M 2016

U.PORTO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

ROLE OF CHROMATIN REMODELERS IN METASTIZATION OF RENAL CELL TUMORS

MARIA JOÃO MARQUES DE CAMPOS FERREIRA
DISSERTAÇÃO DE MESTRADO APRESENTADA
AO INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
DA UNIVERSIDADE DO PORTO EM
ONCOLOGIA – ESPECIALIZAÇÃO EM ONCOLOGIA MOLECULAR

Maria João Marques de Campos Ferreira

ROLE OF CHROMATIN REMODELERS IN METASTIZATION OF RENAL CELL TUMORS

Dissertação de Candidatura ao grau de Mestre em Oncologia – Especialização em Oncologia Molecular submetida ao Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto.

Orientadora

Professora Doutora Carmen de Lurdes Fonseca Jerónimo

Professora Associada convidada com Agregação

Departamento de Patologia e Imunologia Molecular

Instituto de Ciências Biomédicas Abel Salazar

Universidade do Porto

Investigadora Auxiliar e coordenadora do Grupo de Epigenética e Biologia do Cancro

Centro de Investigação

Instituto Português de Oncologia – Porto

Coorientador

Professor Doutor Rui Manuel Ferreira Henrique

Professor Catedrático Convidado

Departamento de Patologia e Imunologia Molecular

Instituto de Ciências Biomédicas Abel Salazar

Universidade do Porto

Diretor do Serviço de Anatomia Patológica

Investigador Sénior do Grupo de Epigenética e Biologia do Cancro

Centro de Investigação

Instituto Português de Oncologia – Porto

“However difficult life may seem, there is always something you can do and succeed at.”

Stephen Hawking

Agradecimentos

Esta tese é o culminar de dois anos de trabalho científico, nos quais desenvolvi novas competências e adquiri conhecimentos que serão muito úteis num futuro próximo. A elaboração deste trabalho não teria sido possível sem o contributo pessoal e profissional de várias pessoas, a quem aqui quero demonstrar o meu profundo agradecimento:

Em primeiro lugar, à minha orientadora, Professora Doutora Carmen Jerónimo, por me ter concedido a oportunidade de integrar o Grupo de Epigenética e Biologia do Cancro. Agradeço o voto de confiança em mim, a paciência, os conselhos e as criticas efetuadas, a exigência e todos os ensinamentos transmitidos que foram fundamentais para a realização deste trabalho. Pela motivação dada e por sempre acreditar neste trabalho mesmo quando ocorreram percalços, um muitíssimo obrigada.

Ao meu coorientador, Professor Rui Henrique, agradeço todas as ideias pertinentes e conselhos que contribuíram em grande parte para um melhor projecto. Agradeço ainda pela contribuição enquanto cientista e patologista e pelas explicações concedidas, sem as quais este trabalho não teria sido bem sucedido.

Ao Professor Manuel Teixeira, por me receber no Centro de Investigação e permitir o usufruto das respectivas instalações.

À Professora Berta Silva, por me ter aceitado no Mestrado em Oncologia, e por todos os conhecimentos transmitidos.

Ainda neste contexto, agradeço ao Serviço de Anatomia Patológica e de Urologia, ao Engenheiro Luís Antunes do Serviço de Epidemiologia, pelas contribuições fundamentais para o presente trabalho.

Agradeço ainda ao ICVS, em particular à Céline Gonçalves e ao Doutor Bruno Costa pela análise de dados do TCGA, e por toda a disponibilidade e amabilidade que demonstraram.

Aos meus colegas do Grupo de Epigenética do Cancro, em especial à Ana Luís, por toda a ajuda prestada e por me ter ensinado todas as técnicas que foram realizadas neste trabalho. Por ter sido um pilar indispensável, e por toda a boa disposição e animação, um grande obrigada! Que a força esteja connosco!

Á Inês, João e Salta, por toda a ajuda e partilha de conhecimentos, por estarem sempre disponíveis a ajudar, pelas técnicas que me ensinaram e também pela boa disposição e brincadeiras, só tenho a agradecer.

Ao meu grupinho Eva e Francisca, agradeço por toda a paciência que tiveram quando as coisas correram menos bem, pela entreajuda, pela animação, e por tanto mais! Sou uma sortuda por ter a vossa amizade e sei que estarão comigo nos bons e nos maus momentos. Sei que o futuro nos trará grandes surpresas!

Às minhas amigas de biologia, Rita, Iris, Lia, Ana, Guida e Ash, por todo o apoio emocional, por estarem sempre dispostas ajudar e por saber que posso sempre contar com vocês. Ao Bruno, por sempre me ouvir quando as coisas correram menos bem, por toda a paciência, por todo o carinho, e acima de tudo, por sempre acreditares que eu era capaz!

Por último, mas não menos importante, à minha família. Em especial aos meus pais, Mauro e Lília. Por todo o apoio que foi muito para além do financeiro. Por toda a paciência e carinho que sempre tiveram comigo, por estarem sempre dispostos a ouvir-me, pelo interesse que demonstraram pelo meu trabalho (mesmo que não entendessem), pelos desabafos, e por sempre acreditarem em mim e nas minhas capacidades. Por isso tudo e muito mais, pela confiança inabalável que sempre tiveram em mim, o meu maior agradecimento. Esta tese é tanto minha quanto vossa.

This study was supported by funding of the Cancer Biology and Epigenetics Group-Research Centre of Portuguese Oncology Institute of Porto (Financiamento Base_CI-IPOP-27).

Resumo

Os tumores de células renais (RCTs) são os mais letais entre os cancros urológicos mais comuns. Devido ao uso generalizado de técnicas de imagem, ocorreu um aumento de deteção de pequenas massas renais, enfatizando a necessidade de uma correta distinção não apenas entre RCTs benignos e malignos mas também dentro dos RCTS malignos, aqueles que vão ser mais agressivos e desenvolver metástases daqueles que terão um crescimento mais indolente e passíveis de tratamento mais conservador. A metilação das histonas tem sido implicada na tumorigénese renal, contudo o seu potencial clínico como biomarcador de metastização em carcinomas de células renais (RCCs) permanece por explorar. Deste modo, o principal objectivo deste estudo foi investigar histonas metiltransferases (HMTs) e histonas desmetilases (HDMs) expressas diferencialmente em RCTs, de modo a avaliar o seu potencial como biomarcadores de metastização. Para tal, *SETDB2* e *MINA* foram validadas numa primeira série em 160 RCTs através de RT-PCR quantitativo. Uma segunda validação numa série de 62 ccRCCs foi efetuada para *MINA*, *SETDB2* e mais três enzimas, *NO66*, *SETD3* e *SMYD2* com o objectivo de aferir o papel destas enzimas na metastização deste subtipo, o mais frequente entre os RCTs. Subsequentemente, uma validação adicional da base de dados do The Cancer Genome Atlas (TCGA) foi efetuada para *MINA*, dado que esta enzima apresentou os resultados mais promissores. Por último, os níveis de mRNA de todas as enzimas foram avaliados nas linhas celulares, e o nível da proteína SETDB2 foi também determinado. Especificamente, na primeira série de 160 RCTs, *SETDB2* e *MINA* estão sobre-expressas em RCTs comparativamente com tecido renal normal (RNTs) e os seus níveis de expressão são mais altos em oncocitomas e carcinomas de células renais cromófobo (chRCC) comparativamente ao carcinoma de células renais de células claras (ccRCC) e o carcinoma de células renais papilar (pRCC). Os níveis de expressão das duas enzimas discriminaram de forma estatisticamente significativa RCTs malignos de benignos. Além disto, a *SETDB2* demonstrou ter níveis de expressão mais elevados e estatisticamente significativos em ccRCCs e pRCCs que não desenvolveram metástases, demonstrando o seu potencial como biomarcador de metastização nestes dois subtipos. A análise de sobrevivência revelou que a combinação de níveis de expressão de *SETDB2* e o Estadio (avançado *vs.* inicial) constituem fatores de prognóstico independentes para a sobrevivência livre de doença. Adicionalmente, na segunda série de 62 ccRCCs, os níveis de expressão do gene *MINA* foram estatisticamente mais elevados em ccRCCs que desenvolveram metástases, o que poderá auxiliar na definição do prognóstico. Contudo, a análise do gene *MINA* na base de dados do TCGA não revelou diferenças estatisticamente significativas entre ccRCCs que desenvolveram metástases e os que não as

desenvolveram, contrariamente aos nossos resultados. Tal pode ser devido a diferenças nos casos (metástases) incluídos nos estudos, dado que na nossa série foram excluídos da análise os casos com metástases identificadas aquando do diagnóstico. Para além destes resultados, os níveis de mRNA das cinco enzimas foram avaliados em linhas celulares (primárias e metastáticas), contudo devido à alta heterogeneidade dos resultados não foi possível retirar conclusões definitivas. Por último, os níveis de proteína SETDB2 foram avaliados nas linhas celulares, demonstrando haver correlação entre os níveis de transcrito e os de proteína. Em conclusão, os resultados do nosso estudo sugerem que os genes *SETDB2* e *MINA* são potenciais biomarcadores de metastização em ccRCCs e pRCCs, sendo requeridos estudos funcionais para melhor compreender o mecanismo biológico subjacente.

Abstract

Renal cell tumors (RCTs) are the most lethal among common urological cancers. Due to the widespread use of imaging there has been an increased detection of small renal masses, emphasizing the need for accurate discrimination not only between benign and malignant RCTs but also among malignant RCTs, specifically between those which will be more aggressive and develop metastases and those that will have a more indolent growth and may be managed more conservatively. Histone methylation has been implicated in renal tumorigenesis, however its potential clinical value as renal cell carcinomas (RCCs) metastization biomarker remains mostly unexplored. Thus, the main goal of this study was to explore differential expression of histone methyltransferases (HMTs) and histone demethylases (HDMs) in RCCs to assess their potential as metastasis biomarker. To achieve this goal, *SETDB2* and *MINA* were validated in a first series in of 160 RCTS by quantitative RT-PCR. Also, a second validation in a series of 62 ccRCC was performed for *SETDB2*, *MINA* and three other enzymes, *NO66*, *SETD3* and *SMYD2*, with the aim of evaluating its potential role in metastization of this subtype, which is the most common among RCTs. Furthermore, an additional validation using the The Cancer Genome Atlas (TCGA) database was performed for *MINA*, because this enzyme displayed the best results. Finally, mRNA levels of all enzymes were assessed in cell lines, as well as SETDB2 protein levels.

Specifically, in the first series of 160 RCTS, *SETDB2* and *MINA* were overexpressed in RCTs compared to renal normal tissues (RNTs) and their expression levels were higher in oncocytomas and chromophobe renal cell carcinoma (chRCC) compared to clear cell renal cell carcinoma (ccRCC) and papillary renal cell carcinoma (pRCC). Moreover, both enzymes expression levels discriminated benign from malignant RCTs. Furthermore, *SETDB2* levels were significantly higher in ccRCCs and pRCCs that did not develop metastases, suggesting a potential as metastization biomarker in these two subtypes. Survival analysis revealed that combined *SETDB2* expression levels and Stage (high vs low) were independent prognostic factors for disease-free survival. Additionally, in the second series of 62 ccRCCs, *MINA* expression levels were statistically higher in ccRCCs that developed metastases, suggesting that it may assist in the assessment of the metastatic potential of ccRCCs. TCGA database analysis for *MINA*, however, did not show statistically significant differences between ccRCCs that developed metastases and those that did not, contrarily to our results in the second validation series. This might be due to differences in case selection as we excluded from analysis the cases that displayed metastases at diagnosis. Furthermore, mRNA levels of all five enzymes were evaluated in

cell lines (primary and metastatic) but due to the heterogeneity of the results no definitive conclusions could be made. Finally, SETDB2 protein levels were evaluated in cell lines and a correlation between transcript and protein levels was depicted. Overall, our results suggest that *SETDB2 and MINA* expression levels are putative biomarkers of metastatic behaviour in ccRCCs and pRCCs. Functional studies are required to unveil the underlying molecular mechanisms.

Table of Contents

Figures Index

Tables Index

List of Abbreviations

APS - Ammonium persulfate

BAP1 – BRCA1 associated protein-1

BHD – Birt-Hogg-Dube

ccRCC – Clear Cell Renal Cell Carcinoma

chRCC – Chromophobe Renal Cell Carcinoma

c-MET – MET proto-oncogene

CpG – Cytosine Phosphate Guanine

DNA – Deoxyribonucleic acid

DNMTs – DNA Methyltransferases

DNMT1 – DNA Methyltransferas 1

DNMT3a – DNA Methyltransferas 3a

DNMT3b – DNA Methyltransferas 3b

FH – Fumarate Hydratase

GUSβ – Glucuronidase Beta

HAT – Histone Acetyltransferases

HDAC – Histone Deacetylases

HDM – Histone Demethylase

HIF – Hypoxia-inducible Factor

HMT – Histone Methyltransferase

IARC – International Agency for Research on Cancer

IL-2 – Interleukin-2

INF-α – Interferon Alpha

MBD – Methyl-CpG-Binding Domain

MDR-1 – Multi Drug Resistence Protein 1

MINA – MYC Induced Nuclear Antigen

miRNA – MicroRNA

mRCC – Metastatic Renal Cell Carcinoma

mRNA – MessengerRNA

mTOR – Mammalian Target of Rapamycin

ncRNA – Non-coding RNA

NO66 – NO66 gene

Onc – Oncocytoma

PBRM1 – Polybromo 1

PDGF – Platelet-derived Growth Factor

pRCC – Papillary Renal Cell Carcinoma

PTEN – Phosphatase and Tensin Homolog

PTM – Post Translational Modifications

RCC – Renal Cell carcinoma

RCT – Renal Cell Tumor

RNA – Ribonucleic Acid

RNT – Renal Normal Tissue

RTCC – Renal Transitional Cell Carcinoma

SETD2 – SET Domain Containing 2

SETD3 – SET Domain Containing 3

SETDB2 – SET Domain Bifurcated 2

SMYD2 – SET and MYND Domain Containing 2

TEMED - Tetramethylethylenediamine

TCGA – The Cancer Genome Atlas

TKIs – Tyrosine Kinase Inhibitors

TNM – Tumor Node Metastasis Classification

TP53 – Tumor Protein p53

VEGF – Vascular Endothelial Growth Factor

VEGF-A – Vascular Endothelial Growth Factor A

VHL – Von Hippel-Lindau

WHO – World Health Organization

INTRODUCTION

1. Kidney Cancer

1.1. Pathologoly

The kidney is an essential organ that is composed of a parenchyma and a collecting system. Its main functions are involved with maintaining the body's homeostatic balance, removing waste products from the blood, regulating blood pressure and secreting hormones [1].

Adult kidney cancers can either arise from renal parenchyma, that includes an outer cortex and an inner medulla, or from the collecting system, which includes renal pelvis and calyces, lined by transitional cells. Those who arise from renal parenchyma are mainly adenocarcinomas, currently known as Renal Cell Carcinomas (RCCs), while the ones that arise from the collecting system are mostly transitional cell carcinomas (RTCC) [2]. RCCs accounts for more than 90% of renal neoplasias, being the most common presentation of kidney cancer. In children, the most frequent renal neoplasia is nephroblastoma (Wilms tumor), which accounts 1.2% of all kidney cancers [2, 3].

1.2. Epidemiology

Kidney cancer is the 14th most common malignancy worldwide and the 8th most prevalent cancer in Europe representing 3.5% of all adult malignancies. Regarding Portugal, in the year of 2012, 1004 new cases of kidney cancer in both sexes have been registered, being the 16th most incident malignancy (Figure 1). Concerning kidney cancer mortality, in 2012, there were 143.406 deaths attributable to this malignancy worldwide. In Europe and Portugal, 49.025 and 368 deaths were caused by kidney cancer, respectively [4, 5].

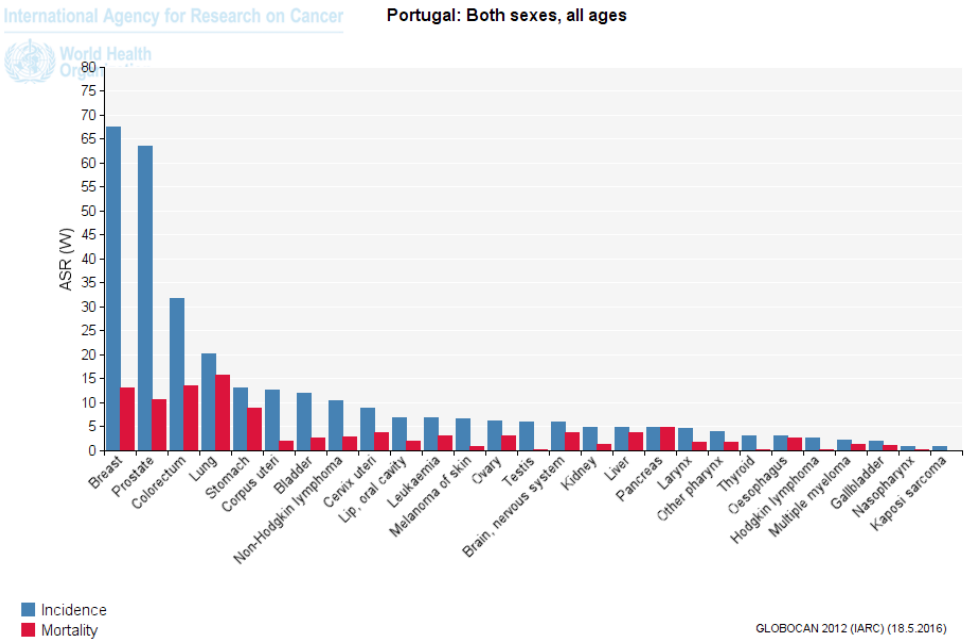


Figure 1- Incidence and Mortality of most prevalent cancers in Portugal for both sexes, in 2012. From [5].

Europe, North America and Oceania displayed the higher incidence rates of kidney cancer, whereas the lowest rates were observed in Asia and South America and Africa , revealing the worldwide variation of kidney cancer incidence (Figure 2) [5].

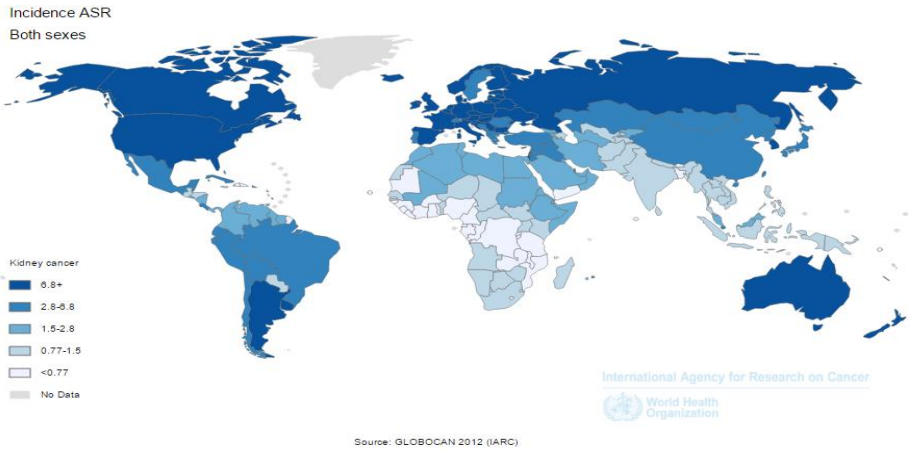


Figure 2 - Incidence for both sexes of kidney cancer worldwide, in 2012. From [5].

Furthermore, kidney cancer incidence varies by gender. Comparing the incidence of kidney cancer between both genders, men have a 2 fold risk ratio higher than women [4].

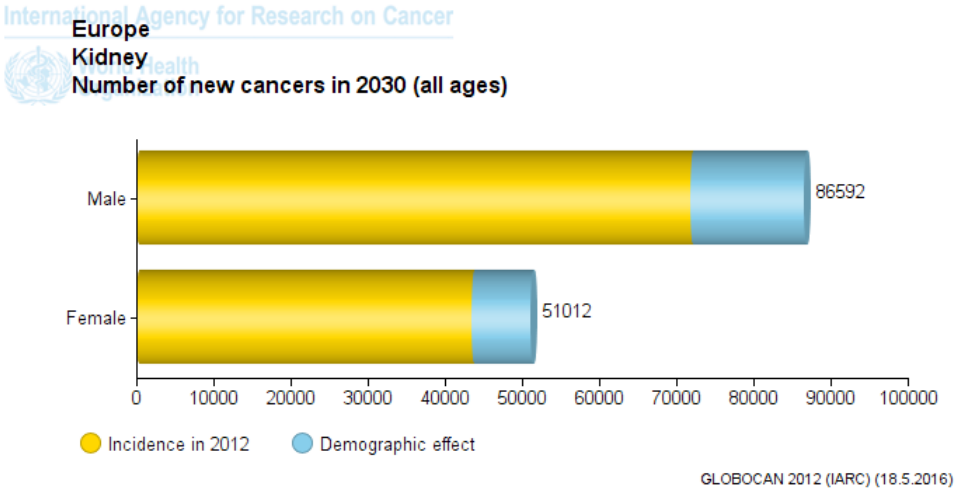


Figure 3 - Predictions of incidence of kidney cancer for the year of 2030, for both sexes in Europe. From [5].

According to Globocan predictions, by the year of 2030, the incidence of kidney cancer in Europe will increase, being more 22.352 new cases of kidney cancer comparing to the year of 2012 (Figure 3) [4].

1.3.Risk Factors

Until now, Renal Cell Carcinomas's aetiology remains mostly elusive. Nonetheles, there are some well-established risk factors, such as age, gender and geographic distribution, smoking, hypertension and obesity [3].

1.3.1. Demographic Risk Factors

RCC's incidence shows differences according to age, sex and race. Typically RCCs are diagnosed in the sixth and seventh decades of life. In Europe and the United States, the incidence of RCCs increases with age, however a plateau is reached at 70–75 years old. This could be due to the less frequent diagnostic testing in this older age group [3, 6]. Age standardised incidences, show that men are at an increased risk of developing RCCs, with a predominance of 3:2 comparing to females. This could be attributable to differences in the prevalence of smoking and occupational exposures. Nonetheless, the incidence of RCC is lower among Asians, which suggests a higher risk of RCC in Caucasians compared to Asians. Although African countries report the lowest incidence rates, African American

display the highest in United States, suggesting that these racial disparities in incidence can be due to differences in frequency of diagnostic and access to health cares [2, 6, 7].

1.3.2. - Lifestyle and Occupational Risk Factors

Cigarette smoking is considered a causal risk factor for RCC by the International Agency for Research on Cancer (IARC). Several studies confirmed that smoking increases the risk of RCC compared to never smoking. The risk of having RCC increases about 50% in male and 20% in females. Cigarette smoking causes chronic tissue hypoxia due to carbon monoxide exposure. Furthermore RCC patients were shown to have higher DNA damage levels in peripheral blood lymphocytes, including deletions in chromosome 3p, induced by a tobacco specific N-nitrosamine [2, 3, 6].

Similarly, body weight excess has been established as a risk factor for RCC in several case-control and cohort studies. Obesity accounts for nearly 30% of RCCs, representing a relative risk in males of 3.3 and in females 2.3. Although the mechanism is not clear yet, it is thought that hormonal changes such as increased levels of endogenous oestrogens may be the mechanism by which oestrogens induce renal cancer. Moreover, high levels of cholesterol and low levels of vitamin D, which are usually seen in obese patients, may favour tumor development by an inhibitory effect on immune cells [2, 3].

Hypertension is also considered a RCC risk factor. Indeed, several studies reported an association with a history of long term hypertension and increasing risk for RCC development. As hypertension is a chronic disease, it is estimated that it affects about 20 to 40 % of the world's population, being an important RCC risk factor [2, 8].

RCC increased risk has also been related with asbestos exposure, organic solvents, copper sulphate, polycyclic aromatic hydrocarbons, radiation, viruses and diuretic analgesics, however the data currently available is rather inconsistant [3].

1.3.3. Inheritance and Acquired Cystic Disease/Chronic Dialysis Risk Factors

The majority of renal cell tumors (RCTs) are believed to be sporadic, however there are some specific types of RCC caused by hereditary genetic defects. Overall, approximately 2–3% of RCCs are familial. Having a first degree relative with RCT is associated with a 2-fold increased risk of developing kidney cancer [6]. There are some hereditary RCC syndromes described.

Von Hippel-Lindau (VHL) disease is a syndrome of hereditary RCC that is inherited through and automosomal dominant trait. It is caused by germline mutations on VHL tumor suppressor gene. Patients with this disease present capillary haemangioblastomas of the central nervous system and retina, ccRCC, phaeiochromocytoma, pancreatic and inner ear tumors [9].

Another RCC hereditary syndrome is hereditary papillary RCC. This syndrome is caused by activating mutations of the MET oncogene mapped on chromosome 7q and is characterized by multiple and bilateral pRCC type 1 [10].

Hereditary RCC leiomyomatosis is an autosomal dominant syndrome which is caused by germline mutations in the FH gene on chromosome 1q. Patients with this syndrome usually present benign leiomyomas of the skin and uterus and occasionally papillary RCC type 2 and uterine leiomyosarcomas [11].

Birt-Hogg-Dube (BHD) syndrome results from mutations at BHD that is mapped at the chromosome 17p and encodes the protein folliculin. It is characterized by benign skin tumors, such as fibrofolliculomas, trochodiscomas and acrchordons, and multiple renal tumors [12].

Approximately 35 to 47% of patients that need dialysis develop acquired cystic disease. Patients with acquired cystic disease can develop a papillary hyperplasia that is a precursor of RCC. In fact, about 9% of patients with this syndrome develop RCC, thus having a higher risk of having RCC than the general population [13].

1.4.Clinical Presentation and Diagnosis

Due to the retroperitoneal localization of the kidney, many renal masses remain asymptomatic until the late stages of the disease. The classical triad of flank pain, haematuria and palpable abdominal mass used to be the typical clinical presentation of RCCs, however, these symptoms are only found in about 6 to 10 % of patients. Bone pain and persistent cough are usually symptoms of patients with metastatic disease [14]. Additionally, paraneoplastic syndromes, such as hypertension, cachexia and weight loss, are present in approximately 30% of patients with symptomatic RCCs [15].

Diagnosis of RCCs may occur through physical examination, laboratory findings and with most relevance imaging approachs. In fact, more than 50% of RCCs are detected incidentally when non-invasive imaging is used in order to investigate other diseases [7].

The proportion of patients with metastatic disease at diagnosis has declined, due to improved imaging techniques, more intense screening and incidental case ascertainment. As a result, these tumors are generally smaller and have a lower category, comparing to symptomatic RCCs. Usually they are small masses (< 4cm diameter), and many clinicians refers to them as having benign behaviour. However, adverse features displayed by small RCCs, such as invasion of the renal capsule, tumor thrombus and lymph node and distant metastasis currently raises concern about the adequacy of management [16, 17].

1.5. Histopathologycal Subtypes of Renal Cell Tumors

According to the current neoplasms classification by World Health Organization (WHO, 2016), there are four major histological renal cell tumors (RCTs) subtypes: Clear Cell Renal Cell Carcinoma (ccRCC), Papillary Renal Cell Carcinoma (pRCC), Chromophobe Renal Cell Carcinoma (chRCC) and Oncocytoma, which is a benign tumor. WHO classification combines morphological and genetic characteristics, recognizing not only these four major subtypes but also some variations or renal cancers with different immunophenotypes or molecular changes with clinical implications [18].

1.5.1. Clear Cell Renal Cell Carcinoma

Clear Cell Renal Cell Carcinoma (ccRCC) is the most frequent Renal Cell Tumor (RCT) subtype, representing 70% to 80% of all RCCs [19]. These tumors are originated from cells of the proximal nephron's tubule and usually have a very vascular stroma, which results in haemorrhagic areas. Due to the high lipid content of the tumoral cells, these tumors have a typical yellow cut surface. Moreover, small cystic necrotic areas are commonly present, being the last one associated with increased aggressive behaviour of the tumors [18-20].

The average size of detection of ccRCCs is 7cm in diameter, however, the detection of smaller lesions is increasing, especially in developed countries, due to the widely use of radiologic imaging techniques. Although size itself is not a determinant of malignancy it is known that a larger tumor size is often associated with higher metastases frequency [18]. The metastatic process in these tumors is commonly hematogeneously, via the vena cava primarily to the lung. However, ccRCC is well known by its late metastasis, even after 10 years or more, and for its metastization to unusual sites [18, 19].

Concerning ccRCC histopathology, these tumours have a diverse architecture. The most usual presentation occurs by solid, alveolar and acinar patterns. It is common a regular network of small thin-walled blood vessels, and the presence of clear cells due to lipid removal during histological processing [18, 21] (Figure 4). Despite this, some tumors may contain minority populations of cells with eosinophilic cytoplasm. This is more often in high grade tumours and adjacent to areas with haemorrhage or necrosis [18].

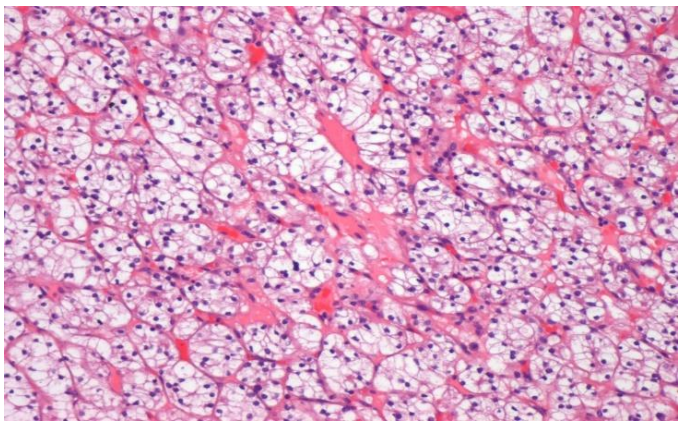


Figure 4 - Microscopic illustration of ccRCC. Original magnification, x100. Haematoxylin and eosin stain used.

Genetically, the deletion of chromosome 3p is considered to be one of the primary events in the carcinogenesis of ccRCC. This genetic alteration occurs in about 70% to 90% of RCC [22, 23]. The regions that are frequently lost or inactivated on this chromosome are 3p12-14, 3p21 and 3p25. The gene that is most commonly involved in the development of ccRCC is the tumor suppressor von Hippel-Lindau (VHL), mapped at 3p25. VHL gene is consistently inactivated in both sporadic and hereditary renal cancers [24, 25]. Indeed, biallelic VHL inactivation is a very high frequency event, that can occur through allelic deletion or loss of heterozygosity along with promoter hypermethylation or gene mutation [26] .

VHL protein functions as a tumor suppressor, since it inhibits growth when it is reintroduced into cultures of renal cell carcinoma [23].This protein plays a major role in the regulation of the transcription factor, hypoxia-inducible factor (HIF), which is a key regulator of hypoxia-inducible genes. VHL protein absence induces HIFs accumulation leading to the transcription of pro-survival and pro-angiogenic factors, such as Vascular Endothelial Growth Factors (VEGF) and Platelet-derived Growth Factor (PDGF) [20, 24]. Moreover, mutations in genes involved in chromatin condensation, such as PBRM1, SETD2, KDM5C9, KDM6A9 and BAP1, were associated with ccRCC subtype demonstrating a major role of the epigenetic deregulation in the development and progression of ccRCC [27].

1.5.2. Papillary Renal Cell Carcinoma

Papillary Renal Cell Carcinoma (pRCC), represents approximately 10% all RCCs, being the second most frequent renal cancer [18]. These tumors frequently contain haemorrhagic areas, necrosis and cystic degeneration. Usually they are well-circumscribed mass enclosed within a pseudo-capsule. Additionally, pRCC subtype is more likely to be bilateral and multifocal than other renal parenchymal malignancies [18, 19].

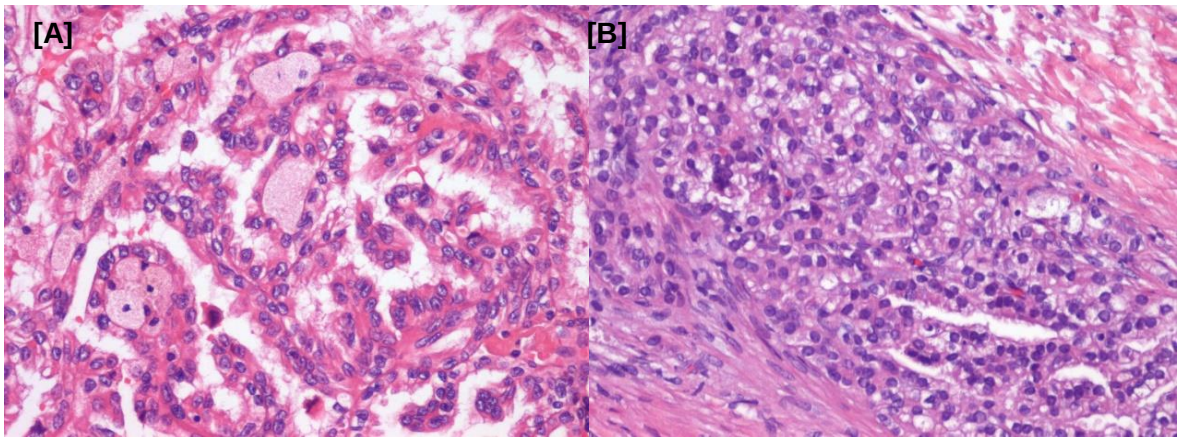


Figure 5 - Microscopic representation of the 2 types of pRCC. [A], pRCC type 1. [B], pRCC type 2. Original magnification, x200. Haematoxylin and eosin stain used.

The histology of pRCC is characterized by malignant epithelial cells that form papillae and tubules in varying proportions. Herein, the tumor papillae contain a delicate fibrovascular core where aggregates of macrophages are common. Moreover, *psammoma* bodies and haemosiderin granules are common [18-21]. There are two morphological types of pRCC described: type 1 and type 2 tumors (Figure 5). Type 1 pRCC are more frequent, accounting for two thirds of all pRCC.

These tumors are often multifocal and composed of papillae covered by single layered small with scanty cytoplasm. Type 2 tumors include more aggressive variants, with cells of higher nuclear grande with eosinophilic cytoplasm and pseudostratified nuclei [18-21].

In addition to these 2 groups, it has been proposed a third group of pRCC, since there are pRCC composed entirely by oncocytes that shows clinicalpathologic features different from type 1 and type 2 tumors. Also, approximately 5% of all pRCC has sarcomatoid dedifferentiation , which is associated with poor prognosis [19].

Regarding pRCC cytogenetics, its characteristic abnormalities include trisomy or tetrasomy of chromosome 7, trisomy of chromosome 17 and chromosome Y's loss. However, other abnormalities were already reported, such as trisomy of 12, 16 and 20, which are thought to be related with tumour progression. Loss of heterozygosity at 9p13

region is also observed being associated with a poorer survival [18]. The mutation of the c-MET proto-oncogene on chromosome 7 is a usual characteristic of hereditary pRCC, however similar somatic mutations were also found in about 13% of sporadic pRCCs. This gene encodes a transmembrane receptor (c-Met) that interacts with hepatocyte growth factor [28, 29].

1.5.3. Chromophobe Renal Cell Carcinoma

Chromophobe Renal Cell Carcinoma (chRCC) accounts for approximately 5% of all RCCs and is originated from the cells of the collecting tubules [18, 30]. Concerning chRCC macroscopy, this tumour is typically a solid circumscribed mass, with slightly lobulated surfaces. The cut surface is usually homogenous, light brown, without haemorrhage and/or necrosis, however it can be seen a central scar in large tumors [18, 21].

Microscopically, these tumors are characterized by a solid growth pattern, sometimes glandular with focal calcifications and thick-walled blood vessels. Classic chRCC histology consists of large polygonal cells with slightly reticulated cytoplasm mixed with smaller cells with granular eosinophilic cytoplasm. Some cells are irregular and multinucleated, having wrinkled nuclei. Also, perinuclear halos are often seen. The eosinophilic variant of chRCC is only composed of eosinophilic cells (Figure 6) [18, 19].

Sarcomatoid variants can also be present. The typical features are spindle-like cells, high cellularity, cellular atypia associated with necrosis and miscrovascular invasion. This particular variant is more common in chRCC than in other RCCs. Overall, having a sarcomatoid variant is a sign of poor prognosis [31].

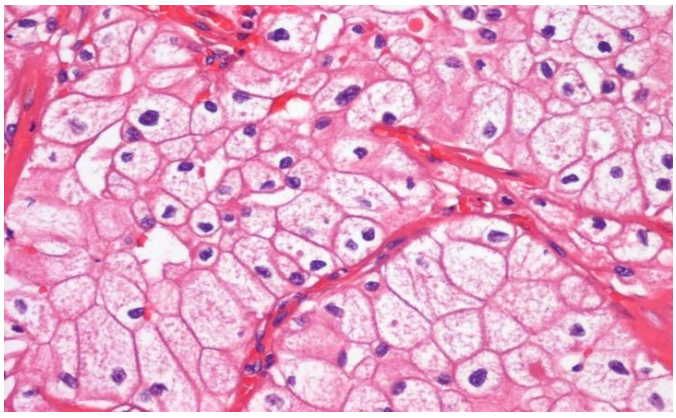


Figure 6 - Microscopic representation of chRCC. Original magnification, x200. Haematoxylin and eosin stain used.

Genetic analysis of chRCC has revealed non-random chromosomal losses regarding chromosomes 1, 2, 6, 10, 13, 17, and 21. In fact, they have been described in roughly 85 to 90 % of chRCCs, showing its potential as a diagnostic marker [31]. These massive chromosomal losses usually lead to a hypodiploid DNA index. Moreover, these genetic losses may lead to tumor suppressor inactivation, promoting carcinogenisis. Indeed, TP53 mutations in 27% of chRCC and loss of heterozygosity (LOH) around the PTEN gene, in chromosome 10 have been reported by others studies [32, 33].

1.5.4. Oncocytoma

Oncocytoma is a benign neoplasm that represents 3 to 5% of all primary epithelial neoplasms of the adult kidney. Macroscopically, these tumours are well-circumscribed and nonencapsulated, displaying a mahogany-brown cut surface. In about 33% of oncocytomas a central scar can be seen, being more common in larger tumors. In fact, oncocytomas can be fairly large at presentation, but the median size is 4 to 5 cm. Besides the central scar, haemorrhagic areas are present in up to 20% of cases [18, 21].

Regarding histopathology, the predominant cell type is called oncocyte and its conformation is round to polygonal with densely granular eosinophilic cytoplasm. The nuclei are round and regular with centrally placed nucleolus. Oncocytoma growth pattern is characterized by solid compact nests, acini, tubules or microcysts of variable sizes. A hypocellularhyalinized stroma is also seen very often (Figure 7) [18, 21].

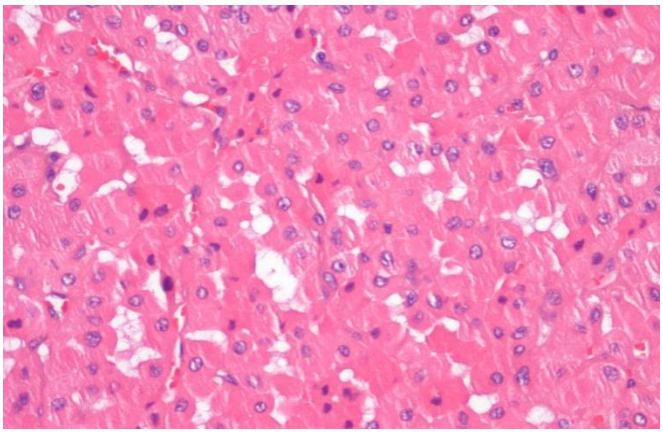


Figure 7- Microscopic representation of Oncocytoma. Original magnification, x200. Haematoxylin and eosin stain used.

At the molecular level, oncocytomas display a mixed population of cells with normal and abnormal karyotypes. Some oncocytomas present translocation of t(5;11) and loss of chromosome 1 and 14 [18, 34].

1.6. Staging

The tumor-node-metastasis (TNM) classification system is generally recommended for clinical and scientific use [35]. TNM system characterizes the degree of tumor local extension at the primary site (T), the involvement of regional lymph nodes (N) and the presence or absence of metastases (M). The latest version of the TNM classification was published in 2010 (Table 1). The prognostic value of the 2010 TNM classification has been confirmed by data from a large multi-centre studies with a good level of evidence [7].

Table 1 - TNM classification of renal cell tumors. From [23].

	T – Primary Tumor
Tx	Primary tumour cannot be assessed
T0	No evidence of primary tumour
T1	Tumour < 7 cm in greatest dimension, limited to the kidney
T1a	Tumour < 4 cm in greatest dimension, limited to the kidney
T1b	Tumour > 4 cm but < 7 cm in greatest dimension
T2	Tumour > 7 cm in greatest dimension, limited to the kidney
T2a	Tumour > 7 cm but < 10 cm in greatest dimension
T2b	Tumours > 10 cm limited to the kidney
T3	Tumour extends into major veins or directly invades adrenal gland or perinephric tissues but not into the ipsilateral adrenal gland and not beyond Gerota's fascia
T3a	Tumour grossly extends into the renal vein or its segmental (muscle-containing) branches or tumour invades perirenal and/or renal sinus (peripelvic) fat but not beyond Gerota's fascia
T3b	Tumour grossly extends into the vena cava below the diaphragm
T3c	Tumour grossly extends into vena cava above the diaphragm or invades the wall of the vena cava
T4	Tumour 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 a single regional lymph node
N2	Metastasis in more than 1 regional lymph node
	M - Distant metastasis
M0	No distant metastasis
M1	Distant metastasis

1.7. Treatment

For localized RCCs, surgery is the standard therapy. Partial nephrectomy is preferred over radical nephrectomy, due to the predisposition to chronic kidney disease caused by radical nephrectomy [36]. Radical nephrectomy consisting in the entire removal of the kidney is performed in patients with locally advanced tumors. Partial nephrectomy entails complete tumor resection leaving the largest possible amount of normal functioning kidney. There are also thermal ablative therapies, such as cryosurgery and radiofrequency ablation, which are alternatives nephron-sparing treatments for patients with localized RCC who are not suitable for conventional surgery [23]. Additionally, active surveillance can be an alternative for patients with small renal masses that do not show progression [7].

For metastatic RCCs, immunotherapy has been the leading treatment. Interferon Alpha (INF-α) and Interleukin-2 (IL-2) are the most common immune modulators used in clinical practice with response rates of approximately 15 and 20 %, respectively. Chemotherapy and hormonal therapy are not standard treatments for RCC, given that the response rates to these agents is very limited [37].

Alongside with immunotherapy, there are novel agents for metastatic RCC (mRCC) treatment. Globally, these agents block important pathways in renal carcinogenesis such as Vascular Endothelial Growth Factor (VEGF), Platelet- derived Growth Factor (PDGF) and Mammalian Target of Rapamycin (mTOR) pathways [38].

VEGFR and PDGFR antagonists include tyrosine kinase inhibitors (TKIs), being sunitinib and sorafetinib the most widely used. A monoclonal antibody against VEGF-A, named bevacizumab is also frequently used. These agents showed a longer progression free survival than INF-α [39]. The most common used mTOR inhibitors are everolimus and temsirolimus. They both showed a higher overall survival when compared to INF-α. The agents mentioned above are all approved by the FDA for mRCC and used in clinical practice. Albeit all these therapies for mRCC, the increase on survival rates are only about 2 to 3 months, showing the need for an improved treatment and the discovery of newer agents for effective treatment of metastatic RCC [40, 41].

1.8. Management of Metastatic Renal Cell Carcinoma

Metastasis is the spread of cells from the primary neoplasm to distant organs or lymph nodes. Most deaths from cancer are due to metastases. The improvements in diagnosis, surgical techniques, patient care and adjuvant therapies does not seem to be enough to improve survival in metastatic cancer [42]. It is known that the major obstacle to effective treatment is the biologic heterogeneity of tumor cells. Furthermore, metastases may occur in lymph nodes and in different organs, and the microenvironment of the specific organ can influence the response of metastatic cells, even their response to therapy [43].

Therefore, one of the current major goals of cancer research is the understanding the pathogenesis of metastasis, on the systemic, cellular and molecular levels [44]. The process of cancer metastasis consists of a long series of sequential, interrelated steps, as explained in Figure 8. Each one of these steps can be limiting, given that a failure at any of the steps can stop the entire process of metastization [42].

The management of mRCC still remains a major challenge to the clinician. The median survival of patients with metastatic disease is very low (6 to 8 months) and the 5-year survival rate is below 10%. One of the reasons for this survival rate is the ineffective effect and response rate of <5% of cytotoxic chemotherapy. The cause of chemoresistence in RCC is due to the expression of Multidrug resistance (MDR-1) gene and its protein, p-glycoprotein [45]. Although RCC is one of the few tumor types that respond to immunotherapy, as mentioned above, the response rate of these agents is only about 20% (for combined interferon-α and interleukin-2), thus, sustaining the need for research effort on metastatic setting of the disease [46].

Albeit the proportion of patients with metastatic disease at diagnosis has declined, due to improved imaging techniques, more intense screening and incidental case ascertainment, a not negligible number of small RCCs (< 4cm diameter) may present renal capsule invasion, tumor thrombus or lymph node and distant metastasis [16, 17].

.

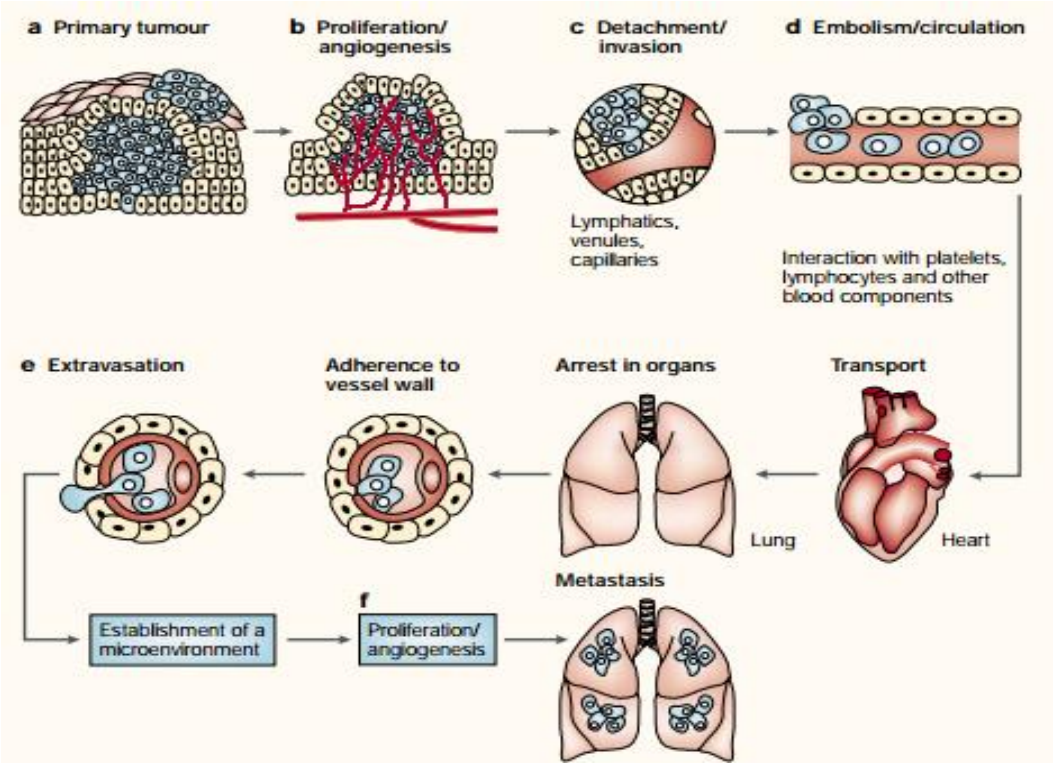


Figure 8 – Representation of the metastatic process: **a** - Represents the growth of neoplastic cells. **b** - Extensive vascularization must happen to give tumor nutrients for its growth. Also, secretion of angiogenic factors establish a capillary network needed for the tumor. **c** - Local invasion of the stroma tumour cells. As lymphatic channels are thin-walled, they offer very little resistance to penetration by tumour cells providing this way the most common route for tumour-cell entry into the circulation. **d** - Detachment and embolization of single tumour cells. Most of the tumor cells are destroyed, however some can escape and survive in the circulatory system. Next, they become trapped in the capillary of distant organs. **e** - Extravasation of the tumor cells. **f** - Proliferation in organ parenchyma. In order to continue growing, the tumor cells need to develop a vascular network and evade destruction by host defences. Adapted from [42].

Indeed, approximately one-third of patients with RCC will eventually develop metastasis and as said before, the long term prognosis for these patients is poor [47]. The current best therapy for mRCC is still inadequate. Therefore, a better understanding of RCC metastization and the identification of new players involved that may be target by therapeutic agents is urgently required.

2. Epigenetics

The “Epigenetics” field was first described by Waddington in 1942, who defined it as “the causal interactions between genes and their products, which bring the phenotype into being”. But the word currently refers specifically to heritable changes in gene expression that are not due to any alteration in the DNA sequence [48].

Currently, four major epigenetic mechanisms are recognized: DNA methylation, non-coding RNAs, histone variants and histones’ posttranslational modifications. As these four mechanisms are dynamic, they work together and interact with which other in order to regulate gene expression [49].

2.1. DNA Methylation

DNA methylation is the most studied epigenetic modification. In humans, DNA methylation occurs in dinucleotide CpGs which are cytosines that precede guanines. It consists in the addition of a methyl group at the 5’ position of a cytosine ring within CpG dinucleotides. These CpG sites are not randomly distributed in the genome, instead there are CpG-rich regions named CpG islands [48]. CpG islands are characterized by a CG content of at least 50% and a ratio of observed/expected CpG dinucleotides of at least 0.6 and span at the 5’ end of the regulatory region of many genes (about 60%). This alteration is catalysed by enzyme DNA methyltransferases (DNMTs), which catalyse the transfer of a methyl group from S-adenosyl methionine to DNA. There are three main DNMTs: DNMT1, which maintains the existing methylation patterns following DNA replication, and DNMT3A and DNMT3B, *de novo* enzymes that target previously unmethylated CpGs [50]. Methylation of CpG islands is associated with gene silencing. This can happen by recruitment of methyl-CpG-binding domain (MBD) proteins that recruit histone modifying and chromatin remodelling complexes to methylated sites. Besides this, DNA methylation can also directly inhibit transcription by blocking the recruitment of DNA binding proteins from their target sites [49]. Hence, DNA methylation is an important regulatory mechanism of gene transcription [51].

2.2. Non-coding RNAs

In recent years, attention has been focused in non-coding RNAs (ncRNAs). Non-coding RNAs are a class of RNAs that do not encode proteins but has several functions in

the cells. In fact, It has become increasingly evident that the portion of the genome that does not code proteins has a crucial function both for normal development and physiology and for disease [52]. Specifically, non-coding RNAs have functional relevance in many cellular pathways such as splicing, chromosome dynamics, RNA editing, inhibition of translation, mRNA destruction, X-chromosome silencing in females and DNA imprinting. Depending on their length, function and interactions, ncRNAS can be distributed in different classes, being the most widely studied microRNAS (miRNAS) [53].

MicroRNAs are small non-coding RNAs, with 18 to 25 nucleotides (nt) in length, which are synthesized and processed in the nucleus and then exported to the cytoplasm where they regulate gene expression. Indeed, it is estimated that miRNAs control approximately 30% of human genes [54]. Initially, miRNAs have been reported as negative regulators of mRNA expression: repress the mRNA translation either by degradation or inhibition of mRNA, depending on the accuracy of matching miRNA-mRNA [55]. However, in recent years, it was suggested that they could also act as positive transcription regulators. Moreover, one miRNA can target many mRNAs, which in turn can be targeted by several miRNAs [56].

In cancer, as in other diseases, aberrant miRNAs expression has been widely reported. Although a trend to a global miRNAs downregulation is seen in cancer, upregulation has also been described [52, 57]. Regarding miRNAs' role in tumorigenesis, they can act as an oncogene (oncomiRs) or tumor-supressors, depending on the target genes and the neoplastic context [58].

2.3.Histone Variants

The substitution of canonical histones by sequential similar non-allelic histones variants is the less studied epigenetic mechanism. Histone variants are called "replacement histones" because they substitute the canonical histones during development and differentiation, establishing cell identity [59, 60]. Thus, impacting in nucleosome-DNA stability and in the efficiency of protein complexes responsible for histone disposition and displacement in the nucleosome [61].

2.4.Posttranslational Modifications of Histone Proteins

Each nucleosome, the basic unit that composes chromatin, is composed by an octamer of four core histone proteins (H3, H4, H2A and H2B) around which 146bp of DNA

is wrapped. Histones are small basic proteins that contain a globular C-terminal domain and a unstructured N-terminal tail [62]. The N-terminal tails of histones protrude from the nucleosome and can be altered by different post-translational modifications which are catalysed by various histone-modifying enzymes. Until now, at least 16 different post-translational histone modifications (PTMs) have been reported, being the most well-known, acetylation, methylation, ubiquitination and phosphorylation [63]. These modifications regulate key cellular processes such as transcription, replication and repair, and are performed by enzymes named “chromatin writers”. On the other hand, PTMS are removed by “chromatin erasers” and recognized by “chromatin readers” in a highly regulated manner (Figure 9) [64, 65].

Histone modifications affect the chromatin structure by either changing the accessibility of chromatin (heterochromatin or euchromatin) or by interfering with other proteins’ recruitment to the chromatin. Moreover, PTMs not only determine the accessibility to specific DNA loci but also provide an informative platform for the recruitment of epigenetic regulators [63]. All of these distinct combinations of PTMs of histone tails are named “Histone code” and along with DNA methylation regulate gene activation or inactivation [48].

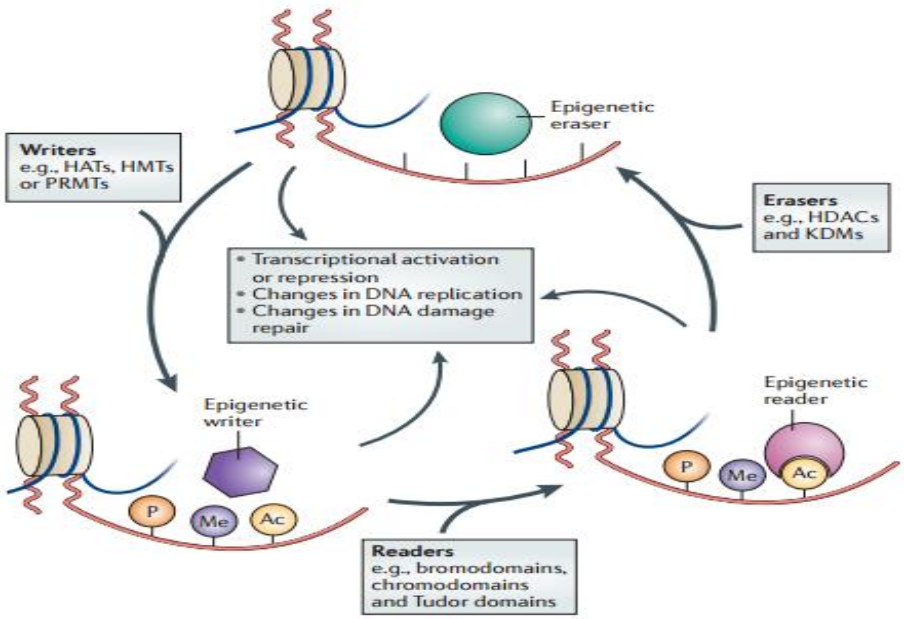


Figure 9 – Schematic representation of chromatin writers, readers and erasers: Epigenetic writers can be histone acetyltransferases (HATs), histone methyltransferases (HMTs), protein arginine methyltransferases (PRMTs) and kinases. Epigenetic readers are bromodomains, chromodomains and Tudor domains. Epigenetic erasers are histone deacetylases (HDACs), lysine demethylases (KDMs) and phosphatases. From [66].

Regarding histone acetylation, this PTM is “written” by histone acetyltransferases (HATs) and “erased” by histone deacetylases (HDACs). The gene transcriptional activity is regulated due to alterations in the electrostatic charge of the nucleosomes [66]. Therefore, states of euchromatin are hyperacetylated which decreases the histone-DNA affinity and allows gene transcription, whereas hypoacetylation is a characteristic of heterochromatin [67].

Concerning histone methylation, the “writers” are histone methyltransferases (HMT) and the “erasers” are histone demethylases (HDM). Histone methylation promote gene activation or repression depending on the residue and the number of methylated molecules added (mono-, di-, or tri-) [68]. For example, trimethylation of lysine 4 on histone H3 (H3K4me3) is very common at transcriptionally active gene promoters. On the other side, trimethylation of H3K9 (H3K9me3) and H3K27 (H3K27me3) is present at gene promoters that are transcriptionally repressed. These two last modifications together constitute the principal silencing mechanism in mammalian cells [66, 69].

PRELIMINARY DATA

1. Brief Contextualization

The work presented in this Master Thesis arises from a previous project developed at the Cancer Epigenetics and Biology Group (GBEC) [70], which aimed to explore the role of HMTs and HDMs in kidney cancer . Thereby, in order to identify what HMTs and HDMs displayed an abnormal expression pattern during renal carcinogenesis, the expression levels of 58 HMTs and 29 HDMs were evaluated in 5 chRCCs, 5 oncocytomas and 5 RNTs. The analysis was performed by comparing RNTs and RCTs as well as chRCCs and oncocytomas. Globally, HMTs upregulation was observed in RCTs compared to RNTs (Figure 10). Conversely, HMTs and HDMs expression levels were downregulated in chRCCs compared to oncocytomas. The mRNA levels of the studied genes were normalized to the betaglucuronidase (*GUSβ*) reference gene and the median value of RNTs and oncocytomas and chRCCs samples was chosen to calculate the fold variation in gene expression between groups, using the comparative Ct method.

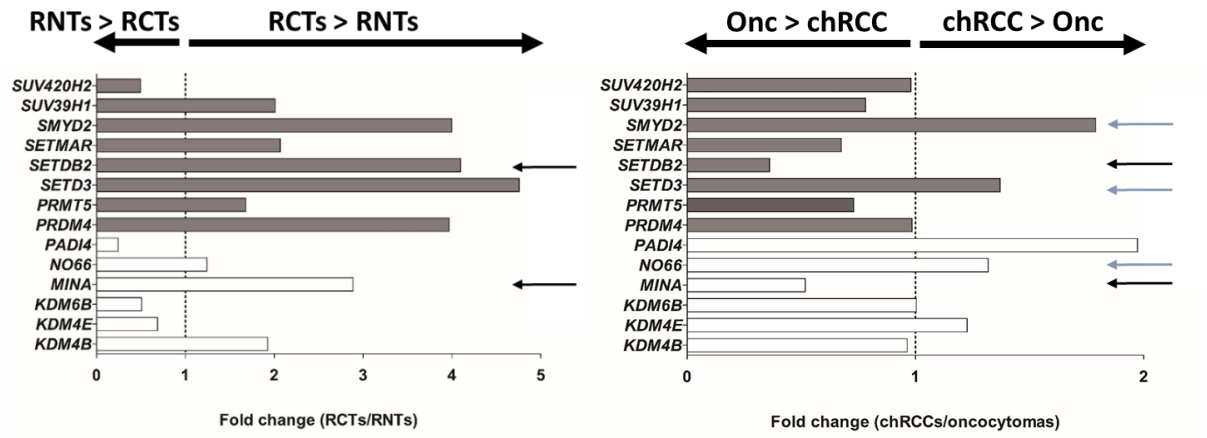


Figure 10 - Fold variation graphics adapted from [70]. Dark bars represent methyltransferases and white bars represent demethylases. Black arrows : Enzymes newly selected; Blue arrows: Previous selected enzymes.

Based on this screening, *MINA* and *SETDB2* were selected for validation in a series of 160 Renal Cell Carcinomas. Additionally, 3 previous validated genes, *SMYD2*, *SETD3* and *NO66* that showed the most differential expression between chRCC and oncocytomas and between RCTs and RNTs were also assessed in a series of 31 ccRCCs that developed metastases comparing to 31 ccRCCs that did not developed metastases.

.

.

AIMS OF THE STUDY

1. Aims

Among common urological cancers, RCTs are the most lethal. Despite modern imaging methods and early diagnosis, one third of RCC patients develop metastatic disease, which is the major cause of mortality. Even with the use of targeted therapies, long-term prognosis for mRCC patients is poor, with a median survival less than two years. Due to the widespread use of imaging, there is an increase in small renal masses detection, requiring accurate tools for discrimination of small tumours that will display aggressive behaviour through metastization and those that will have a more indolent grow. Indeed, although tumours with less than 4cm rarely metastasize to distant organs, some do and carry a dismal prognosis. This emphasizes the need for accurate distinction between those tumor subsets.

Furthermore, recent data implicated chromatin machinery deregulation in renal neoplastic transformation, entailing the selection of five HMTs and HDACs to investigate their role in RCC metastization and their potential as prognostic biomarker in RCCs. To achieve this, several goals were established:

I. Validate *MINA* and *SETDB2* in a large series of 160 RCCs;
II. Ascertain the correlation between *MINA* and *SETDB2* expression with clinicopathological parameters;
III. Assess *MINA*, *SEDTB2*, *NO66*, *SETD3* and *SMYD2* expression levels in a series of ccRCCs including patients that developed metastasis;
IV. Correlate *MINA*, *SETDB2*, *NO66*, *SETD3* and *SMYD2* expression levels with clinicopathological variables;
V. Compare the data obtained in our series with that of TCGA database cohort of patients;
VI. Correlate *SETDB2* transcript with protein levels in kidney cell lines.

MATERIALS AND METHODS

1. Patients and Sample Collection

A series of 160 RCTs comprising 40 cases of each subtype (ccRCCs, pRCCs, chRCCs and oncocytomas) were prospectively collected from patients consecutively diagnosed and submitted to nephrectomy at the Portuguese Oncology Institute of Porto. As controls, 13 renal normal tissue (RNT) samples were collected from patients subjected to nephrectomy due upper urinary tract urothelial carcinoma. All tissues were immediately frozen after surgery and stored at -80ºC. Sampling of more than 70% of malignant cells confirmed by two slides stained with hematoxylin and eosin (H&E) taken before and after frozen section collection for RNA extraction. Relevant clinical data was also collected from clinical charts.

An independent series of 62 ccRCCc comprising 31 ccRCCs that have developed metastasis and 31 ccRCCs that did not progress were also collected. All specimens were immediately frozen after surgery and stored at -80ºC. Sampling of more than 70% of malignant cell was confirmed by the same method as described before. Relevant clinical data was also collected from clinical charts. Samples were paired based on tumor size, gender, age and pathological stage.

This study was approved by institutional ethics review board (CES-IPOPFG-EPE 518/10).

2. RNA Extraction

For RNA extraction, samples were suspended in TRIzol® reagent (Invitrogen™, Cat.#15596018) and chloroform (Merk Milipore, Cat.#MCX10601) was added after the cells were lysed. RNA concentrations and purity ratios were determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). Samples were stored at -80ºc.

3. Validation of Selected Enzymes

MINA and *SEDTB2* mRNA levels were evaluated in two independent series of 160 RCTs. Also in a second validation series of 62 ccRCCs, all five enzymes were evaluated.

For the validation in a series of 160 RCTS a total of 300ng was reverse transcribed and amplified using TransPlex®Whole Transcriptome Amplification Kit (Sigma-Aldrich®, St.

Louis, MO, United States) with subsequent purification using QIAquick PCR Purification Kit (QIAGEN, Germany), according to manufacturer's instructions. *MINA* and *SETDB2* mRNA levels were evaluated using TaqMan® Gene Expression Assays [Applied Biosystems®, Hs99999908 m1 (*GUSβ*), Hs01126272 m1 (*SETDB2*), Hs00262155 m1 (*MINA*)] according to manufacturer's instructions. For each sample, expression levels were normalized using one internal reference gene, *GUSβ*, according to the formula: target gene relative expression = target gene expression level/ *GUSβ* expression level. Each plate included multiple non-template controls and serial dilutions of a cDNA Human Reference Total RNA (Agilent Technologies, Cat.#750500) in order to construct a standard curve.

For the validation in the series of 62 ccRCCs and cell lines, 1 μg of total RNA was reverse transcribed using the High Capacity cDNA Reverse Transcription kit according to the manufacturer's instructions. *MINA*, *SETDB2*, *SMYD2*, *NO66* and *SETD3* mRNA levels were evaluated using TaqMan® Gene Expression Assays [Applied Biosystems®, Cat.# 4331182 Hs00220210 m1 (*SMYD2*), Hs00260120 m1 (*SETD3*), Hs02743012 s1 (*NO66*), Hs99999908 m1 (*GUSβ*), Hs01126272 m1 (*SETDB2*), Hs00262155 m1 (*MINA*)] according to manufacturer's instructions. For each sample, expression levels were normalized using one internal reference gene, *GUSβ*, according to the formula: target gene relative expression = target gene expression level/ *GUSβ* expression level. Each plate included multiple non-template controls and serial dilutions of a cDNA Human Reference Total RNA (Agilent Technologies, Cat.#750500) in order to construct a standard curve.

4. Cell Culture

SMYD2, *NO66*, *SETD3*, *SEDTB2* and *MINA* mRNA levels were also assessed in the renal cell lines, including metastatic and primary tumors cell lines, available in the laboratory. Metastatic cell lines are Caki-1 (metastasis from ccRCC) and ACHN (metastasis from pRCC). Caki-2, 769-P, 786-O and A-498 are primary ccRCCs cell lines. Additionally, HK-2, a normal kidney cell line and HEK 293 an embrionary kidney cell line were also evaluated.

Concerning culture media conditions, Caki-1 and Caki-2 were grown using McCoy's 5A modified Liquid Medium (EMD-Millipore ; 786-O, 769-P, HK-2 and HEK 293 were maintained in RPMI 1640 Liquid Medium (EMD-Millipore); ACHN and A-498 were grown using Eagle's Minimum Essential Medium (EMD-Milipore). All cell lines were supplemented with 10% of Fetal Bovine Serum (FBS) Superior (EMD-Millipore) and 1% Penicillin/Streptomycin (GIBCO®, Carlsbad, CA, USA) and maintained in cell culture flaks

at 37ºC and 5% CO_2 at a humidifying chamber. The cells were harvested using TrypLE™ Express (GIBCO®). All cell lines were routinely tested for *Mycoplasma spp.* contamination.

5. Protein Extraction and Quantification

Total protein was extracted from cell lines using Kinexus Lysis Buffer with Lysis Buffer Cocktail (Kinexus Bioinformatics Corporation, Vancouver, British Columbia, Canada).In short, the protocol begins with removal of the growth medium from cell culture flasks. Next, the cells were washed two times with PBS. The cells were scrapped from the flasks with a cell scrapper (Santa Cruz Biotechnology Inc.) and then transferred to a 1.5 mL tube. Later, the tubes were sonicated on ice for 6 cycles of 15 seconds, having a 15 seconds gap between each cycle. Then the tubes were centrifuged for 30 mins at 13,000 rpm at 4ºC and the supernatant was transferred to a new tube. Protein concentration was assessed using Pierce BCA Protein assay Kit (Thermo Scientific, Waltham, MA, USA), following manufacturer's instructions. All protein samples were stored at -80ºC.

6. SDS-PAGE and Western Blot

SETDB2 Protein expression was assessed in kidney cell lines by Western Blot analysis. Summarily, loading buffer was added in 30µg of total protein and then denatured for 5 minutes at 95ºC. After centrifugation, samples were loaded in a polyacrylamide gel composed by a 10% running gel [10% (w/v) acrylamide/bis-acrylamide solution, 0.375M Tris-HCl pH=8.8, 0.1% (w/s) SDS, 0.1% (w/s) APS and 0.04% (v/v) TEMED] and a 4% stacking gel [4% (w/v) acrylamide/bisacrylamide solution, 0.062M Tris-HCl pH=6.8, 0.1% (w/s) SDS, 0.1% (w/s) APS and 0.25% (v/v) TEMED]. Protein separation was performed in a drive Mini-Protean 3 Eletrophoresis System (BioRad, Hercules, CA, USA) at 120V in a running buffer (0,025M Tris, 0192M glycine and 0.1% SDS, pH=8.3). After SDS-PAGE, proteins were blotted in PVDF membranes (Bio-Rad Laboratories Inc., Hercules, CA, USA) that were previously activated in 20% (v/v) methanol. The membranes and filter papers were incubated for 20 minutes and the gel for 10 minutes at room temperature in transfer buffer [0.025m Tris, 0.192M glycine and 20% (v/v) methanol. Protein blotting was made in Trans-Blot® Turbo™ Transfer System (BioRad, Hercules, CA, USA) for 8 minutes at 25V. After incubation for an hour in a blocking solution 5% (w/v) BSA (ChemCruz™ Biochemicals, Santa Cruz Biotechnology Inc.,) in 0.01M Tris-buffered saline containing 0.1% (v/v) Tween 20], membranes were incubated overnight at 4ºC with primary antibody

SETDB2 (1:1000; # 05-1952 ; EMD-Millipore) in blocking solution. Membranes were washed in TBS with Tween and incubated for 1 hour with Goat Anti-Mouse IgG (1:4000; BioRad, Hercules, CA, USA). The blots were developed using Western BrightTM ECL-spray (Advansta Corporation, Menlo Park, CA, USA) and exposed to AmershamTM Hyperfilm ECL (GETM Healthcare, Buckinghamshire, United Kingdom). All experiments were performed in triplicate.

7. TCGA Dataset Analysis in pRCC, chRCC and ccRCCs Patients

The Cancer Genome Atlas (TCGA) was used to obtain data on *MINA* expression and clinical information, when available, from ccRCCs, pRCCs and ccRCCs patients. All expression data from samples hybridized by the University of North Carolina, Lineberger Comprehensive Cancer Center, using Illumina HiSeq 2000 RNA Sequencing version 2 analysis, were downloaded from TCGA data matrix (http://tcga-data.nci.nih.gov/tcga/tcgaDownload.jsp). This dataset included 533 ccRCC, 290 pRCC and 66 chRCC. The provided value was pre-processed and normalized according to “level 3” specifications of TCGA (see http://cancergenome.nih.gov/dataportal/ for details). Biospecimen Core Resources (BCRs) provided the clinical data of each patient. This data is available for download through TCGA data matrix (http://tcga-data.nci.nih.gov/tcga/dataAccesMatrix.htm).

8. Statistical Analysis

In this work, non-parametric tests were used to ascertain statistical significance for comparisons made. Mann-Whitney U test (MW) was used in pairwise comparisons and Kruskal-Wallis test (KW) was used for comparisons between multiple groups (more than 2). These tests were used both in clinical samples and *in vitro* studies.

The prognostic significance of available clinical variables (histological subtype, pathological stage, Fuhrman grade, age, gender and also HMTs and HDMs expression levels) was assessed by constructing disease-specific and disease-free survival curves using the Kaplan-Meier method with log-rank test (univariable test). The expression levels of *SETDB2* and *MINA* were classified as low or high based on the cutoff value of 25th percentile for *SETDB2* expression and 75th percentile for MINA. A Cox-regression model

using ENTER method comprising the different variables (multivariable test) was also constructed. For this analysis 120 RCC patients were included, which comprised all RCCs subtypes.

Statistical significance was set at $p<0.05$. Bonferroni correction was applied for pairwise comparisons following multiple groups' analyses, dividing p-value by number of groups evaluated (p value $< 0.05/n$). Statistical analysis was performed using SPSS software for Windows, version 22.0 (IBM-SPSS Inc.), and graphs were built using GraphPad Prism 6.0 software for Windows (GraphPad Software Inc.).

RESULTS

1. Validation of *MINA* and *SETDB2* Expression in a Series of 160 RCTs

The validation of *MINA* and *SETDB2* expression levels was performed by quantitative RT-PCR in a series of 160 RCTs and 13 RNTs. The results were fully concordant with those of the TaqMan® Array from preliminary data. Indeed, both enzymes were significantly overexpressed in RCTs compared to RNTs ($p<0.0001$ for *SETDB2* and $p<0.05$ for *MINA*; Figure 11 A-B).

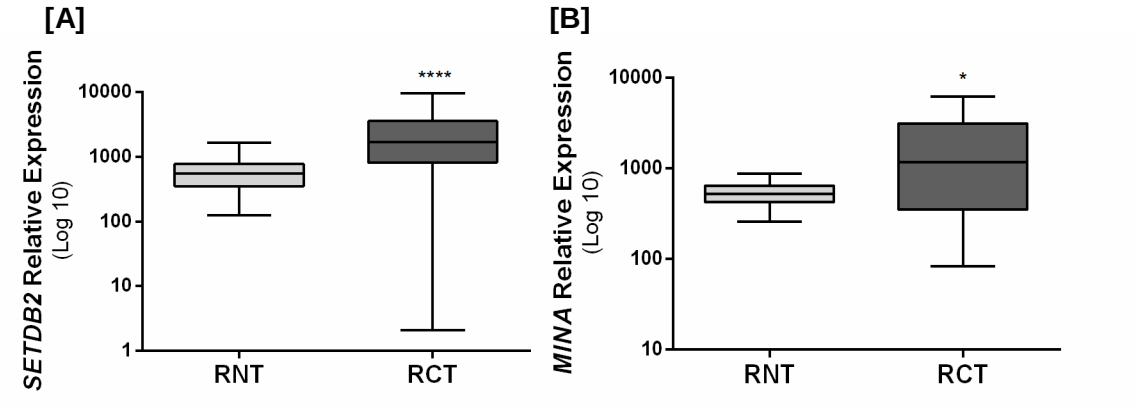


Figure 11 - Expression levels of *SETDB2* (A) and *MINA* (B) in a series of renal cell tumors (RCTs ; n=160) and renal normal tissues (RNTs ;n=13) (**** $p<0.0001$; * $p<0.05$).

Moreover, *MINA* and *SETDB2* expression levels differed significantly between benign and malignant RCTs (Figure 12 A-B).

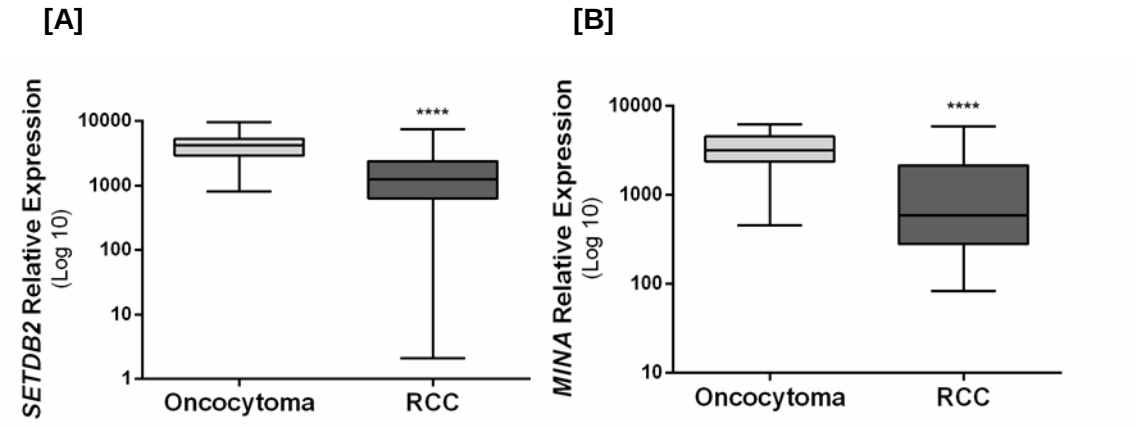


Figure 12 - Expression levels of *SETDB2* (A) and *MINA* (B) in a series of benign tumors (Oncocytomas, n=40) and malignant tumors (Renal Cell Carcinomas [RCCs], n=120) (**** $p<0.0001$).

Moreover, expression levels of both enzymes differed significantly among the four RCT subtypes (Table 2). The highest expression levels of *SETDB2* and *MINA* were displayed by oncocytomas, followed by chRCC (Figure 13 A-B and Table 2). Conversely, pRCC and ccRCC showed the lowest expression levels.

Pairwise comparisons using Mann-Whitney U test, demonstrated for both enzymes that chRCCs significantly differed from pRCCs and ccRCCs. Moreover, *SETDB2* and *MINA* discriminated pRCCs from chRCCs and oncocytomas. Furthermore, *SETDB2* transcript levels differed significantly between chRCCs and oncocytomas (Figure 13 A-B and Table 2).

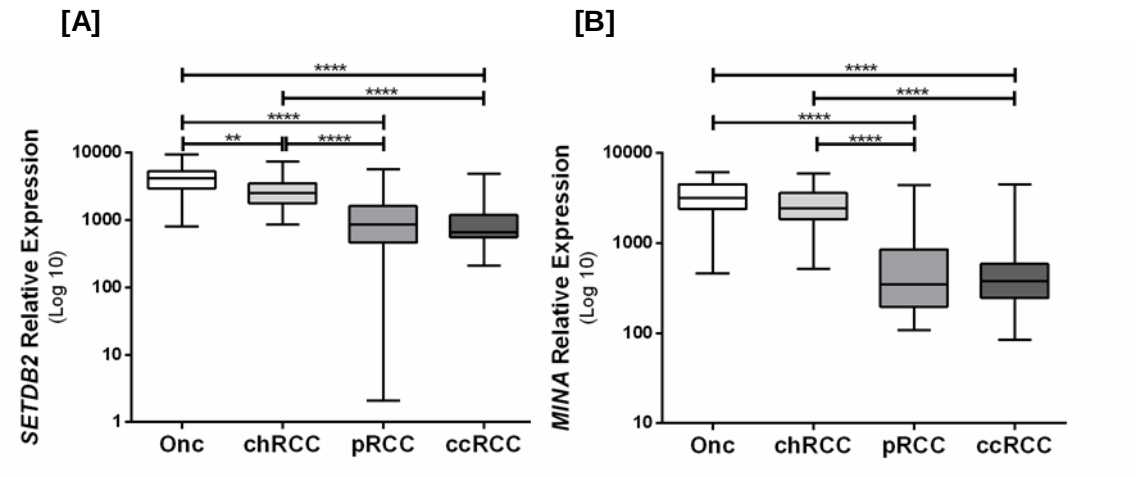


Figure 13 -Distribution of *SETDB2* (A) and *MINA* (B) expression levels among renal cell tumor subtypes. Bonferroni correction was applied and p value was adjusted to 0.0125 (**** p<0.0001;** p<0.01).

Table 2 - Pairwise comparison of *SETDB2* and *MINA* expression among renal cell tumor subtypes using Mann-Whitney test (M-W test). The values were statistically significant when p<0.0125 (Bonferroni's correction) represented in bold.

	***SETDB2* (p value)**	***MINA* (p value)**
ccRCCs *vs* pRCCs	0.391775	0.658
ccRCCs *vs* chRCCs	**2.5822×10^{-10}**	**1.0701×10^{-10}**
ccRCCs *vs* Oncocytomas	**3.0302×10^{-12}**	**6.3992×10^{-12}**
pRCCs *vs* chRCCs	**3.7112×10^{-8}**	**5.6421×10^{-11}**
pRCCs *vs* Oncocytomas	**1.0701×10^{-10}**	**7.8313×10^{-12}**
chRCCs vs Oncocytomas	**0.000332**	0.130856

2. Association Between *SETDB2* and *MINA* Expression Levels and Clinicopathological Features

The clinicopathological characteristics of patients included in this validation series are depicted in Table 3.

Table 3 - Clinical and Pathological data of patients included in the present study.

	RCT	**RNT**
Number of Patients, n	160	13
Age at diagnosis, median (min-max)	61 (29-86)	67.5 (20-83)
Gender, n (%)		
Male	92 (57.5)	7 (53.8)
Female	68 (42.5)	3 (23.1)
N.A.		3 (23.1)
Histological Subtype, n (%)		
ccRCC	40 (25)	N.A.
pRCC	40 (25)	
chRCC	40 (25)	
oncocytoma	40 (25)	
Pathological Stage, n (%)		
pT1	68 (42.5)	N.A.
pT2	23 (14.4)	
pT3	29 (18.1)	
pT4	0 (0)	
N.A.	40 (25)	
Fuhrman Grade, n (%)		
1	3 (1.9)	N.A.
2	41 (25.6)	
3	58 (36.3)	
4	18 (11.3)	
N.A.	40 (25)	
Metastasis, n (%)		
Clear cell RCC	10 (52.6)	N.A.
Papillary RCC	7 (36.8)	
Chromophobe RCC	2 (10.5)	

No significant differences in gender and age were apparent between patients and controls. In malignant tumors, no statistically significant associations were disclosed between *SETDB2* and *MINA* expression levels and Fuhrman categories or pathological stage (data not shown). In RCTs, expression levels of both enzymes were significantly higher in females (Figure 14). Moreover, *MINA* expression levels significantly associated with patient's age (p=0.015)] using 75 years as cut-off.

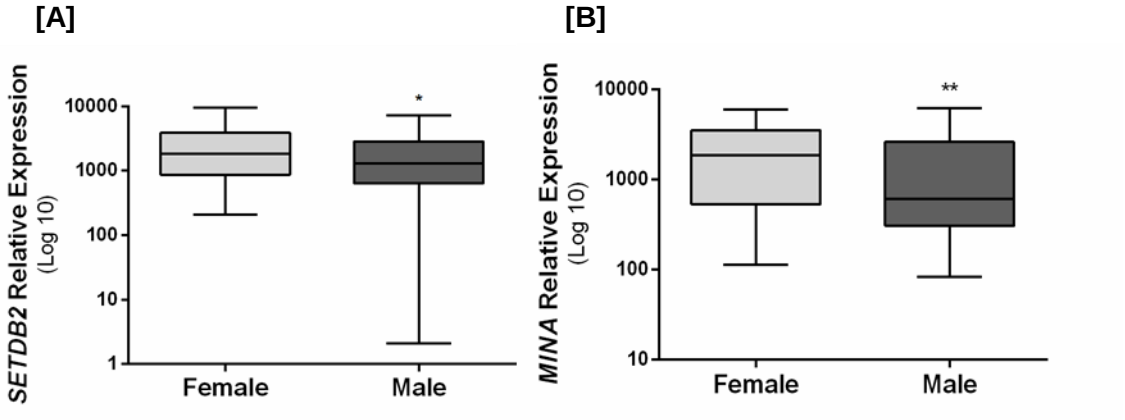


Figure 14 - Expression levels of *SETDB2* (A) and *MINA* (B) in a series of 160 RCTs distributed by gender (** p<0.001; * p<0.05).

In ccRCCs and pRCCs, *SETDB2* expression levels differed significantly between patients that developed metastases (YES) from those patients that did not (NO) (Figure 15).

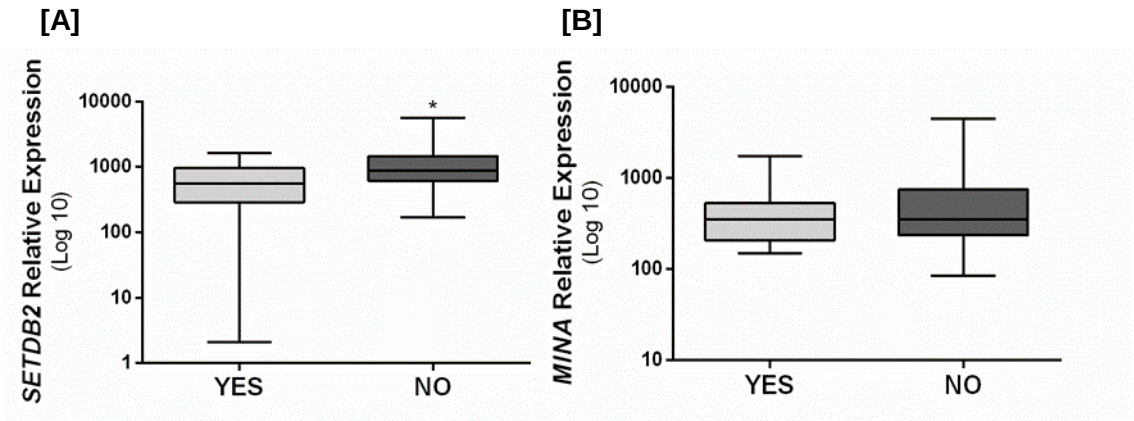


Figure 15 - Expression levels of *SETDB2* (A) and *MINA* (B) in a series of 80 ccRCCs and pRCCs distributed by presence or absence of metastases after diagnosis. No statistical association was found for *MINA* expression levels (* p<0.05).

3. Assessment of *SETDB2* and *MINA* Expression Levels as Prognostic Markers

The median follow-up of RCC patients was 175 months (range: 2-375 months). A total of 15 patients have died from RCC during this period. Disease-specific survival (DSS) analysis showed that low *SETDB2* and *MINA* levels were significantly associated with worse outcome ($p<0.01$ and $p<0.05$ respectively; supplementary Figure 1). Concerning disease-free survival (DFS) analysis, low *SETDB2* levels significantly associated with shorter time to progression ($p<0.0001$). The same trend was observed for *MINA*, but statistical significance was not reached ($p=0.055$; Figure 16). Because DFS endpoint is time to recurrence, which in RCC refers to development of metastases, subsequent analysis focused on this clinical variable.

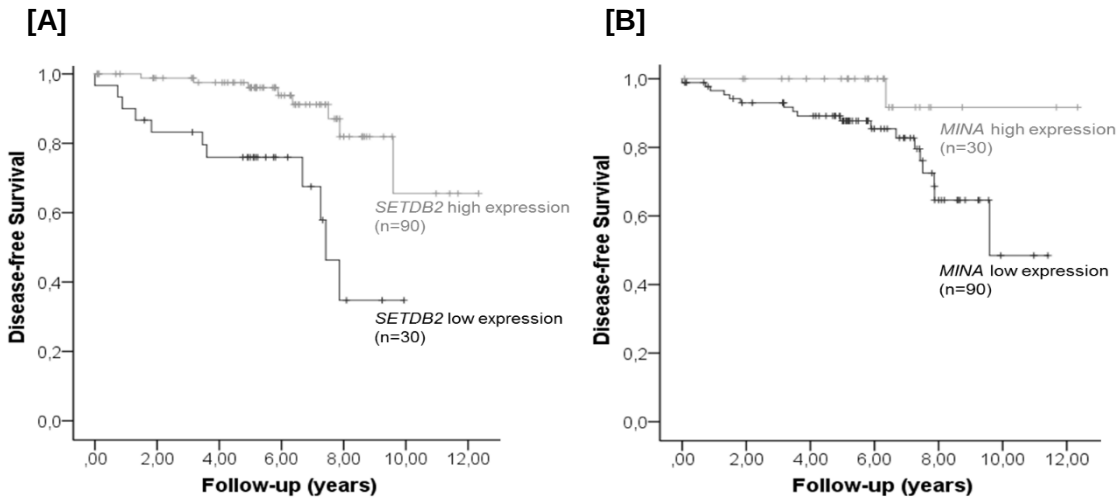


Figure 16 - Kaplan-Meier estimated disease-specific survival curves of 120 RCC patients according to expression levels of *SETDB2* (A) and *MINA* (B). The results of RT-qPCR presented were categorized using first quartile (25th percentile) value as cutoff for *SETDB2* ($p=0.00009$) and using 75th percentile value as cut off for *MINA* ($p=0.055$).

In univariable analysis, higher pathological stage (>pT3) associated with shorter survival (Figure 17), whereas gender, age, subtype and Furhman grade did not disclose any prognostic value within the available follow-up time.

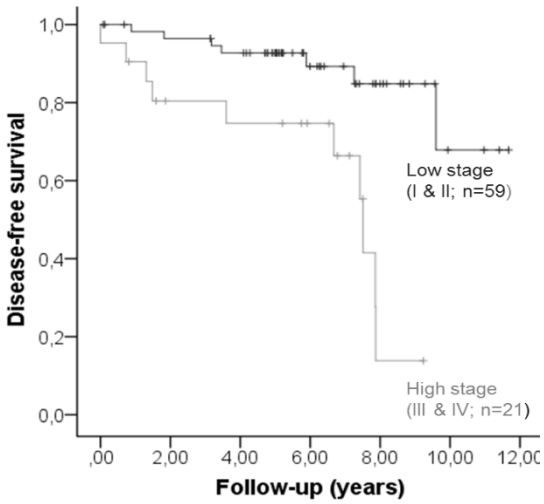


Figure 17 - Kaplan-Meier estimated disease-free survival curve of 120 RCC patients according to pathological stage High *vs* Low (p=0.0013).

In multivariable analysis, a model for assessment of prognostic value, based on an automatic variable selection algorithm (ENTER), was performed. Thus, a final model including *SETDB2* expression levels and pathological stage (High *vs* Low) was predictive of disease-free survival. Moreover a higher risk of recurrence (metastization) was defined for patients with higher pathological stage and lower *SETDB2* expression levels (Table 4).

Table 4 - Prognostic factors in Disease-free Survival obtained by Cox regression multivariable analysis using ENTER method. The high and low levels of *SETDB2* mRNA expression were categorized using 25th percentile value as cut-off.

Prognostic Factor	**Hazard Ratio (HR)**	**95% CI for HR**	**Cox regression p value**
***SETDB2* expression (low levels *vs* high levels)**	4.053	1.604 – 10.241	0.003
Stage (high stage *vs* low stage)	2.895	1.154 – 7.262	0.024

4. Exploring the role of *MINA*, *SETDB2*, *SETD3*, *SMYD2* and *NO66* in Metastasis Development in a Series of 62 ccRCCs

The clinical and pathological characteristics of patients included in this additional ccRCC validation series are depicted in Table 5.

Table 5 - Clinical and Pathological data of patients included in the additional ccRCC validation series.

	ccRCCs that developed metastases (YES)	**ccRCCs that did not developed metastases (NO)**
Number of Patients, n	31	31
Age at diagnosis, median (min-max)	66 (33-82)	62 (39.79)
Gender, n (%)		
Male	21 (67.7)	24 (77.4)
Female	10 (32.3)	7 (22.6)
Histological Subtype, n (%)		
ccRCC	31 (100)	31 (100)
pRCC	0 (0)	0 (0)
chRCC	0 (0)	0 (0)
oncocytoma	0 (0)	0 (0)
Pathological Stage*, n (%)		
pT1a	0 (0)	0 (0)
pT1b	5 (16.1)	7 (22.6)
pT2a	5 (16.1)	5 (16.1)
pT2b	2 (6.5)	4 (12.9)
pT3a	19 (61.3)	15 (48.4)
pT3b	0 (0)	0 (0)
pT4a	0 (0)	0 (0)
Fuhrman Grade, n (%)		
1	0 (0)	0 (0)
2	8 (25.8)	11 (35.5)
3	15 (48.4)	14 (45.2)
4	8 (25.8)	6 (19.4)

*At diagnosis

No significant differences were apparent between the two groups (patients with ccRCCs that developed metastases and patients with ccRCCs that did not develop metastases) for gender and age (p=0.570 and p=0.402 respectively). Furthermore, no statistically significant associations were disclosed between *SETDB2* and *MINA* expression levels and Fuhrman categories or pathological stage.

Moreover, only *MINA* expression levels differed significantly between metastasized ccRCCs and non-metastasized ccRCCs (Figure 18).

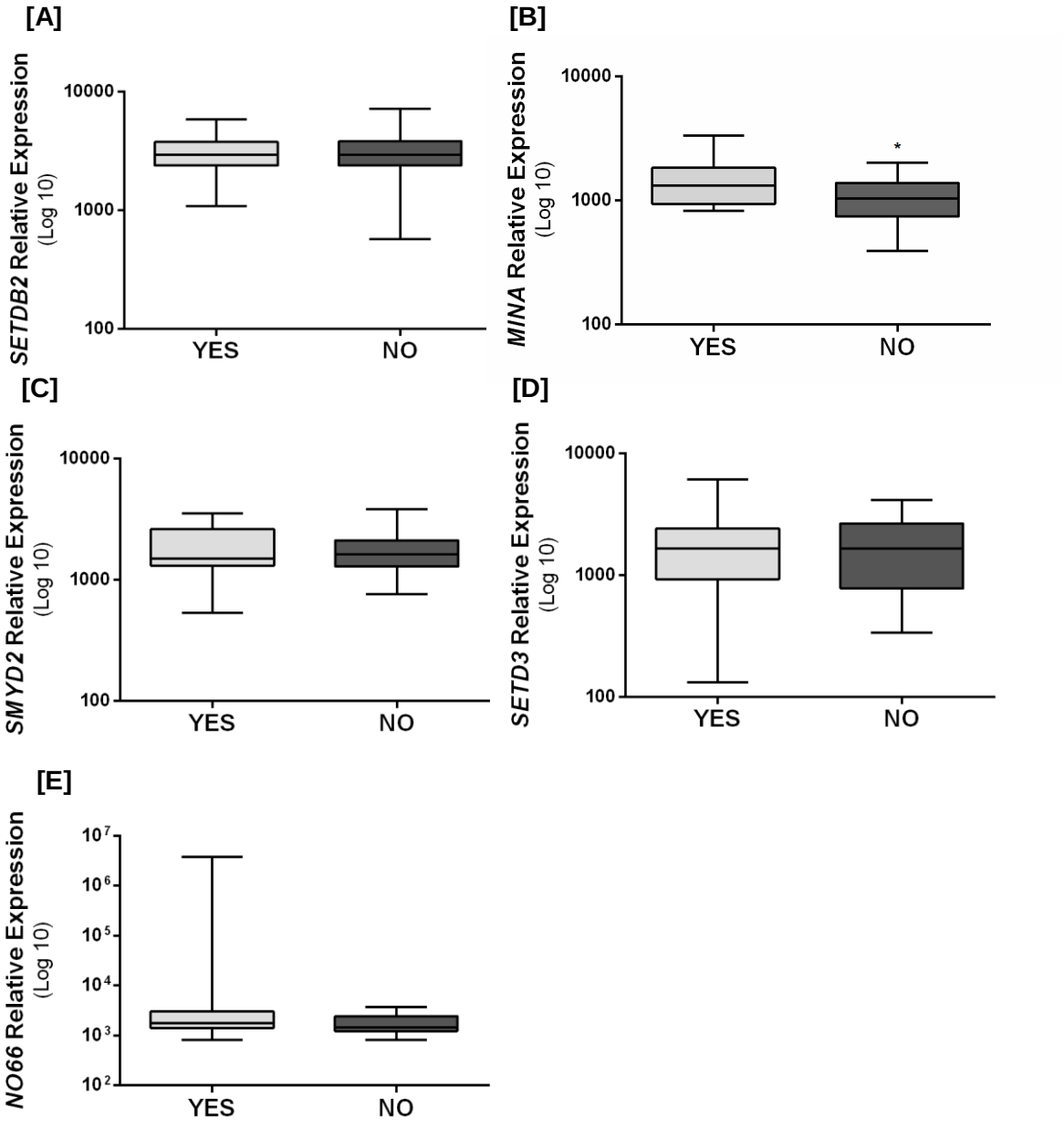


Figure 18 - *SETDB2* (A), *MINA* (B), *SMYD2* (C), *SETD3* (D), *NO66* (E) expression levels in a series of 62 ccRCCs according by development or absence of metastases during follow-up. No statistical association was found for SEDTB2, SEDT3, SMYD2 and NO66 and absence of metastasis formation (* p<0.05).

5. *MINA* Expression Analysis in the TCGA Dataset of RCC Patients

Further validation of the results depicted for *MINA* expression was performed in a larger and independent dataset from TCGA, including RNAseq expression data from 889 RCC patients (533 ccRCC, 290 pRCC, and 66 chRCC) – Table 7.

Table 6 - Clinical and Pathological data of patients included in TCGA database.

	RCT	**RNT**
Number of Patients, n	889	129
Age at diagnosis, median (min-max)	60 (17-90)	61 (28-90)
Gender, n (%)		
Male	579 (65.1)	
Female	286 (32.2)	
N.A.	24 (2.7)	
Histological Subtype, n (%)		
ccRCC	533 (60)	N.A.
pRCC	290 (32.6)	
chRCC	66 (7.4)	
oncocytoma	0	
Pathological Stage, n (%)		
pT1	454 (51.1)	N.A.
pT2	103 (11.6)	
pT3	189 (21.2)	
pT4	101 (11.4)	
N.A.	42 (4.7)	
Metastasis, n (%)		
Clear cell RCC	79 (87.8)	N.A.
Papillary RCC	9 (10)	
Chromophobe RCC	2 (2.2)	

RCT: Renal cell tumors; RNT: Renal normal tissue

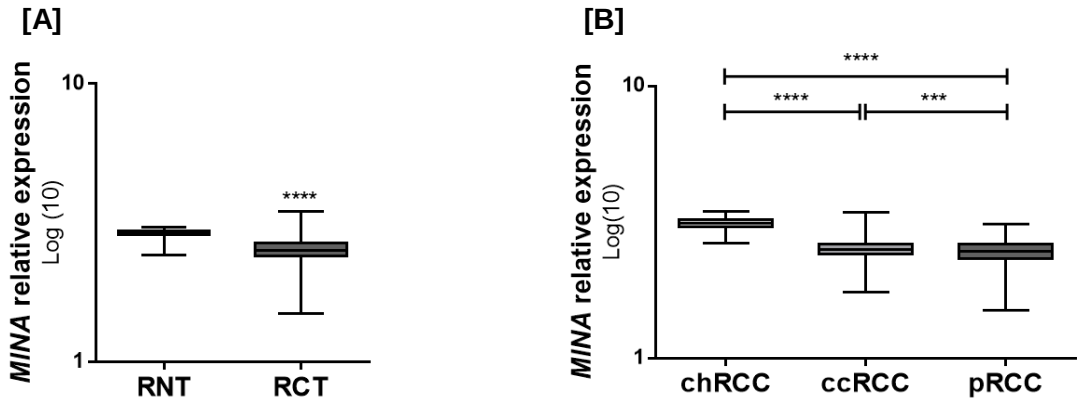


Figure 19 - Expression levels of *MINA* in a series of RCTs (n=889) and RNTs (n=129) (**** p<0.0001) (A). Distribution of *MINA* expression levels across renal cell carcinoma subtypes. Bonferroni correction was applied and p value was adjusted to 0.0125 (**** p<0.0001;*** p<0.001) (B).

Contrarily to our results, significantly lower expression levels were found in RCT compared to RNT (Figure 19A). Similar to our results, however, pairwise comparisons demonstrated that *MINA* expression levels were significantly higher in chRCCs compared to ccRCCs and pRCCs. Furthermore, *MINA* expression levels differed significantly between chRCC and ccRCC, chRCC and pRCC and ccRCC and pRCC (Figure 19B).

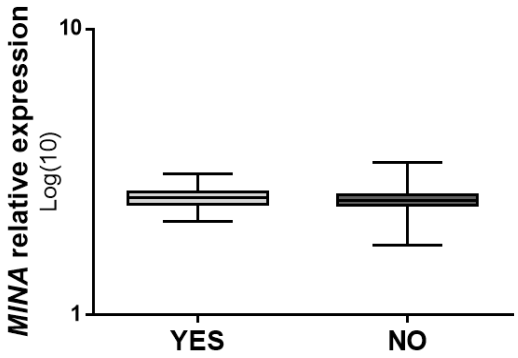


Figure 20 - *MINA* expression levels in a series of 533 ccRCCc distributed by presence or absence of metastases after diagnosis. No statistical association was disclosed.

In the ccRCCs of the TCGA database, no statistically significant difference was disclosed for *MINA* expression levels between the group of patients that developed metastases (YES) and the patients that did not develop metastases (NO) (Figure 20).

6. Validation in Renal Cell Lines

SETDB2, MINA, SETD3, SMYD2 and *NO66* expression profile was also assessed in RCC cell lines (Caki-1, 786-O, Caki-2, ACHN, 769-P, A-498) and normal kidney cell lines (HK-2 and 293-HEK) (Table 7) by quantitative RT-PCR (Figure 21).

Table 7 - Cell lines subtypes

Cell line	Disease
Caki-1	Metastasis of Clear cell RCC
Caki-2	Clear cell RCC
ACHN	Metastasis of papillary RCC
A-498	Clear cell RCC
786-O	Clear cell RCC
769-P	Clear cell RCC
HK-2	Normal kidney
293-HEK	Embrionary kidney

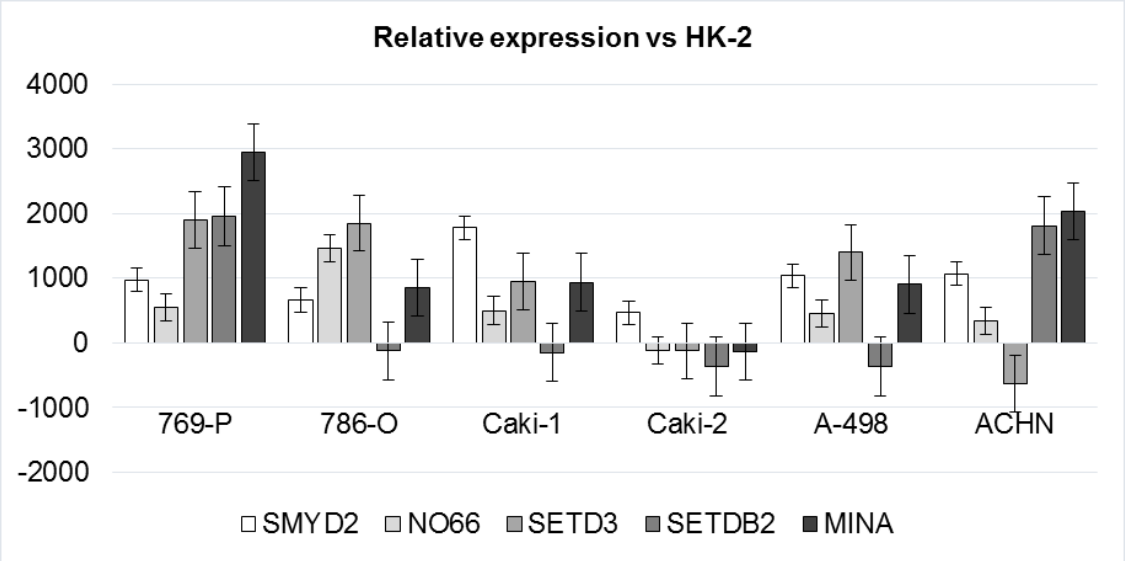


Figure 21 - Relative expression of *SMYD2*, *NO66*, *SETD3*, *SEDTB2*, *MINA* in all cell lines compared to the normal kidney cell line (HK-2).

SMYD2 mRNA levels were higher in Caki-1 and ACHN cells, which are both derived from metastatic cells. NO66 displayed higher mRNA levels in 786-O cell line. The highest transcript levels of *SETD3*, *SETDB2* and *MINA* were displayed by 769-P cell line. Globally all enzymes displayed lower mRNA levels in Caki-2 cell line, a primary cell line from ccRCC, comparing to the normal kidney cell line, HK-2.

7. Assessment of SETDB2 Protein Levels in Cell Lines

SETDB2 mRNA levels and SETDB protein levels, determined by western blot, were compared. 769-P and ACHN cells displayed higher mRNA levels (Figure 21). Furthermore lower *SETDB2* mRNA transcript was displayed by Caki-1 and Caki-2 cell lines. The same trend was observed for SETDB2 protein (Figure 22) as Caki-1 and Caki-2 cell line did not show any protein expression (undetected).

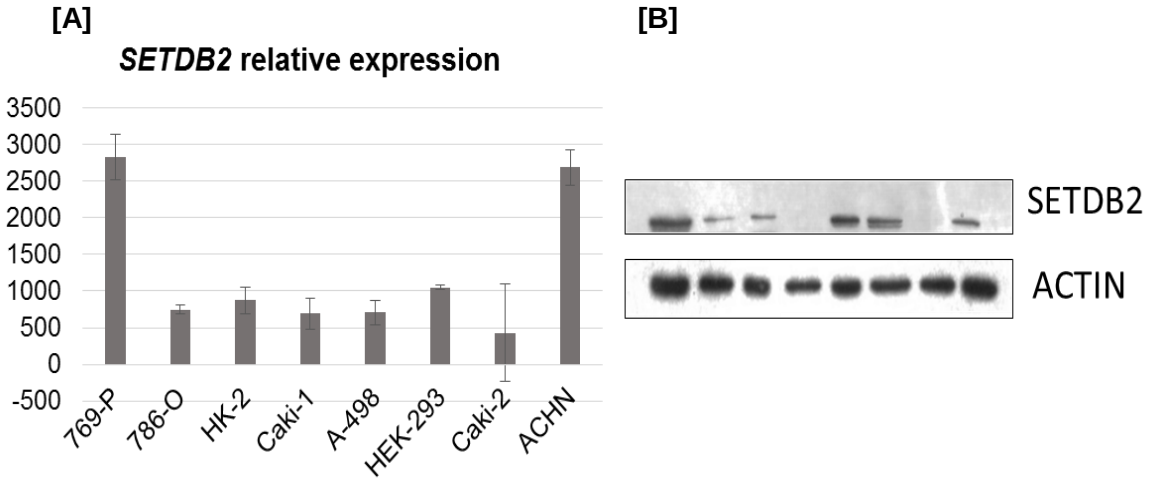


Figure 2 - *SETDB2* transcript (A) and protein (B) expression levels in kidney cell lines. Transcript expression was calculated using *GUSβ* as endogenous control gene and protein expression using Actin as endogenous control. Three biological independent samples were performed, each of them run in triplicate. The order of the Western blot is the same of the relative expression graphic.

The evaluation of the protein levels of the SMYD2, NO66 and SETD3 was not made since it was already assessed in a previous work. For *MINA*, this process is still ongoing. *SETDB2* presented the best results in our first series of validation, and for this reason we have focused in this enzyme.

DISCUSSION

Over the last decade, due to the widespread of imaging methods, the frequency of incidental detection of RCTs has significantly increased. Indeed, the incremented use of Computed Tomography (CT), Magnetic Resonance Imaging (MRI) or Ultrasound (US) allowed for the detection of small and early stage renal tumours, thought to be less aggressive. However, lymph node and distant metastases occur even in small RCCs, supporting the need of proper evaluation and adequate treatment of these small RCCs [16]. Distant metastases are the main cause of the high mortality rate of RCCs. In fact, these carcinomas are the most lethal among the common urologic cancers [71]. Thus, there is an urgent need of biomarkers capable of accurately discriminate tumours that will metastasize from those that will not, especially among those of small dimension (mostly pT1). Epigenetic-based biomarkers may assist in diagnosis, prognosis stratification and prediction of response to targeted therapy [72]. In fact, histone modifications and chromatin modulators have been demonstrated to play an important role in cancer progression [73]. In RCC, certain histone modifications associate with progression-free survival and also correlate with pathological characteristics of tumors [74]. In addition, defects in epigenetic enzymes, including chromatin remodelers and chromatin packaging have been also implicated in the development of RCTs, reflecting the role of these mechanism in the development of renal cancer [75]. In this context, we investigated whether HMTs or HDMs expression could be metastization biomarkers in RCC, and unveil the role of histone methylation in renal tumorigenesis, especially in the metastatic process. To achieve that goal, *MINA* and *SETDB2* were selected for validation in a series of 160 ccRCCs. An additional assessment in a series of 62 ccRCCs and in kidney cell lines was performed for all enzymes. Furthermore, an external validation, using TCGA dataset was carried out for *MINA*, since this enzyme displayed the best results in our second validation series.

In the first validation series (160 RCCs), *SETDB2* and *MINA* expression levels were significantly upregulated in RCTs, among which oncocytomas displayed the highest expression levels. These results are in line with those of the previously published array and is in agreement with their putative role in cancer [70]. Both enzymes showed lower expression levels in ccRCC, however, it did not significantly differ from pRCCs. This finding might be due to the common origin of ccRCC and pRCC [30]. In fact, tumors with different tubular origin (ccRCC and pRCC: proximal convoluted tubule vs Oncocytomas and chRCC: cortical collecting duct) showed the most significant differences. However, *SETDB2* expression levels, were significantly different between chRCCs and oncocytomas. Indeed, this is a relevant finding, as these two histological subtypes display variable degree of morphological overlap, rendering differential diagnosis problematic, particularly between oncocytoma and the eosinophilic variant of chRCCs. Moreover, these results are in line with

a previous publication [76] although little information is available regarding *SETDB2* role in tumorigenesis. Recently, it has been shown to be involved in leukemogenesis since it is mapped to chromosome 13q14, a region commonly affected by structural aberration in B-cell chronic lymphocytic leukemia (B-CLL) [77]. Nonetheless, no data is available concerning a putative role in metastization of solid tumors. Concerning *MINA*, high transcript or protein levels were reported as a feature of several cancers, including renal cell carcinomas, and was associated with poor prognosis [78-84]. Regarding the role of *MINA* in metastization, several other studies showed an association between high *MINA* expression levels and the development of metastasis (lymph node or distant metastases) in other cancer models, such as cholangiocarcinoma, gastric adenocarcinoma and pancreatic carcinoma [82, 83, 85].

We found that *SETDB2* expression levels were significantly higher in ccRCC and pRCC that developed distant metastases, contrarily to *MINA*. Only these two subtypes were evaluated concerning metastatic behaviour as they represent the most aggressive subtypes of RCC, entailing the need for the identification of biomarkers predictive of tumor dissemination.

The prognostic significance of altered *SETDB2* and *MINA* mRNA expression levels was assessed through disease-specific and disease-free survival analysis. Interestingly, in univariable analysis, both genes low expression levels were associated with worse disease-specific survival. This result is opposed to previous reports in esophageal and renal cancers [76, 84]. Nevertheless, in lung cancer, *MINA* overexpression was associated with favourable outcome [86]. It should be emphasized that our results are mostly influenced by pRCC and ccRCC which displayed the lowest expression levels among RCCs. Regarding disease-free survival, the results were similar to those of disease-specific survival, although only *SETDB2* expression levels significantly associated with development of metastases.

When relevant clinicopathological variables were included in the analysis, only *SEDTB2* expression levels and Stage retained statistical significance for disease free survival. In fact, *SETDB2* expression levels could independently predict disease-free survival in RCC patients. However, these results should be analysed with caution since only 19 events (metastases) were found within the follow-up period (7 pRCC, 10 ccRCC and 2 chRCC). Thus, *SETDB2* and *MINA* expression was evaluated in a second series of RCTs. Because ccRCC metastasizes more commonly than any other RCC subtype, the analysis was restricted to 62 ccRCCs.

Three additional enzymes, *SMYD2*, *SETD3* and *NO66*, known to play a role in RCC and previously associated with poor prognosis [70] were also selected. Significantly higher

MINA expression levels were found in ccRCCs that developed metastasis, whereas no differences were found for *SETDB2* expression. This could be due to the differences in the validation series (in the first series pRCCs were also analysed), and thus *SETDB2* could be a biomarker of metastases in pRCCs but not in ccRCCs. Because similar analyses were not performed specifically in pRCCs, no definitive conclusions can be made at this time.

To further validate these results in a completely independent dataset, the TCGA basis was surveyed. Although in this dataset statistically significant differences in MINA expression levels were also depicted between RCTs and RNTs, higher *MINA* expression levels were found NRTs. This discrepancy might derive from the nature of normal renal tissue analysed. Indeed, whereas in the TCGA dataset normal renal tissue was procured from morphologically normal renal parenchyma adjacent to tumor, we used RNT from cases without RCC. Indeed, we have already demonstrated that normal tissues adjacent to RCC harbour epigenetic alterations that may precede neoplastic transformation [87]. Concerning, MINA, no differences were apparent between ccRCCs that developed metastases and ccRCCs that did not, among the 533 ccRCCs from the TCGA dataset. This result was somewhat unexpected but it may be due to differences in follow-up time and enrolment criteria, as we excluded from analysis cases that presented metastasis at diagnosis and only analysed cases in which metastases developed after an apparently curative surgical treatment.

Based on the previous results, we started to unveil the biological significance of altered expression of the five histone modifying enzymes analysed. Thus, we assessed mRNA expression levels in kidney cell lines, including benign and malignant. Nevertheless, results were very heterogeneous and a clear trend was not apparent, precluding definitive conclusions. Because *SETDB2* displayed the most promising results as biomarker of metastatic behaviour, we further assessed whether an association exists between transcript and protein levels. We found that in cell lines there is, indeed, a similar trend between mRNA and protein expression. Thus, the assessment of *SETDB2* expression might be accomplished through immunohistochemistry, allowing for its use on a routine basis, although its high expression levels in normal tissues may constitute a challenge. Despite several array-based biomarkers have been proposed for RCC, most of them have not been validated or have just been validated in a single and limited series of patients [88-91]. Nevertheless, prognostic and predictive biomarkers are clearly needed for RCC, allowing for improved risk stratification and identification of molecular targets.

CONCLUSION AND FUTURE PERSPECTIVES

In this study we found that altered expression levels of enzymes involved in histone methylation are associated with renal tumorigenesis and metastization. In particular, it was found that *SETDB2* and *MINA*, a histone methyltransferase of lysines 4 and 36 of H3 and a histone demethylase, were upregulated in RCTs compared to RNTs. Moreover, those two genes displayed significantly higher expression levels in oncocytomas and were able to distinguish malignant from benign RCTs.

Additionally, ccRCCs and pRCCs that did not develop metastasis displayed significantly higher *SETDB2* expression levels than tumors that progressed and disseminated, thus suggesting that this enzyme might be a metastasis predictor in these RCC subtypes. Moreover, survival analysis revealed that combined *SETDB2* expression levels and Stage (high vs low), were independent prognostic factors for disease-free survival. Conversely, *MINA* was significantly overexpressed in ccRCCs that developed metastasis

Overall, the present study demonstrated that *SETDB2* and *MINA* are involved in the development of metastasis in renal tumorigenesis and might be useful prognostic biomarkers.

In the near future:

We will assess *SETDB2* expression in a series of pRCCs to determine whether it is also a metastases biomarker of this particular subtype

We intend not only consolidate some of the results already obtained but also explore other mechanisms that might clarify the role of those histone methylation enzymes (MINA and SETDB2) in renal tumorigenesis. Thus, genes that might be regulated by these two enzymes will be investigated.

Additionally, because 769-P cells presented the highest *MINA* expression levels, this cell line will be selected for phenotypic assays, such as migration and invasion after transfection with sh-MINA, to assess the functional significance of our findings.

REFERENCES

1. Raissian, Y. and J.P. Grande, *Embryology and Normal Anatomy of the Kidney.* Practical Renal Pathology, A Diagnostic Approach: A Volume in the Pattern Recognition Series, 2012: p. 1.
2. Chow, W.H., L.M. Dong, and S.S. Devesa, *Epidemiology and risk factors for kidney cancer.* Nat Rev Urol, 2010. **7**(5): p. 245-57.
3. Pascual, D. and A. Borque, *Epidemiology of Kidney Cancer.* Advances in Urology, 2008. **2008**: p. 7.
4. Ferlay, J., et al., *Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012.* European Journal of Cancer, 2013. **49**(6): p. 1374-1403.
5. Ferlay, J., et al., *Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012.* International Journal of Cancer, 2015. **136**(5): p. E359-E386.
6. Ljungberg, B., et al., *The epidemiology of renal cell carcinoma.* European urology, 2011. **60**(4): p. 615-621.
7. Ljungberg, B., et al., *EAU guidelines on renal cell carcinoma: 2014 update.* European urology, 2015. **67**(5): p. 913-924.
8. Shapiro, J.A., et al., *Hypertension, antihypertensive medication use, and risk of renal cell carcinoma.* American journal of epidemiology, 1999. **149**(6): p. 521-530.
9. Stolle, C., et al., *Improved detection of germline mutations in the von Hippel-Lindau disease tumor suppressor gene.* Human mutation, 1998. **12**(6): p. 417.
10. Coleman, J.A., *Familial and hereditary renal cancer syndromes.* Urologic Clinics of North America, 2008. **35**(4): p. 563-572.
11. Pfaffenroth, E.C. and W.M. Linehan, *Genetic basis for kidney cancer: opportunity for disease-specific approaches to therapy.* Expert opinion on biological therapy, 2008. **8**(6): p. 779-790.
12. Khoo, S.K., et al., *Inactivation of BHD in sporadic renal tumors.* Cancer Research, 2003. **63**(15): p. 4583-4587.
13. Matson, M.A. and E.P. Cohen, *Acquired cystic kidney disease: occurrence, prevalence, and renal cancers.* Medicine, 1990. **69**(4): p. 217-226.
14. Patard, J.J., et al., *Correlation between symptom graduation, tumor characteristics and survival in renal cell carcinoma.* Eur Urol, 2003. **44**(2): p. 226-32.
15. Kim, H.L., et al., *Paraneoplastic signs and symptoms of renal cell carcinoma: implications for prognosis.* J Urol, 2003. **170**(5): p. 1742-6.
16. Guðmundsson, E., et al., *Metastatic potential in renal cell carcinomas≤ 7cm: Swedish Kidney Cancer Quality Register data.* European urology, 2011. **60**(5): p. 975-982.

17. Jewett, M.A., et al., *Active surveillance of small renal masses: progression patterns of early stage kidney cancer.* European urology, 2011. **60**(1): p. 39-44.
18. Eble, J., G. Sauter, and J. Epstein, *Pathology and Genetics of Tumours of the Urinary System and Male Genital Organs. Lyon, France: IARC; 2004.* World Health Organization Classification of Tumours.
19. Moch, H., *An overview of renal cell cancer: pathology and genetics.* Semin Cancer Biol, 2013. **23**(1): p. 3-9.
20. Cheng, L., et al., *Molecular and cytogenetic insights into the pathogenesis, classification, differential diagnosis, and prognosis of renal epithelial neoplasms.* Hum Pathol, 2009. **40**(1): p. 10-29.
21. Abrahams, N.A. and P. Tamboli, *Epithelial Neoplasms of the Renal Cortex.* Practical Renal Pathology, A Diagnostic Approach: A Volume in the Pattern Recognition Series, 2012: p. 227.
22. Kim, W.Y. and W.G. Kaelin, *Role of VHL gene mutation in human cancer.* J Clin Oncol, 2004. **22**(24): p. 4991-5004.
23. Cohen, H.T. and F.J. McGovern, *Renal-cell carcinoma.* N Engl J Med, 2005. **353**(23): p. 2477-90.
24. Kaelin, W.G., *The von Hippel-Lindau tumor suppressor protein and clear cell renal carcinoma.* Clinical Cancer Research, 2007. **13**(2): p. 680s-684s.
25. Kaelin, W.G., *Molecular basis of the VHL hereditary cancer syndrome.* Nature Reviews Cancer, 2002. **2**(9): p. 673-682.
26. Sato, Y., et al., *Integrated molecular analysis of clear-cell renal cell carcinoma.* Nature genetics, 2013. **45**(8): p. 860-867.
27. Network, C.G.A.R., *Comprehensive molecular characterization of clear cell renal cell carcinoma.* Nature, 2013. **499**(7456): p. 43-49.
28. Peruzzi, B. and D.P. Bottaro, *Targeting the c-Met signaling pathway in cancer.* Clinical Cancer Research, 2006. **12**(12): p. 3657-3660.
29. Klatte, T., et al., *Cytogenetic and molecular tumor profiling for type 1 and type 2 papillary renal cell carcinoma.* Clinical Cancer Research, 2009. **15**(4): p. 1162-1169.
30. Bodmer, D., et al., *Understanding familial and non-familial renal cell cancer.* Human molecular genetics, 2002. **11**(20): p. 2489-2498.
31. Vera-Badillo, F.E., E. Conde, and I. Duran, *Chromophobe renal cell carcinoma: a review of an uncommon entity.* International Journal of Urology, 2012. **19**(10): p. 894-900.
32. Contractor, H., et al., *Mutation of the p53 tumour suppressor gene occurs preferentially in the chromophobe type of renal cell tumour.* The Journal of pathology, 1997. **181**(2): p. 136-139.

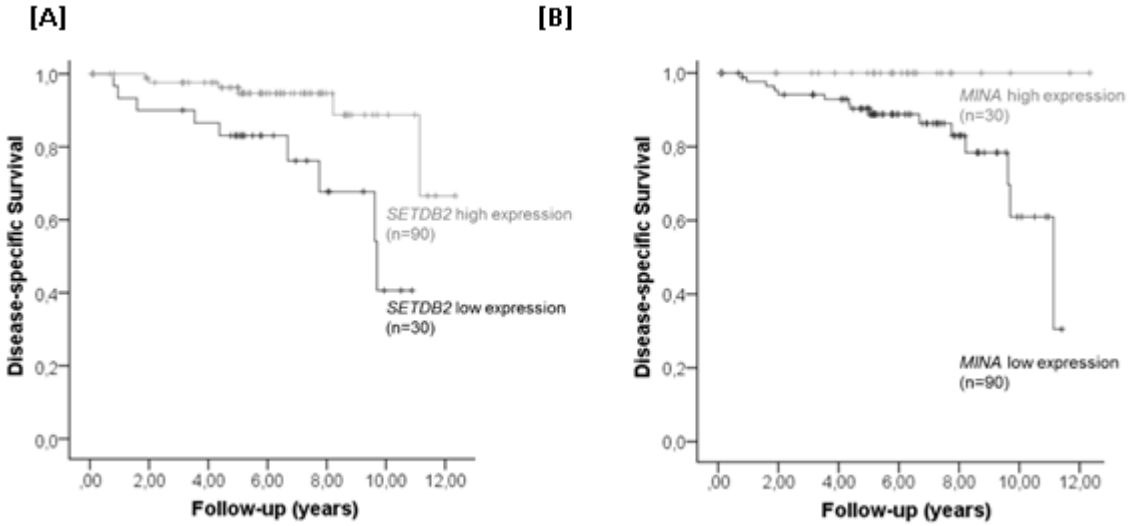
33. Süküsd, F., et al., *Allelic loss at 10q23. 3 but lack of mutation of PTEN/MMAC1 in chromophobe renal cell carcinoma.* Cancer genetics and cytogenetics, 2001. **128**(2): p. 161-163.
34. Junker, K., et al., *Genetic subtyping of renal cell carcinoma by comparative genomic hybridization*, in *Molecular Staging of Cancer*. 2003, Springer. p. 169-175.
35. Sobin, L.H., M.K. Gospodarowicz, and C. Wittekind, *TNM classification of malignant tumours*. 2011: John Wiley & Sons.
36. Ljungberg, B., et al., *EAU Guidelines on Renal Cell Carcinoma: 2014 Update.* European Urology, 2015. **67**(5): p. 913-924.
37. Fisher, R., M. Gore, and J. Larkin. *Current and future systemic treatments for renal cell carcinoma*. in *Seminars in cancer biology*. 2013. Elsevier.
38. Conti, A., et al., *Progress of molecular targeted therapies for advanced renal cell carcinoma.* BioMed research international, 2013. **2013**.
39. Patel, P.H., et al., *Targeting von Hippel-Lindau pathway in renal cell carcinoma.* Clinical Cancer Research, 2006. **12**(24): p. 7215-7220.
40. Dabestani, S., et al., *Local treatments for metastases of renal cell carcinoma: a systematic review.* The Lancet Oncology, 2014. **15**(12): p. e549-e561.
41. Srinivasan, R. and W.M. Linehan, *Treatment of advanced renal cell carcinoma.* Campbell-Walsh Urology, 2012. **2**: p. 1477-1478.
42. Fidler, I.J., *The pathogenesis of cancer metastasis: the'seed and soil'hypothesis revisited.* Nature Reviews Cancer, 2003. **3**(6): p. 453-458.
43. Fidler, I.J., *The organ microenvironment and cancer metastasis.* Differentiation, 2002. **70**(9-10): p. 498-505.
44. Gupta, G.P. and J. Massagué, *Cancer metastasis: building a framework.* Cell, 2006. **127**(4): p. 679-695.
45. Kirkali, Z., E. Tuzel, and M. Mungan, *Recent advances in kidney cancer and metastatic disease.* BJU international, 2001. **88**(8): p. 818-824.
46. Bukowski, R. *Cytokine combinations: therapeutic use in patients with advanced renal cell carcinoma*. in *Seminars in oncology*. 2000.
47. Vasselli, J.R., et al., *Predicting survival in patients with metastatic kidney cancer by gene-expression profiling in the primary tumor.* Proceedings of the National Academy of Sciences, 2003. **100**(12): p. 6958-6963.
48. Esteller, M., *Epigenetics in cancer.* New England Journal of Medicine, 2008. **358**(11): p. 1148-1159.
49. Portela, A. and M. Esteller, *Epigenetic modifications and human disease.* Nature biotechnology, 2010. **28**(10): p. 1057-1068.

50. Rodríguez-Paredes, M. and M. Esteller, *Cancer epigenetics reaches mainstream oncology.* Nature medicine, 2011: p. 330-339.
51. Suvà, M.L., N. Riggi, and B.E. Bernstein, *Epigenetic reprogramming in cancer.* Science, 2013. **339**(6127): p. 1567-1570.
52. Esteller, M., *Non-coding RNAs in human disease.* Nature Reviews Genetics, 2011. **12**(12): p. 861-874.
53. Costa, F.F., *Non-coding RNAs, epigenetics and complexity.* Gene, 2008. **410**(1): p. 9-17.
54. Garzon, R., G.A. Calin, and C.M. Croce, *MicroRNAs in cancer.* Annual review of medicine, 2009. **60**: p. 167-179.
55. Mattick, J.S. and I.V. Makunin, *Non-coding RNA.* Human molecular genetics, 2006. **15**(suppl 1): p. R17-R29.
56. Vasudevan, S., Y. Tong, and J.A. Steitz, *Switching from repression to activation: microRNAs can up-regulate translation.* Science, 2007. **318**(5858): p. 1931-1934.
57. Melo, S.A. and M. Esteller, *Dysregulation of microRNAs in cancer: playing with fire.* FEBS letters, 2011. **585**(13): p. 2087-2099.
58. Esquela-Kerscher, A. and F.J. Slack, *Oncomirs—microRNAs with a role in cancer.* Nature Reviews Cancer, 2006. **6**(4): p. 259-269.
59. Vardabasso, C., et al., *Histone variants: emerging players in cancer biology.* Cellular and molecular life sciences, 2014. **71**(3): p. 379-404.
60. Kamakaka, R.T. and S. Biggins, *Histone variants: deviants?* Genes & development, 2005. **19**(3): p. 295-316.
61. Maze, I., et al., *Every amino acid matters: essential contributions of histone variants to mammalian development and disease.* Nature reviews Genetics, 2014. **15**(4): p. 259-271.
62. Sharma, S., T.K. Kelly, and P.A. Jones, *Epigenetics in cancer.* Carcinogenesis, 2010. **31**(1): p. 27-36.
63. Dawson, M.A., T. Kouzarides, and B.J. Huntly, *Targeting epigenetic readers in cancer.* New England Journal of Medicine, 2012. **367**(7): p. 647-657.
64. Waldmann, T. and R. Schneider, *Targeting histone modifications—Epigenetics in cancer.* Current opinion in cell biology, 2013. **25**(2): p. 184-189.
65. Falkenberg, K.J. and R.W. Johnstone, *Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders.* Nature reviews Drug discovery, 2014. **13**(9): p. 673-691.
66. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.

67. Lennartsson, A. and K. Ekwall, *Histone modification patterns and epigenetic codes.* Biochimica et Biophysica Acta (BBA)-General Subjects, 2009. **1790**(9): p. 863-868.
68. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-1080.
69. Mellor, J., *The dynamics of chromatin remodeling at promoters.* Molecular cell, 2005. **19**(2): p. 147-157.
70. Pires-Luís, A.S., et al., *Expression of histone methyltransferases as novel biomarkers for renal cell tumor diagnosis and prognostication.* Epigenetics, 2015. **10**(11): p. 1033-1043.
71. Chin, A.I., et al., *Surveillance strategies for renal cell carcinoma patients following nephrectomy.* Reviews in urology, 2006. **8**(1): p. 1.
72. Bianco-Miotto, T., et al., *Global levels of specific histone modifications and an epigenetic gene signature predict prostate cancer progression and development.* Cancer Epidemiology Biomarkers & Prevention, 2010. **19**(10): p. 2611-2622.
73. Chi, P., C.D. Allis, and G.G. Wang, *Covalent histone modifications—miswritten, misinterpreted and mis-erased in human cancers.* Nature reviews cancer, 2010. **10**(7): p. 457-469.
74. Ramakrishnan, S., L. Ellis, and R. Pili, *Histone modifications: implications in renal cell carcinoma.* Epigenomics, 2013. **5**(4): p. 453-462.
75. Larkin, J., et al., *Epigenetic regulation in RCC: opportunities for therapeutic intervention?* Nature Reviews Urology, 2012. **9**(3): p. 147-155.
76. Ishizaki, H., et al., *Overexpression of the myc target gene Mina53 in advanced renal cell carcinoma.* Pathology international, 2007. **57**(10): p. 672-680.
77. Mabuchi, H., et al., *Cloning and Characterization of CLLD6, CLLD7, and CLLD8, Novel Candidate Genes for Leukemogenesis at Chromosome 13q14, a Region Commonly Deleted in B-Cell Chronic Lymphocytic Leukemia.* Cancer Research, 2001. **61**(7): p. 2870-2877.
78. Teye, K., et al., *Increased expression of a Myc target gene Mina53 in human colon cancer.* The American journal of pathology, 2004. **164**(1): p. 205-216.
79. Ogasawara, S., et al., *Accelerated expression of a Myc target gene Mina53 in aggressive hepatocellular carcinoma.* Hepatology Research, 2010. **40**(3): p. 330-336.
80. Komiya, K., et al., *Mina53, a novel c-Myc target gene, is frequently expressed in lung cancers and exerts oncogenic property in NIH/3T3 cells.* Journal of cancer research and clinical oncology, 2010. **136**(3): p. 465-473.
81. Zhang, Q., et al., *Expression of Mina53 and its significance in gastric carcinoma.* The International journal of biological markers, 2007. **23**(2): p. 83-88.

82. Tan, X.-P., et al., *Upregulated expression of Mina53 in cholangiocarcinoma and its clinical significance.* Oncology letters, 2012. **3**(5): p. 1037-1041.
83. Tan, X.-p., et al., *Potential effects of Mina53 on tumor growth in human pancreatic cancer.* Cell biochemistry and biophysics, 2014. **69**(3): p. 619-625.
84. Tsuneoka, M., et al., *Mina53 as a potential prognostic factor for esophageal squamous cell carcinoma.* Clinical Cancer Research, 2004. **10**(21): p. 7347-7356.
85. Xing, J., et al., *Mina53, a novel molecular marker for the diagnosis and prognosis of gastric adenocarcinoma.* Oncology reports, 2014. **31**(2): p. 634-640.
86. Komiya, K., et al., *Expression of Mina53, a novel c-Myc target gene, is a favorable prognostic marker in early stage lung cancer.* Lung Cancer, 2010. **69**(2): p. 232-238.
87. Costa, V.L., et al., *Quantitative promoter methylation analysis of multiple cancer-related genes in renal cell tumors.* BMC cancer, 2007. **7**(1): p. 1.
88. Jerónimo, C. and R. Henrique, *Epigenetic biomarkers in urological tumors: A systematic review.* Cancer letters, 2014. **342**(2): p. 264-274.
89. Junker, K., et al., *CD70: a new tumor specific biomarker for renal cell carcinoma.* The Journal of urology, 2005. **173**(6): p. 2150-2153.
90. Meyer, H.-A., et al., *Identification of stanniocalcin 2 as prognostic marker in renal cell carcinoma.* European urology, 2009. **55**(3): p. 669-678.
91. Redova, M., et al., *Circulating miR-378 and miR-451 in serum are potential biomarkers for renal cell carcinoma.* Journal of translational medicine, 2012. **10**(1): p. 1.

SUPLEMENTARY DATA



Supplementary Figure 1 - Kaplan-Meier estimated Disease-specific survival curves of 120 RCC patients according to *SETDB2* expression levels of (A; p=0.002) and *MINA* (B; p=0.021). The RT-qPCR results were categorized using 25th and 75th percentile value as cut-off for *SETDB2* and *MINA*, respectively.