimal residual disease after treatment, require a degree of analytical sensitivity that is often beyond current technical limits of mutation-based ctDNA detection methods. Current methods for the detection of ctDNA involve sequencing

somatic mutations using cfDNA, but the sensitivity of these methods may be low among patients with early-stage cancer given the limited number of recurrent mutations.

By contrast, large-scale epigenetic alterations, which are tissue- and cancer-type specific, are not similarly constrained and therefore potentially have greater ability to detect and classify cancers in patients with early-stage disease. Recent studies demonstrate that cfmeDNA may be useful biomarkers and there are many prospective studies of the clinical utility of cfmeDNA. However, a lot of challenges for cfDNA methylation analysis still remain, and the major challenge is that current approaches are essentially adapted to examining methylation in genomic DNA. Almost all of the cfDNA methylation analysis methods depend on bisulfite sequencing, which can cause a degree of DNA degradation. WGBS provides whole methylation information in cfDNA, but this

CfMeDIP-seq is a sensitive immunoprecipitation-based protocol to analyze the methylome of small quantities of cfDNA with the ability to detect large-scale DNA methylation changes that are enriched for tumor-specific patterns. The methodology also demonstrated robust performance in cancer detection and classification across an extensive collection of plasma samples from several tumor types in a cost-effective manner, reducing technical artefacts introduced during sequencing.

In our study, for the first time we tested the cfMeDIP-seq methodology in an independent set of 48 ccRCC plasma samples collected before RN at different pathologic stages – pT1, pT2, pT3 and metastatic disease. We showed that the methylation pattern of cfDNA is able to distinguish healthy patients from ccRCC patients. Then we clearly demonstrate the presence of ccRCC methylation pattern in the cfDNA of patients before RN at different stages, suggesting that cfDNA could be able to detect early stage of disease and might detect minimal residual disease after RN.

In particular, the elastic net models designed to detect cancer in contrast to normal samples, showed high performance in discriminating between cancers and normal samples across the validation cohort (AUROC = 0.89). The ROC analysis comparing the samples of each pathological stage versus normal samples obtained similar AUROC values among the different pathologic stages, suggesting the high sensitivity of the methodology across the different stages.

This is a new way to approach the study cfDNA. Most of the known methods have struggled to analyze ctDNA because of the sensitivity given to the limited number of recurrent mutations available to properly distinguish between tumor and normal cfDNA in a cost-effective manner and the technical artefacts as a result of sequencing.

A major limitation of the previous analysis is that we don't have a small number of healthy patients in our validation dataset. This prohibits an accurate estimation of both the sensitivity and specificity which need to be assessed to evaluate the clinical utility of our approach.

A strong point of our study is that the cfMeDIP-seq approach has a large methylation 'target size', because it is not focused on identify individually somatically acquired genetic alterations, but the whole methylation pattern, improving dramatically the sensitivity and dynamic range. Although the amount of cfDNA was significantly higher in RCC patients compared to healthy donors ($p=0.03$), cannot be considered a reliable diagnostic tool. In fact, the level of cfDNA is not sufficiently robust to distinguish between cancer and non-cancer, because it is influenced by concomitant diseases and other clinical factors, and therefore it.

In addition, our protocol requires 1 to 10 ng input DNA. The amount of DNA is very compare to the gold-standard MeDIP–seq (100 ng), reduced representation bisulfite sequencing (RRBS) (1,000 ng) and WGBS (2,000 ng).

Because the methylation status differs between normal and malignant tissues and it is tissue-specific, the cfMeDIP–seq technology could be used to detect various cancer. In fact, our research team at the Dana-Farber Cancer Institute, in collaboration with the De Carvalho laboratory has effectively tested the reproducibility of cfMeDIP–seq in a cohort of patients with muscle-invasive bladder cancer (MIBC). We selected 12 pts who underwent RC for MIBC - 6patients who had recurrent disease within 2-3yrs after radical cystectomy RC (C1) and 6 pts who did not (C2). 119 healthy pts without bladder cancer were controls. Also in this case, cfDNA was isolated from 1ml of plasma samples collected after RC and before recurrence (C1) or during the follow-up (C2) and was analyzed by the cfMeDIP-seq using 10ng input cfDNA. The amount of cfDNA isolated from 1ul of plasma was very similar to what we found in RCC cohorts, 13.1ng (6.4-19.7) in C1 and 17.1ng (13.6-21.2) in C2.

In this case, we identified approximately 137,000 peaks which were present at least in one sample and the supervised classification identified 61 DMR (FDR<0.050), predominantly located in

intergenic region, which distinguished C1 from C2. The randomized sample tests proved the discriminatory power of the identified set. Moreover, the supervised analysis comparing the status of the identified DMRs relative to healthy controls showed 28 regions were differentially methylated (logFC > +/- 1, FDR < 0.05), clearly demonstrating that cfmeDNA can be readily harvested from MIBC patients to detect cancer-specific methylation patterns and predict recurrence of MIBC post-RC.

10. Conclusion

Although cfMeDIP–seq method awaits further validation in completely independent datasets, our findings in RCC patients and the preliminary results in MIBC patients underscore the potential utility of cfDNA methylation profiles as a basis for non-invasive, cost-effective, sensitive and accurate early tumor detection for cancer interception, and for multi-cancer classification.

Table 1. American Joint Committee on Cancer (AJCC) TNM Staging System for Kidney Cancer 8th ed., 2017). [NCCN Guidelines Version 2.2019- Kidney Cancer].

T	Primary Tumor
TX	Primary tumor cannot be assessed
T0	No evidence of primary tumor
T1	Tumor ≤7 cm in greatest dimension, limited to the kidney
T1a	Tumor ≤4 cm in greatest dimension, limited to the kidney
T1b	Tumor >4 cm but ≤7 cm in greatest dimension, limited to the kidney
T2	Tumor >7 cm in greatest dimension, limited to the kidney
T2a	Tumor >7 cm but ≤10 cm in greatest dimension, limited to the kidney
T2b	Tumor >10 cm, limited to the kidney
T3	Tumor extends into major veins or perinephric tissues, but not into the ipsilateral adrenal gland and not beyond Gerota's fascia
T3a	Tumor extends into the renal vein or its segmental branches, or invades the pelvicalyceal system, or invades perirenal and/ or renal sinus fat but not beyond Gerota's fascia
T3b	Tumor extends into the vena cava below the diaphragm
T3c	Tumor extends into the vena cava above the diaphragm or invades the wall of the vena cava
T4	Tumor invades beyond Gerota's fascia (including contiguous extension into the ipsilateral adrenal gland)

N	Regional Lymph Nodes
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in regional lymph node(s)

M	Distant Metastasis
M0	No distant metastasis
M1	Distant metastasis

Table 2. **AJCC Prognostic Groups**

	T	N	M
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T1-T2	N1	M0
	T3	N0-N1	M0
Stage IV	T4	Any N	M0
	Any T	Any N	M1

Table 2. Summary of the studies evaluating cfDNA in RCC.

Study	Type of study	Purpose of the study	Patients (*n*)	Control (*n*)	Source Material	Isolated Method	Detection method
Goessl et al.[116], 1998	Prospective	Diagnostic	40 RCC	10 healthy individuals	Plasma, 1ml	Qiamp Blood Kit (Qiagen, Hilden, Germany).	Fluorescent PCR
Eisenberger et al.[111], 1999	Prospective	Diagnostic	25 RCC 1 AML 1 MN 3 OCT	8 individualsa with nephrolithiasis 8 healthy individuals	Serum, NR	Digestion with proteinase K (Boehringer Mannheim GmbH, Mannheim, Germany) in the presence of sodium dodecyl sulfate at 48 °C overnight, followed by phenol–chloroform extraction and ethanol precipitation.	PCR
Knobloch et al.[112], 2002	Prospective	Diagnosic	53 RCC 1 renal B cell lymphoma 6 TCC	20 healthy individuals	Serum, 2-4ml	Qiamp Midi-Kit (Qiagen, Hilden, Germany)	Fluorescent PCR
Hoque et al.[113], 2005	Prospective	Diagnostic	18 RCC	30 healthy individuals	Serum, nr	Digestion with 50 µg/ml proteinase K (Boehringer, Mannheim, Germany) in the presence of 1% SDS at 48°C overnight, followed by phenol/chloroform extraction and ethanol precipitation.	Fluorescence-based RT-PCR (Taqman)
Perego et al.[107], 2008	Prospective	Diagnostic, Prognostic	48 RCC 1 TCC 5 OCT	41 healthy individuals	Plasma, 1ml	QIAamp DNA Mini kit (Qiagen, Italy)	qRT-PCR (β-globin)
Hauser et al.[110], 2010	Prospective	Diagnostic, Prognostic	35 RCC	54 healthy individuals	Serum, 1ml	ChargeSwitch gDNA Kit (Invitrogen, Paisley,Scotland)	qRT-PCR (actin-β)
Feng et al.[148], 2010	Prospective	Diagnostic	78 RCC	42 healthy individuals	Serum, 400µl	QIAamp DNA Blood Mini Kit (Qiagen, Courtaboeuf,France)	qRT-PCR (actin-β)
De Martino et al.[106], 2011	Prospective	Diagnostic Prognostic	157 RCC	43 benign renal tumors	Serum, 1ml	QIAamp Ultrasens Virus kit (Qiagen, Hilden, Germany)	qRT-PCR (RNF185)
Hauser et al.[110], 2013	Prospective	Diagnostic	35 RCC	54 healthy individuals	Serum,1ml	ChargeSwitch gDNA Kit (Invitrogen, Paisley,Scotland)	qRT-PCR (actin-β)
Skrypkina et al.[114], 2016	Prospective	Diagnostic	27 RCC	15 healthy individuals	Plasma, 2ml	Proba Na kit (DNA-Technology, Russia)	qRT-PCR (actin-β)

AML angiomyolipoma; MN metanephric nephroma; NR, not reported; OCT oncocytoma; RCC, renal cell carcinoma; TCC, transitional cell carcinoma.

Table 3. Diagnostic information of microsatellite alterations (loss of heterozygosity and/or microsatellite instability) in RCC.

Study	Microsatellite markers	Patients (*n*)	Control (*n*)	Sensitivity	Specificity
Goessl et al.[116], 1998	D3S1307(3p), D3SI560(3p), D3SI289(3p), D3SI300(3p)	40 RCC	10 healthy individuals	63% (at least one MA) 35% (more than one MA)	100%
Eisemberger et al.[111], 1999	D1S251 (1pq), HTPO(2p), D3S1317(3p), D3S587(3p), D3S1560(3p), D3S1289(3p), D3S1286 (3p), D3S1038(3p) , D4S243(4pq), FGA(4)(4q), CSF(5q), ACTBP2(5p), D8S348(8q), D8S307(8p), D9S747(9p), D9S242(9p), IFNa(9p), D9S162(9p), D11S488(11q), THO(11p), vWA(12p), D13S802(13q), MJD(14q), D17S695(17p), D17S654(17p), D18S51(18q), MBP(18q), D21S1245(21q).	25 RCC 1 AML 1 MN 3 OCT	8 individualsa with nephrolithiasis 8 healthy individuals	60% (at least one MA)	100%
Knobloch et al.[112], 2002	D3S1560(3p), D3S2450(3p), D3S3666(3p), D3S2408(3p), D3S1259(3p), D5S1720(5p), D5S1480(5p), D5S476(5p), D5S818(5p), D7S1796(7p), D7S1807(7p), D8S261(8p), D8S560(8p), D9S925(9p), D13S153(13p), D17S799(17p), D17S1306 (17p), D17S783(17p), D17S1298(17p), D17S807(17p)	53 RCC 1 renal B cell lymphoma 6 TCC	20 healthy individuals	74% (using 9 MA) 87% (using 20 MA)	85%
Perego et al.[107], 2008	D3S1566(3p), D3S1285(3p), D3S1300(3p), D3S1289(3p), D3S1597(3p)	48 RCC 1 TCC 5 OCT	41 healthy individuals	55.6% (at least one MA)*	NR

AML angiomyolipoma; MN metanephric nephroma; NR, not reported; OCT oncocytoma; RCC, renal cell carcinoma; TCC, transitional cell carcinoma; MA, microsatellite alterations (loss of heterozygosity and/or microsatellite instability)
*MA was possible only for 9 patients whose preoperative plasma DNA was available.

Table 4. Diagnostic information of cfDNA methylation in RCC and controls.

Study	Markers	Number of methylation positive/number of total RCC patients (%)	Number of methylation positive/number of total controls case (%)	Sensitivity (%)	Specificity (%)	Cutoff value	AUC	Source	Detection method
Hoque et al.[113], 2005	***APC***	1/18 (5.5)	1/30 (3.3)	5.5	96.7	4.5	NR	Serum,NR	Fluorescence-based real-time PCR (Taqman)
	ARF	1/18 (5.5)	1/30 (3.3)	5.5	96.7	0	NR		
	CDH1	6/18 (33.3)	2/30 (6.6)	33.3	93.4	0.3	NR		
	GSTP1	1/18 (5.5)	0/30 (0)	5.5	100	0	NR		
	MGMT	0/18 (0)	1/30 (3.3)	0	96.7	0	NR		
	p16	4/18 (22.2)	0/30 (0)	22.2	100	0	NR		
	RAR-82	1/18 (5.5)	0/30 (0)	5.5	100	0.1	NR		
	RASSF1A	2/18 (11.1)	1/30 (3.3)	11.1	96.7	0.1	NR		
	TIMP3	3/18 (16.6)	0/30 (0)	17	100	1	NR		
De Martino et al.[106], 2011	***RASSF1A***	72/157 (45.9)	3/43 (7)	45.9	93	0	0.694	Serum, 1ml	Restriction endonuclease q-PCR
	PTGS2	60/157 (38.2)	15/43 (34.9)	38.2	65.1	0	0.517		
	P16	73/157 (46.5)	19/43 (44.2)	46.5	55.8	0	0.512		
	VHL	79/157 (50.3)	4/43 (8.3)	50.3	90.7	0	0.705		
Hauser et al.[115], 2013	***APC***	19/35(54.3)	5/54 (9.3)	54.3	90.7	0.37	0.72	Serum, 1ml	Methylation-sensitive restriction enzymes
	GSTP1	6/35 (17.1)	1/54 (1.9)	17.1	98.1	0.75	0.57		
	p14(ARF)	5/35 (14.3)	0/54 (0)	14.3	100	0.26	0.57		
	P16	9/35 (25.7)	9/54 (16.7)	25.7	83.3	0	NR		
	PTGS2	8/35 (22.9)	2/54 (3.7)	22.9	96.3	0.47	0.59		
	RAR-B	14/35 (40)	8/54 (14.8)	40.0	85.2	0.19	0.61		
	RASSF1A	8/35 (22.9)	1/54 (1.9)	22.9	98.2	0.09	0.60		
	TIMP3	20/35 (57.1)	21/54 (38.9)	57.1	61.1	0	NR		
Skrypkina et al.[114], 2016	**APC**	14/27 (51.9)	1 /15(6.7)	51.9	93.3	0	NR	Plasma, 2ml	Methylation-specific polymerase chain reaction (MS-PCR)
	FHIT	15/27 (55.6)	0/15 (0)	55.6	100	0	NR		
	ITGA9	0 /27(0)	0/15 (0)	0	100	0	NR		
	LRRC3B	20/27 (74.1)	5/15 (33.3)	74.1	66.7	0	NR		
	RASSF1	17/27 (63.0)	1/15 (6.7)	62.9	93.3	0	NR		
	VHL	0/27 (0)	0/15 (0)	0	100	0	NR		

APC, adenomatosis-poliposis-coli gene; ARF, ARF tumor suppressor protein gene; CDH1, cadherin-1 gene; FHIT, fragile histidine triad gene; GSTP1, gluthation-a-transferase-protein 1 gene; ITGA9, integrin subunit alpha 9 gene; LRRC3B , leucine rich repeat containing 3B gene; MGMT, O-6-methylguanine-DNA methyltransferase gene; NR, not reported; p16, cyclin-dependent kinase inhibitor 2A; PTGS2, prostaglandin-endoperoxidase synthase; RAR-82, retinoic acid receptor beta; RCC, renal cell carcinoma; RAR-B retinoid-acid-receptor-beta gene; RASSF1A Ras association domain family member 1A; TIMP3, tissue inhibitor of metalloproteinase-gene, VHL, von Hippel-Lindau.

Table 5. Main characteristics of study patients (N = 46)

Variable	Subjects, n. (n=46)
Age (years)	
≤65	35
>65	11
Gender	
Male	33
Female	13
Race	
White	53
Hispanic	1
pT Category	
pT1	25
pT2	7
pT3	6
M1	8
Fuhrman Grade	
G1	2
G2	24
G3	21
G4	5
Hystological type	
RCCcc	46
Surgery	
Radical nephrectomy	36
Partial nephrectomy	13

Figure. 1. The dynamic regulation of DNA methylation and aberrant DNA methylation in cancer. DNA methyltransferases (DNMTs) convert unmodified cytosine to 5-methylcytosine (5 mC). Ten-eleven translocation (TET) enzymes can oxidize 5 mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). Thymine DNA glycosylase (TDG)/base excision repair (BER) pathway excises 5fC/5caC and regenerates unmodified cytosine. In cancer cells, the whole genome levels of 5mC and 5hmC are decreased.

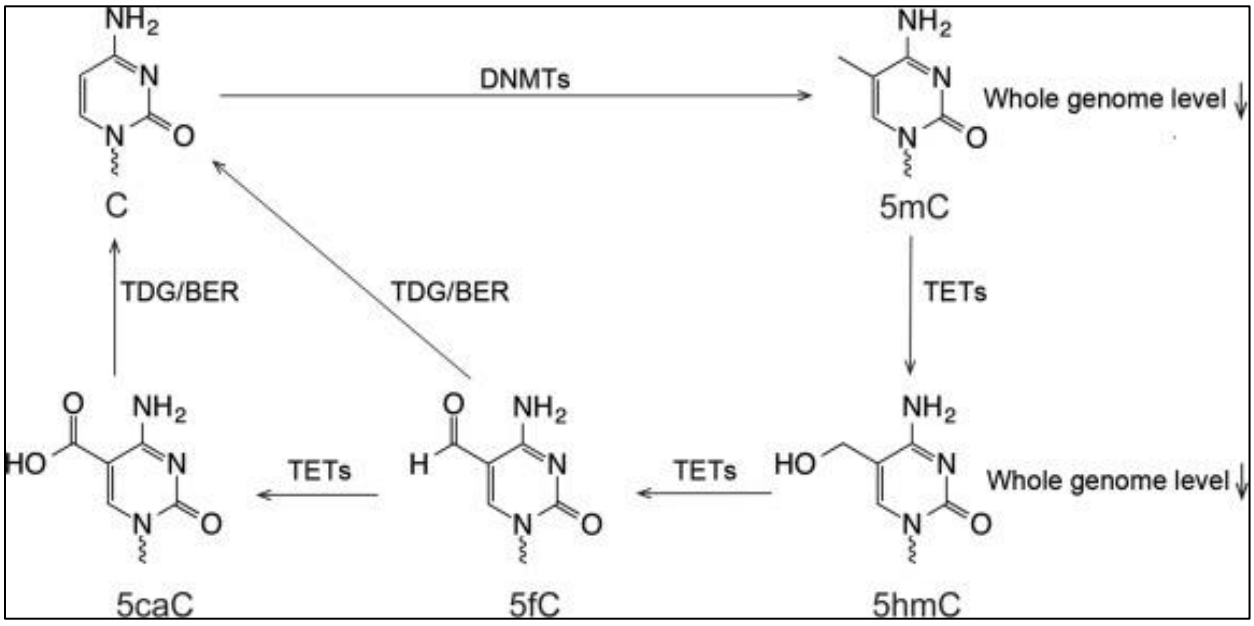
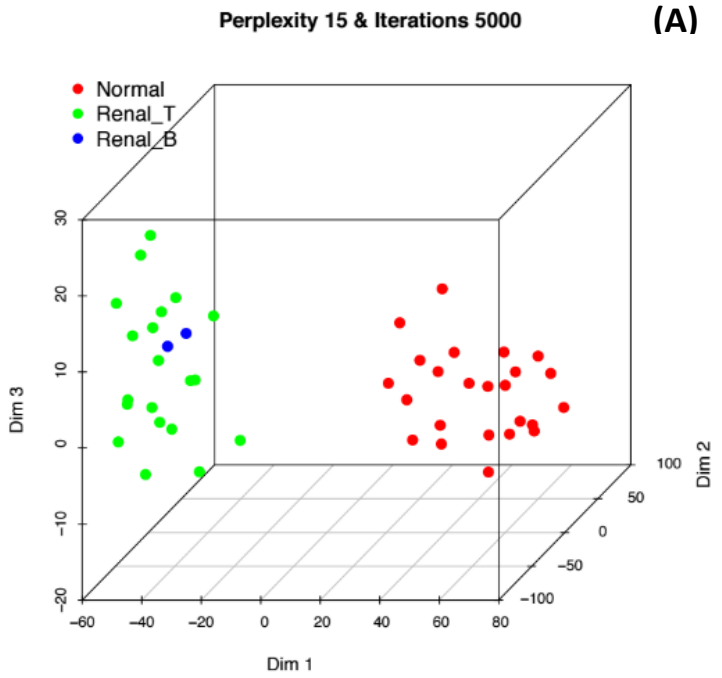


Figure 2. Differentially methylated regions (DMRs) were identified by comparing: the two Renal_B (Boston/blue dots) samples to a set of 24 normal samples (red) and the 20 Renal_T (Toronto/green) to normal **(A)** and the six Renal_B (Boston/blue dots) samples to a set of 24 normal samples (red) and the 20 Renal_T (Toronto/green) to normal **(B).** Using a dimensional reduction technique, t-distributed stochastic neighbor embedding (t-SNE), the samples can be distinguished and co-cluster with the Toronto samples

.



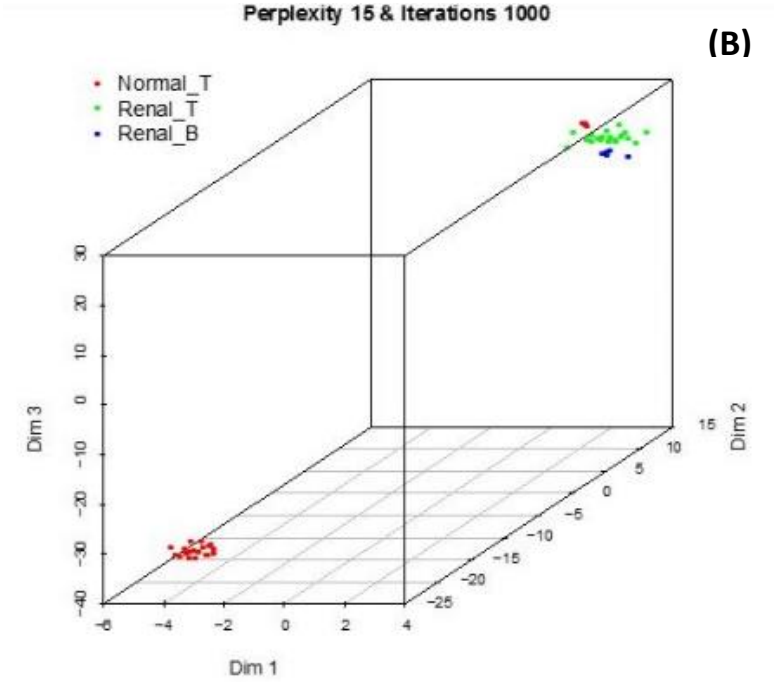


Figure. 3. Schematic representation of the cfMeDIP–seq protocol.

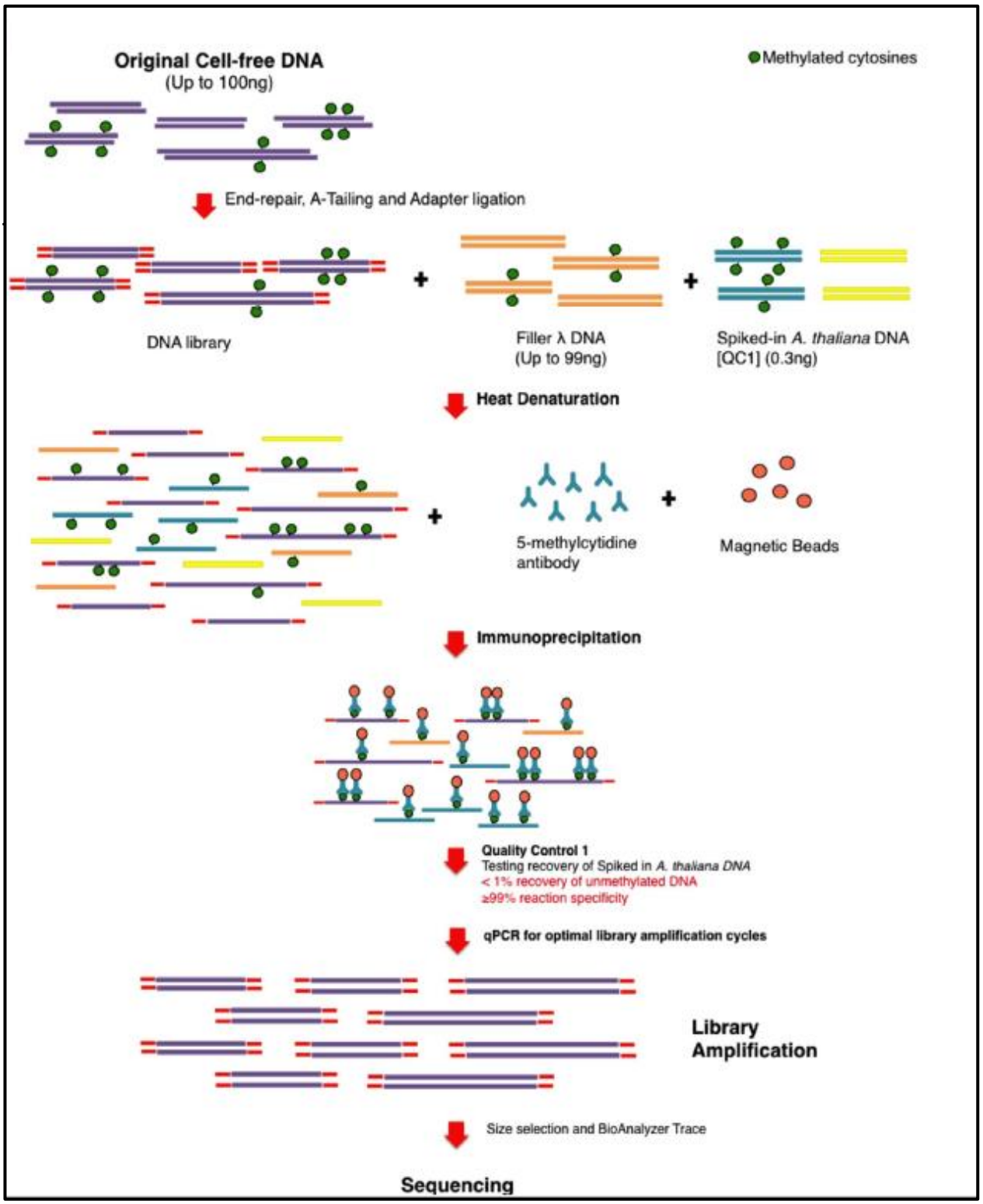
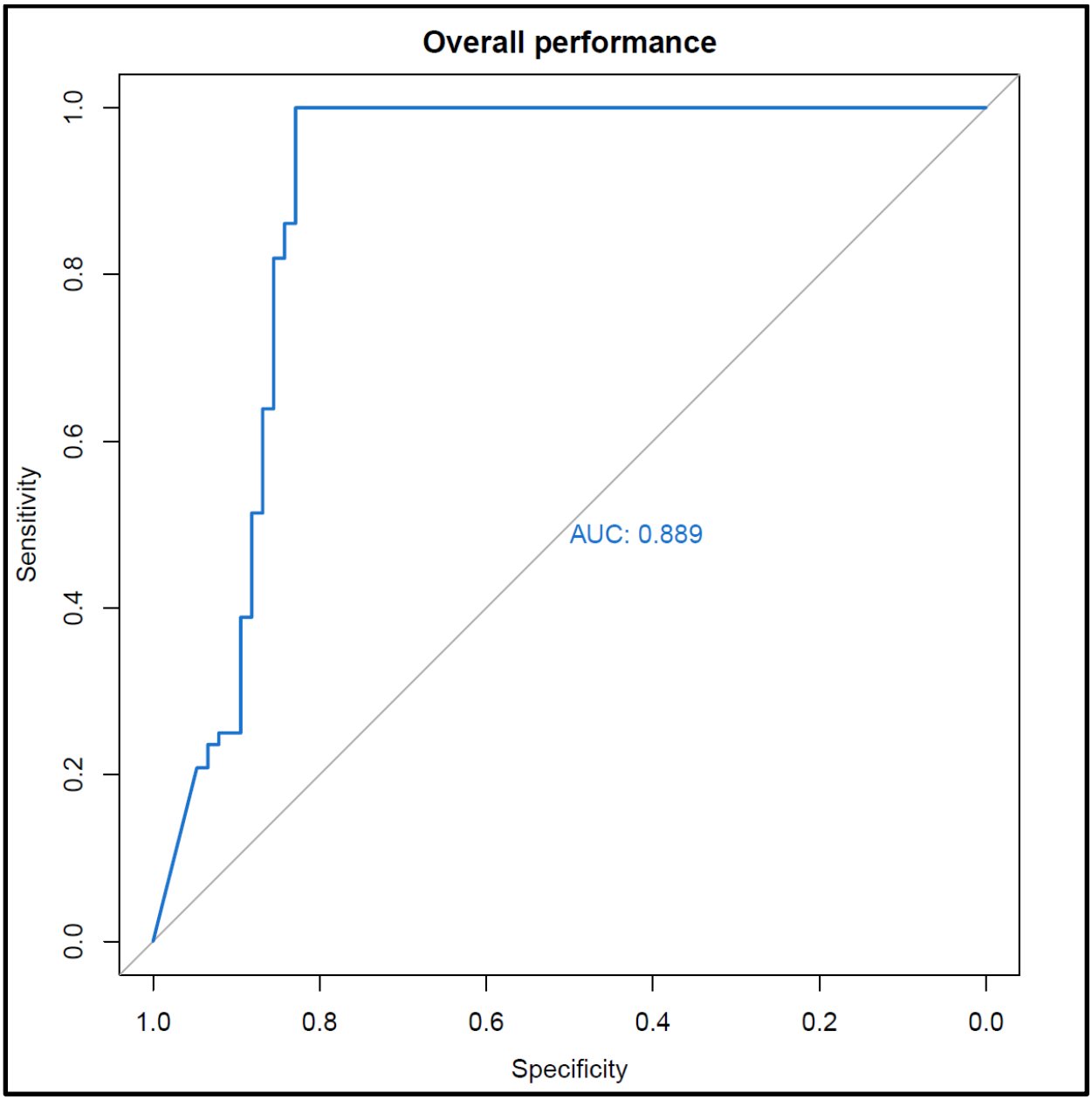
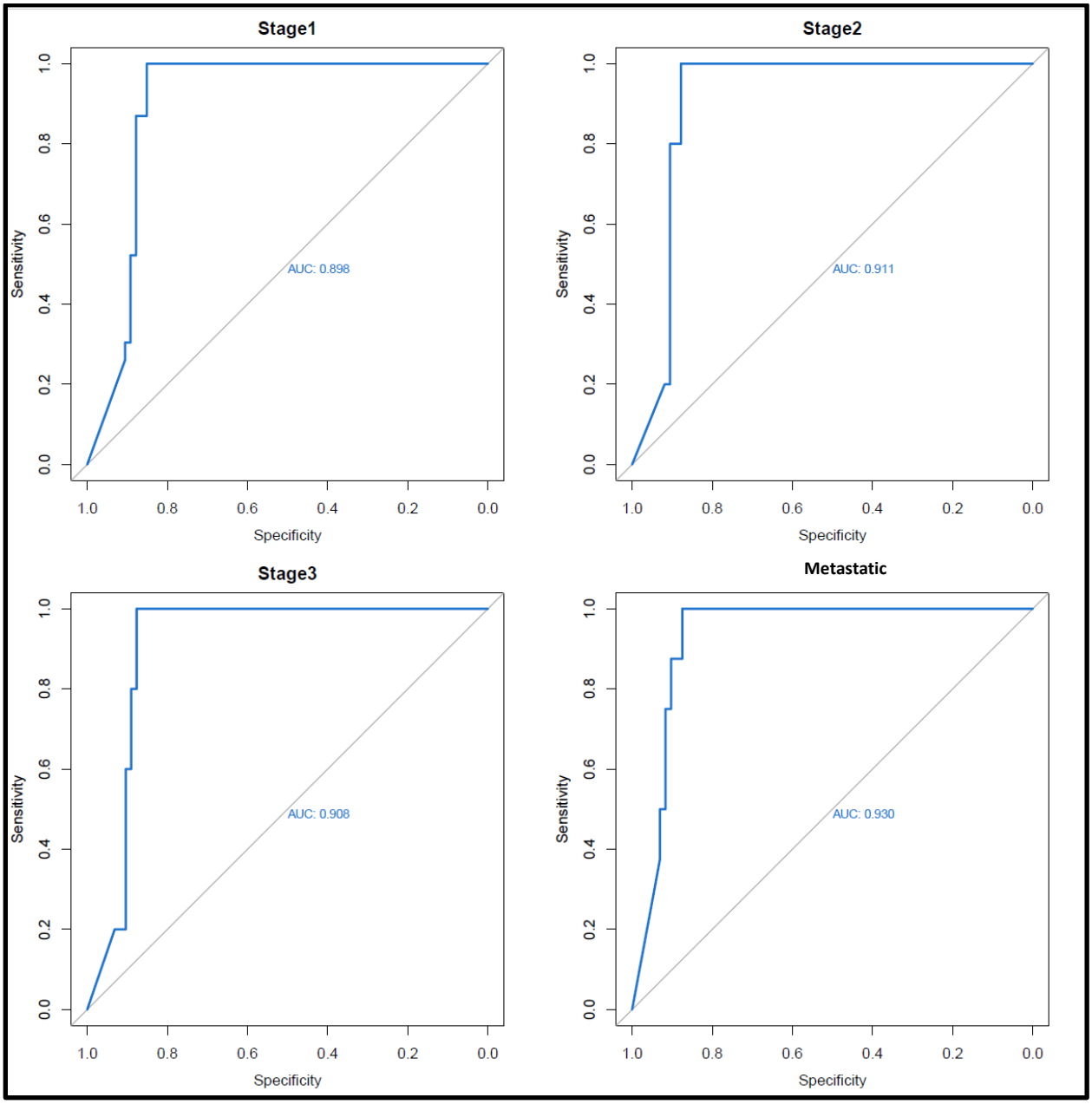


Figure 4. ROC curves – y axis = sensitivity, x axis = specificity. Curves were constructed to distinguish between normal and RCC samples in a cohort of 118 samples. AUROC values are printed (a). Further stratification was also performed to evaluate performance for samples by stage (b).

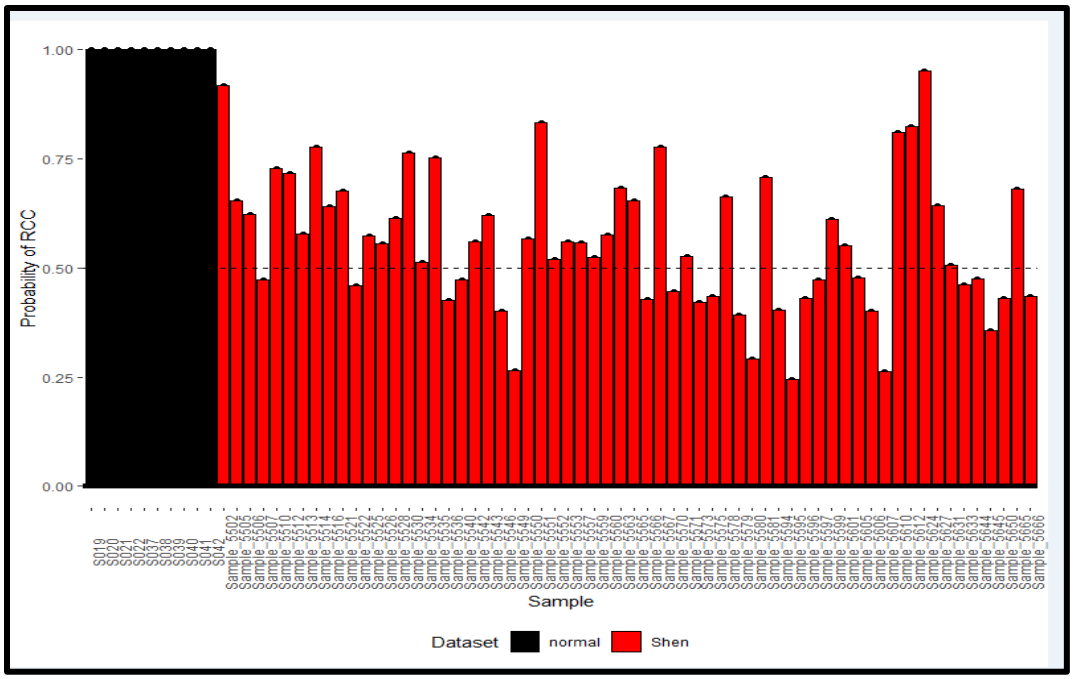
(a)



(b)


Stage1
AUC: 0.898
Stage2
AUC: 0.911
Stage3
AUC: 0.908
Metastatic
AUC: 0.930
Sensitivity
Specificity
1.0
0.8
0.6
0.4
0.2
0.0

Figured 5. Barplots – y axis = probability of the sample being classified as malignant based on an ensemble of classifiers trained on 20 RCC and 24 normals in Shen paper[16]. The barplots for the tumour sample are coloured by stage, and the barplots of the normal samples are coloured by the source of the data.



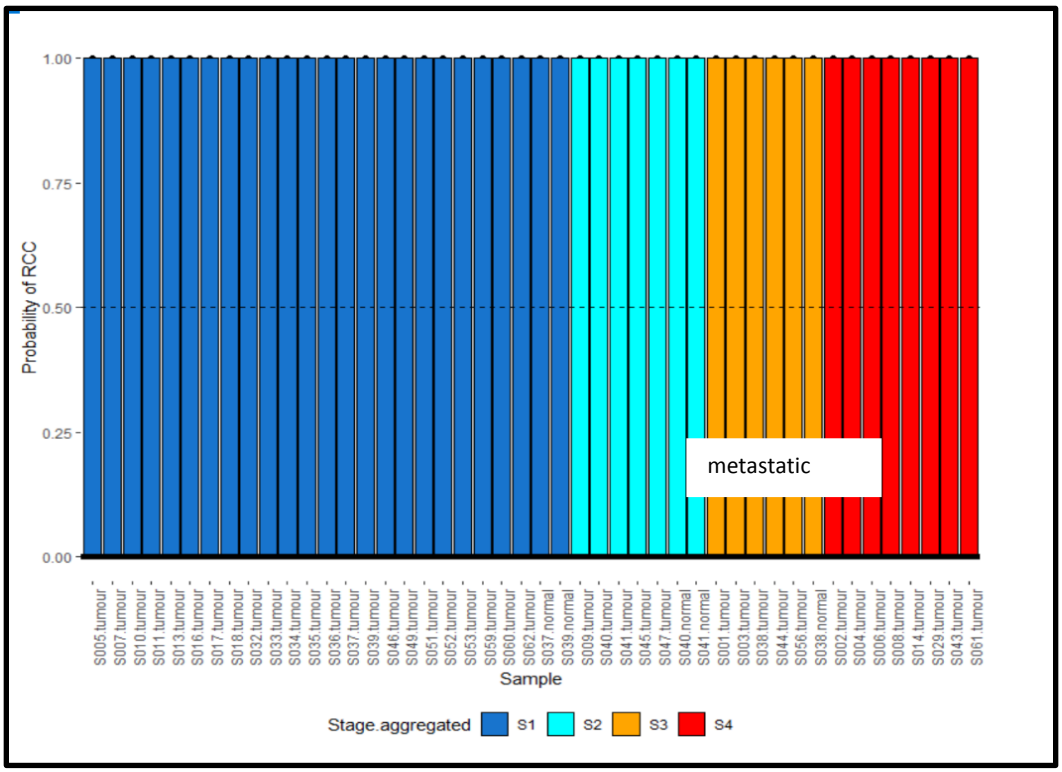
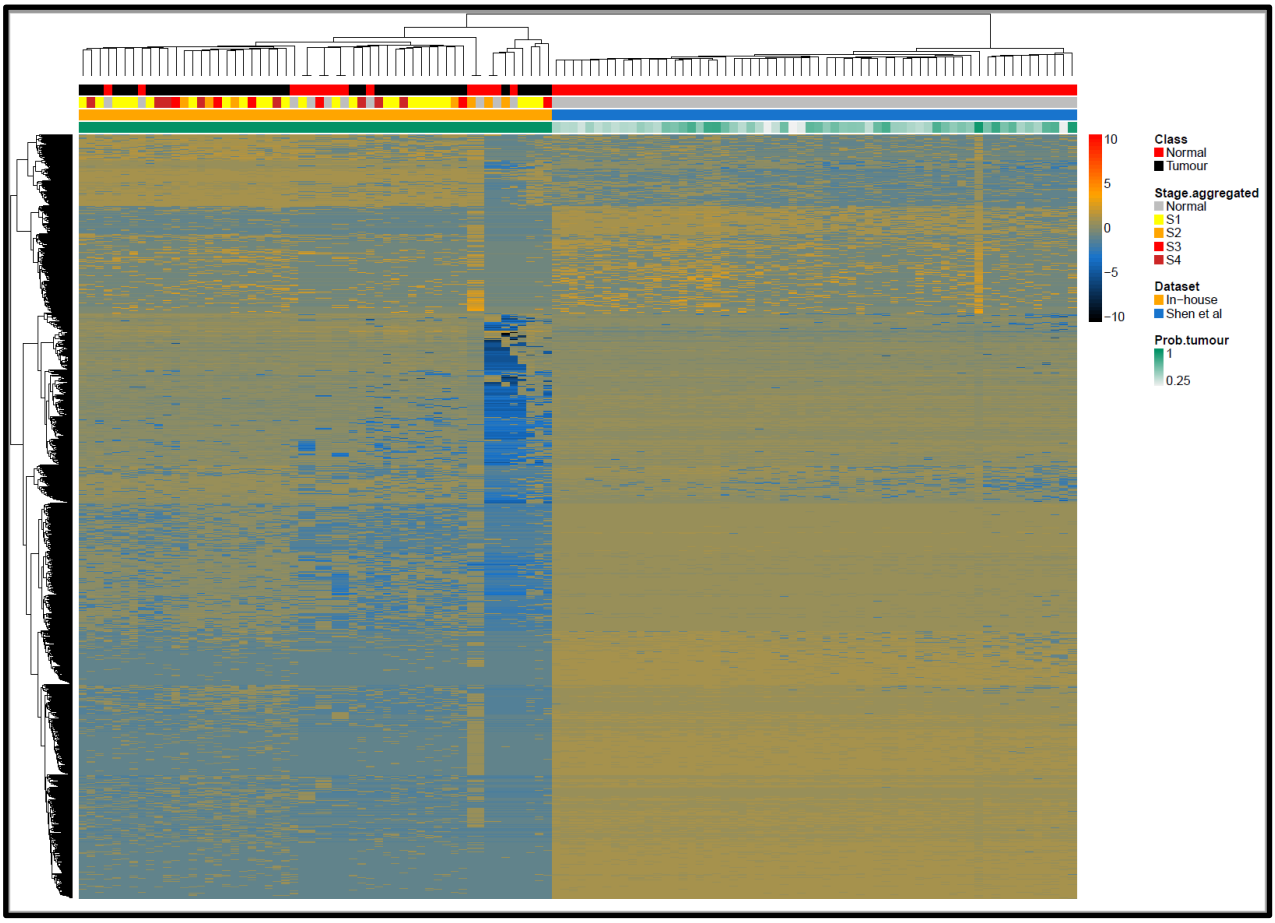


Figure 6. Heatmap. Heatmap shows all windows that were originally selected during the training of an ensemble classifier to discriminate between normal and cancer samples in a 189 sample discovery cohort consisting of eight classes, previously described and published in Shen paper[16]. Rows represent features, columns represent samples. Annotation ribbons highlight cancer/normal status, the probability allocated by the classifier, and the source of the data, respectively. The signal is plotted in the form of Z-scores of log2(counts per million) values.



9. REFERENCES

1. Siegel, R.L., Miller, K.D. & Jemal, A. Cancer statistics, 2018. *CA: A Cancer Journal for Clinicians* **68**, 7-30 (2018).
2. Ferlay, J. *et al.* Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer* **136**, E359-E386 (2014).
3. Luciani, L.G., Cestari, R. & Tallarigo, C. Incidental renal cell carcinoma-age and stage characterization and clinical implications: study of 1092 patients (1982-1997). *Urology* **56**, 58-62 (2000).
4. Jayson, M. & Sanders, H. Increased incidence of serendipitously discovered renal cell carcinoma. *Urology* **51**, 203-205 (1998).
5. DeCastro, G.J. & McKiernan, J.M. Epidemiology, Clinical Staging, and Presentation of Renal Cell Carcinoma. *Urologic Clinics of North America* **35**, 581-592 (2008).
6. Berger, A. *et al.* Laparoscopic Radical Nephrectomy for Renal Cell Carcinoma: Oncological Outcomes at 10 Years or More. *Journal of Urology* **182**, 2172-2176 (2009).
7. Chung, S.-D. *et al.* Long-Term Follow-up of Hand-Assisted Laparoscopic Radical Nephrectomy for Organ-Confined Renal Cell Carcinoma. *Urology* **69**, 652-655 (2007).
8. Eggener, S.E. *et al.* Renal Cell Carcinoma Recurrence After Nephrectomy for Localized Disease: Predicting Survival From Time of Recurrence. *Journal of Clinical Oncology* **24**, 3101-3106 (2006).
9. Konety, B.R. Postoperative surveillance protocol for patients with localized and locally advanced renal cell carcinoma based on a validated prognostic nomogram and risk group stratification system. *Urologic Oncology: Seminars and Original Investigations* **24**, 87-88 (2006).
10. Mekhail, T.M. *et al.* Validation and Extension of the Memorial Sloan-Kettering Prognostic Factors Model for Survival in Patients With Previously Untreated Metastatic Renal Cell Carcinoma. *Journal of Clinical Oncology* **23**, 832-841 (2005).
11. Schwarzenbach, H., Hoon, D.S.B. & Pantel, K. Cell-free nucleic acids as biomarkers in cancer patients. *Nature Reviews Cancer* **11**, 426-437 (2011).
12. Wan, J.C.M. *et al.* Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nature Reviews Cancer* **17**, 223-238 (2017).
13. Esteller, M. Epigenetics in cancer. *N Engl J Med* **358**, 1148-59 (2008).
14. Roadmap Epigenomics, C. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317 (2015).
15. Slieker, R.C. *et al.* Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. *Epigenetics & Chromatin* **6**, 26 (2013).
16. Shen, S. *et al.* Sensitive tumor detection and classification using plasma cell-free DNA methylomes. *Nature* **(in press)**(2018).
17. Taiwo, O. *et al.* Methylome analysis using MeDIP-seq with low DNA concentrations. *Nature Protocols* **7**, 617 (2012).
18. Znaor, A., Lortet-Tieulent, J., Laversanne, M., Jemal, A. & Bray, F. International Variations and Trends in Renal Cell Carcinoma Incidence and Mortality. *European Urology* **67**, 519-530 (2015).

19. Ferlay, J. *et al.* Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. *European Journal of Cancer* **49**, 1374-1403 (2013).
20. Thompson, R.H. *et al.* Renal Cell Carcinoma in Young and Old Patients—Is There a Difference? *Journal of Urology* **180**, 1262-1266 (2008).
21. Pantuck, A.J., Zisman, A. & Belldegrun, A.S. THE CHANGING NATURAL HISTORY OF RENAL CELL CARCINOMA. *Journal of Urology* **166**, 1611-1623 (2001).
22. Hollingsworth, J.M., Miller, D.C., Daignault, S. & Hollenbeck, B.K. Rising Incidence of Small Renal Masses: A Need to Reassess Treatment Effect. *JNCI: Journal of the National Cancer Institute* **98**, 1331-1334 (2006).
23. Tsivian, M., Moreira, D.M., Caso, J.R., Mouraviev, V. & Polascik, T.J. Cigarette Smoking Is Associated With Advanced Renal Cell Carcinoma. *Journal of Clinical Oncology* **29**, 2027-2031 (2011).
24. Hidayat, K., Du, X., Zou, S.-Y. & Shi, B.-M. Blood pressure and kidney cancer risk. *Journal of Hypertension* **35**, 1333-1344 (2017).
25. Adams, K.F. *et al.* Body Size and Renal Cell Cancer Incidence in a Large US Cohort Study. *American Journal of Epidemiology* **168**, 268-277 (2008).
26. Choi, Y. *et al.* Body mass index and survival in patients with renal cell carcinoma: A clinical-based cohort and meta-analysis. *International Journal of Cancer* **132**, 625-634 (2012).
27. Albiges, L. *et al.* Body Mass Index and Metastatic Renal Cell Carcinoma: Clinical and Biological Correlations. *Journal of Clinical Oncology* **34**, 3655-3663 (2016).
28. Truong, L.D., Krishnan, B., Cao, J.T.H., Barrios, R. & Suki, W.N. Renal neoplasm in acquired cystic kidney disease. *American Journal of Kidney Diseases* **26**, 1-12 (1995).
29. Mandel, J.S. *et al.* International renal-cell cancer study. IV. Occupation. *International Journal of Cancer* **61**, 601-605 (1995).
30. Cho, E. Prospective Evaluation of Analgesic Use and Risk of Renal Cell Cancer. *Archives of Internal Medicine* **171**, 1487 (2011).
31. Gnarra, J.R., Lerman, M.I., Zbar, B. & Linehan, W.M. Molecular Studies of the VHL Tumor Suppressor Gene in Renal Cell Carcinoma. in *Biology of Renal Cell Carcinoma* 33-39 (Springer New York, 1995).
32. Cohen, A.J. *et al.* Hereditary Renal-Cell Carcinoma Associated with a Chromosomal Translocation. *New England Journal of Medicine* **301**, 592-595 (1979).
33. Zucchetto, A. *et al.* History of treated hypertension and diabetes mellitus and risk of renal cell cancer. *Annals of Oncology* **18**, 596-600 (2006).
34. Torres, V.E., Keith, D.S., Offord, K.P., Kon, S.P. & Wilson, D.M. Renal ammonia in autosomal dominant polycystic kidney disease. *Kidney International* **45**, 1745-1753 (1994).
35. DellaValle, C.T. *et al.* Dietary intake of nitrate and nitrite and risk of renal cell carcinoma in the NIH-AARP Diet and Health Study. *British Journal of Cancer* **108**, 205-212 (2012).
36. Lindblad, P. *et al.* International renal-cell cancer study. V. Reproductive factors, gynecologic operations and exogenous hormones. *International Journal of Cancer* **61**, 192-198 (1995).
37. Bellocco, R. *et al.* Alcohol drinking and risk of renal cell carcinoma: results of a meta-analysis. *Annals of Oncology* **23**, 2235-2244 (2012).

38. Kabat, G.C., Silvera, S.A.N., Miller, A.B. & Rohan, T.E. A cohort study of reproductive and hormonal factors and renal cell cancer risk in women. *British Journal of Cancer* **96**, 845-849 (2007).
39. Störkel, S. & van den Berg, E. Morphological classification of renal cancer. *World Journal of Urology* **13**, 153-158 (1995).
40. Patard, J.-J. *et al.* Prognostic Value of Histologic Subtypes in Renal Cell Carcinoma: A Multicenter Experience. *Journal of Clinical Oncology* **23**, 2763-2771 (2005).
41. Zisman, A. *et al.* Improved Prognostication of Renal Cell Carcinoma Using an Integrated Staging System. *Journal of Clinical Oncology* **19**, 1649-1657 (2001).
42. Yao, M. VHL Tumor Suppressor Gene Alterations Associated With Good Prognosis in Sporadic Clear-Cell Renal Carcinoma. *CancerSpectrum Knowledge Environment* **94**, 1569-1575 (2002).
43. Gnarra, J.R. *et al.* Mutations of the VHL tumour suppressor gene in renal carcinoma. *Nature Genetics* **7**, 85-90 (1994).
44. Nickerson, M.L. *et al.* Improved Identification of von Hippel-Lindau Gene Alterations in Clear Cell Renal Tumors. *Clinical Cancer Research* **14**, 4726-4734 (2008).
45. Skinner, D.G., Colvin, R.B., Vermillion, C.D., Pfister, R.C. & Leadbetter, W.F. Diagnosis and management of renal cell carcinomaA clinical and pathologic study of 309 cases. *Cancer* **28**, 1165-1177 (1971).
46. Gudbjartsson, T. *et al.* Effect of incidental detection for survival of patients with renal cell carcinoma: Results of population-based study of 701 patients. *Urology* **66**, 1186-1191 (2005).
47. Aldaabil, R. & Peereboom, D. Paraneoplastic Syndromes in Renal Cell Carcinoma. in *Renal Cell Carcinoma* 135-145 (Humana Press).
48. Johnson, C.D., Dunnick, N.R., Cohan, R.H. & Illescas, F.F. Renal adenocarcinoma: CT staging of 100 tumors. *American Journal of Roentgenology* **148**, 59-63 (1987).
49. Rini, B.I. *et al.* Kidney. in *AJCC Cancer Staging Manual* 747-756 (Springer International Publishing, 2016).
50. Kay, F.U. *et al.* Diagnostic Performance and Interreader Agreement of a Standardized MR Imaging Approach in the Prediction of Small Renal Mass Histology. *Radiology* **287**, 543-553 (2018).
51. Sánchez-Ortiz, R.F. *et al.* A renal mass in the setting of a nonrenal malignancy. *Cancer* **101**, 2195-2201 (2004).
52. Dechet, C.B. *et al.* Prospective Analysis of Computerized Tomography and Needle Biopsy With Permanent Sectioning to Determine the Nature of Solid Renal Masses in Adults. *Journal of Urology* **169**, 71-74 (2003).
53. Koga, S. *et al.* THE DIAGNOSTIC VALUE OF BONE SCAN IN PATIENTS WITH RENAL CELL CARCINOMA. *The Journal of Urology*, 2126-2128 (2001).
54. Studer, U.E. *et al.* Enlargement of Regional Lymph Nodes in Renal Cell Carcinoma is Often not Due to Metastases. *Journal of Urology* **144**, 243-245 (1990).
55. Semelka, R.C. *et al.* Renal cancer staging: Comparison of contrast-enhanced CT and gadolinium-enhanced fat-suppressed spin-echo and gradient-echo MR imaging. *Journal of Magnetic Resonance Imaging* **3**, 597-602 (1993).

56. H, W. *et al.* Comparing whole body 18F-2-deoxyglucose positron emission tomography and technetium-99m methylene diphosphate bone scan to detect bone metastases in patients with renal cell carcinomas - a preliminary report. *Journal of Cancer Research and Clinical Oncology* **128**, 503-506 (2002).
57. Mandel, P. & Metais, P. [Not Available]. *C R Seances Soc Biol Fil* **142**, 241-3 (1948).
58. Leon, S.A., Shapiro, B., Sklaroff, D.M. & Yaros, M.J. Free DNA in the serum of cancer patients and the effect of therapy. *Cancer Res* **37**, 646-50 (1977).
59. Stroun, M. *et al.* Neoplastic Characteristics of the DNA Found in the Plasma of Cancer Patients. *Oncology* **46**, 318-322 (1989).
60. Tamkovich, S.N. *et al.* Circulating DNA and DNase Activity in Human Blood. *Annals of the New York Academy of Sciences* **1075**, 191-196 (2006).
61. Peters, D.L. & Pretorius, P.J. Origin, translocation and destination of extracellular occurring DNA — A new paradigm in genetic behaviour. *Clinica Chimica Acta* **412**, 806-811 (2011).
62. Holdenrieder, S. *et al.* Nucleosomes in serum of patients with benign and malignant diseases. *International Journal of Cancer* **95**, 114-120 (2001).
63. Tickner, J.A., Urquhart, A.J., Stephenson, S.-A., Richard, D.J. & Oâ€™Byrne, K.J. Functions and Therapeutic Roles of Exosomes in Cancer. *Frontiers in Oncology* **4**(2014).
64. Gahan, P.B. & Stroun, M. The virtosome-a novel cytosolic informative entity and intercellular messenger. *Cell Biochemistry and Function* **28**, 529-538 (2010).
65. Rykova, E.Y. *et al.* Cell-free and cell-bound circulating nucleic acid complexes: mechanisms of generation, concentration and content. *Expert Opinion on Biological Therapy* **12**, S141-S153 (2012).
66. Laktionov, P.P. *et al.* Cell-Surface-Bound Nucleic Acids: Free and Cell-Surface-Bound Nucleic Acids in Blood of Healthy Donors and Breast Cancer Patients. *Annals of the New York Academy of Sciences* **1022**, 221-227 (2004).
67. Suzuki, N., Kamataki, A., Yamaki, J. & Homma, Y. Characterization of circulating DNA in healthy human plasma. *Clinica Chimica Acta* **387**, 55-58 (2008).
68. Stroun, M., Lyautey, J., Lederrey, C., Olson-Sand, A. & Anker, P. About the possible origin and mechanism of circulating DNA. *Clinica Chimica Acta* **313**, 139-142 (2001).
69. Diehl, F. *et al.* Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proceedings of the National Academy of Sciences* **102**, 16368-16373 (2005).
70. Diehl, F. *et al.* Circulating mutant DNA to assess tumor dynamics. *Nature Medicine* **14**, 985-990 (2008).
71. Bettegowda, C. *et al.* Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. *Science Translational Medicine* **6**, 224ra24 (2014).
72. Jahr, S. *et al.* DNA Fragments in the Blood Plasma of Cancer Patients: Quantitations and Evidence for Their Origin from Apoptotic and Necrotic Cells. *Cancer Research* **61**, 1659 (2001).
73. Schwarzenbach, H. *et al.* Detection of tumor-specific DNA in blood and bone marrow plasma from patients with prostate cancer. *International Journal of Cancer* **120**, 1465-1471 (2007).
74. Opatrna, S., Wirth, J., Korabecna, M. & Sefrna, F. Cell-Free Plasma DNA during Hemodialysis. *Renal Failure* **31**, 475-480 (2009).

75. Beck, J. *et al.* Serum DNA Motifs Predict Disease and Clinical Status in Multiple Sclerosis. *The Journal of Molecular Diagnostics* **12**, 312-319 (2010).
76. Rainer, T.H. Plasma DNA, prediction and post-traumatic complications. *Clinica Chimica Acta* **313**, 81-85 (2001).
77. Atamaniuk, J., Kopecky, C., Skoupy, S., Saemann, M.D. & Weichhart, T. Apoptotic cell-free DNA promotes inflammation in haemodialysis patients. *Nephrology Dialysis Transplantation* **27**, 902-905 (2011).
78. Gornik, I., Wagner, J., Gašparović, V., Lauc, G. & Gornik, O. Free serum DNA is an early predictor of severity in acute pancreatitis. *Clinical Biochemistry* **42**, 38-43 (2009).
79. Jylhävä, J. *et al.* Aging is associated with quantitative and qualitative changes in circulating cell-free DNA: The Vitality 90+ study. *Mechanisms of Ageing and Development* **132**, 20-26 (2011).
80. Garcia Moreira, V., Prieto Garcia, B., Baltar Martin, J.M., Ortega Suarez, F. & Alvarez, F.V. Cell-Free DNA as a Noninvasive Acute Rejection Marker in Renal Transplantation. *Clinical Chemistry* **55**, 1958-1966 (2009).
81. Koffler, D., Agnello, V., Winchester, R. & Kunkel, H.G. The Occurrence of Single-Stranded DNA in the Serum of Patients with Systemic Lupus Erythematosus and Other Diseases. *Journal of Clinical Investigation* **52**, 198-204 (1973).
82. Destouni, A. *et al.* Cell-free DNA levels in acute myocardial infarction patients during hospitalization. *Acta Cardiologica* **64**, 51-57 (2009).
83. Cui, M. *et al.* Cell-Free Circulating DNA: A New Biomarker for the Acute Coronary Syndrome. *Cardiology* **124**, 76-84 (2013).
84. Beutler, E., Gelbart, T. & Kuhl, W. Human red cell glucose-6-phosphate dehydrogenase: All active enzyme has sequence predicted by the X chromosome-encoded cDNA. *Cell* **62**, 7-9 (1990).
85. Lam, N.Y.L. EDTA Is a Better Anticoagulant than Heparin or Citrate for Delayed Blood Processing for Plasma DNA Analysis. *Clinical Chemistry* **50**, 256-257 (2004).
86. Jung, M. Changes in Concentration of DNA in Serum and Plasma during Storage of Blood Samples. *Clinical Chemistry* **49**, 1028-1029 (2003).
87. Angert, R.M. Fetal Cell-free Plasma DNA Concentrations in Maternal Blood Are Stable 24 Hours after Collection: Analysis of First- and Third-Trimester Samples. *Clinical Chemistry* **49**, 195-198 (2003).
88. Chiu, R.W. *et al.* Effects of blood-processing protocols on fetal and total DNA quantification in maternal plasma. *Clin Chem* **47**, 1607-13 (2001).
89. Lui, Y.Y., Chik, K.W. & Lo, Y.M. Does centrifugation cause the ex vivo release of DNA from blood cells? *Clin Chem* **48**, 2074-6 (2002).
90. Chan, K.C.A. Effects of Preanalytical Factors on the Molecular Size of Cell-Free DNA in Blood. *Clinical Chemistry* **51**, 781-784 (2005).
91. Umetani, N. Increased Integrity of Free Circulating DNA in Sera of Patients with Colorectal or Periampullary Cancer: Direct Quantitative PCR for ALU Repeats. *Clinical Chemistry* **52**, 1062-1069 (2006).
92. Lee, T.-H., Montalvo, L., Chrebtow, V. & Busch, M.P. Quantitation of genomic DNA in plasma and serum samples: higher concentrations of genomic DNA found in serum than in plasma. *Transfusion* **41**, 276-282 (2001).

93. Steinman, C.R. Free DNA in serum and plasma from normal adults. *Journal of Clinical Investigation* **56**, 512-515 (1975).
94. Lui, Y.Y. *et al.* Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clin Chem* **48**, 421-7 (2002).
95. Beutler, E., Gelbart, T. & Kuhl, W. Interference of heparin with the polymerase chain reaction. *Biotechniques* **9**, 166 (1990).
96. Holdenrieder, S., Von Pawel, J., Nagel, D. & Stieber, P. Long-term stability of circulating nucleosomes in serum. *Anticancer Res* **30**, 1613-5 (2010).
97. Sozzi, G. *et al.* Effects of Prolonged Storage of Whole Plasma or Isolated Plasma DNA on the Results of Circulating DNA Quantification Assays. *JNCI: Journal of the National Cancer Institute* **97**, 1848-1850 (2005).
98. Tse, C. & Capeau, J. [Real time PCR methodology for quantification of nucleic acids]. *Ann Biol Clin (Paris)* **61**, 279-93 (2003).
99. Bernard, P.S. & Wittwer, C.T. Real-time PCR technology for cancer diagnostics. *Clin Chem* **48**, 1178-85 (2002).
100. Hindson, B.J. *et al.* High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Analytical Chemistry* **83**, 8604-8610 (2011).
101. Metzker, M.L. Sequencing technologies — the next generation. *Nature Reviews Genetics* **11**, 31-46 (2009).
102. Kircher, M. & Kelso, J. High-throughput DNA sequencing - concepts and limitations. *BioEssays* **32**, 524-536 (2010).
103. Forshew, T. *et al.* Noninvasive Identification and Monitoring of Cancer Mutations by Targeted Deep Sequencing of Plasma DNA. *Science Translational Medicine* **4**, 136ra68-136ra68 (2012).
104. Newman, A.M. *et al.* An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nature Medicine* **20**, 548-554 (2014).
105. Leary, R.J. *et al.* Detection of Chromosomal Alterations in the Circulation of Cancer Patients with Whole-Genome Sequencing. *Science Translational Medicine* **4**, 162ra154-162ra154 (2012).
106. de Martino, M., Klatte, T., Haitel, A. & Marberger, M. Serum cell-free DNA in renal cell carcinoma. *Cancer* **118**, 82-90 (2011).
107. Perego, R.A. *et al.* Concentration and microsatellite status of plasma DNA for monitoring patients with renal carcinoma. *European Journal of Cancer* **44**, 1039-1047 (2008).
108. Wan, J., Zhu, L., Jiang, Z. & Cheng, K. Monitoring of Plasma Cell-Free DNA in Predicting Postoperative Recurrence of Clear Cell Renal Cell Carcinoma. *Urologia Internationalis* **91**, 273-278 (2013).
109. Gang, F. *et al.* Prediction of Clear Cell Renal Cell Carcinoma by Integrity of Cell-free DNA in Serum. *Urology* **75**, 262-265 (2010).
110. Hauser, S. *et al.* Cell-free circulating DNA: Diagnostic value in patients with renal cell cancer. *Anticancer Res* **30**, 2785-9 (2010).
111. Eisenberger, C.F. *et al.* Diagnosis of Renal Cancer by Molecular Urinalysis. *JNCI Journal of the National Cancer Institute* **91**, 2028-2032 (1999).

112. von Knobloch, R. *et al.* High frequency of serum DNA alterations in renal cell carcinoma detected by fluorescent microsatellite analysis. *International Journal of Cancer* **98**, 889-894 (2002).
113. Hoque, M.O. *et al.* Quantitative Detection of Promoter Hypermethylation of Multiple Genes in the Tumor, Urine, and Serum DNA of Patients with Renal Cancer. *Cancer Research* **64**, 5511-5517 (2004).
114. Skrypkina, I. *et al.* Concentration and Methylation of Cell-Free DNA from Blood Plasma as Diagnostic Markers of Renal Cancer. *Disease Markers* **2016**, 1-10 (2016).
115. Hauser, S., Zahalka, T., Fechner, G., Muller, S.C. & Ellinger, J. Serum DNA hypermethylation in patients with kidney cancer: results of a prospective study. *Anticancer Res* **33**, 4651-6 (2013).
116. Goessl, C. *et al.* Microsatellite analysis of plasma DNA from patients with clear cell renal carcinoma. *Cancer Res* **58**, 4728-32 (1998).
117. Ellinger, J. *et al.* DNA hypermethylation in papillary renal cell carcinoma. *BJU International* **107**, 664-669 (2010).
118. Bird, A. DNA methylation patterns and epigenetic memory. *Genes & Development* **16**, 6-21 (2002).
119. Kriaucionis, S. & Heintz, N. The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science* **324**, 929-930 (2009).
120. Jones, P.A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics* **13**, 484-492 (2012).
121. Feng, S., Jacobsen, S.E. & Reik, W. Epigenetic Reprogramming in Plant and Animal Development. *Science* **330**, 622-627 (2010).
122. Smith, Z.D. & Meissner, A. DNA methylation: roles in mammalian development. *Nature Reviews Genetics* **14**, 204-220 (2013).
123. Baylin, S.B. & Jones, P.A. A decade of exploring the cancer epigenome — biological and translational implications. *Nature Reviews Cancer* **11**, 726-734 (2011).
124. Esteller, M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nature Reviews Genetics* **8**, 286-298 (2007).
125. Baylin, S.B. & Jones, P.A. Epigenetic Determinants of Cancer. *Cold Spring Harbor Perspectives in Biology* **8**, a019505 (2016).
126. Cimmino, L., Abdel-Wahab, O., Levine, Ross L. & Aifantis, I. TET Family Proteins and Their Role in Stem Cell Differentiation and Transformation. *Cell Stem Cell* **9**, 193-204 (2011).
127. Lorsbach, R.B. *et al.* TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia* **17**, 637-641 (2003).
128. Larson, R. Faculty of 1000 evaluation for Mutation in TET2 in myeloid cancers. in *F1000 - Post-publication peer review of the biomedical literature* (Faculty of 1000, Ltd., 2009).
129. Abdel-Wahab, O. *et al.* Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood* **114**, 144-147 (2009).
130. Dalton, S.R. & Bellacosa, A. DNA demethylation by TDG. *Epigenomics* **4**, 459-467 (2012).
131. Diaz, L.A. & Bardelli, A. Liquid Biopsies: Genotyping Circulating Tumor DNA. *Journal of Clinical Oncology* **32**, 579-586 (2014).
132. Crowley, E., Di Nicolantonio, F., Loupakis, F. & Bardelli, A. Liquid biopsy: monitoring cancer-genetics in the blood. *Nature Reviews Clinical Oncology* **10**, 472 (2013).

133. Warton, K., Mahon, K. & Samimi, G. Methylated circulating tumor DNA in blood: power in cancer prognosis and response. *Endocrine-related cancer* **23**, R157-171 (2016).
134. Baylin, S. & Jones, P. A decade of exploring the cancer epigenome- biological and translational implications. *Nature Reviews Cancer* **11**, 726-734 (2011).
135. Chimonidou, M. *et al.* DNA Methylation of Tumor Suppressor and Metastasis Suppressor Genes in Circulating Tumor Cells. *Clinical Chemistry* **57**, 1169 (2011).
136. Powrózek, T., Krawczyk, P., Kucharczyk, T. & Milanowski, J. Septin 9 promoter region methylation in free circulating DNA—potential role in noninvasive diagnosis of lung cancer: preliminary report. *Medical Oncology* **31**, 917 (2014).
137. Shivapurkar, N. & Gazdar, A. DNA methylation based biomarkers in non-invasive cancer screening. *Curr mol med* **10**, 123-132 (2010).
138. Cairns, P. Gene methylation and early detection of genitourinary cancer: the road ahead. *Nature Reviews Cancer* **7**, 531 (2007).
139. Dulaimi, E., Hillinck, J., de Caceres, I.I., Al-Saleem, T. & Cairns, P. Tumor Suppressor Gene Promoter Hypermethylation in Serum of Breast Cancer Patients. *Clinical Cancer Research* **10**, 6189 (2004).
140. Newman, A.M. *et al.* An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nature Medicine* **20**, 548 (2014).
141. Lehmann-Werman, R. *et al.* Identification of tissue-specific cell death using methylation patterns of circulating DNA. *Proceedings of the National Academy of Sciences* **113**, E1826 (2016).
142. Visvanathan, K. *et al.* Monitoring of Serum DNA Methylation as an Early Independent Marker of Response and Survival in Metastatic Breast Cancer: TBCRC 005 Prospective Biomarker Study. *Journal of Clinical Oncology* **35**, 751-758 (2016).
143. Potter, N.T. *et al.* Validation of a Real-Time PCR–Based Qualitative Assay for the Detection of Methylated SEPT9 DNA in Human Plasma. *Clinical Chemistry* **60**, 1183 (2014).
144. Chan, K.C.A. *et al.* Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number aberrations by plasma DNA bisulfite sequencing. *Proceedings of the National Academy of Sciences* **110**, 18761 (2013).
145. Sun, K. *et al.* Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. *Proceedings of the National Academy of Sciences* **112**, E5503 (2015).
146. Grunau, C., Clark, S. & Rosenthal, A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res* **29**, e65 (2001).
147. Sharma, S., Kelly, T.K. & Jones, P.A. Epigenetics in cancer. *Carcinogenesis* **31**, 27-36 (2010).
148. Feng, G. *et al.* Quantification of Plasma Cell-Free DNA1in Predicting Therapeutic Efficacy of Sorafenib on Metastatic Clear Cell Renal Cell Carcinoma. *Disease Markers* **34**, 105-111 (2013).