Sílvia Fernandes Rocha
Metabolomic studies in human renal cell carcinoma cell lines using gas chromatography/mass spectrometry

Metabolomic studies in human renal cell carcinoma cell lines using gas chromatography/mass spectrometry
Sílvia Fernandes Rocha

MESTRADO EM ONCOLOGIA

ESPECIALIZAÇÃO EM ONCOLOGIA LABORATORIAL

Metabolomic studies in human renal cell carcinoma cell lines using gas chromatography/mass spectrometry

Sílvia Fernandes Rocha

2019

SÍLVIA FERNANDES ROCHA

Metabolomic studies in human renal cell carcinoma cell lines using gas chromatography/mass spectrometry

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

Orientadora – Doutora Maria Paula do Amaral Alegria Guedes de Pinho
Categoria – Investigadora Coordenadora
Afiliação – Laboratório Associado UCIBIO - REQUIMTE, Laboratório de Toxicologia, do Departamento de Ciências Biológicas da Faculdade de Farmácia da Universidade do Porto

Coorientadora – Professora Doutora Márcia Cláudia Dias de Carvalho
Categoria – Professora Associada
Afiliação – Faculdade de Ciências da Saúde da Universidade Fernando Pessoa

This work was financed by FEDER - Fundo Europeu de Desenvolvimento Regional funds through the COMPETE 2020 - Operational Programme for Competitiveness and Internationalisation (POCI), and by Portuguese funds through FCT - Fundação para a Ciência e a Tecnologia in the framework of the project POCI-01-0145-FEDER-030388 - “ACCuseD - renAl Cell Carcinoma detection renal cancer Detection: a translational metabolomics research based on Volatile Organic Compounds fingerprinting” - PTDC/SAU-SER/30388/2017.

Acknowledgments

My acknowledgments will be written in Portuguese, because only then all the special people who accompanied me during this journey will be able to read them.

Dr.ª Paula Guedes e Prof.ª Dr.ª Márcia Carvalho, o meu imenso obrigada por todo o suporte, atenção, disponibilidade e orientação. Sem a vossa ajuda e positivismo esta dissertação não teria sido concretizada.

Obrigada ao Laboratório de Toxicologia do Departamento de Ciências Biológicas da Faculdade de Farmácia da Universidade do Porto por toda a atenção, disponibilidade e dedicação.

Estendo os meus agradecimentos à Prof.ª Carmen Jerónimo e ao grupo de Epigenética e Biologia do Cancro do IPO-Porto, bem como à Faculdade de Ciências da Saúde da Universidade Fernando Pessoa, por toda a colaboração e suporte prestados.

A toda a minha família agradeço profundamente por todo o amor e carinho que me deram durante esta jornada.

Avôs padrinhos, sempre preocupados com a minha felicidade e bem-estar, foram os meus melhores conselheiros e pilares de sabedoria. Sou-vos eternamente grata.

Tias Anabela e Alexandra muito obrigada pela vossa ajuda e partilha de experiências universitárias.

Mãe, Pai, e Maninhas do meu coração, obrigada por nunca terem deixado a distância desvanecer a harmonia familiar, por todo o amor e carinho e por nunca me terem deixado desistir.

Juliana e Inês, as minhas segundas irmãs e melhores confidentes, obrigada por todos os telefonemas e desabafos, em particular pelas lágrimas, abraços e gargalhadas que demos juntas.

O meu último agradecimento e mais especial é feito a ti Margarida, a minha “Boss” e anjo da guarda do laboratório. Obrigada por me teres acompanhado e ajudado em todas as ocasiões, até mesmo ao fim-de semana. Obrigada por teres deixado o teu doutoramento em segundo

plano para te dedicares a mim e a este projeto. Obrigada por nunca teres duvidado de mim e das minhas capacidades. Obrigada por teres aturado todas as minhas loucuras, e principalmente por teres sido a minha boia de salvação nos momentos menos bons. Essencialmente agradeço-te por nunca teres desistido de mim. Sem ti eu nunca teria conseguido chegar ao fim, e por isso te serei eternamente grata.

A todos vocês,

Muito obrigada!

Abstract

Metabolomics is defined as the systematic study of all metabolites (the metabolome) within a biological system. This '-omic' science involves the qualitative and quantitative analysis of a huge diversity of low molecular weight metabolites (LMWM) present in several biological matrices, such as cells, tissues or biofluids. Metabolomics has been employed in several clinical areas (e.g. diagnosis, prognosis, and therapy) due to the key idea that particular disease or pathological condition may present an exclusive metabolic profile.

Metabolites are the final products of the biochemical pathways and the change in their normal levels can affect the normal functioning of the organism. Their identification and measurement can be performed in matrices whose collection is less invasive than biopsies, so, over the years, metabolomics has proved to be particularly valuable in the research of biomarkers for cancer diagnosis. The application of a metabolomics approach in cancer research arises from the fact that cancer cells are metabolically reprogrammed to control the energy required by the rapid growth and development of the tumor, producing a specific "metabolic signature".

Renal Cell Carcinoma (RCC) constitute approximately 90-95% of all kidney neoplasms and is a very common and fatal group of malignancies whose incidence continues to rise. In addition, current diagnostic techniques are not very sensitive and specific, and access to the RCC subtype requires the use of extremely invasive approaches. For these reasons, it is clearly necessary the identification of appropriate biomarkers to develop faster and more sophisticated as well as less invasive diagnostic techniques.

The project developed in this dissertation aimed to evaluate the potential of the metabolic signature, particularly the volatile organic compounds (VOCs) and volatile carbonyl compounds (VCCs) in the exometabolome, to discriminate RCC cells from normal cells and between two histological subtypes (clear cell RCC and papillary RCC). In order to achieve these goals, Headspace-Solid Phase Microextraction/Gas Chromatography-Mass Spectrometry (HS-SPME/GC-MS)-based metabolomics was applied for the metabolic profiling of five different tumorigenic cell lines, namely three clear cell (769-P, 786-O and Caki-1) and two papillary RCC (Caki-2 and ACHN), and one non-tumorigenic cell line (HK-2). Aiming at the classification of the disease according to histological grade and subtype, this study comprised the analysis of two RCC subtypes, the clear cell and the papillary, in both metastatic and non-metastatic stages. After statistical analysis, it was observed that VOCs and VCCs collected from the

exometabolome allowed good separation between tumorigenic and normal cell lines, as well as the discrimination between subtypes (clear cell *vs* papillary) and between low- and high-grade papillary RCC, but only in the case of VCCs.

The obtained results are promising and reaffirm the importance of studying potential VOC- and VCC-biomarkers for the diagnosis of kidney cancer (KC), in particular in the distinction of RCC subtypes. Nevertheless, the metabolic pathways that lead to VOCs and VCCs production are not yet clearly understood, hindering an accurate interpretation of the results. Ultimately, validation studies and application of the method to clinical samples (particularly urine) from RCC patients are required, not only to increase knowledge about the metabolism of VOCs and VCCs, but also to evaluate and confirm the clinical translatability of these potential biomarkers.

Keywords: Renal Cell Carcinoma (RCC); Metabolomics; Volatile Organic Compounds (VOCs); Volatile Carbonyl Compounds (VCCs); Biomarkers; HS-SPME/GC-MS

Resumo

A metabolómica é definida como o estudo sistemático de todos os metabolitos (o metaboloma) de um sistema biológico. Esta ciência '-ómica' envolve a análise quantitativa e qualitativa de uma grande diversidade de metabolitos de baixo peso molecular presentes em várias matrizes como células, tecidos ou biofluidos. A metabolómica tem sido aplicada em várias áreas clínicas (por exemplo, diagnóstico, prognóstico e terapêutica) devido à ideia chave de que uma determinada doença ou condição patológica podem apresentar um perfil metabólico exclusivo.

Os metabolitos são produtos finais das vias bioquímicas e a alteração dos seus níveis normais pode afetar o normal funcionamento do organismo. A sua identificação e medição podem ser realizadas em várias matrizes cuja recolha é menos invasiva do que as biópsias, pelo que ao longo dos anos a Metabolómica mostrou ser particularmente valiosa na pesquisa de biomarcadores para o diagnóstico do cancro. A aplicação de uma abordagem metabolómica na investigação do cancro prende-se ao fato de que as células cancerígenas são metabolicamente reprogramadas de forma a controlar as necessidades energéticas exigidas pelo rápido crescimento e desenvolvimento do tumor, produzindo uma "assinatura metabolómica" específica dessas alterações.

O Carcinoma de Células Renais (CCR) constitui aproximadamente 90-95% de todas as neoplasias do rim, sendo considerado um grupo de neoplasias malignas extremamente fatal, cuja incidência continua a aumentar. Além disso, as técnicas de diagnóstico atuais são pouco sensíveis e específicas, e o acesso ao subtipo de RCC requer o uso de abordagens extremamente invasivas, pelo que é claramente necessário a identificação de biomarcadores apropriados para desenvolver técnicas diagnósticas mais rápidas e sofisticadas, bem como menos invasivas.

O projeto desenvolvido nesta dissertação teve como objetivo avaliar o potencial da assinatura metabólica, particularmente dos Compostos Orgânicos Voláteis (COVs) e Compostos Carbonílicos Voláteis (CCVs) no exometaboloma, para distinguir células cancerígenas de células normais, e discriminar dois subtipos histológicos (*clear cell* CCR e *papillary* CCR). Para alcançar esses objetivos, foi aplicada uma abordagem metabolómica com base na Microextração em Fase Sólida e Cromatografia Gasosa acoplada de Espectrometria de Massa (HS-SPME/GC-MS) de forma a analisar o perfil metabólico de cinco linhas celulares tumorais, nomeadamente três *clear cell* (769-P, 786-O e Caki-1) e duas *papillary* (Caki-2 e

ACHN), e uma linha celular não tumorigénica (HK-2). Com o objetivo de classificar a doença de acordo com o grau e subtipo histológico, este estudo compreendeu a análise de dois subtipos de CCR, o *clear cell* e o papilar, ambos no estadio metastático e não metastático. Após análise estatística, observou-se que os COVs e CCVs recolhidos a partir do exometaboloma permitiram uma boa separação entre as linhas celulares tumorigénicas e a linha celular normal, bem como a discriminação entre os dois subtipos examinados, e entre metastático papilar *vs* não metastático papilar, mas apenas no caso dos CCVs.

Os resultados obtidos são promissores e reafirmam a importância do estudo de potenciais biomarcadores COV e CCV no diagnóstico do cancro renal, em particular na distinção dos subtipos de CCR. No entanto, as vias metabólicas que levam à produção de COVs e CCVs ainda não estão completamente compreendidas, dificultando a interpretação dos resultados. Em suma, são necessários estudos de validação e aplicação do método a amostras clínicas (particularmente urina) de pacientes com CCR, quer para aumentar o conhecimento sobre o metabolismo dos COVs e CCVs, quer para avaliar e confirmar a translabilidade clínica desses potenciais biomarcadores.

Palavras-chave: Carcinoma de Células Renais (CCR); Metabolómica; Compostos Orgânicos Voláteis (COVs); Compostos Carbonílicos Voláteis (CCVs); Biomarcadores; HS-SPME/GC-MS

General index

Acknowledgments i

Abstract vi

Resumo viii

General index xi

Index of figures xiii

Index of tables xvii

List of abbreviations xviii

1 Introduction to Kidney Cancer 1

1.1 Epidemiology 1

1.2 Types of neoplasms 5

1.2.1 Benign neoplasms 6

1.2.2 Renal Cell Carcinoma 6

i Symptoms 7

ii RCC epidemiology 8

iii RCC classification 9

iv RCC and genetics 11

v Diagnosis of RCC 13

vi Staging of RCC 14

vii Prognosis of RCC 16

viii Follow-up of RCC 17

ix Treatment options for RCC 17

2 Metabolomics in Cancer Research 21

2.1 Definition and metabolomics workflow 21

2.2 *In vitro* models for cancer metabolomic studies 24

2.3 Metabolome and volatile compounds 26

2.3.1 Volatile organic compounds 26
2.3.2 Volatile carbonyl compounds 28
2.4 Analytical techniques for metabolomic studies 29
2.5 Quality control 29
2.6 Bioinformatics and statistical analysis 30
2.7 Metabolomics in early diagnosis of RCC 31
3 Aims of the experimental work 33
4 Experimental work – material and methods 35
4.1 Chemicals 35
4.2 Cell lines and culture conditions 35
4.3 Sample collection 37
4.4 Sample preparation 39
4.5 HS-SPME/GC-MS analysis: equipment and conditions 39
4.6 GC-MS data pre-processing 40
4.7 Statistical analysis 42
5 Results 44
5.1 VOCs signature profile in RCC cell lines vs normal kidney cell line 44
5.2 VCCs signature profile in RCC cell lines vs normal kidney cell line 54
6 Discussion 66
6.1 Discriminative VOCs in RCC vs normal cell lines 66
6.2 Discriminative VCCs in RCC vs normal cell lines 68
7 Conclusion and future perspectives 74
8 References 76
9 Appendix 106

Index of figures

Figure 1. Estimated age-standardized cancer incidence and mortality rates in 2018, worldwide, both genders, all ages. Data source: Globocan 2018. Available from: http://gco.iarc.fr, accessed on 24/05/2019. 2

Figure 2. Estimated age-standardized cancer incidence and mortality rates of KC distribution by continents in 2018, both genders, all ages. Data source: Globocan 2018. Adapted from: http://gco.iarc.fr, accessed on 24/05/2019. 2

Figure 3. (a) Estimated number of new cases in 2018, Europe, Portugal, all cancers, both genders, all ages. **(b)** Estimated number of deaths in 2018, Europe, Portugal, all cancers, both genders, all ages. Data source: Globocan 2018. Adapted from: http://gco.iarc.fr, accessed on 24/05/2019. 3

Figure 4. Percent of cases by stage at diagnosis: kidney and RPC, worlwide, 2009-2015, all races, both genders. Data Source: SEER Summary Stage 2000. Adapted from: https://seer.cancer.gov/statfacts/html/kidrp.html, accessed on 24/05/2019. 4

Figure 5. 5-year relative survival by stage at diagnosis: kidney and RPC, worlwide, 2009-2015, all races, both genders. Data Source: SEER Summary Stage 2000. Adapted from: https://seer.cancer.gov/statfacts/html/kidrp.html, accessed on 24/05/2019. 4

Figure 6. Percentage of new cases by age group: kidney and RPC, worlwide, 2012-2016, all races, both genders. Data Source: SEER Summary Stage 2000. Adapted from: https://seer.cancer.gov/statfacts/html/kidrp.html, accessed on 24/05/2019. 5

Figure 7. Kidney and nephron anatomy. Adapted from: https://picswe.net/pics/nephron-illustration-43.html, accessed on 22/06/2019. 7

Figure 8. The biological organization of the '-omic' sciences. 21

Figure 9. Procedure of the cell culture. 38

Figure 10. Procedure of the sample preparation. 39

Figure 11. Representation of a typical HS-SPME performed in the present work. **(a)** Culture medium + internal standard + NaCl; **(b)** Incubation and HS generation; **(c)** Sample extraction from the HS; **(d)** Sample injection through GC injector. 40

Figure 12. PCA-X score scatter plot obtained for the HS-SPME/GC-MS chromatograms of all samples and QCs, namely the extracellular medium of HK-2 (n=10, ●), 769-P (n=10, ●), 786-O (n=10, ●), Caki-1 (n=10, ●), ACHN (n=10, ●), Caki-2 (n=10, ●) and QCs (n=10, ●). The ellipses indicate the 95% confidence limit of the model. 44

Figure 13. (a) PCA-X and **(b)** PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of all samples, namely the extracellular medium of the normal cell line (HK-2, n=10, ●) and the cancer cell lines (n=50, ●). The ellipses indicate the 95% confidence limit of the model. 45

Figure 14. (a), (c), (e), (g) and **(i)** PCA-X and **(b), (d), (f), (h)** and **(j)** PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of the normal cell line (HK-2, n=10, ●) and the cancer cell lines 769-P (n=10, ●), 786-O (n=10, ●), Caki-1 (n=10, ●), ACHN (n=10, ●) and Caki-2 (n=10, ●), respectively. The ellipses indicate the 95% confidence limit of the model. 46

Figure 15. Boxplots of six VOCs found significantly altered between HK-2 (●) and 769-P (●) cell lines, namely 2-ethyl-1-hexanol, 2,4-dimethyl-1-heptene, 2,6-di-tert-butylquinone, acetophenone, dodecane and 1,3-dimethyl-benzene. 49

Figure 16. Boxplots of three VOCs found significantly altered between HK-2 (●) and 786-O (●) cell lines, namely ethylbenzene, dodecane and 1,3-dimethyl-benzene. 50

Figure 17. Boxplots of five VOCs found significantly altered between HK-2 (●) and Caki-1 (●) cell lines, namely 2,4-dimethyl-1-heptene, 2,6-di-tert-butylquinone, dodecane, ethylbenzene and 1,3-dimethyl-benzene. 51

Figure 18. Boxplots of six VOCs found significantly altered between HK-2 (●) and Caki-2 (●) cell lines, namely 1,3-di-tert-butylbenzene, 2-ethyl-1-hexanol, 2,4-dimethyl-1-heptene, 2,6-di-tert-butylquinone, acetophenone and cyclohexanol. 52

Figure 19. Boxplots of four VOCs found significantly altered between HK-2 (●) and ACHN (●) cell lines, namely 2-ethyl-1-hexanol, decane, dodecane and 1,3-dimethyl-benzene. 53

Figure 20. PCA-X score scatter plot obtained for the HS-SPME/GC-MS chromatograms of all samples, QCs and MM namely the extracellular medium of HK-2 (n=10, ●), 769-P (n=10, ●), 786-O (n=10, ●), Caki-1 (n=10, ●), ACHN (n=10, ●), Caki-2 (n=8, ●), QCs (n=10, ●) and ME (n=5, ●). The ellipses indicate the 95% confidence limit of the model..................................54

Figure 21. (a) PCA-X and **(b)** PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of all samples, namely the extracellular medium of the normal cell line (HK-2, n=10, ●) and the cancer cell lines (n=48, ●). The ellipses indicate the 95% confidence limit of the model..55

Figure 22. (a), (c), (e), (g) and **(i)** PCA-X and **(b), (d), (f), (h)** and **(j)** PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of the normal cell line (HK-2, n=10, ●) and the cancer cell lines 769-P (n=10, ●), 786-O (n=10, ●), Caki-1 (n=10, ●), ACHN (n=10, ●) and Caki-2 (n=8, ●), respectively. The ellipses indicate the 95% confidence limit of the model. 56

Figure 23. Boxplots of four VCCs found significantly altered between HK-2 (●) and 769-P (●), namely acetaldehyde, acetone, cyclohexanone and methylglyoxal. ME - metabolomic medium incubated under the same conditions as the samples (●)...58

Figure 24. Boxplots of four VCCs found significantly altered between HK-2 (●) and 786-O (●) cell lines, namely acetaldehyde, cyclohexanone, glyoxal and methylglyoxal. ME (●). 59

Figure 25. Boxplots of two VCCs found significantly altered between HK-2 (●) and Caki-1 (●) cell lines, namely acetaldehyde and cyclohexanone. ME (●)...60

Figure 26. Boxplots of four VCCs found significantly altered between HK-2 (●) and Caki-2 (●) cell lines, namely acetaldehyde, cyclohexanone, glyoxal and methylglyoxal. ME (●). 61

Figure 27. Boxplots of two VCCs found significantly altered between HK-2 (●) and ACHN (●) cell lines, namely acetaldehyde and methylglyoxal. ME (●)..62

Figure 28. (a), (c) and **(e)** PCA-X and **(b), (d)** and **(f)** PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of the ccRCC (n=30, ●), pRCC (n=18, ●), non-metastatic ccRCC (769-P and 786-O, n=20, ●), metastatic ccRCC (Caki-1, n=10, ●), non-metastatic pRCC (Caki-2, n=8, ●) and metastatic pRCC (ACHN, n=10, ●), respectively. The ellipses indicate the 95% confidence limit of the model..63

Figure 29. Identified metabolites that are responsible for the separation between metastatic pRCC (ACHN, ●) vs non-metastatic pRCC (Caki-2, ●) cell lines, namely acetaldehyde, cyclohexanone, glyoxal and methylglyoxal. ME (●). 64

Index of tables

Table 1. Types of neoplasms in the kidney. 5

Table 2. Human kidney epithelial neoplasms. While ccRCC is the most common histologic type, there exist rarer forms each associated with known genetic mutations and corresponding syndromes. 13

Table 3. Stages of RCC. This system was developed by the AJCC [105] and adopted by the Union for International Cancer Control (UICC) [106]. 15

Table 4. General characteristics of the normal cell lines. 36

Table 5. Passages used during the experiment for each cell line. *n=1 and n=6 are independent passages. 38

Table 6. Data pre-processing parameters for VOCs and VCCs analysis 41

Table 7. List of VOCs selected as important in discriminating the cancer cell lines (769-P, 786-O, Caki-1, Caki-2, and ACHN) from the normal cell line (HK-2). The *p-value*, ES (± ES_{SE}) and % variation (± uncertainty) are represented for each VOC. - VOC only present in the ccRCC; - VOC only present in the pRCC; - VOC present in RCC subtypes; ↓ - decreased VOC in the cancer cell line; ↑ - increased VOC in the cancer cell line. 47

Table 8. List of VCCs selected as important in discriminating the cancer cell lines (769-P, 786-O, Caki-1, Caki-2, and ACHN) from the normal cell line (HK-2), and in discriminating de metastatic pRCC from the non-metastatic pRCC. The *p-value*, ES (± ES_{SE}) and % variation (± uncertainty) are represented for each VCC. - VCC only present in the ccRCC; - VCC only present in the pRCC; - VCC present in RCC subtypes; ↓ - decreased VCC in the cancer cell line; ↑ - increased VCC in the cancer cell line. 57

List of abbreviations

AGE	Advanced Glycation End Product
AJCC	American Joint Committee on Cancer’s
ATCC	American Type Culture Collection
BHD	Birt Hogg Dubé
ccRCC	Clear Cell Renal Cell Carcinoma
chRCC	Chromophobe Renal Cell Carcinoma
CRC	Colorectal Cancer
CT	Computed Tomography
CYP450	Cytochrome P450
EBRT	External Beam Radiotherapy
ES	Effect Size
ES_{SE}	Effect Size Standard Error
FBS	Fetal Bovine Serum
GC-MS	Gas Chromatography-Mass Spectrometry
GLO1	Glyoxalase 1
GLO2	Glyoxalase 2
GSH	Glutathione
HCl	Hydrochloric Acid
HLRCC	Hereditary Leiomyomatosis and Renal Cell Carcinoma
HMBD	Human Metabolome Database
HPRC	Hereditary Papillary Renal Carcinoma
HS	Headspace
HS-SPME	Headspace-Solid Phase Microextraction
IL-2	Interleukin-2
KC	Kidney Cancer
LMWM	Low Molecular Weight Metabolites
LC	Liquid Chromatography
LC-MS	Liquid Chromatography-Mass Spectrometry
MG	Methylglyoxal
MM	Metabolomic Medium
ME	Incubated Metabolomic Medium

MRI	Magnetic Resonance Imaging
MS	Mass Spectrometry
NaCl	Sodium Chloride
NIST	National Institute of Standards and Technology
NMR	Nuclear Magnetic Resonance
NMRI	Nuclear Magnetic Resonance Imaging
OPLS-DA	Orthogonal Partial Least Squares-Discriminant Analysis
pRCC	Papillary Renal Cell Carcinoma
PCA	Principal Component Analysis
PFBHA	O-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine hydrochloride
PLS-DA	Partial Least Squares-Discriminant Analysis
PTFE	Polytetrafluoroethylene
Q^2	Predictive capability of the model
QC	Quality Control
R^2	Fraction of the original data explained by the model
R^2X	Variance explained by X matrix
R^2Y	Variance explained by Y matrix
RCC	Renal Cell Carcinoma
RECIST	Response Evaluation Criteria in Solid Tumors
ROS	Reactive Oxygen Species
RPC	Renal Pelvis Cancer
RT	Radiation Therapy
SEER	Surveillance, Epidemiology, and End Results
SRM	Small Renal Mass
SPME	Solid Phase Microextraction
TA	Total Area
UICC	Union for International Cancer Control
VIP	Variable Importance in Projection
VCCs	Volatile Carbonyl Compounds
VEGF	Vascular Endothelial Growth Factor
VHL	von Hippel-Lindau
VOCs	Volatile Organic Compounds

1 Introduction to Kidney Cancer

Kidney cancer (KC) is a heterogeneous malignancy frequently found during the initial nephrologic workup incidentally. A great part of the tumors is diagnosed by a clinician that is familiarized with paraneoplastic symptoms and physical cancerous warning signs [1, 2], while the rest is diagnosed in patients with diseases unrelated to cancer, such as during the workup of acute renal failure [3, 4].

Comparatively to other tumors, KC is considered a rare malignancy, however, metastatic disease is associated with decreased rate survival and one-third of the patients require conventional medical therapies[5]. Another treatment option is surgical resection, generally used with curative intent, once it is an effective treatment in the localized illness. Current research efforts have unveiled many key molecular pathways involved in the oncogenic transformation, contributing to the development of novel therapies and consequently amplifying patients' survival rate. The presence of a diverse set of systemic signs and symptoms help clinicians to determine the better treatment options and also to preview illness prognosis, in order to improve the patients' outcome [5].

Almost all renal pelvis cancers (RPC) are transitional cell carcinomas, whereas a vast majority of KCs originated in the renal parenchyma are adenocarcinomas/renal cell carcinomas (RCC) [6].

1.1 Epidemiology

According to Globocan 2018 age-standardized estimates, KC is 16th cancer with the most incident and mortal rates worldwide, for both genders and all ages (Figure 1). This pathology is also more incidental in North America and more deadly in Europe (Figure 2).

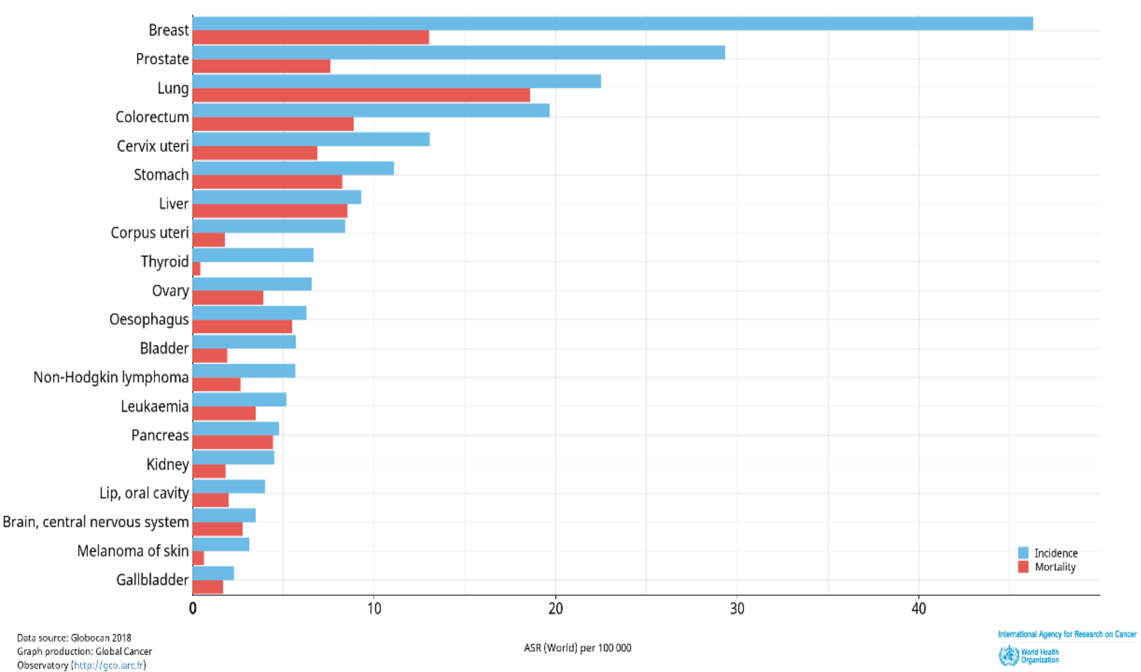


Figure 1. Estimated age-standardized cancer incidence and mortality rates in 2018, worldwide, both genders, all ages. Data source: Globocan 2018. Available from: http://gco.iarc.fr, accessed on 24/05/2019.

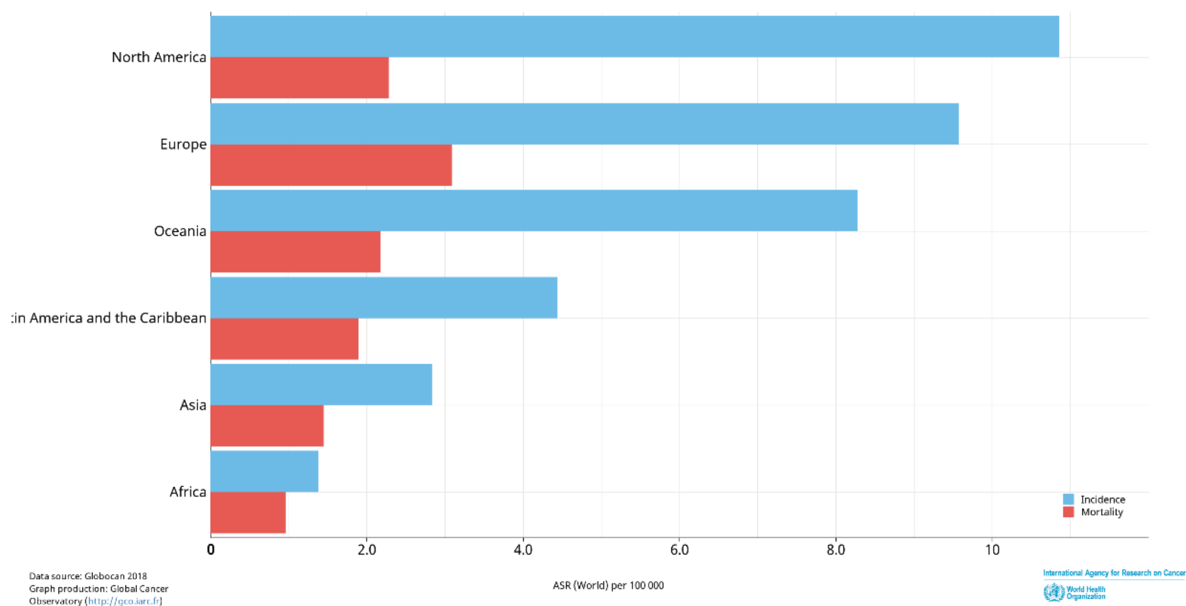


Figure 2. Estimated age-standardized cancer incidence and mortality rates of kidney cancer distribution by continents in 2018, both genders, all ages. Data source: Globocan 2018. Adapted from: http://gco.iarc.fr, accessed on 24/05/2019.

Approximately 3.2% of new cancer cases in Portugal in 2018 were caused by renal cancer (Figure 3a) and in a total of 1 972 438 deaths from cancer, 55 216 (2.8%) were due to the same neoplasm (Figure 3b).

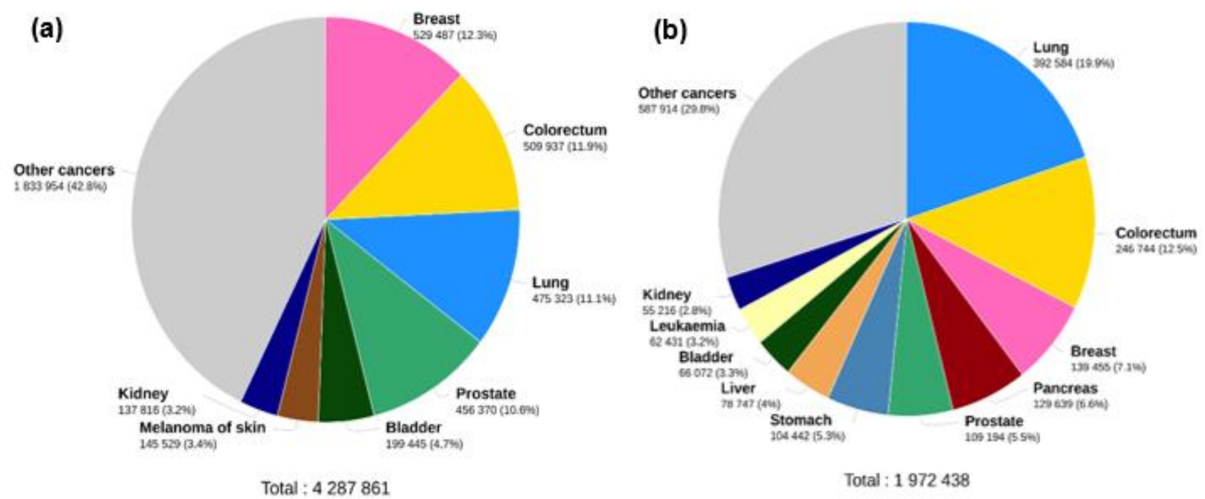


Figure 3. (a) Estimated number of new cases in 2018, Europe, Portugal, all cancers, both genders, all ages. **(b)** Estimated number of deaths in 2018, Europe, Portugal, all cancers, both genders, all ages. Data source: Globocan 2018. Adapted from: http://gco.iarc.fr, accessed on 24/05/2019.

Survivorship includes crucial domains as diagnosis details, treatment toxicities/complications, risk of relapse, lifestyle alterations and identification of psychosocial necessities and resources [7].

Surveillance, Epidemiology and End Results Program (SEER) 2018 statistics presents a percentage of 74.8 for 5 years' relative survival (comparison between the survival of patients diagnosed with cancer and the survival of all population who are the same age, gender, race and that have not been diagnosed with kidney and RPC). According to the same entity, kidney and RPC patients have a better chance of surviving 5 years when it is diagnosed in early stages. A major part of patients (65.1%) are diagnosed at the local stage (where cancer started) (Figure 4), in which the 5-year survival for localized kidney and RPC is 92.5% (Figure 5). However, when cancer has spread to a different part of the body it can be classified as a regional or distant disease. Thus, the extent of disease is a determining factor in the choice of the best therapy for the patient and consequently has a major impact on survival length. The

regional disease has a 5-year survival of 69.6%, while the distant one accounts for 12.0%. Kidney and RPC affects particularly older people (65-74 years) (Figure 6).

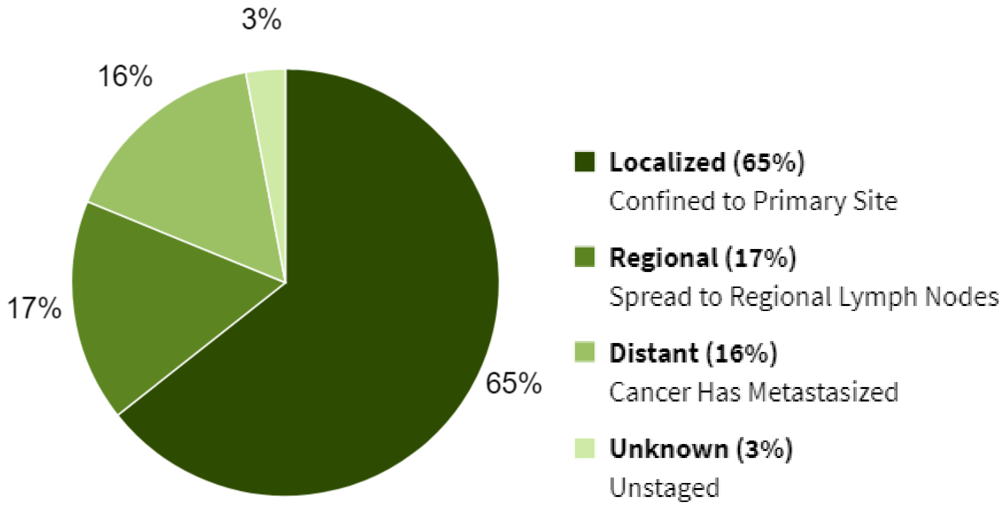


Figure 4. Percent of cases by stage at diagnosis: kidney and RPC, worlwide, 2009-2015, all races, both genders. Data Source: SEER Summary Stage 2000. Adapted from: https://seer.cancer.gov/statfacts/html/kidrp.html, accessed on 24/05/2019.

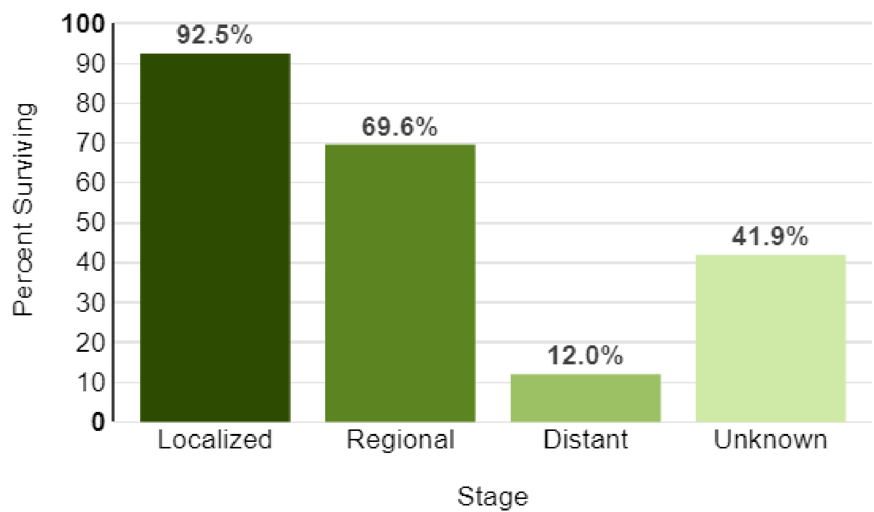


Figure 5. 5-year relative survival by stage at diagnosis: kidney and RPC, worlwide, 2009-2015, all races, both genders. Data Source: SEER Summary Stage 2000. Adapted from: https://seer.cancer.gov/statfacts/html/kidrp.html, accessed on 24/05/2019.

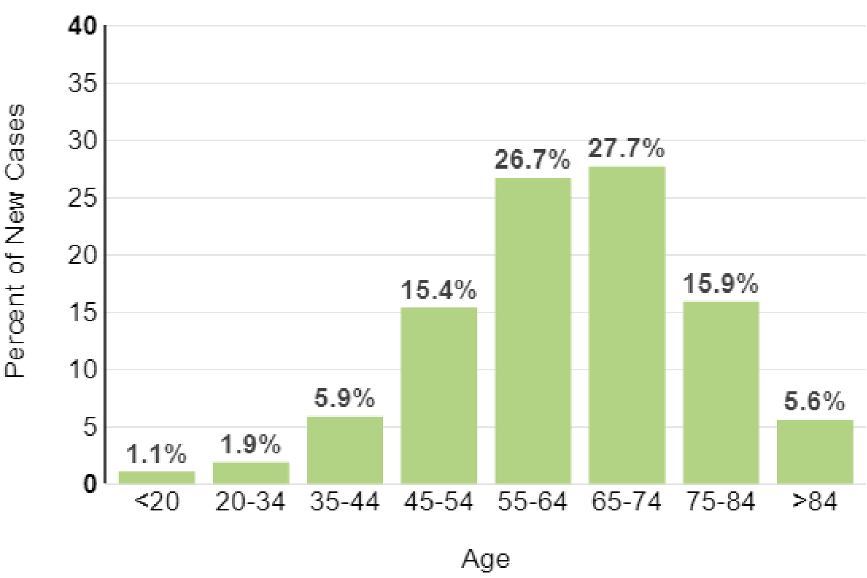


Figure 6. Percentage of new cases by age group: kidney and RPC, worlwide, 2012-2016, all races, both genders. Data Source: SEER Summary Stage 2000. Adapted from: https://seer.cancer.gov/statfacts/html/kidrp.html, accessed on 24/05/2019.

1.2 Types of neoplasms

The benign neoplasms include papillary adenoma, renal oncocytoma, and metanephric adenoma and adenofibroma, while the malignant tumors are subdivided in 5 main groups: clear cell RCC (ccRCC), papillary RCC (pRCC), chromophobe RCC (chRCC), rare types of RCC and unclassified RCC [5, 8] (Table 1).

Table 1. Types of neoplasms in the kidney.

Benign neoplasms	Renal cell carcinomas
Papillary adenoma	Clear cell renal cell carcinoma
Renal oncocytoma	Papillary renal cell carcinoma
Metanephric adenoma and adenofibroma	Chromophobe renal cell carcinoma
	Rare types of renal cell carcinoma
	Unclassified renal cell carcinoma

1.2.1 Benign neoplasms

Papillary adenoma is the term that defines any microscopic appearance of low-grade pRCC with a lesion less than 5 mm [8]. This is the major benign tumor in the renal cortex and includes genetic abnormalities as loss of chromosome Y and a combination of genetic alterations including trisomy of chromosomes 7 and 17, analogous to but less extensive than the modifications in pRCC [8, 9]. These tumors are frequently constituted by small 'blue' cells or large eosinophilic cells with solid-tubular-papillary structures [9].

The renal oncocytoma is a typically large tumor with a central stellate scar and mahogany-brown appearance [8] that comprises about 5% of renal cortical neoplasms [8-10]. Microscopically, oncocytoma cells are organized in cysts, islands, solid sheets, tubules, and abundantly composed by eosinophilic cytoplasm filled with mitochondria [8, 11]. The genetic abnormalities comprise translocation between chromosome arm 11q13 and other chromosomes or loss of chromosome arms 1p, 14q and Y chromosome [12-14].

The metanephric adenoma and adenofibroma are rare benign cortical tumors constituted by small uniform tubules with cuboidal epithelium [8, 9].

1.2.2 Renal Cell Carcinoma

The RCC, entitled 'the internist's tumor', is frequently found during nephrology screening exams in patients with unrelated disorders [4, 5]. The development of human RCCs is presumed to occur along the extent of the nephron, whereby ccRCC and pRCC are thought to emerge from the epithelial cells of the proximal tubule, and chRCC, oncocytoma, and collecting duct RCC from the distal nephron [15] (Figure 7). Each type has a different genetic, biology and behavior.

Some studies determined that 15-48% of RCC cases would be detected incidentally [16], and that at the time of diagnosis 25-30% of the subjects would develop metastatic disease, including them into a dismal prognostic group [5]. The presence of symptoms is a common feature of tumor progression, particularly in high grade [4]. Contrary to symptomatic tumors, the incidental RCC has a lower aggressiveness, metastasis rates and recurrence, and better prognosis and survival [4, 16].

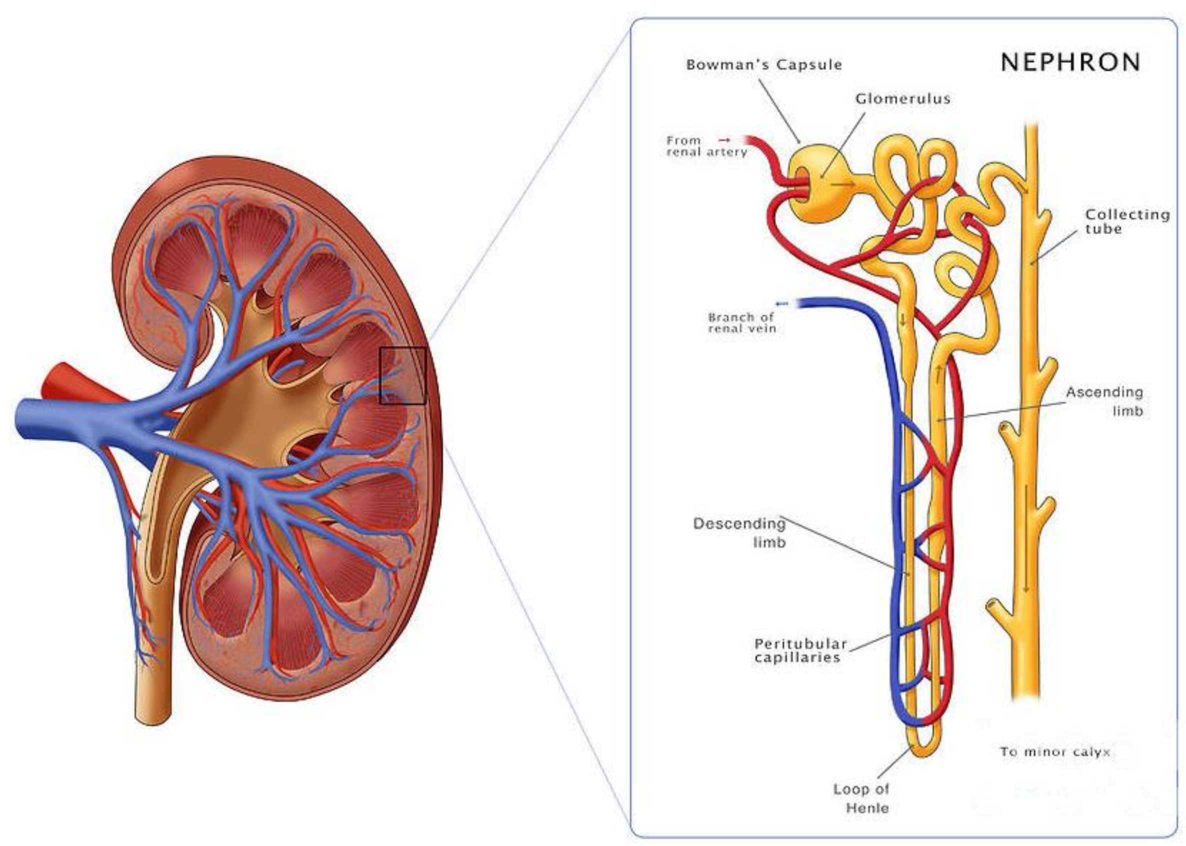


Figure 7. Kidney and nephron anatomy. Adapted from: https://picswe.net/pics/nephron-illustration-43.html, accessed on 22/06/2019.

i Symptoms

Over the past two decades, there has been a considerable increase in the incidence of RCC, partly due to more unruly use of imaging techniques [6]. This event accounts for 48-66% of RCC diagnosis [17], 13-17% of benign tumors [18, 19] and introduced the concept "small renal mass" (SRM) [20]. The SRM is usually diagnosed in the absence of symptoms or other problems [21, 22] and is defined as a renal mass (benign or malign) with a diameter of less than 4 cm [20, 23].

Before the introduction of noninvasive abdominal imaging studies, asymptomatic tumors were detected later or not diagnosed and 67-74% of RCC remained undiscovered until death [17]. However, with the widespread use of these techniques the majority of renal tumors have been incidentally diagnosed with smaller dimensions and at an earlier stage [4, 24, 25].

Some possible signs and symptoms of KC include: hematuria (the main symptom in symptomatic RCC diagnosis [26]), anemia, back pain, a mass on the side or lower back,

fatigue, fever (not attributed by an infection and that does not go away), loss of appetite and weight loss [20].

ii RCC epidemiology

RCC incidence has raised 126% in the United States since 1950 [21], with a similar increasing pattern worldwide, somehow due to the overemployment of techniques for renal tumor detection, such as ultrasonography, computed tomography scan (CT scan), and magnetic resonance imaging (MRI)/nuclear resonance imaging (NMRI) [6, 21, 27]. The increase in incidence is observed at all clinical stages [27], with particular contribution from low-stage tumors [6]. Therefore, it would be expected a decrease in distant stage, however, that is not observed since there must be other unidentified factors that are associated with RCC development [27].

The identified risk factors that may contribute to RCC development are [28]:

- Age - renal tumors diagnosed in younger age are commonly characterized by lower tumor stages, grades and more favorable histologic patterns compared with tumors in older patients [29].
- Smoking - raises the risk of developing RCC and appears to be associated to how much cigarettes are smoked [30, 31]. Despite the fact that smoking cessation decreases the risk of RCC, it takes several years to reach the level of an individual who never smoked [30-32].
- Alcohol - heavy drinking of alcoholic beverages such as wine, liquor, and beer is associated with RCC [33], however, some studies demonstrated that moderate alcohol consumption is associated with reduced RCC risk [34, 35].
- Obesity - due to changes in certain hormones, being overweight can increase the chance of developing this neoplasm [36-39], while an inverse relation of physical activity to the risk of renal cancer is observed [40]. Translational studies demonstrated that RCC could be promoted by extended consumption of a high-fat diet [41].
- Hypertension - a positive association between KC and hypertension was demonstrated [37, 42]; several studies suggested that some medications for high blood pressure treatment (e.g. diuretics) may increase the risk of KC [43-45], however, it is extremely difficult to be sure if it is the disease, the medicine, or both that may be promoting a higher risk for cancer developing.

- Advanced kidney disease - individuals affected by advanced kidney disease, in particular patients requiring dialysis, are included in the highest risk group of developing RCC [37, 46, 47].
- Environmental factors/workplace exposure - some studies suggested that workplace environment may increase the risk for RCC, due to the exposure to certain substances as cadmium [48], some herbicides, and organic solvents, particularly trichloroethylene [49, 50].
- Family history - people with any inherited conditions and a strong family history of RCC, particularly those who have a brother/sister with this disease, have a higher risk to develop it [51]. It remains unclear if this occurrence is due to common genes and/or something that both individuals were exposed to.
- Genetic abnormalities - some people inherit a tendency to develop certain types of cancer [5, 10, 15, 52].
- Gender - even knowing that RCC is about twice as common in men as in women [53], gender is considered an unreliable risk factor. This doubt is due to the fact that women are less likely to be smokers and/or susceptible to the exposure to carcinogenic chemicals at work than men [54]. These factors may account for some of the observed differences.
- Race - another non-established risk factor is ethnicity, even though the incidence and mortality being greater among blacks than whites [6], the reasons for this racial disparity remains unclear [55, 56].

Individuals affected by the criteria registered above constitute only a small portion of all cases, however, they have a considerably higher risk to develop KC [15]. Nevertheless, some recent studies show evidence of possible RCC prevention through rich diets in fruit and vegetables [57, 58] as well as elevated levels of vitamin D [59, 60].

iii RCC classification

RCC is not a single disease but, rather, a heterogeneous group of tumors owing to their distinct clinical and histological features [61, 62]. Different RCC subtypes induce distinct outcomes, even when the patients are stratified by tumor stage and grade [62]. Thus, it becomes extremely useful the RCC classification into subtypes, in order to easily predict the most accurate and precise diagnosis [10] and prognosis [61], and also to counseling patients

into the best-individualized treatment plan, once therapeutic strategies may differ from patient to patient [63, 64].

The most valuable parameter for RCC classification is the degree of differentiation, however, the development pattern, the absence or presence of calcification (more frequent in pRCC and chRCC) and the tumor-spreading patterns can give additional information to identify the RCC subtype [61].

The ccRCC is the most frequent malignancy in adult kidney [65], and also the most prevalent tumor in the renal cortex in clinical practice, comprising approximately 70-75% of all renal cancers [5, 8, 10, 62]. Most ccRCC contains a clear cytoplasm, while others present eosinophilic or granular appearance. The most common genetic abnormality is the loss of genetic material in chromosome 3p [8]. Half the malignancies express somatic mutations in the von Hippel-Lindau (VHL) gene, however, 20% of these cases have this gene inactivated by hypermethylation [8].

The second most prevalent histologic pattern is the pRCC that comprises 15% of the disease cases and is classified into type 1 and type 2 [8, 10, 62]. This malign tumor is affected by Y chromosome loss and also by trisomies in several chromosomes (3q, 7, 12, 16, 17, and 20) [8]. The designation papillary carcinoma is applied in the presence of these genetic alterations, even if *papillae* are not very visible [8].

The chRCC subtype encompasses about 5% of cases of RCC [8, 10, 62]. This cancer has an evident blue appearance when stained with Hale's colloidal iron, due to numerous microvesicles contained in the cytoplasm [8, 66]. Chromophobe cases are affected by hypodiploidy and monosomy of chromosomes 1, 2, 6, 10, 13, 17, and 21 [8].

The rare subtype can be differentiated in collecting duct carcinoma (Bellini duct carcinoma), RCC associated with chromosome's short arm translocation, medullary RCC, RCC associated with neuroblastoma, multilocular cystic RCC and mucinous tubular and spindle RCC [67]. The collecting duct carcinoma is the most frequent in the rare subtype and comprises less than 1% of surgical cases, the reason why it is considered a rare high-grade RCC [8, 10]. This malignancy contains dysplastic changes in collecting ducts and affinity for *Ulex europaeus* lectin, however, its low incidence hinders its genetic study, as well as its assertive recognition [8, 9].

The unclassified subtype encloses less than 2% of RCC cases, often presenting high-grade with genetic and morphological variability [8]. All cases that do not fit into the categories of neoplasm described above are included in this group [8, 9].

iv RCC and genetics

Recent advances and improvements in the genetic field in knowledge and techniques underlying the pathogenesis of renal cell neoplasms have led to the categorization of distinct types of tumors [9, 21]. It is currently known that genetic mutations transferred during cell division are intrinsically tangled to the neoplastic transformation [9]. These genetic alterations are essential determinants for tumors behavior, biology (e.g. cell adhesion, differentiation, death and proliferation) and morphology [9, 21].

The base of this malignancy derives from a hereditary/inherited form or from a sporadic/non-inherited form. Patients who have hereditary form develop the multifocal and bilateral disease at a young age [65], while the sporadic form tends to appear as a single, unilateral tumor at a later age [68].

For instance, sporadic cancer requires two independent genetic events, while in hereditary tumors patients inherit a mutation that predisposes them to cancer development, however, an additional mutation is necessary to trigger tumorigenesis (two-hit theory) [68, 69].

The inherited tumors are related to already identified genetic mutations, syndromes and other diseases [5, 10], whereby encompass different histological types of RCC (e.g. VHL is a hereditary form of ccRCC, Hereditary Papillary Renal Carcinoma (HPRC) is a hereditary form of pRCC type 1, Birt Hogg Dubé (BHD) is a hereditary form of chRCC and oncocytoma, and Hereditary Leiomyomatosis RCC (HLRCC) is a hereditary form of pRCC type 2) (Table 2) [10].

The same tumor suppressor gene can be involved in the origin of both inherited and non-inherited variants of a particular malignancy [69]. VHL disease is an inherited cancer syndrome that affects numerous organs such as the adrenal gland, cerebellum, epididymis, eyes, kidney, pancreas and spinal cord [65, 70-74]. The basis of this syndrome is a tumor suppressor gene [75-77] located in the chromosome 3 [76], in which the inactivation of one or of both alleles by partial deletion or mutations that shorten the VHL protein promote tumorigenesis [65, 78, 79].

This gene encodes the protein pVHL, a key element at oncogenic mechanisms involved at tissue response to hypoxia, and also the hyperangiogenic phenotype of KC [5], two cancer hallmarks [80].

The two-hit theory of Knudson [69] foretells that in familial cancer syndromes such as VHL, the genotype of each neoplasm should involve an allele with an inherited germline mutation and loss of function of the inherited wild-type allele through chromosomal deletion or point mutation [78].

VHL gene abnormalities are exclusively found at ccRCC tumors, in which nearly 35-45% of the individuals with VHL disease develop multiple bilateral tumors and cysts [65]. This gene is

mutated in the germline of individuals with the hereditary form as well as in sporadic nonhereditary ccRCC [79, 81-83]. Inactivation of the VHL gene results in HIF1-α and HIF2-β overexpression [84, 85], activating the expression of genes that are implicated in angiogenesis, hypoxia, and other signaling pathways that require the vascular endothelial growth factor (VEGF) [21]. Thus, increased transcription of growth factors levels as VEGF [86, 87], GLUT-1, PDGF, TGF-α and erythropoietin, explain the huge vascularization and glucose transport in this RCC subtype [5, 10].

Like ccRCC, the papillary subtype is a metastatic tumor and occurs in both, inherited and non-inherited forms [10]. It also has similar genetic and histological characteristics to collecting duct carcinoma [10]. HPRC is an inherited form of pRCC with an autosomal dominant transmission with reduced penetrance, in which the patients are predisposed to develop bilateral and multifocal tumors [10, 88-90]. The genetic abnormality responsible for this disease is localized in chromosome 7 at the c-Met gene [89]. Mutations in this oncogene are unusual in sporadic pRCC and not detected in the solid tumors panel [91], however, c-Met is compromised in the germline of HPRC affected individuals by missense mutations in its tyrosine kinase domain [5, 10, 52]. The association between HPRC and type 1 pRCC is characteristic of trisomy 7 with a nonrandom duplication of the mutated c-Met allele [92].

The hereditary cancer syndrome BHD seems to be related to loss of function of the tumor suppressor gene BHD [10, 93, 94]. The affected individuals are at risk for the development of bilateral and multifocal renal tumors (chromophobe/oncocytic hybrid (50%), chRCC (33%), oncocytoma (7%), or ccRCC (5%)) [95].

The patients that are affected by tumors associated with BHD, HPRC and VHL, mutations are frequently bilateral and multifocal and with a relatively stable growth, whereby these individuals are submitted to the surgical procedure that better minimizes the chance of metastasis and that preserves the kidneys for as long as possible [10]. Generally, this approach requires strategical management in multiple kidney operations before the tumor reaches 3 cm of size [10].

Contrarily to the VHL, HPRC or BHD based malignancies, patients with HLRCC are often solitary tumors associated with the type 2 pRCC [10] associated with a genetic abnormality in the fumarate hydratase (FH) gene [5]. This malignancy can be extremely aggressive and lethal presenting early metastization, whereby it is recommended early surgical intervention [10].

Table 2. Human kidney epithelial neoplasms. While ccRCC is the most common histologic type, there exist rarer forms, each associated with known genetic mutations and corresponding syndromes.

	RCC				Benign neoplasm
Type	ccRCC	pRCC		chRCC	Oncocytoma
		Type 1	Type 2		
Gene mutation	VHL	Met	FH	BHD	
Disease	VHL	HPRC	HLRCC	BHD	

Despite the knowledge improvement inside the genetic (e.g. genetic alterations) and biochemical (e.g. key signaling pathways) fields, the incidence and mortality of KC have aggravated during the last decades [5]. Thus, the development of new drugs and other therapies approaches based on RCC subtypes are particularly essential [10] in order to increase the overall survival and to improve patient's quality of life. However, it is indispensable to biopsy the tumor to obtain a feasible diagnosis and subsequently to choose the most appropriate treatment before it begins [5].

v Diagnosis of RCC

Early diagnosis is crucial to effective treatment and patient survival [5], however, it is necessary to understand the genetic basis of these tumors, in order to provide the development of better methods of diagnosis [10].

For instance, the implementation of screening programs is a promising methodology to prevent histological and clinical aggressiveness, while detecting and treating favorable prognosis tumors before any symptomatology [4].

In the absence of other symptoms, nephrology clinicians overemploy sophisticated imagiological techniques (such as kidney ultrasonography) to evaluate renal insufficiency, proteinuria, or hematuria, even at slighter cases of renal insufficiency, whereby classic Virchow triad of gross hematuria, flank mass and pain are unusually observed at the first diagnosis [5].

After RCC diagnosis, tests are performed to detect metastasis to other organs. Metastization is one of the main causes of death in affected patients and occurring from the primary tumor to other parts of the body, and through the lymph system and/or the blood. The process used to analyze the spread within the kidney and to other organs is denominated staging [96].

The detection and diagnosis of RCC is mainly based on abdominal and renal examination using: physical exam of the body in order to check overall health and to find uncommon variations; history of the patients' health, lifestyle, past illnesses and treatments; urinalysis and blood chemistry studies; CT scan; ultrasound exam; MRI and/or biopsy [96].

The evolution of imaging procedures lead to its widespread use in cancer field [25, 97], however, during the last years, these imaging techniques have been used unconsciously, contributing to an increase in the diagnosis of completely asymptomatic patients [25, 97].

vi Staging of RCC

The staging system for RCC is based on the extent of the cancer beyond the kidney [98-100]. The spread in the blood vessels may not indicate poor prognostic signs if the tumor is confined to the principal organ, the kidney.

The RCC pathological stage is the main prognostic indicator since it is based on tumor size and its clinical behavior and also on the extent of the invasion [3, 15, 101, 102]. The local extension of the primary tumor (T), the involvement of regional lymph nodes (N), and the presence of distant metastases (M) define the TNM classification [15]. It is essential to identify the disease stage to apply the most effective treatment and to predict the disease outcome. The staging of RCC can be made through a CT scan, MRI/NMRI, chest x-ray, and bone scan.

According to previous studies, lower stages are more common in incidental than in symptomatic RCC [103, 104].

RCC staging is defined by the American Joint Committee on Cancer's (AJCC) TNM classification system (Table 3) [105].

Table 3. Stages of RCC. This system was developed by the AJCC [105] and adopted by the Union for International Cancer Control (UICC) [106].

Stage	TNM classification	
I	○ T1, N0, M0	"The tumor is 7 centimeters or smaller and is found in the kidney only."
II	○ T2, N0, M0	"The tumor is larger than 7 centimeters and is found in the kidney only."
III	○ T1, N1, M0 ○ T2, N1, M0 ○ T3, N0, M0 ○ T3, N1, M0	"The cancer in the kidney is any size and cancer has spread to a) nearby lymph nodes, b) the blood vessels in or near the kidney (renal vein or vena cava), c) the structures in the kidney that collect urine, or d) the layer of fatty tissue around the kidney."
IV	○ T4, any N, M0 ○ Any T, any N, M1	"Cancer has spread a) beyond the layer of fatty tissue around the kidney and may have spread into the adrenal gland above the kidney with cancer, or b) to other parts of the body, such as the brain, lung, liver, adrenal gland, bone, or distant lymph nodes."

Tumor grade, as well as other tumor markers, may be promising approaches for future RCC evaluation [4], whereby the former is considered the second most important variable with independent prognostic value [3, 102]. Tumor grading is based on cancer differentiation and a variety of morphological parameters that can be obtained through light microscopy. In fact, the grade is an alternative marker for a determined number of cellular processes, with ploidy and cell proliferative activity being paramount [107].

Fuhrman nuclear grade is based on the microscopic morphology of a neoplasm with hematoxylin and eosin staining and its use is more successful in ccRCC cases than in other histologic types [11]. Oncocytomas are not graded and the chromophobe disease is usually low grade [15]. Low- and high-grade classification is one of the criteria to separate pRCC tumors into types 1 and 2 [15]. Fuhrman grade may help delineate categories of survival time better than clinical judgment alone.

vii Prognosis of RCC

Prognostic prediction is based on factors such as the histological cell type, the stage of the tumor, grade within the stage and the clinical information [15], however, clinical guidelines support a stage-specific stratification of patient prognosis [106, 108]. The stage is considered the best predictor of survival since it specifies cancers' size and kidney spread. In this regard, the identification of molecular alterations that are related to neoplasm behavior as well as to the clinical outcome within organ-confined, locally advanced, or metastatic disease, will contribute to better management of RCC [15].

Better prognosis has been assigned to incidental diagnosis, small size, early stage [20] and localized disease [10]. Even with the involvement of the regional lymphatic or blood vessels, a considerable number of patients can achieve prolonged survival and cure [109]. A great part of cases is diagnosed when the tumor is localized and when it is relatively advantageous surgical removal. However, as expected, patients with tumors in advanced stages (e.g. presence of metastases) have a poor disease-free survival [10]. Nevertheless, a few patients survive after complete surgical resection of the tumor, while, occasionally, others have late tumor relapse, even several years after initial treatment. Curiously, a few well-documented cases of RCC report spontaneous tumor regression in the absence of therapy, however, these events are not very common and may not lead to long-term survival [110-112].

RCC prognosis also varies widely according to the histological subtype [113]. ccRCC is associated with a rich vascular network and alveolar architecture promoting a strong progression and aggressive behavior [62]. This subtype is also strongly associated with hematogenous metastases which affect particularly the lungs, liver, and bones [114]. Lymph node metastases are found in up to 15% of cases [67].

ccRCC has a relatively unfavorable prognosis compared to other subtypes, however, the collecting duct carcinoma and the unclassified category have the worst prognosis [67, 115]. The collecting duct carcinoma, a high aggressive lesion [116, 117], presents a quite unfavorable prognosis since at diagnosis about 30% of the cases already have metastases and 60–70% of them die within two years [67]. The unclassified subtype is a histologically heterogeneous group of tumors, commonly rated as high-grade lesions [67] and consequently with poor survival [11].

pRCCs are usually solid and well defined lesions with slow growth [90]. Papillary and chRCC are associated with a more indolent behavior and better prognosis when compared to other subtypes [20, 62, 115, 116, 118-120]. The former presents two different biological behaviors of the lesion, with a quite distinctive prognosis, in which type 2 pRCC is associated

with a worse prognosis due to be detected at higher grades and later stages than type 1 pRCC [67]. Thus, type 2 pRCC frequently appears associated with ganglial metastasis and occasionally with venous invasion [67]. In the case of the chRCC subtype, only 7% of the tumors are metastatic [67].

Almost all stage IV RCC cases are not curable and the treated patients with progressing, recurring, or relapsing disease are often associated with poor prognosis. Prognostic markers as carbonic anhydrase IX have been the main targets of immunohistochemically studies in the context of advanced RCC [15].

viii Follow-up of RCC

A better understanding of tumor biology may promote an accurate individualization of patient prognosis and thus a more personalized follow-up treatment [108]. So far, the early treatment of metastasis is not associated with an improved outcome when compared with the delayed one, and any specific follow-up protocol has impact in early or advanced RCC prognosis [106]. CT scans of thorax and abdomen are routinely carried out in order to offer a long-term follow-up since there is a possibility of late relapse, yet the time interval depends on risk factors [121]. The principal method used to evaluate drug efficacy is the Response Evaluation Criteria in Solid Tumors (RECIST), however, there is no clinical evidence that RECIST used to quantify disease progression is a clinical valid endpoint in the decision to interrupt or change treatment [106].

The clinical guidelines of major international societies are based either on tumor stage or integrated prognostic tools, however, there are significant differences in their recommended follow-up protocols for RCC due to lack of accurate information on patient surveillance after primary treatment [108].

ix Treatment options for RCC

The standard treatments used in individuals affected by RCC are surgery (the main treatment of RCC), chemotherapy, radiation therapy (RT), immunotherapy, and targeted therapy. However, new types of treatment are being tested in clinical trials and the patients can enter clinical trials before, during, or after starting their cancer treatment [96].

The first treatment for a disease is also called first-line, induction therapy or primary therapy or primary treatment, and is chosen as being the most appropriate for the patient (e.g. radical nephrectomy, cytoreductive nephrectomy [122-125], Cabozantinib for patients with

intermediate or poor-risk disease [126, 127], Pazopanib [128-130], Interferon-alpha [131, 132], Temsirolimus [133], Bevacizumab with or without Interferon-alpha [134-137], Sunitinib [128, 129, 138, 139], Interleukin-2 (IL-2) [131, 140, 141], Ipilimumab plus Nivolumab [142], palliative external beam radiotherapy (palliative EBRT)). When the initial treatment doesn't work or stops working, it is chosen a second-line therapy (e.g. Everolimus (for patients who have previously been treated with Sunitinib and/or Sorafenib) [143], Axitinib [144], Cabozantinib (for patients who have previously been treated with Sunitinib, Pazopanib, Sorafenib, or Axitinib) [145], Sorafenib [146], Nivolumab (for patients who have previously been treated with Sunitinib, Pazopanib, Sorafenib, and/or Axitinib) [147], palliative EBRT), and if the chosen drug is not efficient in eliminating the tumor or causes severe side effects, another medicine may be administered concomitantly or as a substitute.

Surgery to remove part or all of the kidney is frequently used to treat RCC. The types of surgery listed below may be used: radical nephrectomy (removal of the kidney, adrenal gland, surrounding tissue, and usually nearby lymph nodes); partial nephrectomy (removal of tumor within the kidney and some of the surrounding tissue; used primarily to prevent loss of kidney function in patients with the other kidney damaged or removed), simple nephrectomy (kidney removal only). After surgery, some patients receive adjuvant treatment, such as RT and chemotherapy, in order to exterminate the remaining cancer cells and thereby decrease the risk of recurrence [96].

RT is used to eliminate cancer cells or to stop their proliferation. This treatment can be done through external RT (machine outside the body radiates de tumor) or internal RT (radioactive substance sealed in seeds, catheters, wires or needles, placed directly into or near the tumor), depending on the stage and type of cancer. External RT is also used as palliative therapy to relieve symptoms and improve quality of life [96].

Chemotherapy is used to kill cancer cells or to stop them from dividing through drug administration (e.g. vinblastine, floxuridine, 5-fluorouracil, capecitabine, and gemcitabine). The way the chemotherapy is given, systemic chemotherapy or regional chemotherapy, depends on the stage and type of tumor. In the former, the drug enters the bloodstream reaching cancer cells throughout the body, while the regional administration is directly applied to the cerebrospinal fluid, an organ, or a body cavity affecting mainly the cancer cells [96].

Immunotherapy, also called biologic therapy or biotherapy, is used to boost or restore patients' natural defenses against cancer through their immune system. This treatment can be administered using immune checkpoint inhibitor therapy, which blocks checkpoint proteins and increases the ability to eliminate cancer cells (e.g. Ipilimumab plus Nivolumab [142] and

Nivolumab [147, 148]); Interferon, which slow tumor growth by affecting the proliferation of cancer cells [131, 132, 149, 150]; and IL-2, which boosts the growth and activity of several immune cells [131, 132, 140, 141, 149-153]. The former is particularly used to treat advanced cancers that can't be treated with surgery and is divided into two types: CTLA-4 inhibitor and PD-1 inhibitor [96].

The targeted therapy is used to recognize and attack specific cancer cells without damaging the normal ones. To this purpose, antiangiogenic agents that will stop tumor growth by blocking the formation of blood vessels are administered. There are two types of antiangiogenic treatment, kinase inhibitors, and monoclonal antibody therapy. Kinase inhibitors are divided in VEGF inhibitors (e.g. Sunitinib [135, 136, 138, 139], Pazopanib [128-130], Cabozantinib [126, 127, 145, 154, 155], Axitinib [144, 156], Sorafenib [146]) and mammalian target of rapamycin (mTOR) inhibitors (e.g. Everolimus [143] and Temsirolimus [133]) [96].

All of these treatments may cause side effects, therefore this parameter should also be taken into account when choosing the best treatment for each patient. After treatment has ended, it is necessary to perform follow-up tests from time to time, since cancer may relapse years after initial treatment, in the kidney or in other parts of the body [96].

The available and suitable treatments for patients with cancer in stage I are: simple nephrectomy, partial nephrectomy, or radical nephrectomy [157, 158]; RT and arterial embolization as palliative therapy [157, 159]; and clinical trial of a new therapy [96].

For stage II treatments are considered partial nephrectomy or radical nephrectomy [157]; nephrectomy before or after RT; RT and arterial embolization [157] as palliative therapy; and clinical trial of a new therapy [96].

For stage III treatments are considered radical nephrectomy [160]; arterial embolization and subsequently radical nephrectomy [157, 159]; RT before or after radical nephrectomy [157]; RT, arterial embolization [157, 159] and nephrectomy as palliative therapy; and clinical trial of immunotherapy following surgery [96].

For stage IV treatments are considered nephrectomy to reduce the size of the tumor [122, 124, 125]; radical nephrectomy; targeted therapy; immunotherapy; and RT as palliative therapy [96].

2 Metabolomics in Cancer Research

2.1 Definition and metabolomics workflow

Metabolomics is the study of the complete metabolome of the cell, organ or organism. One of the best descriptions for this '-omic' science had been presented by Steve Oliver (Manchester University), who used the term 'metabolomics' to describe "the complete set of metabolites/low-molecular-weight intermediates, which are context-dependent, varying according to the physiology, developmental or pathological state of the cell, tissue, organ or organism" [161].

Of the various '-omics' sciences (Figure 8), metabolomics is a relatively new approach that aims to quantify all small molecules (molecular weight less than 1500 Daltons) existent in any particular physiological state in any biological system [162]. Through the measurement and interpretation of the data obtained from the analysis of metabolites in a biofluid (e.g. plasma, urine) or tissue, it is provided a unique signature that can be useful to compare subtle to huge metabolic differences between normal and diseased states [163].

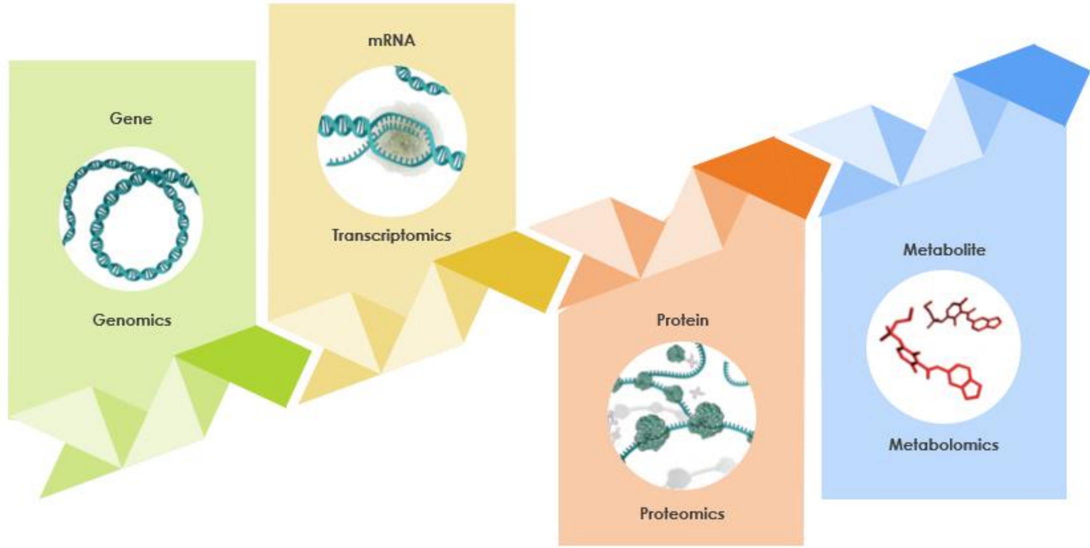


Figure 8. The biological organization of the '-omic' sciences.

In short, the metabolic signature improves our understanding of toxicological and non-toxicological effects at the end of a biological cascade driving some conclusions about the establishment and development of the disease itself [162].

The metabolome is defined as all metabolites within a cell or secreted by a cell [164, 165], and is directly associated with the genome, once a considerable amount of genes may be related with a unique metabolite biosynthesis or degradation [161].

Information flows from DNA to RNA to protein to function, suggesting that gene expression regulates enzyme concentration through mRNA sequestration, transcription, translation or degradation (hierarchical regulation). Nevertheless, there isn't a proportional correlation between mRNA levels and the concentrations of the corresponding biochemical enzymes (function), dominant regulation. Furthermore, transcriptome and proteome may be an incomplete monitor to follow cell function regulation, whereby the metabolome (whole complement of metabolites in a cell localized downstream of both transcription and translation) [166] may be an important source of information to understand dysregulations at the enzymatic level [167].

Metabolome analysis is a more clinically translatable method to identify disturbed pathways in specific pathologies when compared with transcriptomics and proteomics [168]. These '-omic' tools complement each other, however, measuring metabolite concentrations is more useful than following the rates of chemical reactions directly [168].

Alterations in metabolic pathways (i.e. set by the complement and properties of the cellular enzymes) have a major impact on metabolite concentration, and flux control is related to the absolute amount of the present enzyme, however, its activity and concentration variation have a small impact on metabolic fluxes [169, 170]. Environmental and genetic alterations induce an abnormal response in biologic systems, thus, regarding the changes at metabolites levels, it is possible to predict which regulatory pathway is altered in cell/biological systems [171].

In cancer field, several metabolites have been shown to be powerful tools to distinguish healthy tissue from tumors, providing important information about tumorigenesis in monitoring cellular activities such as cell-cycle progression or apoptosis [168]. The 'metabolomic' methodologies encompass the combination of multivariate statistics and an analytical technique such as gas chromatography-mass spectrometry (GC-MS) or liquid chromatography-mass spectrometry (LC-MS), and nuclear magnetic resonance (NMR) [168]. The metabolic activity can also be quantified measuring concentrations of metabolites that represent a wide range of compounds, namely low molecular weight metabolites (LMWM) [168]

The identification and quantification of all the metabolites present in the sample is not a precise and accurate single analytical process [162]. Thus, metabolomic techniques can be divided into two main groups accordingly to the stage the metabolite identification is performed during data processing, targeted and non-targeted methodologies [162].

The non-targeted approach comprises a global analysis of the metabolome, analyzing all the measurable analytes, including unknown signals. Targeted metabolomics analyzes only specific metabolite signals previously identified and is, therefore, the most useful approach to determine precisely and accurately relative concentrations of a limited number of endogenous metabolites [162].

Compared to transcriptomic and proteomic, analytical techniques are particularly more robust and mature for metabolite detection and quantification [172]. Four different strategies can be applied to answer general and more specific questions:

- Metabolite targeted analysis is principally used for screening goals and for highly sensitive analyses [171, 172]. This approach may require a considerable sample clean-up in order to avoid interferences [171].
- Metabolic profiling focus in identifying and quantifying a pre-selected subclass of metabolites (such as polar lipids, isoprenoids, or carbohydrates) or specific pathway elements, in a biological sample [171], creating a metabolism 'bar code'. This technique claims a sample preparation and data acquisition centered on the chemical properties of the metabolites of interest, in order to reduce matrix effects [171, 172];
- Metabolic fingerprinting or metabonomics is a suitable strategy for fast sample sorting according to the origin/biological relevance, maintaining a high throughput. This approach is particularly advantageous to predict the tumors' behavior to specific therapies, however, it must be considered that important metabolic effects might be covered during data acquisition due to irreproducible matrix effects, instigating insufficient analytical resolution and low sensibility in this technique [171, 172];
- Metabolite untargeted analysis is the ideal approach to quantify metabolite concentrations, secretions, and fluxes in cells and tissues in which there is a direct correlation between the metabolic activity and gene expression and proteome [164, 173]. The comprehensive and quantitative analysis of all metabolome explored in this methodology has great applicability, since it can improve the understanding of such systems [171, 172], reveal possible novel

therapeutic targets and, subsequently, evaluate their efficacy, safety and toxicity [168, 173]. Metabolomic data sets are complex to evaluate, thus, it is necessary a well-designed procedure for sample preparation, storage, normalization and data acquisition [171]. These experiments must also include tools that ensure high sensitivity, selectivity, applicability and matrix independence, considering a strategy with utmost importance in identifying unknown metabolites, and possibly compare with theoretical models [171].

In addition, metabolomics uses an analytical tool combined with pattern-recognition methods to increase the understanding of pathogenic mechanisms, to monitor tumor growth and regression, and consequently to improve treatment procedures [168].

Metabolomic approaches produce highly reproducible data sets, however, only a few metabolites are measured, because the amount of compounds present in a mammalian tissue is possibly far lesser than the quantity of transcripts present in the mammalian transcriptome [168]. Nevertheless, this technique has a considerable disadvantage, a particular metabolite can reveal genomic changes but still remains hard to accurately determine the altered metabolic pathway [168].

This specific 'omic' is therefore a promising approach to identify potential biomarkers and also to develop anticancer therapies, mainly apoptosis inducers [163, 168]. Simultaneous detection of all metabolites require a previous experience design establishing which technique is most suitable for data acquisition and also the adequate procedure for sample preparation [171].

Ultimately, metabolomics research may be more clinically translatable than other '-omic' sciences [174], however, the chemical and structural identification of the metabolites require an additional time-consuming approach [175].

2.2 *In vitro* models for cancer metabolomic studies

Each research platform has its own benefits and limitations in terms of study design and expected outcome and the choice of the best model for cancer research must take these aspects into account [176].

In vitro models are crucial platforms for the study of cellular pathways, epigenetic, genetic [177], cancer progression, proliferation deregulation and apoptosis [178], in order to define

potential molecular markers [179] and to the screening, development, and improvement of cancer therapies [180].

Immortalized cancer cell lines have been extensively used for '-omics' research and proved to be an excellent tool for the study of tumor biological mechanisms [177]. This model allowed the detection of deregulated genes and signaling pathways in this disease [178, 181] and was also in the origin of the development and testing of new therapies [177] and anticancer drugs presently used [182].

In vitro metabolomic studies present more controllable experiment designs and easier interpretation of the results due to the absence of some confounding factors such as age, gender, race, diet and others that affect the metabolic profile [183]. Moreover, this experimental model is relatively inexpensive and is easily maintained in culture and stored when compared to animal models, since cancer cell lines have a high proliferation rate that allows an increase in the number of experiments that can be performed [178, 179, 184]. Cell lines results are highly reproducible and easily comparable with the extensive literature that is currently available [179]. Contrary to tissue samples from patients where there are many cell types present, the cells are genetically similar and very straightforward to sequence [178, 179, 184]. However, working with cell lines can bring some disadvantages such as the fact that sometimes it is not possible to know the cancer subtype or tumor stage and grade from which the cell line was established. The major drawback of using this model is the easy contamination that is generated by microorganisms or other cells [179, 184]. This problem is potentiated by the transfer of aliquots interlaboratorially and erroneous characterization of the cell lines. This model is also susceptible to genetic and phenotypic mutations, as the cells tend to adapt to the surrounding artificial environment and may undergo changes that make them very different from their parental origin [178, 179, 184]. Cell lines are usually established from metastatic lesions and then adapted to 2-dimensional monolayer culture [185], whereby the samples are associated with inherent variability. Last but not least, the results obtained in cancer research using cell lines are commonly extrapolated to *in vivo* human tumors [179, 186], yet cell-cell and cell-matrix interactions seen *in vivo* are lost *in vitro* assays, making these cells less clinically relevant and less useful in translational research.

By comparing cancer cell lines with a normal cell line, differences in metabolome composition can be identified, and this data can be utilized to the diagnosis and therapeutic studies. The cancer cells are distinguished from non-malignant based on their extremely active metabolomic profile and also on their characteristic microenvironment [163]. It is also important

to mention that the analysis of the extracellular metabolites does not require cell disruption, and may allow the monitorization of the metabolic changes over time [183].

2.3 Metabolome and volatile compounds

The cell metabolome can be divided into two portions, the endometabolome, which comprises all metabolites present inside the cell, and the exometabolome, which represents all metabolites in the extracellular culture medium that surrounds the cell [183, 187].

The normal or abnormal state of metabolic activity can be revealed through the analysis of the exometabolome and/or endometabolome profiles, which are directly dependent on the exchange of metabolites (consumed and released) between cells and the extracellular medium [183, 188]. Thus, due to a close relationship between intra- and extracellular metabolism, the metabolic footprint allows the understanding of the mechanism of action of different stimuli in a non-invasive way, providing information about the intracellular metabolic status [188, 189].

Volatile metabolites constitute only a small portion of the metabolome, whereby their accurate identification is essential for a conclusive understanding of living systems. They are the end products of numerous biochemical pathways and also perfect translators deleterious changes in the metabolism of a given organism [190]. The study of volatile compounds has been successfully applied in the plant biology area [191-193], the food industry [194, 195] and also in the cancer field [196-198] as well as other pathologies [199, 200].

2.3.1 Volatile organic compounds

Biochemical processes such as oxidative stress and lipid peroxidation generate LMWM known as volatile organic compounds (VOCs) [201]. VOCs are carbon-based molecules that can be produced and released in exhaled breath or in body fluids such as urine, blood, sweat or feces [202-205]. These compounds can be collected from the headspace (HS) of those matrices and cells' extracellular medium [206], however, it is important to take into account that they present high vapor pressure at room temperature being easily volatilized. VOCs are classified according to their molecular weight (alkanes, alkenes, alcohols, aldehydes, ketones, dialdehydes and furfurals) and boiling point (50-260 ºC) [187].

Curiously, different VOCs patterns have been correlated with distinct diseases, particularly cancer, in which it was revealed that dogs are able to diagnose bladder [207], colorectal [208], melanoma [209] and lung and breast cancers [210] by sensing patients. These studies support

the idea that different diseases have a characteristic “smellprint” [196], whereby certain compounds may be potential diagnosis biomarkers [202, 205, 206, 208, 211-215].

The interest in studying the cancer volatilome relies on the fact that VOCs are end products of cellular metabolism [196] and are released through cell membranes. Evidently, cancer cells will produce a different pattern of VOCs due to their genetic mutations and consequent adaptations and modifications in biochemical pathways. Cancer tissues are also associated with hypoxic and anoxic environments [216]. These conditions may induce differences in the oxidative status and consequently lead to different VOCs profile.

The analysis of VOCs requires a simple and fast sample preparation when compared to other compounds (e.g. amino acids and fatty acids that need a derivatization step prior to analysis), preventing loss of metabolites. The study of these molecules has other advantages as a non-invasive collection and the analysis of numerous compounds across different samples [217].

VOC-biomarker studies in the exometabolome are more reliable than the studies done in the endometabolome, due to the fact that the endometabolomic techniques require additional procedures that may lead to substantial loss of VOCs (e.g. cell membranes disruption) [183].

The collection of VOCs in *in vitro* assays can be performed in two main matrices, namely the headspace (HS) of cell-free culture medium or HS of medium containing cells [187]. These two approaches present some differences in the extraction procedures, such as the addition of salts to improve the efficiency of the analysis in the cell-free culture medium.

Over the years, some improvements have been made in the extraction and collection procedures of these molecules. Solid-phase microextraction (SPME) is one of the best options among the possible techniques that can be employed since it enables minimal sample treatment and, consequently, minor alteration of sample metabolites [218, 219]. This technique comprises the use of a fiber coated with an extracting phase, which extracts volatile and non-volatile compounds. SPME is also a simple, economical, solvent-free and low time-consuming method during VOC extraction. A similar technique is HS-SPME, in which compounds present in the gas space above the sample are analyzed. However, this approach efficiency depends on the chemical nature of the compounds, type of fiber coating, extraction temperature and time, salting-out effect and the type of sample [220, 221]. Currently, HS-SPME is commonly coupled with GC-MS due to the benefits of the union of these two techniques, such as the rapid, efficient, automated extraction that is followed by immediate identification.

The utmost limitation of these investigations lies in the difficulty of comparing the studies, due to several factors that comprise the lack of standardized extraction to discrepancies in cell

density, the period of culture, different collection procedures and sample storage [187]. Another drawback in VOCs analysis is that the metabolites that are altered in cancer cell lines and deliberated potential biomarkers may not be found in *in vivo* or *ex vivo* analysis, hindering clinical translatability. In addition, VOCs collected in *in vitro* assays can be artefacts originated from the culture flasks, extraction and collection conditions [222].

2.3.2 Volatile carbonyl compounds

Volatile carbonyl compounds (VCCs), namely aldehydes and ketones, are intermediates and end products of metabolic pathways. These molecules are sensitive to metabolic disturbs and may be used as biomarkers of these metabolic changes that are found in several pathological conditions, such as cancer. A great part of diseases is caused by an overproduction of free radicals and reactive oxygen/nitrogen species [223-225], leading to oxidative stress and subsequent oxidation of nucleic acids, proteins, and lipids. These toxic processes are related to increased levels of LMWC and DNA strand damage [226-228].

Considered as the major molecular contributor involved in oxidative stress [229], lipid peroxidation is a complex process that produces lipid peroxides and 'oxidative stress second messengers' (aldehydes and ketones) [230]. The peroxidation of polyunsaturated fatty acids generates reactive VCCs [231-233], which consequently during the glycoxidation may produce methylglyoxal [234]. Lipid peroxidation can form more stable products, such as ketones, when induced by oxidative stress [235-237].

Formaldehyde, glyoxal, benzaldehyde, hexanal, and heptanal, are VCCs that have already been suggested as potential cancer biomarkers [238-242]. Therefore, the evaluation of the VCCs levels is essential to understand and endorse the mechanisms behind homeostasis and dysregulation of biological systems [229, 235, 243]. However, the analysis of VCCs can be very challenging due to their low concentrations in samples, the biological matrices complexity, and the non-volatile or semi-volatile profile [239, 243-246]. In order to facilitate volatilization and consequently the analysis of these molecules, a derivatization step is added to the sample preparation procedure. This process is based on a chemical reaction, between the compound of interest and the derivatizing agent (e.g. O-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine hydrochloride (PFBHA)), which increases the volatility and allows the evaluation of those compounds on analytical platforms such as GC-MS.

GC-MS is the main technique to analyze aldehydes, ketones and other endogenous metabolites due to its sensitive detection and high chromatographic resolution [247].

2.4 Analytical techniques for metabolomic studies

The complete analysis of the metabolome cannot be fulfilled accomplished with the use of a single analytical platform since its application would only allow the evaluation of a portion of the metabolome.

The study of metabolomics depends on the use of complex and expensive technology such as GC-MS, LC-MS, or NMR, in order to separate with high-resolution and identify the metabolites within a biofluid or tissue [163].

GC is a separation technique limited to easily volatilized, thermally stable and non-polar analytes, such as some metabolites, in which they are separated according to their volatility [163]. To circumvent this limitation, derivatization procedures are used to increase the volatility of the compounds present in the sample [248].

The GC platform is based on both retention time in the column and fragmentation pattern produced by the MS, making GC-MS an extremely advantageous analytical system. In addition, given the existence of many commercially available libraries for metabolite identification, GC-MS is regularly implemented in approaches of analysis of target metabolites, whereby a single compound or a subset of known metabolites are analyzed [248].

Throughout the years, other methodologies emerged in order to simplify sample preparation. SPME is an example of these techniques that offers many advantages compared to classical ones and is commonly used to study the volatile portion of the metabolome. This approach has the ability to combine extraction and concentration into a single solvent-free step preventing artifacts, is fast, inexpensive and easily automated, requires low samples volume, and has high sensitivity and reproducibility [249].

2.5 Quality control

During metabolomic studies is essential to ensure that the detection of metabolite differences among samples is not hindered by undesired sources of instrumental deviations, such as contamination peaks [183, 250]. Thus, the validation of metabolites and the minimization of false results can be achieved with the inclusion of quality control (QC) samples. These samples together with standardization, guarantee and optimize the consistency and reproducibility of the data obtained during metabolomic analysis, contribute to the analytical precision determination and provide the possibility of performing the signal correction, reducing signal variations[183]. QCs are also useful to evaluate the quality of samples, whether they are degraded or contaminated.

2.6 Bioinformatics and statistical analysis

After the data processing is chosen the ideal method for statistical analysis, depending on the study design and the variables involved [163].

The final data in untargeted metabolomic studies is massive, making it necessary to analyze the data with specialized software. Prior to statistical analysis, the data is processed and normalized to lessen the variations caused by experimental factors inherent to the analytical platforms (e.g., daily alterations in sensitivity) or to sample preparation [251]. The correction of the data includes peak normalization and alignment, baseline correction, noise reduction and deconvolution of peaks [162, 252].

Nowadays, there are plenty of helpful bioinformatics software available (e.g. MZmine, XCMS, etc.) to quickly achieve this purpose [251]. Nevertheless, since cellular density affects the metabolite concentration, there are some more traditional normalization methods used in *in vitro* metabolomic studies. These methods include cell counting, protein or DNA quantification [253].

Subsequently to data pre-processing, it is possible to compare the metabolite profile among samples of highly proliferative cells [183, 254], however, this step outcome is constituted by a great number of challenging variables to interpret [252], whereby statistical analysis takes place.

Multivariate analysis has two different types of approach, the unsupervised and the supervised. The purpose of the unsupervised approach lies in easily identifying the existence of similarities or differences between the samples and the detection of outliers. Therefore, this statistical analysis separates the samples according to the metabolites in their composition, without any information related to the class of the samples [255, 256]. In this sequence is commonly obtained the Principal Component Analysis (PCA-X). Subsequently, is applied the supervised analysis to uncover the discriminative features between different classes of samples [255, 257]. The Partial Least Squares-Discriminant Analysis (PLS-DA) and the orthogonal PLS-DA (OPLS-DA) are the two principal outcomes of this analysis [255]. Once the potential metabolic signatures are identified as discriminative biomarkers, the metabolites must be identified. Metabolomic studies can be very challenging, since most peaks may not be identified during the analysis. To solve this problem, scientists use databases that help identify these peaks (e.g. Human Metabolome Database (HMBD) and National Institute of Standards and Technology (NIST)) [250-252] and standards for the respective analyte to obtain a more

accurate identification [250]. Lastly, it is important to evaluate the functional and biological relevance of the identified metabolites, for instance, through metabolic pathway analysis.

2.7 Metabolomics in early diagnosis of RCC

Metabolomics methodologies have a crucial role in improving the understanding of cancer biology and it is becoming an important component in early diagnosis, which has the potential to significantly improve patient survival [5, 168]. This paradigm reflects KC reality, once the diagnosis prior to metastatic spread has the likelihood to improve survival odds from 10% to 90% [5].

A screening test for the general population to detect RCC at an earlier and curative stage is the main topic for future metabolomics research on KC. In this regard, it has increased the number of projects with a particular interest in discovering biomarkers that may be amenable detected in a non-invasive way using body fluids, differentiating benign tumors from malignant neoplasms and/or predicting growth rate [15]. Urine is an example of a biofluid used for this purpose, given the fact that most biochemical processes present a urinary metabolic "signature" [175]. In fact, these processes are closely related and important in the study of KC, because the most RCC tumors occur within the tubule epithelium and result in aberrant secretion of metabolites into the tubular lumen (urinary space) [163]. Following these pathways, it is possible to exploit the final result, the metabolome, and thus to distinguish the urinary metabolome of healthy individuals from RCC patients. High concentrations of metabolites in the neoplasm compared to normal levels in the control may indicate increased tumor secretion or a systemic response from the body to the tumor itself [163].

The ultimate goal in biomarker identification is to determine the molecular structure for each of the significant compounds of the analysis performed [175]. Ultimately, these potential biomarkers will be identified and validated in large human cohort studies in order to conceive a fast and inexpensive clinical method for early diagnosis and to get access to the responses after specific therapeutic interventions [175].

3 Aims of the experimental work

Concerning RCC, there have been great efforts to improve diagnostic approaches over the years and, in this regard, metabolomic studies have shown promising results. Therefore, this dissertation focused on performing an *in vitro* untargeted metabolomic approach with the hope of contributing to the RCC biomarkers research and to increase the knowledge about RCC metabolism. To this end, the volatile metabolic signatures of five RCC immortalized cell lines, comprising the two most common subtypes of RCC: ccRCC (769-P, 786-O and Caki-1) and pRCC (Caki-2 and ACHN), and one normal human kidney cell line (HK-2) were analyzed. Thus, the expected outcomes of the present study are as follow:

- To obtain a comprehensive metabolomic profiling of volatile metabolites (VOCs and VCCs) in the culture medium of normal and RCC cell lines using HS-SPME/GC-MS based analytical platform;
- Demonstrate the suitability of metabolomics for differentiating the metabolic pathways of normal and RCC cells in such a way that it will be possible to identify specific biomarkers or variation of levels of compounds able to distinguish different subtypes of RCC (ccRCC and pRCC);
- Elucidate metabolites and biochemical pathways that will allow the recognition of biomarkers for RCC.

To the best of our knowledge, there are no studies addressing the exchange of both VOCs and VCCs by human RCC cell lines, making this work pivotal in that matter.

4 Experimental work – material and methods

4.1 Chemicals

RPMI-1640 medium, trypsin-EDTA solution (0.25%), sodium chloride (NaCl, ≥ 99.5%), PFBHA, ≥ 99%, thymol (≥ 98.5%) were obtained from Sigma-Aldrich (St. Louis, Missouri, USA). Heat-inactivated fetal bovine serum (FBS) and antibiotic mixture of penicillin/streptomycin (10.000 U/mL/10.000 mg/mL) were obtained from GIBCO Invitrogen (Barcelona, Spain). Hydrochloric acid (HCl) and sodium hydrogen carbonate were obtained from Merck (Darmstadt, Germany). Methanol (≥ 99.9%) was obtained from VWR (Leuven, Belgium). All other chemicals and reagents used were of analytical grade and were dissolved in deionized water unless otherwise indicated. Acetaldehyde (≥ 99%), acetone (≥ 99%), cyclohexanone (≥ 99%) and methylglyoxal (40% aqueous solution) were purchased from Sigma-Aldrich (Madrid, Spain). Glyoxal (≥ 95%) was purchased from Fluka (Madrid, Spain).

4.2 Cell lines and culture conditions

Five human immortalized RCC cell lines (769-P, 786-O, ACHN, Caki-1, and Caki-2) and one human normal kidney epithelial immortalized cell line (HK-2) were obtained from American Type Culture Collection (ATCC; Manassas, VA, USA). General characteristics of all cell lines are summarized in Table 4.

769-P and 786-O cell lines carry the mutated VHL [258, 259] and have high VEGF production [260], suggesting a ccRCC phenotype [261].

Caki-1 is a metastatic ccRCC cell line model that harbors the wild-type VHL. It was shown that this cell line produces ccRCC tumors in nude mice [262] and that has high VEGF production [263] (especially in hypoxic conditions [264]).

ACHN cell line was established from a pleural effusion derived from metastatic disease. Initial studies described this cell line as poorly differentiated clear cell [265], however, a more recent genomic study suggests the existence of papillary features [266] as c-Met polymorphism that is specific for pRCC [91].

Table 4. General characteristics of normal and RCC cell lines used in this study.

	HK-2	786-O	769-P	Caki-1	Caki-2	ACHN
Organism	*Homo sapiens*, human	*Homo sapiens*, human	*Homo sapiens*, human	*Homo sapiens*, human	*Homo sapiens*, human	*Homo sapiens*, human
Gender	Male	Male	Female	Male	Male	Male
Age	adult	58 years	63 years	49 years	69 years	22 years
Ethnicity	NA	Caucasian	Caucasian	Caucasian	Caucasian	Caucasian
Tissue	Kidney, cortex/proximal tubule	Kidney	Kidney	Kidney; derived from metastatic site: skin	Kidney	Kidney; derived from metastatic site: pleural effusion
Disease	Papilloma	ccRCC	ccRCC	ccRCC	pRCC	pRCC
Morphology	Epithelial	Epithelial	Epithelial	Epithelial	Epithelial	Epithelial
Growth Properties	Adherent	Adherent	Adherent	Adherent	Adherent	Adherent
Tumorigenic	No	Yes	Yes	Yes	Yes	Yes

Caki-2 was established from a primary kidney tumor, and was primarily defined as a cell line that belonged to the ccRCC subtype (expresses wild-type VHL but does not express HIF-2α), however, this cell line has a low expression of HIF-1α for unknown reasons [260], which led to a more accurate evaluation of the tumors formed by Caki-2 in nude mouse. The analysis of the tumors revealed cystic papillary formations with microvilli and microfilaments,

multilamellar bodies, lysosomes or lipid droplets, and few mitochondria [267-269]. Although it has been treated as being of the ccRCC subtype, a growing number of studies with the Caki-2 cell line have suggested that it belongs to the papillary subtype [270]. These cells express high levels of c-Met [271] and harbor chromosome 8 aberrations [272] characteristic of pRCC, however, reports on VHL gene status are inconsistent, some researchers did not detect genetic alterations [260], while others [273, 274] reported some mutations.

The selection of cell lines was performed to allow the study of the most common RCC subtypes (ccRCC and pRCC) at different stages of the disease. All cell lines were cultured in RPMI-1640 medium in order to eliminate possible confounding factors.

The growth medium was prepared as indicated by the manufacturer and supplemented with 10% FBS and 1% of penicillin/streptomycin. All cell lines were incubated at 37 °C and 5% CO_2. To minimize possible effects of the used medium on cellular metabolic profiles, the culture conditions were maintained throughout the entire assay.

4.3 Sample collection

All cell lines were routinely tested for *Mycoplasma spp.* contamination (PCR Mycoplasma Detection Set, TaKaRa) and cultured in plastic T25 culture flasks at least 3 passages before use. After 24h a confluence of 30-40% was observed, whereby the culture medium was discarded and replaced with 6.3 mL of fresh metabolomic medium (MM). This medium had exactly the same composition as the initial culture medium, however, the preparation of all required MM was performed only once, in order to submit all cell lines to the same experimental conditions, thereby reducing potential confounding factors and possible effects on metabolic footprints. Prior to sample collection, the culture flasks were maintained in the incubator for another 48h, reaching 90% confluence (Figure 9). The experiments were carried out during 10 independent passages, resulting in a total number of 60 experiments (Table 5). Concomitantly, 5 flasks were incubated with only MM, in order to compare the consumption and excretion of the discriminant metabolites with the compounds that were present in the medium.

After cells' growth to approximately 90% of confluence, the extracellular medium from each culture flask was collected to falcons on ice and centrifuged for 5 minutes at 2000 g at 4 ºC. The supernatant was divided into two 2.5 mL aliquots for further samples analysis of the metabolic profiles of VOCs and VCCs, and in two 0.5 mL aliquots for the QCs of VOCs and VCCs and stored immediately at -80 ºC until analysis (Figure 10). The same procedure was applied to all experimental assays.

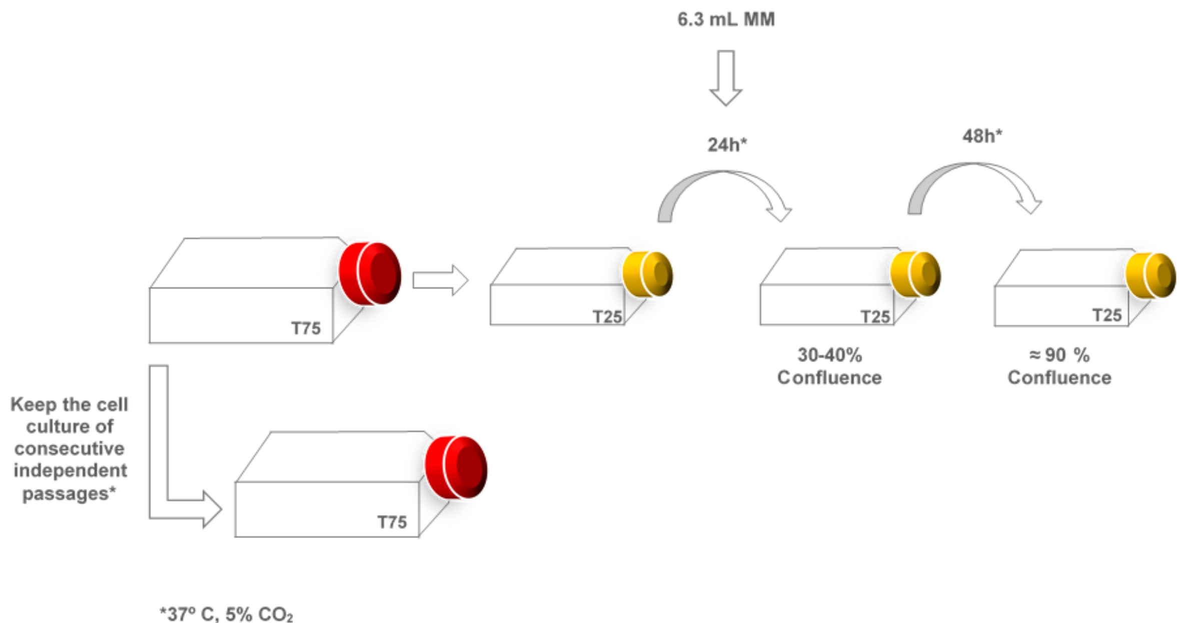


Figure 9. Procedure of the cell culture.

Table 5. Passages used during the experiment for each cell line. *n=1 and n=6 are independent passages.

	HK-2	786-O	769-P	ACHN	Caki-1	Caki-2
n=1*	9	8	7	7	8	9
n=2	10	9	8	8	9	10
n=3	11	10	9	9	10	11
n=4	12	11	10	10	11	12
n=5	13	12	11	11	12	13
n=6*	10	9	8	8	9	10
n=7	11	10	9	9	10	11
n=8	12	11	10	10	11	12
n=9	13	12	11	11	12	13
n=10	14	13	12	12	13	14

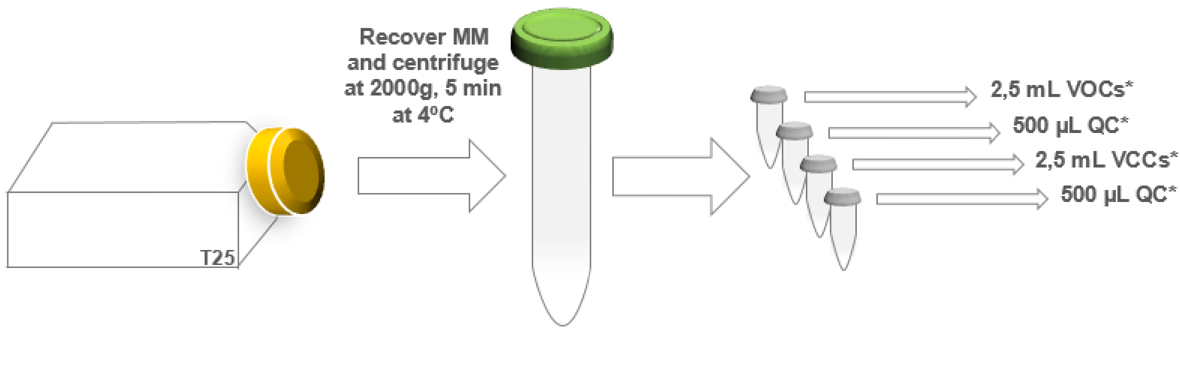


Figure 10. Procedure of the sample preparation.

4.4 Sample preparation

To obtain an extensive range of volatile compounds present in the exometabolome of the cell, two different approaches based on HS-SPME coupled to GC–MS were performed. The HS of the cell culture medium allowed the direct analysis of VOCs, while the VCCs determination was carried out after a derivatization step with PFBHA, in order to increase their volatility and their limit of detection. Prior to sample preparation, the samples were slowly thawed on ice to minimize the loss of volatile compounds and then prepared according to a GC-MS-based protocol previously optimized by our group, with some modifications [275].

Briefly, for VOCs analysis, each sample (2 mL) was placed into a 10 mL glass vial, capped with a polytetrafluoroethylene (PTFE) septum and a screw cap containing 0.43 g of NaCl and 100 µL of thymol (final concentration 5 µg/mL), the corresponding internal standard to monitor GC-MS performance and method reproducibility. For VCCs analysis, each sample (1 mL) was placed into a 10 mL glass vial, capped with a PTFE septum and a screw cap containing 23.5 µL of 40 g/L PFBHA. For each procedure, QCs were prepared as a pool of all samples and divided into aliquots to avoid repeated freezing and thawing [183]. Sample preparation was randomized in both protocols in order to avoid analytical bias.

4.5 HS-SPME/GC-MS analysis: equipment and conditions

The chromatographic analysis of VOCs was performed using CombiPAL autosampler (Varian Pal Autosampler, Switzerland) and the Cycle Composer software (CTC Analytics

System Software, Switzerland), according to HS-SPME procedure (Figure 11) optimized by Araújo et al. (2018) [275]. VOCs detection was accomplished using the EVOQ 436 GC system (Bruker Daltonics, Fremont, CA) coupled to a SCION Single Quadrupole mass detector and the Bruker Daltonics MS workstation software 8.2, while VCCs data were acquired using a SCION Triple Quadrupole mass detector.

For both methodologies, the chromatographic separation was carried by high-purity helium C-60 (Gasin, Portugal) at a constant flow rate of 1.0 mL/min through a capillary column Rxi-5Sil MS (30 m × 0.25 mm × 0.25 μm) from RESTEK. The oven temperature was held for 1 min at 40 °C, followed by an increase of 5 °C/min until reaching 250 °C. After holding the temperature for 5 min at 250° C, it was increased at a rate of 5 °C/min to 300 °C. The injection was made in splitless mode, the injector was held for 20 min at 250 °C, and the run time to obtain total separation was 58 min. The MS detector was operated in electron impact mode (70 eV) at 250º C, the transfer line temperature was 230 °C and the manifold temperature was 40 °C. Data acquisition was performed in full scan mode with a mass range between 40 and 500 *m/z* at a scan rate of 6 scans/s and randomized to avoid analytical bias. A QC sample was repeatedly analyzed under the same conditions, one every six samples, to assess the analytical reproducibility of both methods.

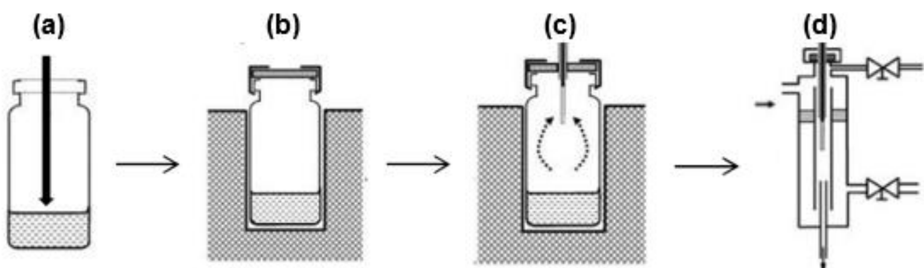


Figure 11. Representation of a typical HS-SPME performed in the present work. **(a)** Culture medium + internal standard + NaCl; **(b)** Incubation and HS generation; **(c)** Sample extraction from the HS; **(d)** Sample injection through GC injector.

4.6 GC-MS data pre-processing

All raw data files obtained from GC-MS were converted into a universal format (net.cdf files). Data were subsequently pre-processed to convert instrumental data sets into a manageable format for data analysis and remove any bias from the background noise and retention time fluctuations. Prior to statistical analysis, all data were pre-processed using

MZmine 2.39 software. This step consisted of baseline correction (used in raw data to remove random noise and baseline shift before peak detection), peak detection (to detect all peaks and concomitantly to avoid possible false positives), and chromatogram deconvolution (to decompose overlapped chromatographic signals) and alignment (to adjust for any minor variation in retention time) [276]. The identification of metabolites was performed by comparison of the MS spectra with standards (whenever available), the NIST database spectral library, and comparison of the experimental and theory (literature) Kovats index.

The parameters to accomplish the VOCs and VCCs data pre-processing are presented in Table 6.

Data were subsequently normalized by total area (TA) by dividing each peak area in the chromatogram by the total peak areas of the respective chromatogram. TA normalization is useful for the correction of small variations derived from sample preparation and analysis [277].

Artifact peaks such as GC contaminants from the chromatographic column (e.g. cyclosiloxanes, siloxanes, and phthalates) detected in comparison with blank samples and chromatographic peaks with a signal-to-noise ratio less than three were also removed from the data matrix. The resulting *m/z*-RT pairs, sample names, and normalized peak areas were imported to SIMCA 15.0.2 software (Umetrics Umea, Sweden) and subjected to multivariate analysis.

Table 6. Data pre-processing parameters for VOCs and VCCs analysis.

Parameter		VOCs	VCCs
Crop filtering	m/z	50 – 250	50 – 500
	Retention time (min)	6 – 30	9 – 40
Peak detection	Noise level	2.0E4	5.0E4
Chromatogram deconvolution	Peak range (min)	0.03 – 0.50	0.01 – 0.20
	Baseline level	1.0E4	1.0E5
Alignment	m/z tolerance	0.1	0.1
	Retention time tolerance (min)	0.05	0.2

4.7 Statistical analysis

Prior to the multivariate statistical analysis, the data set was tested by different methods (Pareto, Unit Variance and centered scaling) for the purpose of minimizing possible differences between samples [278], whereupon Pareto presented the best results. PCA-X, unsupervised analysis, was first applied to detect trends and outliers and to evaluate the distribution of the QC samples, followed by the PLS-DA, supervised analysis, to discriminate classes and to identify the specific metabolic signature of each cancer cell line [279]. The PLS-DA is centered on splitting the systematic variations into two parts, one that expresses the relation with the response and another that expresses the variation not related to the response, making the interpretation of the obtained models cleaner and easier [255]. The output is a score scatter plot, in which each point represents an individual sample, and analogous samples are grouped in clusters. The quality of the models was evaluated using SIMCA 15.0.2 software. The robustness was estimated in terms of R^2 (the fraction of the original data explained by the model) and Q^2 (predictive capability of the model). The former is divided into two parts, R^2X that intends the variance explained by X matrix (e.g. GC-MS data) and R^2Y that represents the variance explained by Y matrix (e.g. sample class). All volatile compounds with VIP (Variable Importance in Projection) value greater than one were considered potential discriminant compounds for the separation between cell lines and were subsequently submitted to univariate analysis [279].

The first step of the univariate analysis was to determine if the data presented a normal distribution using the D'Agostino & Pearson omnibus normality test followed by the unpaired t-test with Welch's correction (for normal distribution) test or the unpaired Mann-Whitney test (for non-normal distribution). The common practice is to interpret a *p*-value lower than 0.05 as an indicator of a statistically significant model [279]. In addition, the percentage of variation, the uncertainty of the variation quotient, as well as the effect size and its uncertainty, were calculated [280]. The univariate analysis was performed in GraphPad 6.0.1 software.

5 Results

5.1 VOCs signature profile in RCC cell lines vs normal kidney cell line

Multivariate analysis was used to evaluate the reproducibility of the analytical method and the discriminant capability of the PCA-X and PLS-DA models.

The analysis of the QCs had the purpose of confirming that analytical variation was minimal and the method was reproducible. The method is considered reproducible as all QC samples are closed projected in PCA-X (Figure 12).

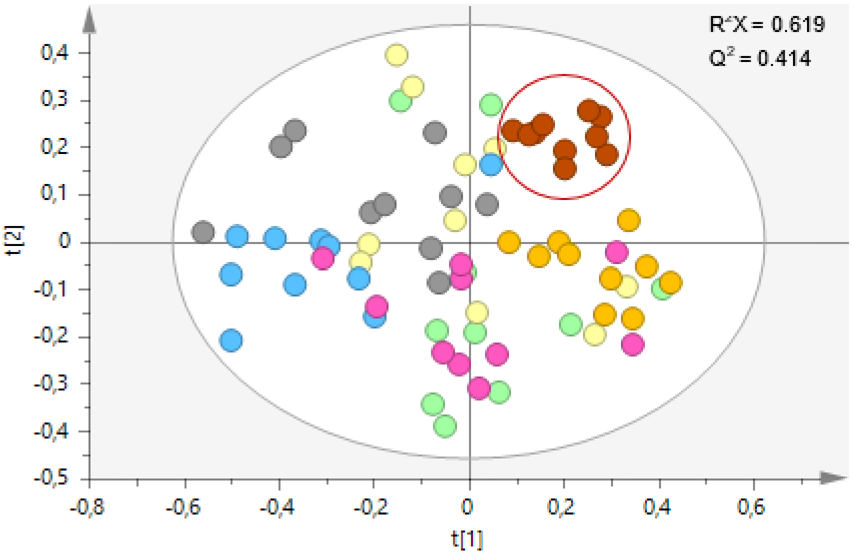


Figure 12. PCA-X score scatter plot obtained for the HS-SPME/GC-MS chromatograms of all samples and QCs, namely the extracellular medium of HK-2 (n=10, ●), 769-P (n=10, ●), 786-O (n=10, ●), Caki-1 (n=10, ●), ACHN (n=10, ●), Caki-2 (n=10, ●) and QCs (n=10, ●). The ellipses indicate the 95% confidence limit of the model.

The study of the cell lines extracellular medium was accomplished through the execution of PCA-X and PLS-DA scores scatter plots (Figure 13a and 13b, respectively). The multivariate statistical models showed a good separation between normal and cancer cell lines, which is confirmed by the Q^2 value (68% of the variation is explained by cell lines differences).

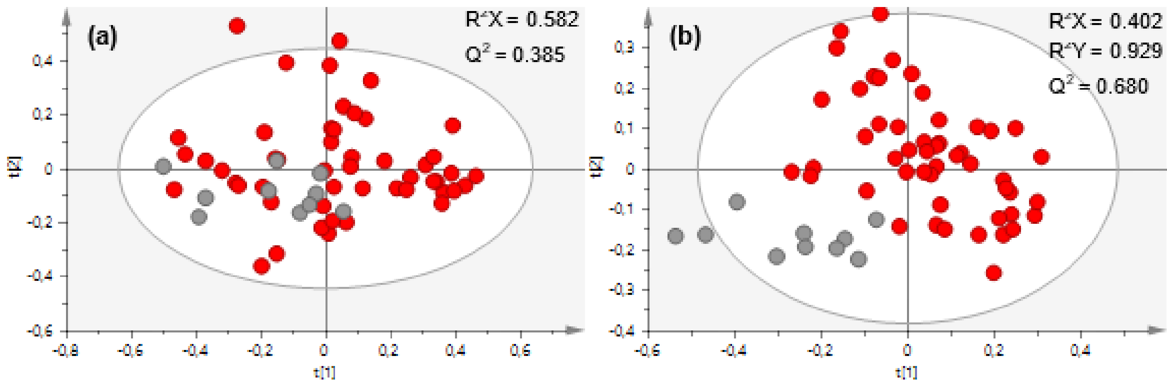
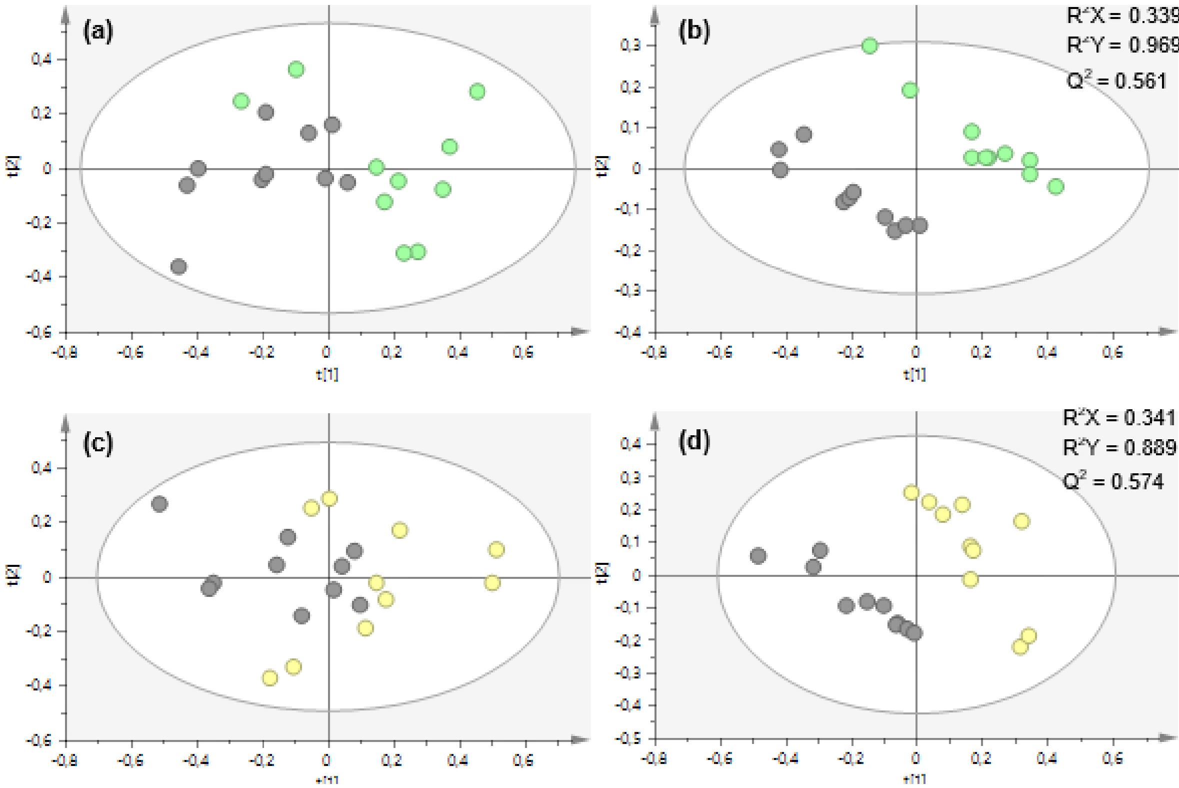


Figure 13. (a) PCA-X and **(b)** PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of all samples, namely the extracellular medium of the normal cell line (HK-2, n=10, ●) and the cancer cell lines (n=50, ●). The ellipses indicate the 95% confidence limit of the model.

PCA-X and PLS-DA scores scatter plots were also performed considering the comparison between the normal cell line and each RCC cell line (Figure 14), in order to study which compounds were responsible for the separation. The separation was maximized in the PLS-DA models, which presented good quality parameters (R^2X, R^2Y, and Q^2) (Figures 14b, 14d, 14f, 14h, and 14j).



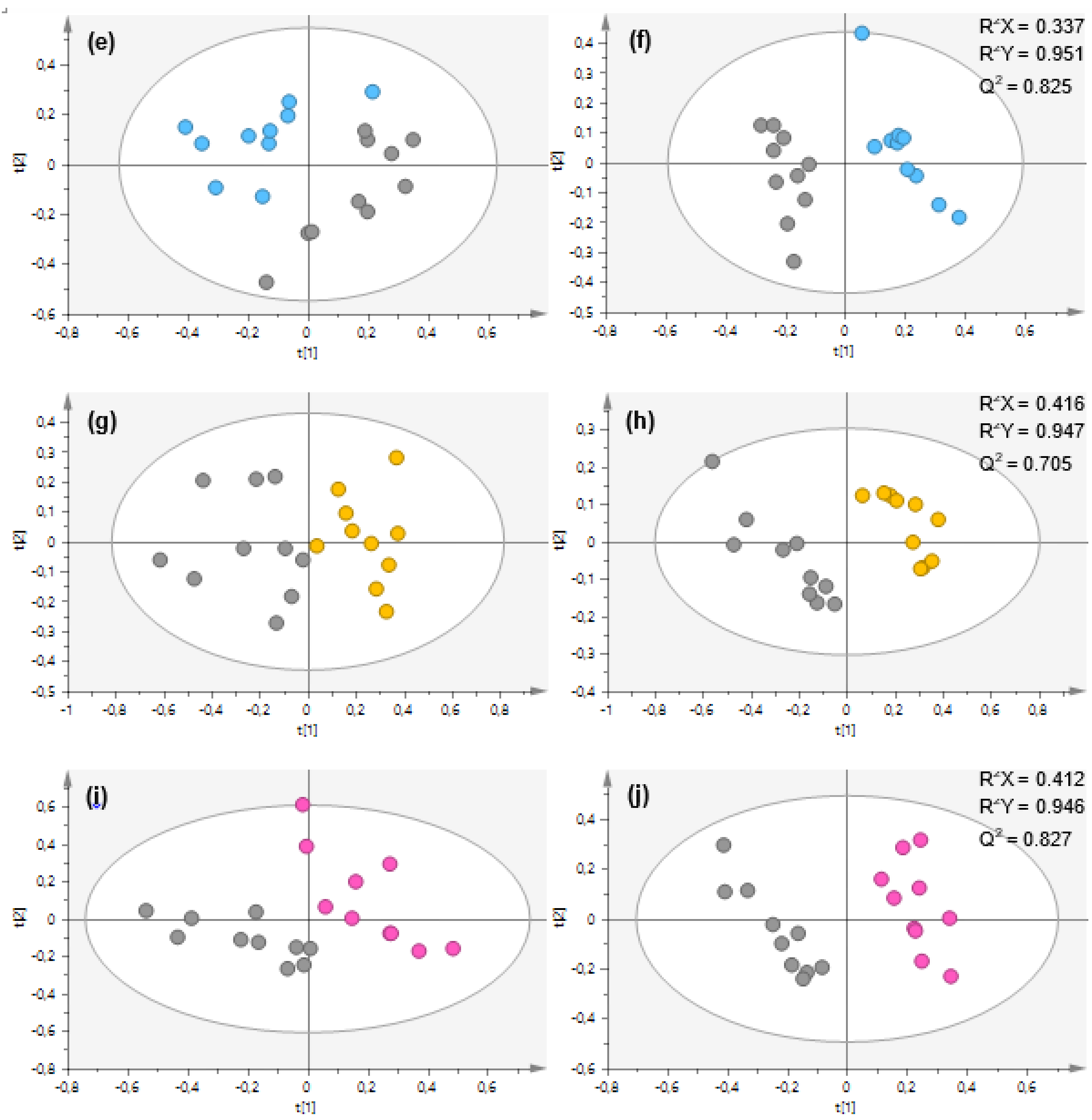


Figure 14. (a), (c), (e), (g) and **(i)** PCA-X and **(b), (d), (f), (h)** and **(j)** PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of the normal cell line (HK-2, n=10, ●) and the cancer cell lines 769-P (n=10, ●), 786-O (n=10, ●), Caki-1 (n=10, ●), ACHN (n=10, ●) and Caki-2 (n=10, ●), respectively. The ellipses indicate the 95% confidence limit of the model.

The univariate analysis was performed to identify metabolites that have significantly different concentrations in RCC vs normal cells. Table 7 presents the *p*-value, effect size (ES) and % of variation values for each metabolite with a VIP greater than 1.

Table 7. List of VOCs selected as important in discriminating the cancer cell lines (769-P, 786-O, Caki-1, Caki-2, and ACHN) from the normal cell line (HK-2). The *p*-value, ES (± ES_{SE}) and % variation (± uncertainty) are represented for each VOC. [pink] - VOC only present in the ccRCC; [green] - VOC only present in the pRCC; [beige] - VOC present in RCC subtypes; ↓ - decreased VOC in the cancer cell line; ↑ - increased VOC in the cancer cell line.

Organic Classification	VOCs	ccRCC									pRCC					
		non-metastatic						metastatic			non-metastatic			metastatic		
		769-P vs HK-2			786-O vs HK-2			Caki-1 vs HK-2			Caki-2 vs HK-2			ACHN vs HK-2		
		p-value [a]	ES (± ES_{SE}) [b]	% variation (± uncertainty)	p-value [a]	ES (± ES_{SE}) [b]	% variation (± uncertainty)	p-value [a]	ES (± ES_{SE}) [b]	% variation (± uncertainty)	p-value [a]	ES (± ES_{SE}) [b]	% variation (± uncertainty)	p-value [a]	ES (± ES_{SE}) [b]	% variation (± uncertainty)
Alcohols	2-Ethyl-1-hexanol	0.0017	1.66 (0.98)	17.08 (4.06) ↑							0.0268	1.09 (0.91)	12.81 (4.73) ↑	< 0.0001	3.08 (1.27)	28.31 (3.44) ↑
	Cyclohexanol										0.0005	2.17 (1.08)	23.55 (4.16) ↑			
Alkanes	Decane													0.0146	-1.22 (0.92)	-26.27 (10.58) ↓
	Dodecane	0.0176	-1.18 (0.92)	-27.95 (11.79) ↓	0.0266	-1.09 (0.91)	-23.60 (10.50) ↓	0.0297	-1.07 (0.90)	-21.31 (9.51) ↓				0.0002	-2.21 (1.08)	-42.41 (10.42) ↓
Alkene	2,4-Dimethyl-1-heptene	0.0487	1.00 (0.89)	84.81 (25.63) ↑				0.0106	1.35 (0.94)	95.34 (20.50) ↑	0.0385	1.03 (0.90)	62.48 (19.81) ↑			
Aromatic Hydrocarbons	1,3-Di-tert-butylbenzene										0.0231	-1.14 (0.91)	-21.87 (9.26) ↓			
	Ethylbenzene				0.0321	-1.05 (0.90)	-44.45 (23.34) ↓	0.0053	-1.53 (0.969	-50.70 (18.97) ↓						
	1,3-Dimethyl-benzene	0.0025	-1.58 (0.97)	-45.93 (16.12) ↓	0.003	-1.55 (0.97)	-49.02 (17.90) ↓	< 0.0001	-2.54 (1.15)	-56.69 (13.34) ↓				0.0247	-1.11 (0.91)	-32.38 (14.95) ↓
Ketones	2,6-Di-tert-butylquinone	0.0007	-2.00 (1.04)	-31.18 (7.90) ↓				0.0211	-1.14 (0.91)	-11.30 (4.50) ↓	0.0012	-1.91 (1.03)	-34.42 (9.33) ↓			
	Acetophenone	0.0085	-1.33 (0.94)	-39.15 (15.64) ↓							0.0004	-2.28 (1.10)	-49.59 (12.39) ↓			

Organic Classification	VOCs (cont.)	ccRCC									pRCC					
		non-metastatic						metastatic			non-metastatic			metastatic		
		769-P vs HK-2			786-O vs HK-2			Caki-1 vs HK-2			Caki-2 vs HK-2			ACHN vs HK-2		
		p-value a	ES (± ES_{SE}) b	% variation (± uncertainty)	p-value a	ES (± ES_{SE}) b	% variation (± uncertainty)	p-value a	ES (± ES_{SE}) b	% variation (± uncertainty)	p-value a	ES (± ES_{SE}) b	% variation (± uncertainty)	p-value a	ES (± ES_{SE}) b	% variation (± uncertainty)
Unknowns	1										< 0.0001	-3.01 (1.26)	-60.22 (12.24) ↓			
	2	0.0299	1.10 (0.91)	62.33 (18.49) ↑	0.0295	1.11 (0.91)	70.97 (20.13) ↑	0.0002	2.40 (1.12)	179.88 (16.87) ↑	0.0185	1.17 (0.91)	41.50 (12.59) ↑			
	3							0.0021	1.77 (1.00)	50.06 (9.69) ↑				0.0005	-1.91 (1.03)	-37.58 (10.36) ↓
	4							< 0.0001	3.08 (1.27)	125.54 (10.74) ↑						
	5	0.0006	1.91 (1.03)	62.02 (10.64) ↑	0.0004	2.04 (1.05)	73.85 (11.30) ↑	< 0.0001	4.37 (1.59)	100.44 (6.56) ↑						
	6							0.0318	1.10 (0.91)	66.22 (12.29) ↑						
	7													0.0326	-1.05 (0.90)	-17.01 (7.61) ↓
	8							0.0007	2.21 (1.08)	69.80 (10.04) ↑	0.0354	1.00 (0.89)	27.23 (10.27) ↑	0.0431	-1.05 (0.90)	-27.12 (12.74) ↓
	9							0.002	1.77 (1.00)	63.31 (11.62) ↑						
	10										< 0.0001	-5.26 (1.83)	-71.93 (9.15) ↓			
	11							0.0175	1.27 (0.93)	52.59 (14.09) ↑				0.0038	-1.51 (0.96)	-29.48 (9.83) ↓
	12				0.0431	-1.02 (0.90)	-61.00 (36.31) ↓				0.0353	-1.08 (0.90)	-39.43 (19.45) ↓	0.0115	-1.16 (0.91)	-67.80 (37.71) ↓
	13				0.0063	1.40 (0.94)	39.29 (10.07) ↑									

NOTE:

a ES determined as described in reference Berben et al. (2012);

b 95% significance level (*p*-value <0.05).

The comparison between 769-P (non-metastatic ccRCC) and HK-2 (normal) cell lines presented eight discriminant compounds (Table 7). Boxplots of six VOCs found significantly altered between these two cell lines are depicted in Figure 15.

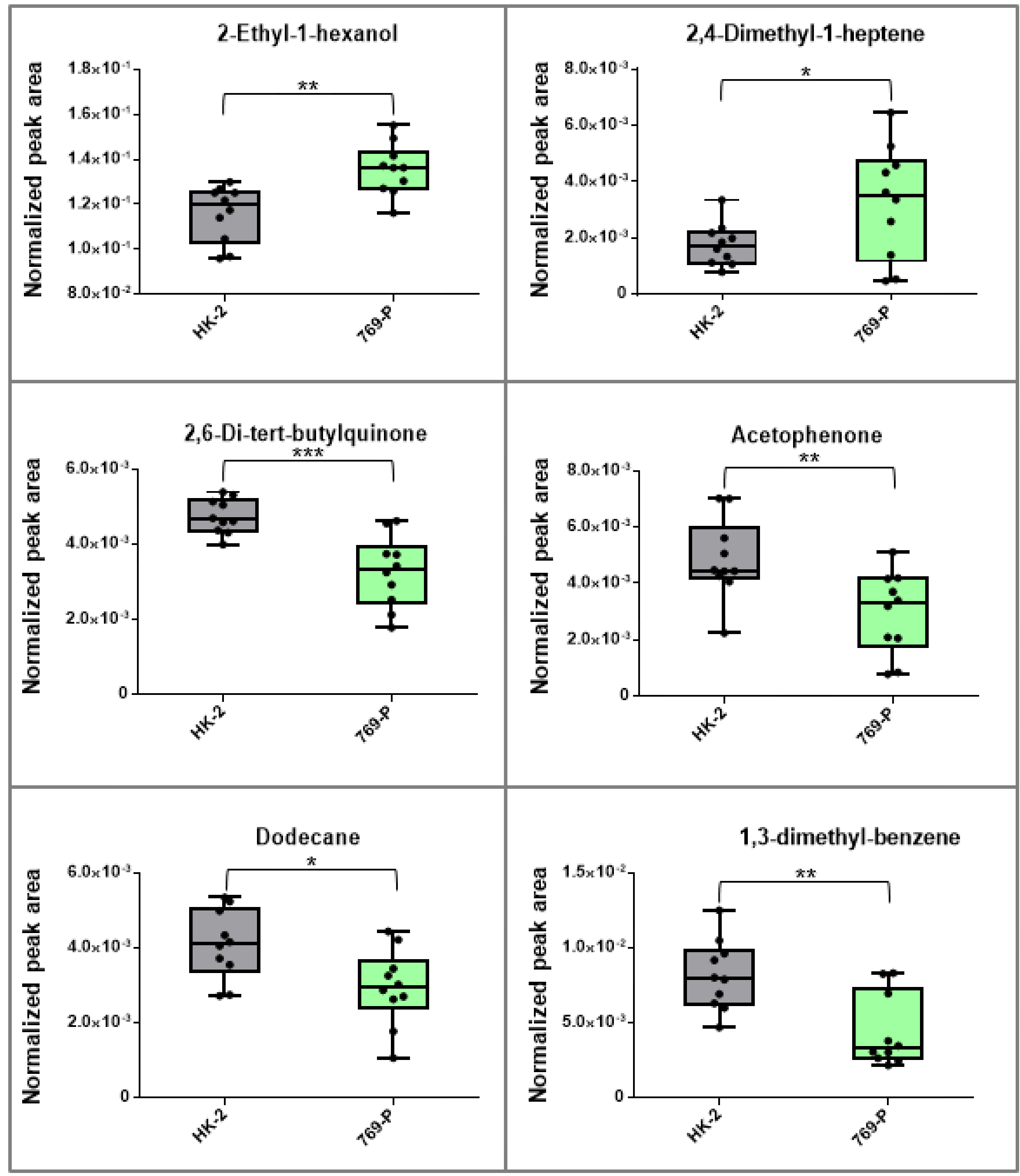


Figure 15. Boxplots of six VOCs found significantly altered between HK-2 (●) and 769-P (●) cell lines, namely 2-ethyl-1-hexanol, 2,4-dimethyl-1-heptene, 2,6-di-tert-butylquinone, acetophenone, dodecane and 1,3-dimethyl-benzene.

The comparison between 786-O (non-metastatic ccRCC) and HK-2 cell lines presented seven discriminant compounds (Table 7). Figure 16 shows the boxplots of three VOCs found significantly altered between these two cell lines.

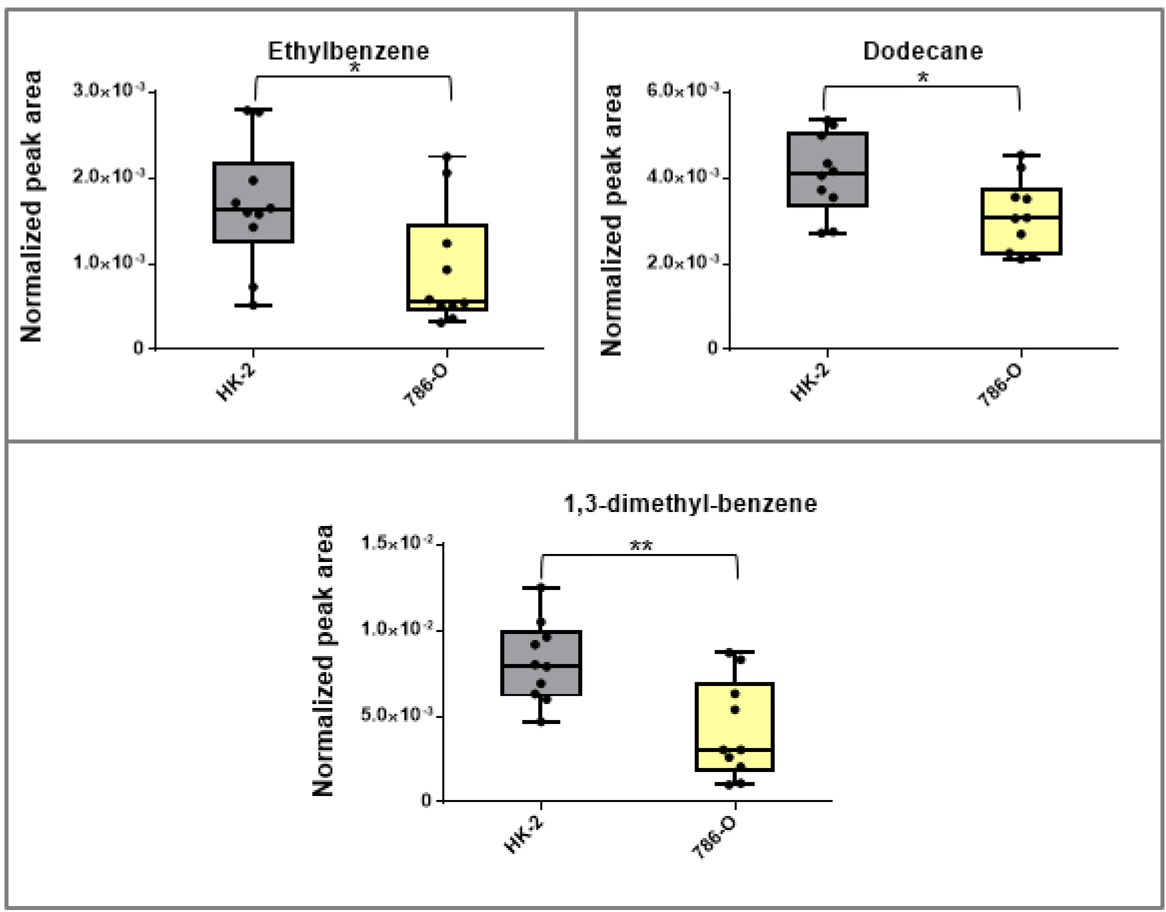


Figure 16. Boxplots of three VOCs found significantly altered between HK-2 (●) and 786-O (○) cell lines, namely ethylbenzene, dodecane and 1,3-dimethyl-benzene.

The comparison between Caki-1 (metastatic ccRCC) and HK-2 cell lines revealed fourteen discriminant compounds (Table 7). Boxplots of five VOCs found significantly altered between these two cell lines are presented in Figure 17.

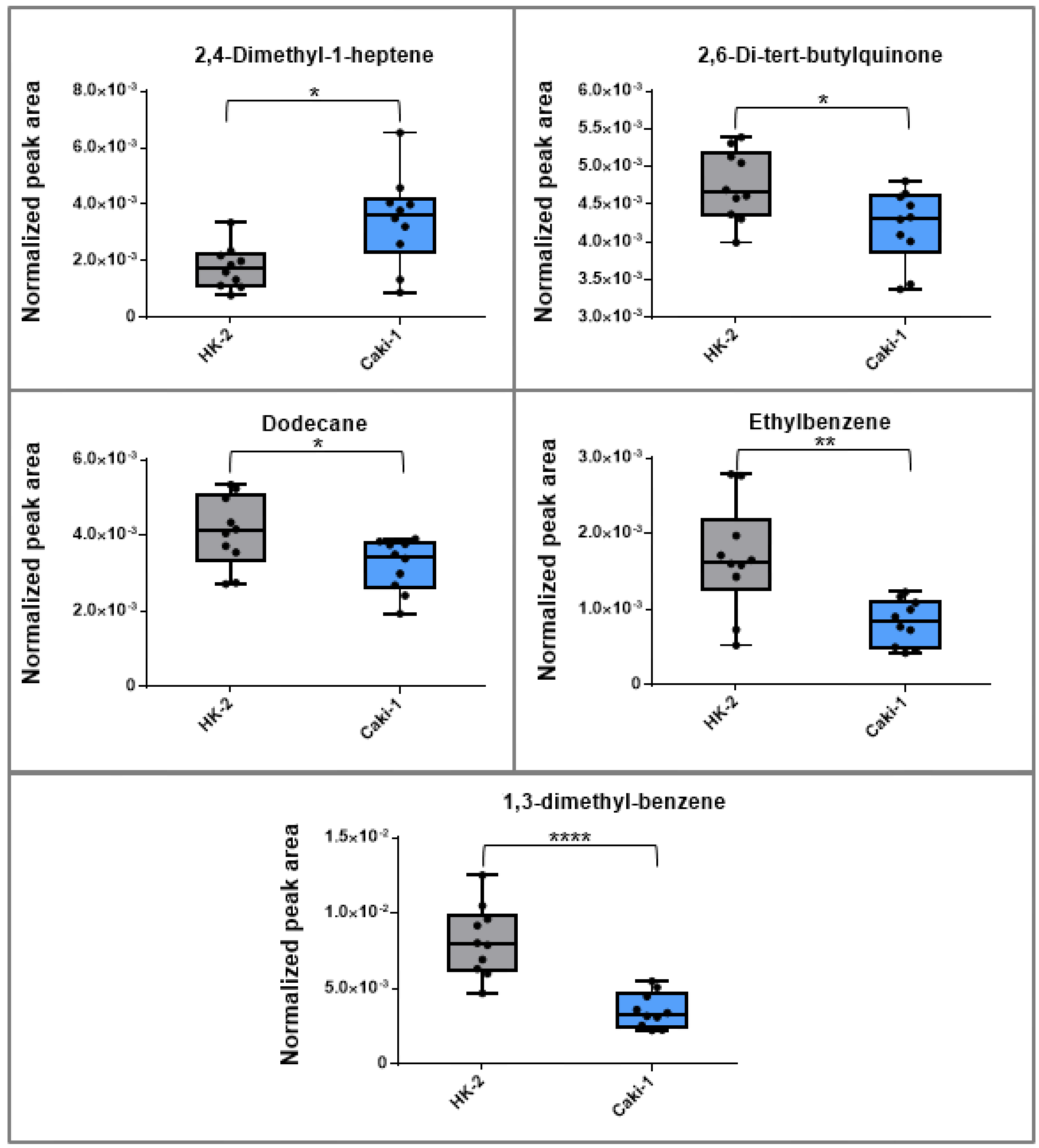


Figure 17. Boxplots of five VOCs found significantly altered between HK-2 (●) and Caki-1 (●) cell lines, namely 2,4-dimethyl-1-heptene, 2,6-di-tert-butylquinone, dodecane, ethylbenzene and 1,3-dimethyl-benzene.

The comparison between Caki-2 (non-metastatic pRCC) and normal HK-2 cell lines presented eleven discriminant compounds (Table 7). Boxplots of six VOCs found significantly altered between these two cell lines are shown in Figure 18.

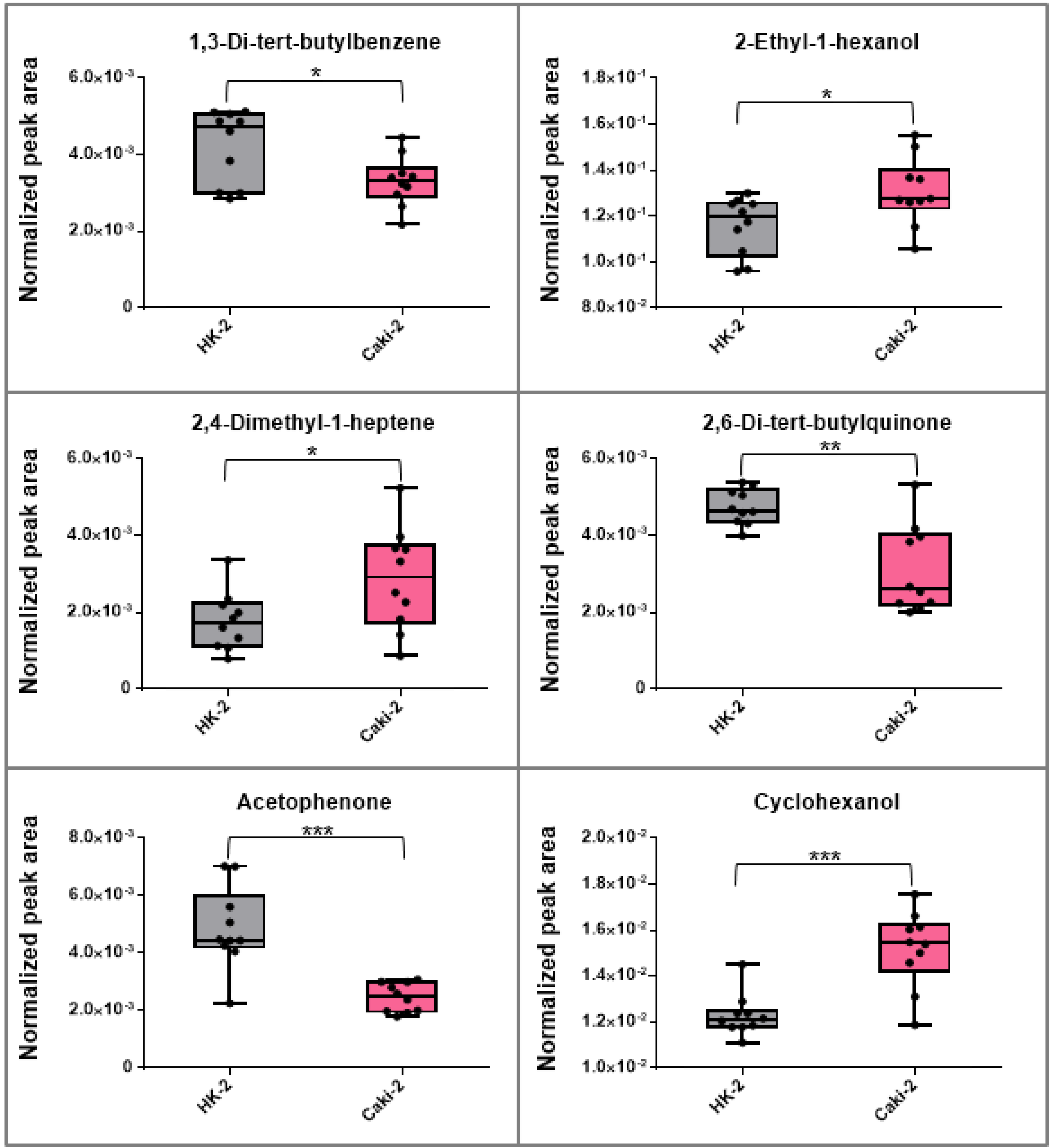


Figure 18. Boxplots of six VOCs found significantly altered between HK-2 (●) and Caki-2 (●) cell lines, namely 1,3-di-tert-butylbenzene, 2-ethyl-1-hexanol, 2,4-dimethyl-1-heptene, 2,6-di-tert-butylquinone, acetophenone and cyclohexanol.

The comparison between ACHN (metastatic pRCC) and HK-2 cell lines unveiled nine discriminant compounds (Table 7). Boxplots of four VOCs found significantly altered between these two cell lines are presented in Figure 19.

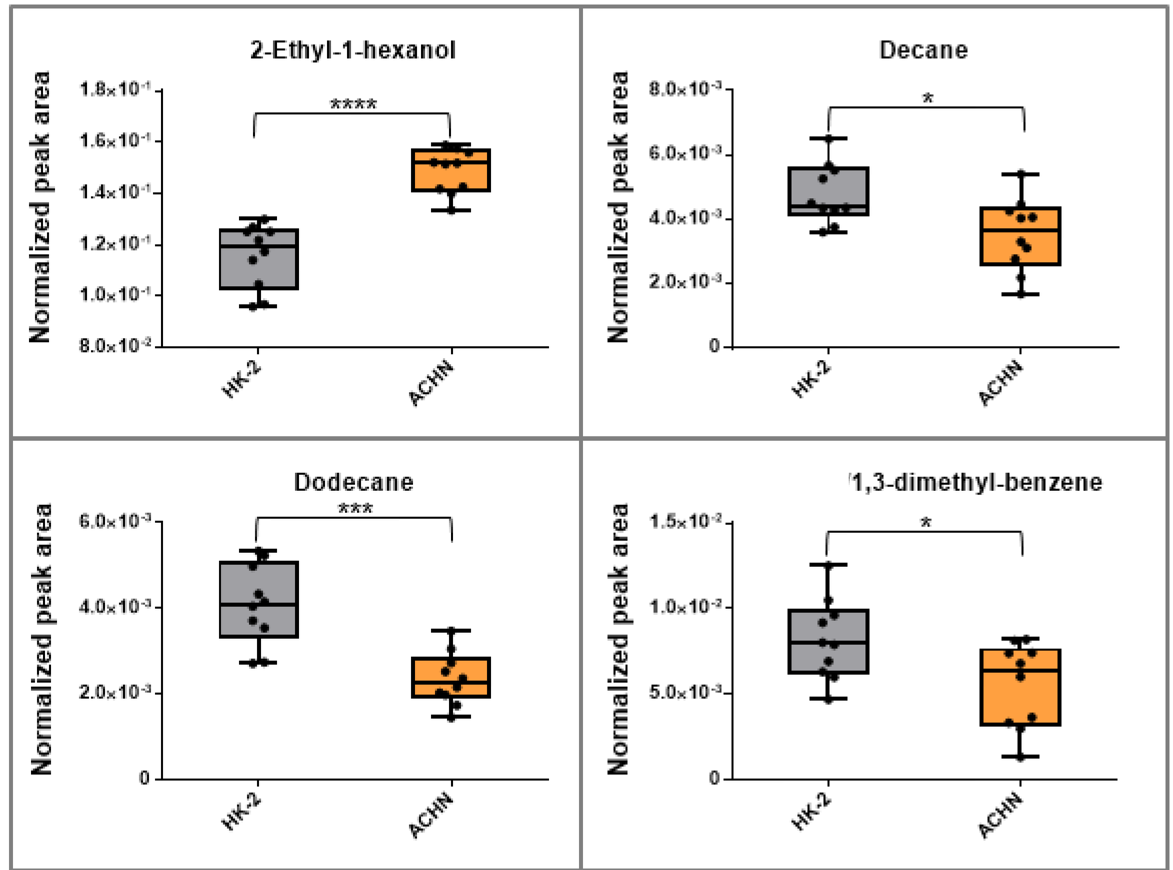


Figure 19. Boxplots of four VOCs found significantly altered between HK-2 (●) and ACHN (●) cell lines, namely 2-ethyl-1-hexanol, decane, dodecane and 1,3-dimethyl-benzene.

5.2 VCCs signature profile in RCC cell lines vs normal kidney cell line

Multivariate analysis performed on VCCs data reveals that QC samples are not closely clustered and centered in the PCA-X (Figure 20). Of note, the analysis of the cell line Caki-2 was achieved only with eight experiments due to an error during the injection of two samples.

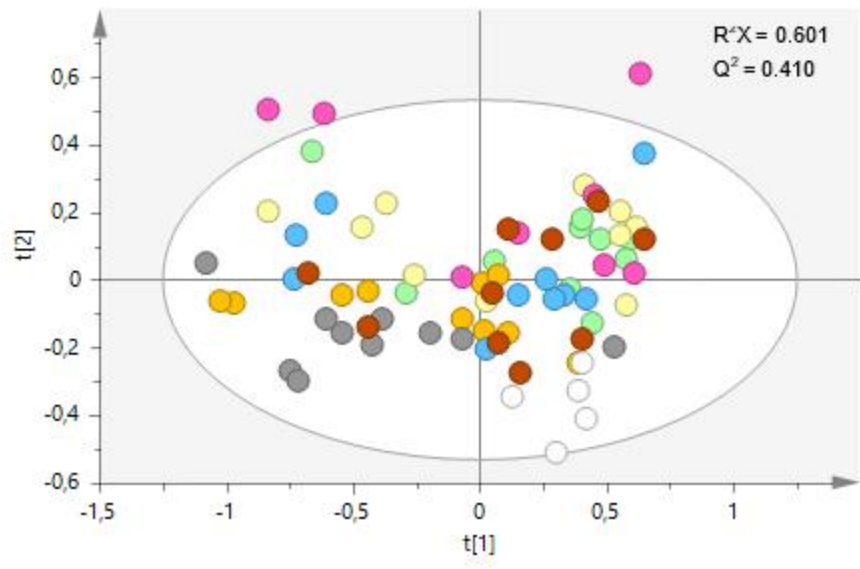


Figure 20. PCA-X score scatter plot obtained for the HS-SPME/GC-MS chromatograms of all samples, QCs and MM namely the extracellular medium of HK-2 (n=10, ●), 769-P (n=10, ●), 786-O (n=10, ○), Caki-1 (n=10, ●), ACHN (n=10, ●), Caki-2 (n=8, ●), QCs (n=10, ●) and ME (n=5, ○). The ellipses indicate the 95% confidence limit of the model.

The study of the cell lines extracellular medium was accomplished through the execution of PCA-X and PLS-DA scores scatter plots (Figure 21a and 21b, respectively). The multivariate statistical models showed a clear separation between normal and cancer cell lines, which is confirmed by the Q^2 value (82% of the variation is explained by cell lines differences).

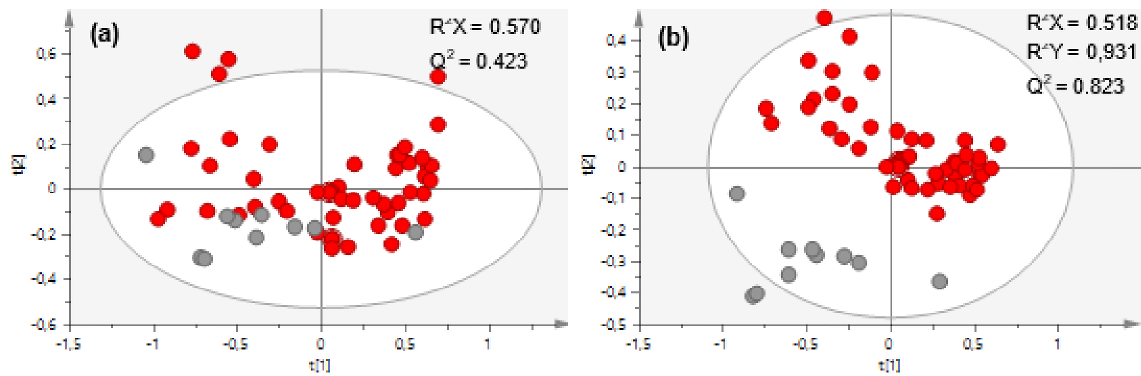
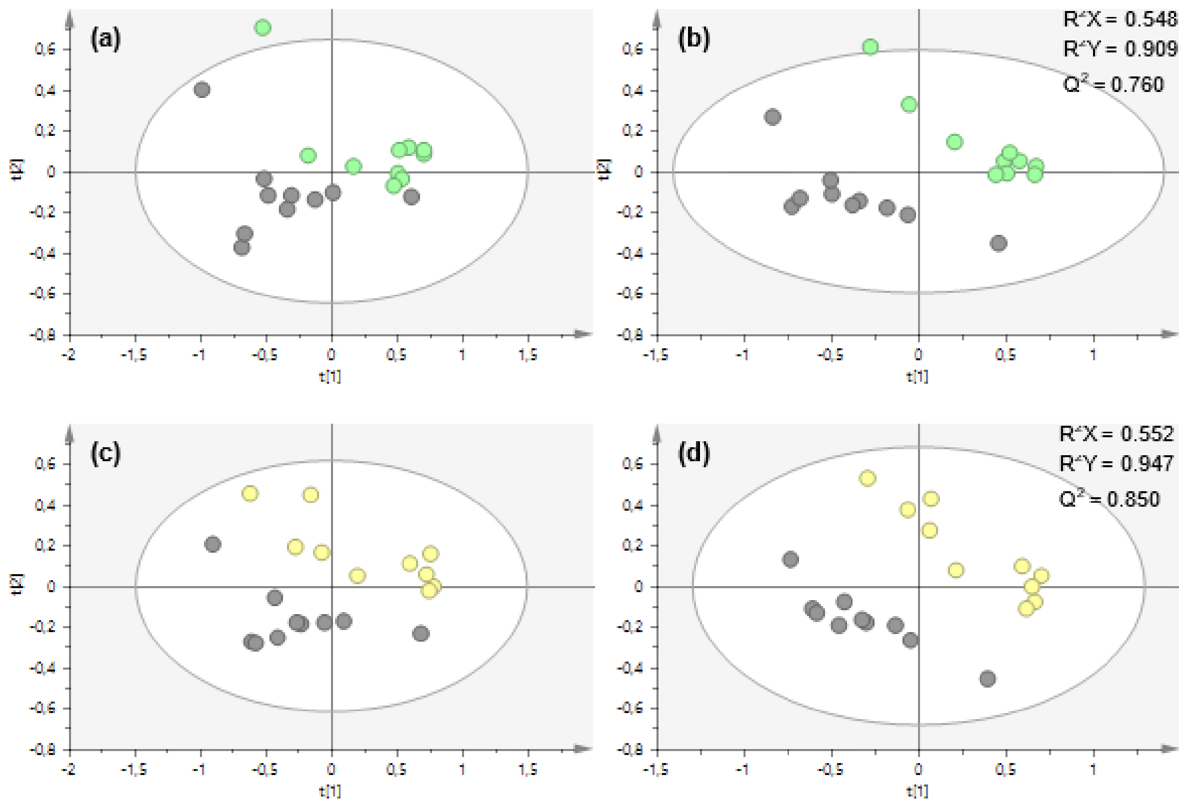


Figure 21. **(a)** PCA-X and **(b)** PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of all samples, namely the extracellular medium of the normal cell line (HK-2, n=10, ●) and the cancer cell lines (n=48, ●). The ellipses indicate the 95% confidence limit of the model.

PCA-X and PLS-DA scores scatter plots were next performed considering only the comparison between the normal cell line and each RCC cell line (Figure 22), in order to study which compounds were responsible for the separation. The separation was maximized in the PLS-DA models, which presented good quality parameters (R^2X, R^2Y, and Q^2) (Figures 22b, 22d, 22f, 22h, and 22j).



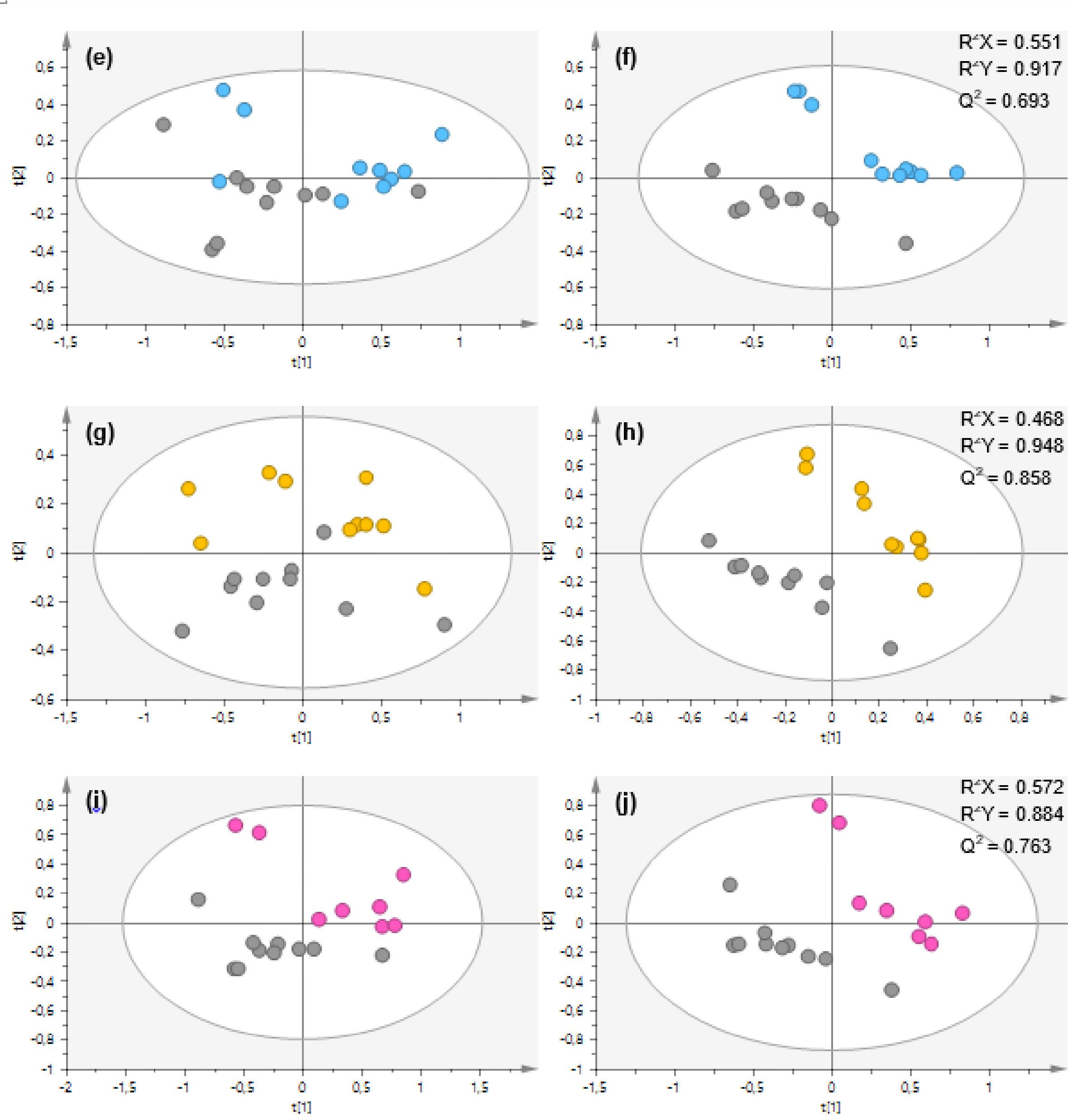


Figure 22. (a), (c), (e), (g) and **(i)** PCA-X and **(b), (d), (f), (h)** and **(j)** PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of the normal cell line (HK-2, n=10, ●) and the cancer cell lines 769-P (n=10, ●), 786-O (n=10, ●), Caki-1 (n=10, ●), ACHN (n=10, ●) and Caki-2 (n=8, ●), respectively. The ellipses indicate the 95% confidence limit of the model.

The univariate analysis was performed for all cell lines comparisons and compared with the metabolomic medium incubated under the same conditions as the samples, in order to determine which metabolites have increased or decreased as a result of cell metabolism. Table 8 presents the p-value, effect size (ES) and % of variation values for each metabolite with a VIP greater than 1.

Table 8. List of VCCs selected as important in discriminating the cancer cell lines (769-P, 786-O, Caki-1, Caki-2, and ACHN) from the normal cell line (HK-2), and in discriminating de metastatic pRCC from the non-metastatic pRCC. The *p*-value, ES (± ES_{SE}) and % variation (± uncertainty) are represented for each VCC. [pink] - VCC only present in the ccRCC; [green] - VCC only present in the pRCC; [beige] - VCC present in RCC subtypes; ↓ - decreased VCC in the cancer cell line; ↑ - increased VCC in the cancer cell line.

Organic Classification	VCCs	ccRCC									pRCC								
		non-metastatic						metastatic			non-metastatic			metastatic			non-metastatic pRCC vs metastatic pRCC		
		769-P vs HK-2			786-O vs HK-2			Caki-1 vs HK-2			Caki-2 vs HK-2			ACHN vs HK-2					
		p-value a	ES (± ES_{SE}) b	% variation (± uncertainty)	p-value a	ES (± ES_{SE}) b	% variation (± uncertainty)	p-value a	ES (± ES_{SE}) b	% variation (± uncertainty)	p-value a	ES (± ES_{SE}) b	% variation (± uncertainty)	p-value a	ES (± ES_{SE}) b	% variation (± uncertainty)	p-value a	ES (± ES_{SE}) b	% variation (± uncertainty)
Aldehydes	Acetaldehyde	< 0.0001	-4.12 (1.53)	-94.35 (18.58) ↓	< 0.0001	-4.06 (1.51)	-93.09 (18.35) ↓	< 0.0001	-3.74 (1.43)	-86.00 (17.26) ↓	< 0.0001	-3.74 (1.51)	-91.20 (16.49) ↓	< 0.0001	-3.07 (1.27)	-75.42 (16.87) ↓	0.0036	-1.62 (1.03)	-64.20 (22.51) ↓
	Glyoxal				0.0029	-1.38 (0.94)	-57.97 (25.32) ↓				0.0133	-1.26 (0.98)	-55.74 (24.62) ↓				0.0472	-1.01 (0.95)	-39.42 (20.93) ↓
	Methlyglyoxal	< 0.0001	-2.52 (1.15)	-65.35 (16.49) ↓	< 0.0001	-3.00 (1.27)	-67.42 (14.10) ↓				< 0.0001	-3.07 (1.34)	-70.23 (14.03) ↓	0.0185	-1.09 (0.91)	-33.78 (15.93) ↓	0.0044	-1.34 (0.99)	-55.05 (22.65) ↓
Ketones	Acetone	0.0122	-1.27 (0.93)	-31.12 (12.41) ↓															
	Cyclohexanone	< 0.0001	-3.35 (1.33)	-63.42 (11.88) ↓	0.0039	-1.50 (0.96)	-33.33 (11.42) ↓	0.0028	-1.56 (0.979	-36.60 (12.27) ↓	< 0.0001	-5.39 (1.97)	-97.56 (12.83) ↓				< 0.0001	-5.37 (1.96)	-97.48 (12.87) ↓
Unknowns	1	0.0086	-1.34 (0.94)	-53.35 (23.19) ↓	0.0089	-1.52 (0.96)	-62.76 (25.81) ↓												
	2	0.0034	-1.58 (0.97)	-42.24 (14.55) ↓							0.0047	-1.51 (1.01)	-43.36 (15.43) ↓				0.0021	-1.68 (1.04)	-47.13 (15.28) ↓
	3	0.0039	-1.66 (0.98)	-74.38 (30.56) ↓	0.0026	-1.73 (1.00)	-79.31 (32.56) ↓	0.027	-1.14 (0.91)	-53.06 (27.23) ↓	0.0052	-1.52 (1.02)	-70.95 (27.53) ↓				0.0117	-1.13 (0.96)	-38.50 (17.70) ↓
	4	0.0354	0.71 (0.87)	34.01 (17.45) ↑	< 0.0001	2.06 (1.06)	-80.79 (11.95) ↓												
	5	0.0038	1.52 (0.96)	57.94 (12.66) ↑															
	6													0.0473	-0.96 (0.89)	-51.26 (30.70) ↓			
	7				0.0019	-1.84 (1.01)	-81.34 (31.94) ↓				0.0062	-1.57 (1.02)	-76.72 (31.09) ↓						

NOTE:

a ES determined as described in reference Berben et al. (2012);

b 95% significance level (*p*-value <0.05).

The comparison between 769-P (non-metastatic ccRCC) and HK-2 cell lines presented nine discriminant compounds (Table 8). Boxplots of four VCCs found significantly altered between these two cell lines are depicted in Figure 23.

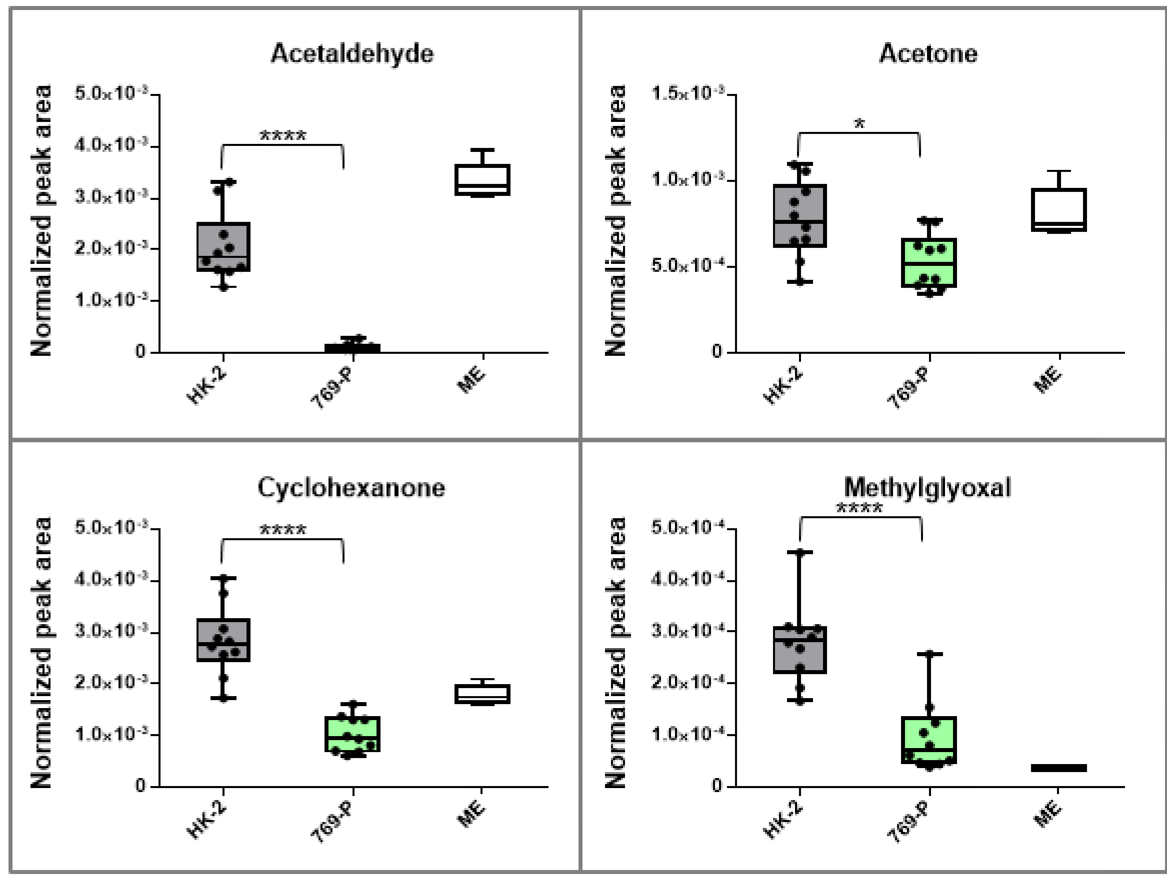


Figure 23. Boxplots of four VCCs found significantly altered between HK-2 (●) and 769-P (●), namely acetaldehyde, acetone, cyclohexanone and methylglyoxal. ME - metabolomic medium incubated under the same conditions as the samples (○).

The comparison between 786-O (non-metastatic ccRCC) and HK-2 cell lines presented eight discriminant compounds (Table 8). Boxplots of four VCCs found significantly altered between these two cell lines are shown in Figure 24.

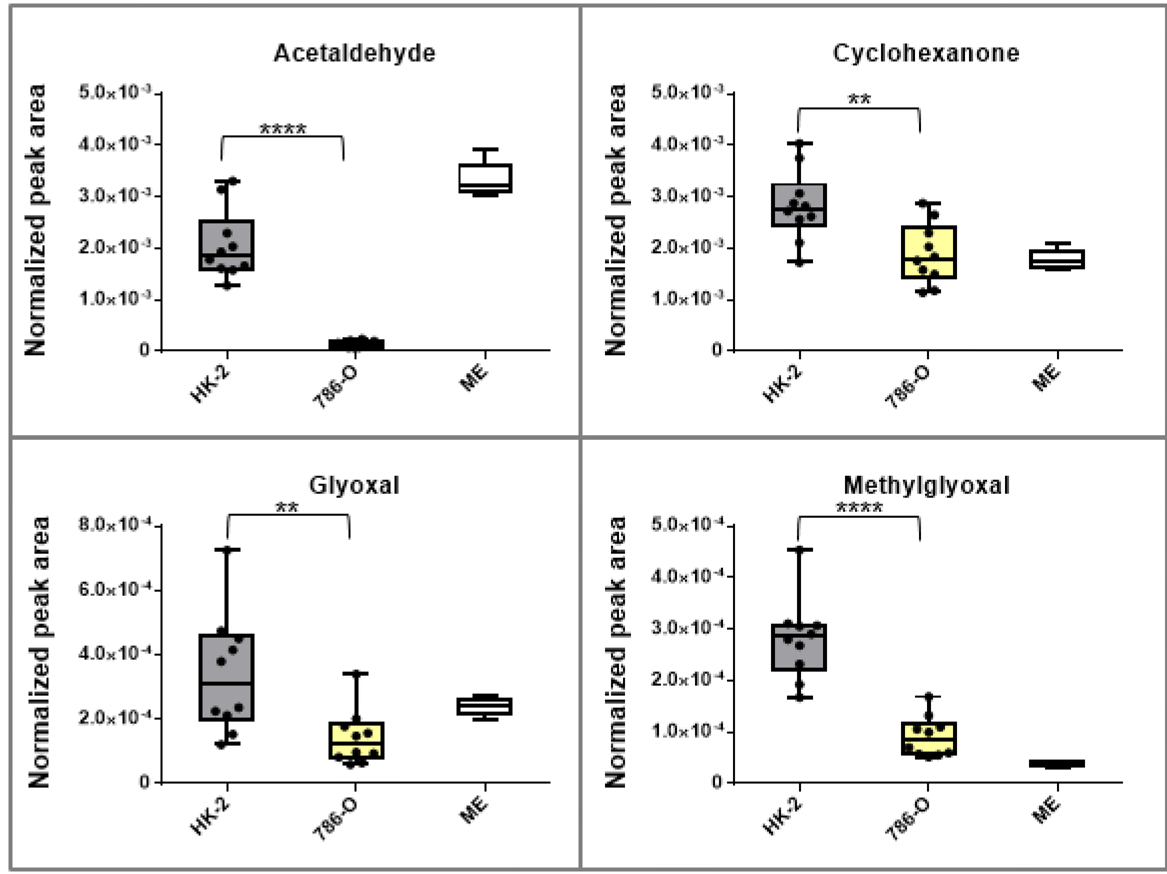


Figure 24. Boxplots of four VCCs found significantly altered between HK-2 (●) and 786-O (○) cell lines, namely acetaldehyde, cyclohexanone, glyoxal and methylglyoxal. ME (○).

The comparison between Caki-1 (metastatic ccRCC) and HK-2 cell lines revealed three discriminant compounds (Table 8). Figure 25 shows the boxplots of two VCCs found significantly altered between these cell lines.

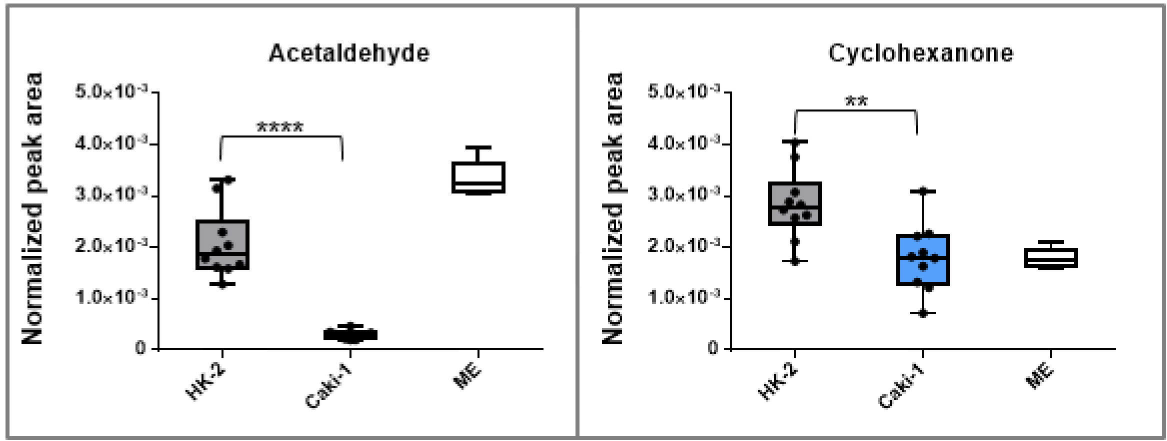


Figure 25. Boxplots of two VCCs found significantly altered between HK-2 (●) and Caki-1 (●) cell lines, namely acetaldehyde and cyclohexanone. ME (○).

The comparison between Caki-2 (non-metastatic pRCC) and HK-2 cell lines presented seven discriminant compounds (Table 8). Boxplots of four VCCs found significantly altered between these two cell lines are illustrated in Figure 26.

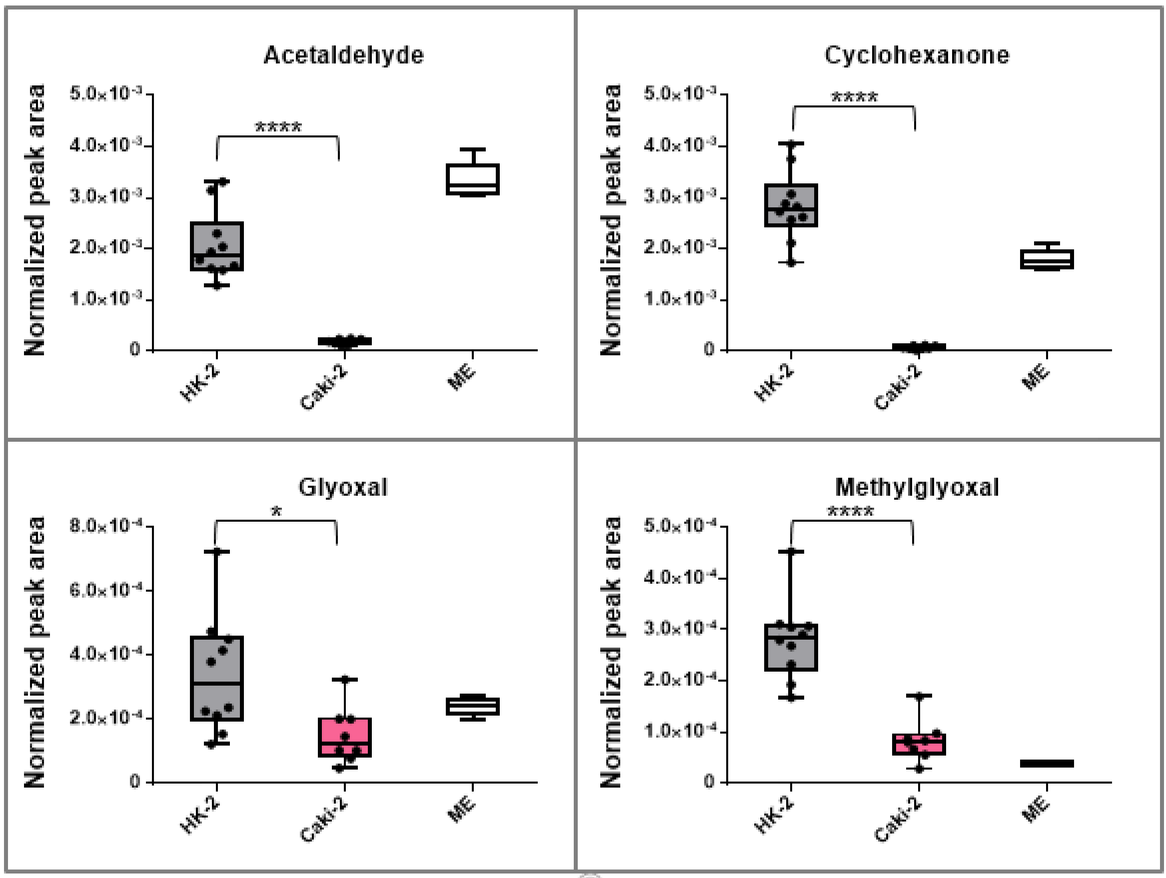


Figure 26. Boxplots of four VCCs found significantly altered between HK-2 (●) and Caki-2 (●) cell lines, namely acetaldehyde, cyclohexanone, glyoxal and methylglyoxal. ME (○).

The comparison between ACHN (metastatic pRCC) and HK-2 cell lines presented three discriminant compounds (Table 8). Boxplots of two VCCs found significantly altered between these two cell lines are represented in Figure 27.

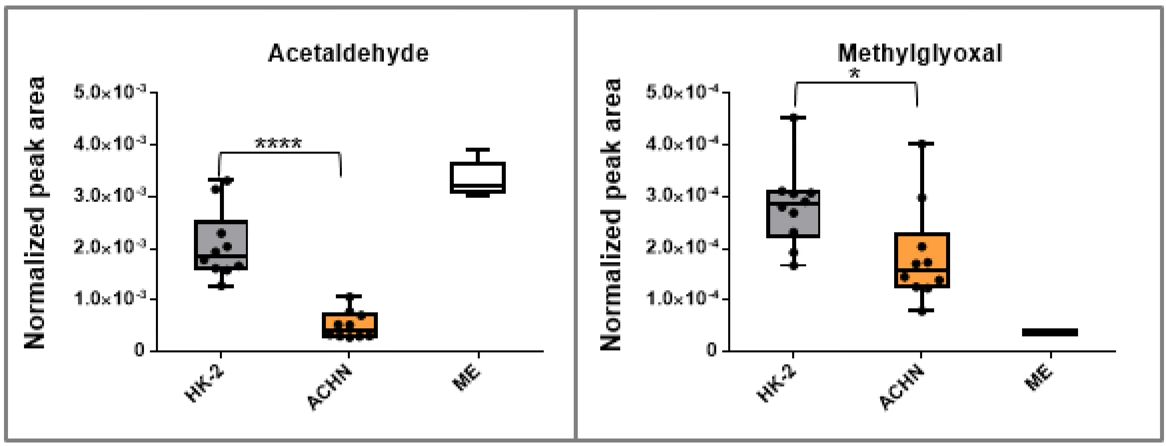


Figure 27. Boxplots of two VCCs found significantly altered between HK-2 (●) and ACHN (●) cell lines, namely acetaldehyde and methylglyoxal. ME (○).

Analysis of VCCs also included the comparison between metastatic *vs* non-metastatic for each RCC subtype and the comparison between ccRCC *vs* pRCC subtypes. PCA-X and PLS-DA scores scatter plots were performed (Figure 28) to verify whether there was a separation between the combinations referred to above and to study which compounds were responsible for such separation. The supervised model showed some separation (VIP>1) only in the comparison between metastatic pRCC (ACHN) and non-metastatic pRCC (Caki-2) on the first principal component (PC1) (Figure 28f). Once the other analyzes obtained low Q^2 values there was no interest in pursuing univariate statistics (Figures 28b and 28d). To guarantee the quality of the model, each PLS-DA presents the values of R^2X, R^2Y, and Q^2.

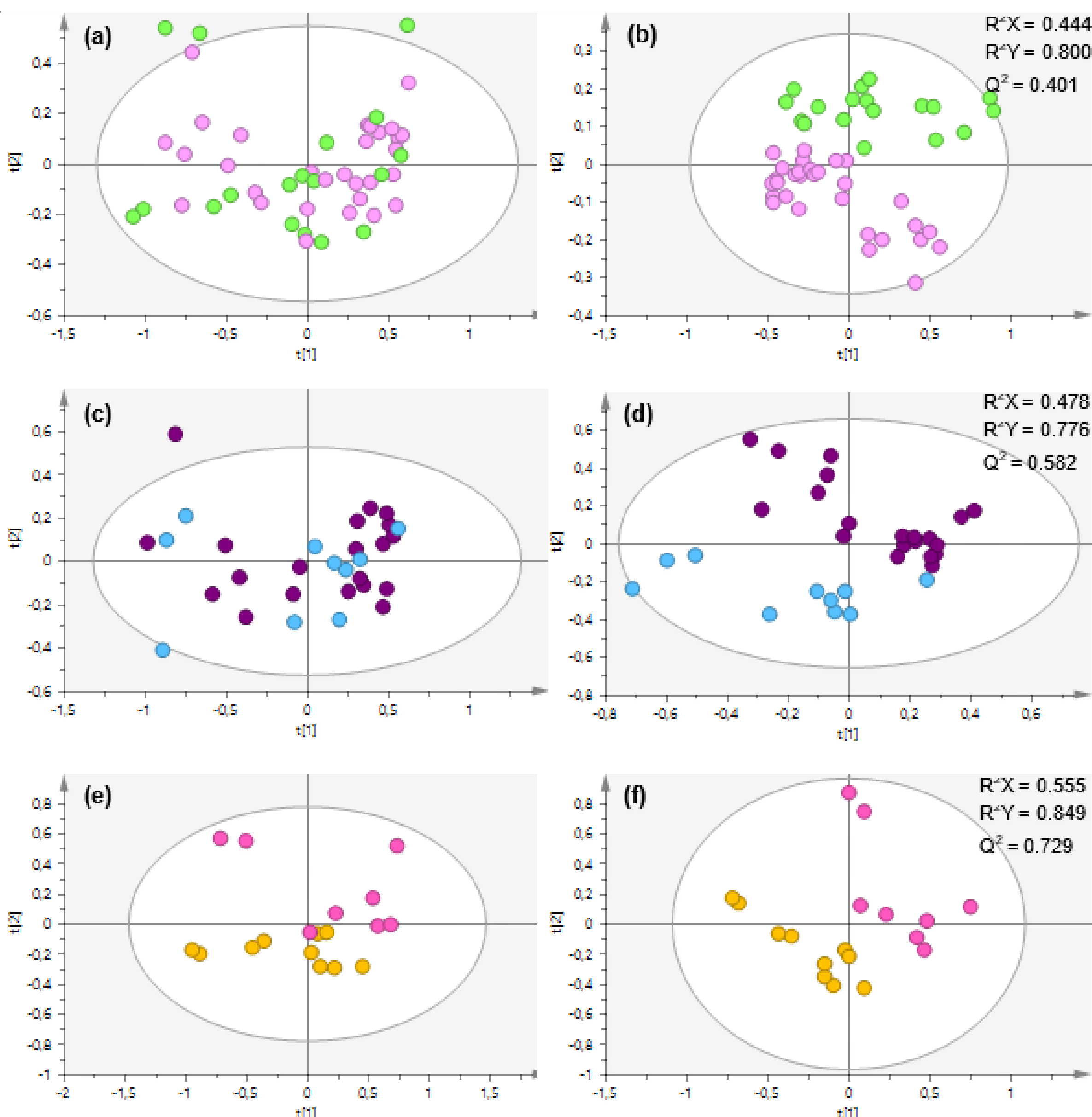


Figure 28. (a), (c) and **(e)** PCA-X and **(b), (d)** and **(f)** PLS-DA scores scatter plots obtained for the HS-SPME/GC-MS chromatograms of the ccRCC (n=30, ●), pRCC (n=18, ●), non-metastatic ccRCC (769-P and 786-O, n=20, ●), metastatic ccRCC (Caki-1, n=10, ●), non-metastatic pRCC (Caki-2, n=8, ●) and metastatic pRCC (ACHN, n=10, ●), respectively. The ellipses indicate the 95% confidence limit of the model.

The comparison between metastatic pRCC (ACHN) *vs* non-metastatic pRCC (Caki-2) cell lines presented six discriminant compounds. Boxplots of some of the most important metabolites responsible for the separation between these two cell lines are represented in Figure 29.

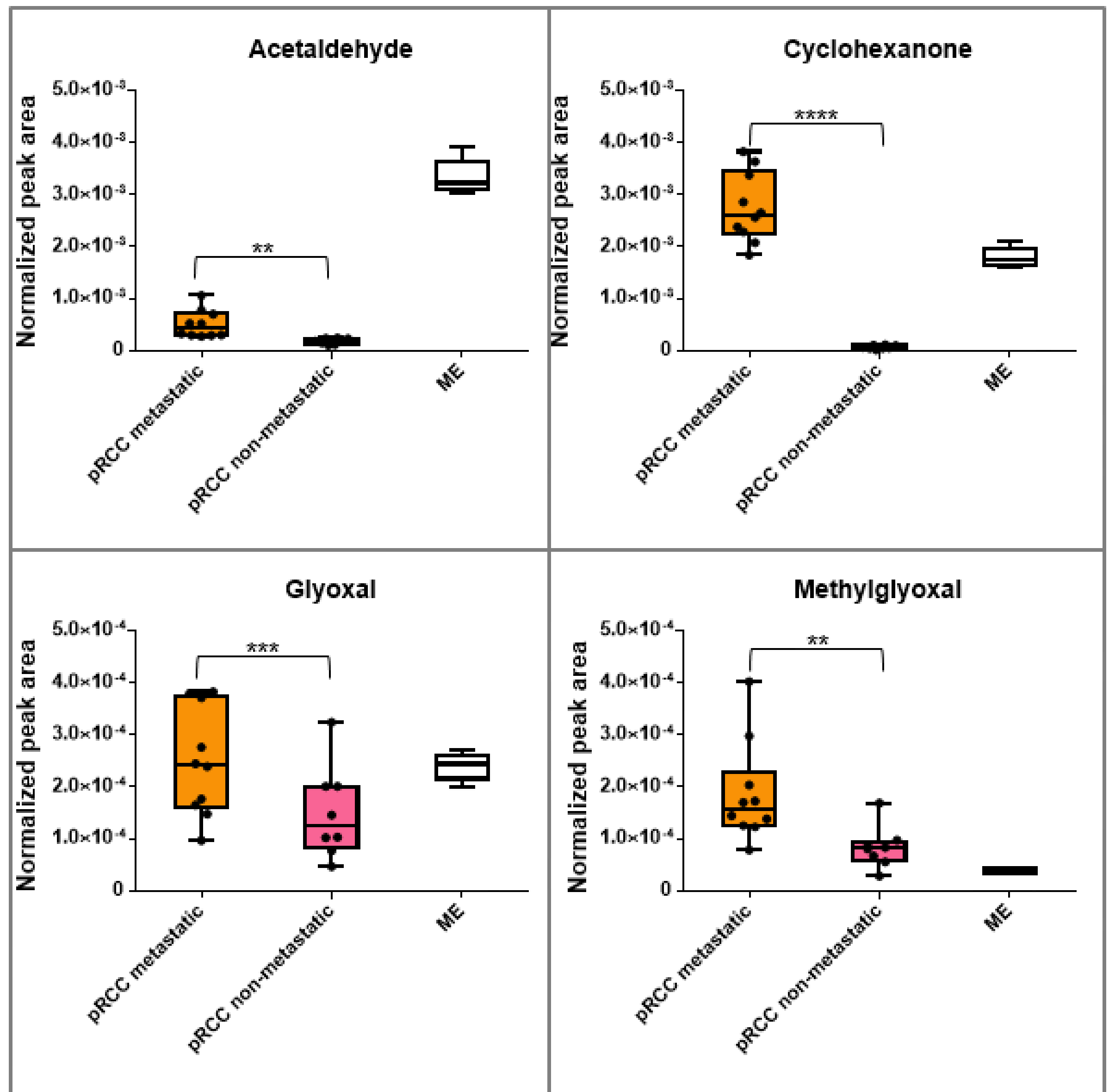


Figure 29. Identified metabolites that are responsible for the separation between metastatic pRCC (ACHN, ●) *vs* non-metastatic pRCC (Caki-2, ●) cell lines, namely acetaldehyde, cyclohexanone, glyoxal and methylglyoxal. ME (○).

6 Discussion

In this work, we presented the potential of volatile metabolome to discriminate several RCC cell lines from a normal kidney cell line. The cancer cell lines used in this project are classified according to histological subtype, with 769-P, 786-O and Caki-1 representing ccRCC, and Caki-2 and ACHN the pRCC. In addition, Caki-1 and ACHN cell lines represent metastatic ccRCC and pRCC, respectively, whereas all the others are non-metastatic, which allowed for extending our investigation on the metabolites responsible for discriminating RCC cell lines according to histological subtype and metastatic potential.

A diversity of VOCs and VCCs was found to be highly discriminative between the normal and RCC cell lines, which will be discussed as follows.

6.1 Discriminative VOCs in RCC vs normal cell lines

Multivariate statistical analysis unveiled 23 discriminant compounds, however, only 10 of them were identified through the database, namely 1,3-di-tert-butylbenzene, 2,4-dimethyl-1-heptene, 2,6-di-tert-butylquinone, 2-ethyl-1-hexanol, acetophenone, cyclohexanol, decane, dodecane, ethylbenzene, 1,3-dimethyl-benzene. The classes of the identified chemical compounds comprise alkanes, alkenes, aromatic hydrocarbons, alcohols and ketones.

The alkane decane was found significantly altered only in the metastatic papillary subtype (ACHN), with decreased levels in this cancer cell line medium. Contrary to our findings, a recent *in vitro* study identified that this VOC appears with higher levels in lung cancer cell lines than in normal one [281]. Decane was also reported as a potential screening biomarker for the hepatocellular carcinoma [282], lung cancer [211, 283] and melanoma [284].

Regarding the alkane dodecane, it was found significantly altered in all comparisons, except for Caki-2 *vs* HK-2. This outcome revealed always the same behavior, a diminished level of this VOC in the medium of cancer cell lines. Serasanambati et al. observed the same condition in lung cancer cell lines [285], however, a major part of previous investigations reflected the opposite. Higher levels of dodecane were reported in exhaled breath samples of lung cancer [206, 281, 286] and colorectal cancer (CRC) patients [287], in serum samples from breast cancer patients [288] and in melanoma tissues [203] when compared to healthy controls. Similar results were observed in *in vitro* studies for bladder cancer [289] and non-Hodgkin's lymphoma [290].

The alkene 2,4-dimethyl-1-heptene was found significantly increased in 769-P, Caki-1 and Caki-2 cell lines when compared with HK-2. 2,4-Dimethyl-1-heptene was already identified in melanoma cell cultures [291] and found increased in lung cancer cell lines [285, 292] and reported as potential biomarker to the same pathology.

The aromatic hydrocarbon 1,3-di-tert-butylbenzene was found significantly decreased in Caki-2 medium compared to HK-2 medium. Importantly, this significant outcome was observed only for the referred comparison, suggesting that this VOC may be a potential biomarker for the papillary subtype. Similar to this result, an *in vitro* study observed decreased levels of 1,3-di-tert-butylbenzene in the extracellular medium of bladder cancer cell lines [289], however, higher levels of this compound were found in non-Hodgkin's lymphoma cell lines [290] and in the blood collected from CRC patients [293] when compared to healthy controls.

The ethylbenzene revealed an interesting outcome since it only appears significantly altered (decreased levels) in two ccRCC cell lines (786-O and Caki-1). Hence, this aromatic hydrocarbon may constitute a potential biomarker for ccRCC, despite being already reported in the urine of bladder cancer patients [294], and in bladder cancer [289] and breast cancer cell lines [295].

The 1,3-dimethyl-benzene presented a decreased level in all cancer cell lines, except Caki-2. However, this aromatic hydrocarbon is not disease specific, having been found decreased in bladder cancer cell lines [289].

The oxidative stress is the main mechanism that affects hydrocarbons emission [296]. These compounds are produced by oxidation of fatty acids by reactive oxygen species (ROS) [297, 298] via lipid peroxidation and their oxidation in cytochrome p450 (CYP450) may be related to the production of alcohols [298].

The alcohol 2-ethyl-1-hexanol was significantly increased in 769-P, Caki-2 and ACHN when compared with HK-2. This compound has been reported as a potential biomarker in several pathologies such as lung [299], prostate [300] and renal cancers [301]. According to previous studies, a higher concentration of this alcohol was found in the exhaled breath of lung cancer patients [302, 303] when compared to healthy individuals. The same behavior was revealed in the analysis of urine from RCC patients [301] and in studies with lung cancer cell lines [296, 304].

The alcohol cyclohexanol presented significant results only for the comparison between Caki-2 and HK-2, with higher levels found in the cancer cell line medium. Thus, cyclohexanol may be a potential specific biomarker for the identification of the papillary subtype. Consistently to our work, this VOC also revealed higher levels in lung cancer cell lines [296]. Besides this

similarity, the opposite was also reported, cyclohexanol was dramatically decreased in lung cancer [290] and bladder cancer cell lines [289].

The high levels of alcohols may be the response of increased oxidative stress and upregulated CYP450 [296]. The alcohols can be further oxidized into aldehydes and carboxylic acids by alcohol dehydrogenase and aldehyde dehydrogenase, respectively [297, 298].

The ketone acetophenone was significantly discriminative only for two comparisons, 769-P *vs* HK-2 and Caki-2 *vs* HK-2, in which it was observed in diminished levels in the culture medium of cancer cell lines. A previous *in vitro* study also found decreased levels of this compound in bladder cancer cell lines [289]. However, higher levels of acetophenone were reported in exhaled breath of lung cancer patients [281, 299, 305] when compared to controls.

The 2,6-di-tert-butylquinone was significantly decreased in the medium of 769-P, Caki-1 and Caki-2 cell lines. This ketone was detected in urine samples from breast cancer patients [306].

Additionally, some unidentified compounds were also significantly altered and relevant to discriminate the two RCC subtypes in study, namely unknowns 1 and 7 present on in papillary RCC and unknowns 4, 5, 6 and 13 present on in clear cell RCC. Thus, it is essential that future investigations focus on the identification of these compounds.

According to the organic classification, the alkene and alcohols were increased in the RCC cell lines, while the alkanes, aromatic hydrocarbons and ketones were decreased.

Observations from the current study are not exact whereby it is necessary further investigations to confirm the importance of the identified VOCs and the exact biochemical pathways that are involved [296].

6.2 Discriminative VCCs in RCC vs normal cell lines

A diversity of VCCs was found to be highly discriminative between the normal and RCC cell lines and between ccRCC and pRCC subtypes. The classes of the identified chemical compounds comprise aldehydes and ketones.

Multivariate statistical analysis unveiled a total of 12 discriminant VCCs, however, only 5 were identified through the database and confirmed with the respective standards, namely acetaldehyde, cyclohexanone, acetone, methylglyoxal and glyoxal.

Reactive carbonyl compounds mediate crucial intracellular signaling pathways that are essential for cell survival [307] and their physiological production by aerobic cells increases under conditions of cell injury [308]. The excessive formation of ROS generates cell damage

and death [309]. Reactive carbonyls such as acetaldehyde, glyoxal and methylglyoxal are very important to analyze since they are the main products of lipid peroxidation, a process that is strongly associated with cancer [310] and mutagenesis [311]. Tumor development induced by carcinogenic aldehydes was demonstrated *in vivo*, showing mutational inactivation of various genes (e.g. tumor suppressor genes) [312].

In the present study, acetaldehyde may be considered the most important compound of this assay, since this VCC shows the most statistically significant result. This aldehyde is present in the composition of the culture medium being consumed by all tested cell lines. However, cancer cells have a higher consumption of acetaldehyde than normal cells. Interestingly, the papillary metastatic cell line (ACHN) has a lower consumption of acetaldehyde when compared to the non-metastatic papillary cell line (Caki-2). The acetaldehyde is a product of physiological cell metabolism [313] and a highly reactive carbonyl compound that forms adducts with proteins, phospholipids, and DNA [314]. This compound is classified in humans as a group 1 carcinogen [315] and inflicts DNA damage and mutations particularly in the form of base damage and DNA-protein crosslinks [316], whereby significantly decreases the activity of the DNA repair enzyme O-6-methylguanine-DNA methyltransferase [317]. This compound affects cell viability, the thiol status [317] and the intracellular Ca^{2+} levels (transient or sustained increases in cytosolic Ca^{2+}) [317, 318]. Thiol depletion in the cellular content of glutathione (GSH) and the inhibition of the DNA repair enzyme increase cells susceptibility to other aldehydes. GSH performs an antioxidant function whereby it plays an important role in several cellular processes such as cell differentiation, proliferation and apoptosis, and its homeostasis disturbance is involved in many human diseases. GSH deficiency increases the susceptibility to oxidative stress implicated in cancer progression [309]. This VCC has been associated with multistep carcinogenesis in cultured human cells studies, in which the exposure to mM concentration levels promotes a wide range of cytopathic effects [317]. Overall, mutagenic and carcinogenic properties of acetaldehyde in several stages in carcinogenesis such as mutations, growth inhibition and accelerated differentiation [319-321], have been found in *in vitro* model systems [312, 322, 323] and in vivo [322].

Acetone was found significantly decreased only for the comparison between 769-P (ccRCC non-metastatic) and HK-2. This VCC is present in the culture medium and is consumed only by the cancer cell line. The acetone is produced by the hepatocytes from excess acetyl-coenzyme A, and other ketone bodies enter the bloodstream to be oxidized via Krebs cycle in peripheral tissue [324]. However, in patients with cancer cells, the metabolism of cells swings to aerobic glycolysis rather than moving to the Krebs cycle, producing the Warburg effect [325].

This compound was also found in the exhaled breath of perioperative lung cancer patients under intensive oral care [326], and it was demonstrated an increased concentration in lung cancer patients when compared to healthy nonsmokers [213]. Contrary to our findings, a previous *in vitro* study has shown significantly higher levels of acetone in the headspace of lung cancer cell lines [292].

Cyclohexanone is present in the culture medium and is consumed by 769-P and Caki-2 cell lines, while the normal cell line (HK-2) releases this ketone. The presence of cyclohexanone was also observed during the comparison of exhaled breath from CRC patients and healthy controls, however, the cancer group showed significantly higher levels of cyclohexanone [287].

Methylglyoxal (MG) is another important compound in this assay, since it was found significantly decreased in the medium of all tested RCC cell lines, except Caki-1, when compared to HK-2 medium. Importantly, this compound is not present in the culture medium and is released by all cancer cell lines. Additionally, the metastatic papillary cell line (ACHN) presents a higher release of methylglyoxal.

Metabolic reprogramming toward aerobic glycolysis in cancer cells, make them accumulate reactive dicarbonyl compounds [327, 328] such as MG. This aldehyde is a highly reactive compound, is produced during glycolysis as a side product [327, 329] and induces the formation of advanced glycation end products (AGEs) [330] that are involved in numerous pathologies including cancer [328, 331, 332]. In all human cells, MG is detoxified by the glyoxalase system, constituted by two enzymes, glyoxalase 1 (GLO1) and glyoxalase 2 (GLO2) [331], which catalyze MG to d-lactate [328, 331]. Decreased expression and enzymatic activity of GLO1 were observed in high stage CRC tumors when compared with low stage disease, suggesting a pro-tumor role for dicarbonyl stress [328]. Furthermore, it was shown that MG modified proteins are more accumulated in tumor protein extracts than in their normal counterparts, indicating that the dicarbonyl stress generated by MG is higher in colon cancer tissue than in normal cells [328]. Regarding the impact of MG-mediated carbonyl stress on tumor progression, it was demonstrated that cancer cells with elevated MG stress induce enhanced growth and metastatic potential *in vivo*. Their study also showed that MG-induced AGEs are commonly present in breast cancer [330]. Another study using an *in vivo* model also demonstrated the pro-tumorigenic and pro-metastatic role of endogenous MG accumulation in breast cancer cells [50]. The analysis of urine samples from RCC patients unveiled significantly higher levels of MG in RCC than in controls, suggesting possible disturbances in RCC biochemical pathways [333]. This result shows that cancer cells release more MG than controls, not being consistent with our findings.

Regarding glyoxal, it was found significantly decreased in 786-O and Caki-2 cell lines. Additionally, the non-metastatic papillary cell line presents a high consumption of glyoxal, while the metastatic one presents a slightly consumption. A study using male outbreed Wistar rats demonstrated that glyoxal presents tumor-promoting activity on rat glandular stomach carcinogenesis [334]. Glyoxal is an endogenously formed compound during normal human metabolism conditions [335]. The principal pathway for glyoxal detoxification is the same mechanism described for MG, the glutathione-dependent glyoxalase system. This mechanism convert glyoxal in glycolate [335]. Like MG, glyoxal is a highly reactive compound that can generate AGEs by binding to proteins [335]. This process is mainly based in glyoxal binding to DNA, promoting the formation of adducts, mutations, chromosomal aberrations, sister chromatid exchanges, and DNA single strand breaks [335]. Another *in vitro* study showed that the aldehyde dehydrogenase is an important detoxification enzyme for glyoxal. In this assay, Min mice were also exposed to different concentrations of glyoxal and the result showed a significant increase in tumor size in the small intestine of the animals, suggesting that this compound promotes tumor growth [335]. According to other experiments in rats, exposure to glyoxal significantly increased the incidence of adenocarcinomas in the glandular stomach, indicating that this VCC may promote tumor activity in the glandular stomach of the animal [334].

The last comparisons of our VCC study, ccRCC *vs* pRCC and ccRCC metastatic *vs* non-metastatic were not responsible for a significant outcome. However, these results are consistent with Schaeffeler et al. (2018), since during the study of metabolic alterations in RCC subtypes predicted through transcriptome analyses, they obtained similar results to clear cell and papillary samples [336]. These two RCC subtypes are developed at the proximal tubule and the presumed cells of origin may be the same. Thus, these findings are consistent with the idea of these two RCC subtypes presenting a similar metabolic profile. Relatively to the comparison ccRCC metastatic vs ccRCC non-metastatic, Schaeffeler et al. (2018) also identified through metabolomic/lipidomic profile of RCC subtypes and metastases that the samples were clustered into subgroups separating non-tumor and chRCC from a group composed by ccRCC, pRCC, and ccRCC-derived metastases [336]. Lastly, these data are in accordance with our study, since it provides evidences that ccRCC-derived metastases and ccRCC primary tumors present the same alterations and consequently similar metabolite levels.

Additionally, some unidentified compounds were also significantly altered and relevant to discriminate the two RCC subtypes in study. Thus, it is essential future investigation developing

the identification of these compounds with the help of standard chemicals. Lastly, all identified aldehydes and ketones were decreased in the cancer cell lines medium.

7 Conclusion and future perspectives

Over the last years, due to non-invasive sampling approach, the interest on VOC and VCC profiles has increased, as well as their association with several pathophysiological states.

In this project we attempted to look for the presence of volatile compounds which may help to discriminate normal from cancer cells. This knowledge will be essential to introduce VOCs and VCCs into routine screening procedures.

Our study provided new insights about the characteristic VOCs and VCCs phenotype in ccRCC and pRCC subtypes, where a comprehensive volatile metabolomic signature of RCC cell lines that covered the profile of VCCs was obtained and reported for the first time.

The VOCs panel is composed by 23 metabolites (1,3-di-tert-butylbenzene, 2,4-dimethyl-1-heptene, 2,6-di-tert-butylquinone, 2-ethyl-1-hexanol, acetophenone, cyclohexanol, decane, dodecane, ethylbenzene, 1,3-dimethyl-benzene and 13 unknowns), and the VCCs panel is composed by 12 metabolites (acetaldehyde, acetone, cyclohexanone, glyoxal, methylglyoxal and 7 unknowns). These results obtained from the *in vitro* prediction demonstrated the potential of VOCs and VCCs in discriminating RCC cell lines from normal one using HS-SPME/GC-MS technique. For instance, merely a few compounds found are unique in a subtype.

The obtained data showed that the levels of these volatile compounds are distinct when comparing the exometabolome of RCC and normal cell lines, and that these differences allow the determination of VOC and VCC biomarkers panels for the early diagnosis of RCC.

Hereafter, since the unknown metabolites contributed for the discrimination of RCC subtypes and cancer *vs* normal cell lines it would be important to identify those compounds using standards, as well as to explore the biological and pathological significance of the identified discriminatory compounds. These approaches might lead to the study of urine samples, which definitely is the best biological sample for RCC biomarkers research. Obviously, since VOCs and VCCs metabolism is not yet completely understood, these outcomes cannot provide comprehensive conclusions. Furthermore, isolated *in vitro* cancer cells do not reproduce the exact environment conditions of tumors in humans, whereby the achieved data might diverge from the data acquired in biological samples. Evidently, it is mandatory *in vivo* translation through analysis of human samples, in order to promote the complete validation of the obtained RCC volatilome. In conclusion, our results are very promising and demonstrate the potential of VOCs and VCCs as diagnosis and prognosis biomarkers.

8 References

1. Gold, P.J., A. Fefer, and J.A. Thompson, *Paraneoplastic manifestations of renal cell carcinoma.* Semin Urol Oncol, 1996. **14**(4): p. 216-22.
2. Palapattu, G.S., B. Kristo, and J. Rajfer, *Paraneoplastic syndromes in urologic malignancy: the many faces of renal cell carcinoma.* Reviews in urology, 2002. **4**(4): p. 163-170.
3. Skinner, D.G., et al., *Diagnosis and management of renal cell carcinoma. A clinical and pathologic study of 309 cases.* Cancer, 1971. **28**(5): p. 1165-77.
4. Tsui, K.H., et al., *Renal cell carcinoma: Prognostic significance of incidentally detected tumors.* Journal of Urology, 2000. **163**(2): p. 426-430.
5. Weiss, R.H. and P.Y. Lin, *Kidney cancer: identification of novel targets for therapy.* Kidney Int, 2006. **69**(2): p. 224-32.
6. Chow, W.H., et al., *Rising incidence of renal cell cancer in the United States.* JAMA, 1999. **281**(17): p. 1628-31.
7. Moretto, P., et al., *Kidney cancer survivorship survey of urologists and survivors: The gap in perceptions of care, but agreement on needs.* Canadian Urological Association journal = Journal de l'Association des urologues du Canada, 2014. **8**(5-6): p. 190-194.
8. Bostwick, D.G. and J.N. Eble, *Diagnosis and classification of renal cell carcinoma.* Urologic Clinics of North America, 1999. **26**(3): p. 627-635.
9. Kovacs, G., et al., *The Heidelberg classification of renal cell tumours.* J Pathol, 1997. **183**(2): p. 131-3.
10. Linehan, W.M., M.M. Walther, and B. Zbar, *The genetic basis of cancer of the kidney.* J Urol, 2003. **170**(6 Pt 1): p. 2163-72.
11. Amin, M.B., et al., *Prognostic impact of histologic subtyping of adult renal epithelial neoplasms: an experience of 405 cases.* Am J Surg Pathol, 2002. **26**(3): p. 281-91.
12. *<2015_Metabolomic profile of glycolysis and the pentose phosphate.pdf>.*
13. Presti, J.C., Jr., et al., *Comparative genomic hybridization for genetic analysis of renal oncocytomas.* Genes Chromosomes Cancer, 1996. **17**(4): p. 199-204.

14. Walter, T.A., et al., *Translocation t(9;11)(p23;q12): a primary chromosomal change in renal oncocytoma.* The Journal of urology, 1989. **142**(1): p. 117-119.
15. Cairns, P., *Renal cell carcinoma.* Cancer Biomark, 2010. **9**(1-6): p. 461-73.
16. Konnak, J.W. and H.B. Grossman, *Renal-Cell Carcinoma as an Incidental Finding.* Journal of Urology, 1985. **134**(6): p. 1094-1096.
17. Volpe, A., et al., *The natural history of incidentally detected small renal masses.* Cancer, 2004. **100**(4): p. 738-45.
18. Pahernik, S., et al., *Small renal tumors: correlation of clinical and pathological features with tumor size.* J Urol, 2007. **178**(2): p. 414-7; discussion 416-7.
19. Frank, I., et al., *Solid renal tumors: an analysis of pathological features related to tumor size.* J Urol, 2003. **170**(6 Pt 1): p. 2217-20.
20. Sánchez-Martín, F.M., et al., *Small renal masses: incidental diagnosis, clinical symptoms, and prognostic factors.* Advances in urology, 2008. **2008**: p. 310694-310694.
21. Pantuck, A.J., A. Zisman, and A.S. Belldegrun, *The changing natural history of renal cell carcinoma.* J Urol, 2001. **166**(5): p. 1611-23.
22. Volpe, A., *The role of surveillance in the management of small renal masses.* ScientificWorldJournal, 2007. **7**: p. 860-8.
23. Duchene, D.A., et al., *Histopathology of surgically managed renal tumors: analysis of a contemporary series.* Urology, 2003. **62**(5): p. 827-30.
24. Porena, M., et al., *Incidentally detected renal cell carcinoma: role of ultrasonography.* J Clin Ultrasound, 1992. **20**(6): p. 395-400.
25. Jayson, M. and H. Sanders, *Increased incidence of serendipitously discovered renal cell carcinoma.* Urology, 1998. **51**(2): p. 203-5.
26. Sugimura, K., et al., *Microscopic hematuria as a screening marker for urinary tract malignancies.* Int J Urol, 2001. **8**(1): p. 1-5.
27. Hock, L.M., J. Lynch, and K.C. Balaji, *Increasing incidence of all stages of kidney cancer in the last 2 decades in the United States: An analysis of Surveillance, Epidemiology and End Results program data.* Journal of Urology, 2002. **167**(1): p. 57-60.

28. Sims, J.N., et al., *Racial Disparities and Preventive Measures to Renal Cell Carcinoma.* Int J Environ Res Public Health, 2018. **15**(6).
29. Verhoest, G., et al., *Relationship between Age at Diagnosis and Clinicopathologic Features of Renal Cell Carcinoma.* European Urology, 2007. **51**(5): p. 1298-1305.
30. Parker, A.S., et al., *Smoking cessation and renal cell carcinoma.* Ann Epidemiol, 2003. **13**(4): p. 245-51.
31. Cumberbatch, M.G., et al., *The Role of Tobacco Smoke in Bladder and Kidney Carcinogenesis: A Comparison of Exposures and Meta-analysis of Incidence and Mortality Risks.* Eur Urol, 2016. **70**(3): p. 458-66.
32. Tsivian, M., et al., *Cigarette smoking is associated with advanced renal cell carcinoma.* J Clin Oncol, 2011. **29**(15): p. 2027-31.
33. Bellocco, R., et al., *Alcohol drinking and risk of renal cell carcinoma: results of a meta-analysis.* Ann Oncol, 2012. **23**(9): p. 2235-44.
34. Song, D.Y., et al., *Alcohol intake and renal cell cancer risk: a meta-analysis.* Br J Cancer, 2012. **106**(11): p. 1881-90.
35. Karami, S., S.E. Daugherty, and M.P. Purdue, *A prospective study of alcohol consumption and renal cell carcinoma risk.* Int J Cancer, 2015. **137**(1): p. 238-42.
36. Bergström, A., et al., *Obesity and renal cell cancer--a quantitative review.* Br J Cancer, 2001. **85**(7): p. 984-90.
37. Macleod, L.C., et al., *Risk factors for renal cell carcinoma in the VITAL study.* J Urol, 2013. **190**(5): p. 1657-61.
38. Liao, L.M., et al., *Circulating levels of obesity-related markers and risk of renal cell carcinoma in the PLCO cancer screening trial.* Cancer Causes Control, 2017. **28**(7): p. 801-807.
39. Gild, P., B. Ehdaie, and L.A. Kluth, *Effect of obesity on bladder cancer and renal cell carcinoma incidence and survival.* Curr Opin Urol, 2017. **27**(5): p. 409-414.
40. Behrens, G. and M.F. Leitzmann, *The association between physical activity and renal cancer: systematic review and meta-analysis.* Br J Cancer, 2013. **108**(4): p. 798-811.

41. Yang, X.F., et al., *Twist2 and CD24 expression alters renal microenvironment in obesity associated kidney cancer.* Eur Rev Med Pharmacol Sci, 2018. **22**(2): p. 358-364.
42. Hidayat, K., et al., *Blood pressure and kidney cancer risk: meta-analysis of prospective studies.* J Hypertens, 2017. **35**(7): p. 1333-1344.
43. Mellemgaard, A., et al., *Risk factors for renal cell carcinoma in Denmark: role of medication and medical history.* Int J Epidemiol, 1994. **23**(5): p. 923-30.
44. Shapiro, J.A., et al., *Hypertension, antihypertensive medication use, and risk of renal cell carcinoma.* Am J Epidemiol, 1999. **149**(6): p. 521-30.
45. Grossman, E., F.H. Messerli, and U. Goldbourt, *Does diuretic therapy increase the risk of renal cell carcinoma?* Am J Cardiol, 1999. **83**(7): p. 1090-3.
46. Chien, C.C., et al., *Epidemiology of cancer in end-stage renal disease dialysis patients: a national cohort study in Taiwan.* J Cancer, 2017. **8**(1): p. 9-18.
47. Mazzucotelli, V., et al., *De novo cancer in patients on dialysis and after renal transplantation: north-western Italy, 1997-2012.* J Nephrol, 2017. **30**(6): p. 851-857.
48. Song, J.k., et al., *Association between cadmium exposure and renal cancer risk: a meta-analysis of observational studies.* Scientific Reports, 2015. **5**: p. 17976.
49. Desimone, M.C., W.K. Rathmell, and D.W. Threadgill, *Pleiotropic effects of the trichloroethylene-associated P81S VHL mutation on metabolism, apoptosis, and ATM-mediated DNA damage response.* J Natl Cancer Inst, 2013. **105**(18): p. 1355-64.
50. Moore, L.E., et al., *Occupational trichloroethylene exposure and renal carcinoma risk: evidence of genetic susceptibility by reductive metabolism gene variants.* Cancer Res, 2010. **70**(16): p. 6527-36.
51. Clague, J., et al., *Family history and risk of renal cell carcinoma: results from a case-control study and systematic meta-analysis.* Cancer Epidemiol Biomarkers Prev, 2009. **18**(3): p. 801-7.
52. Linehan, W.M., et al., *Genetic basis of cancer of the kidney: Disease specific approaches to therapy.* Clinical Cancer Research, 2004. **10**(18): p. 6282s-6289s.

53. Aron, M., et al., *Impact of Gender in Renal Cell Carcinoma: An Analysis of the SEER Database.* European Urology, 2008. **54**(1): p. 133-142.
54. Lucca, I., et al., *Gender differences in incidence and outcomes of urothelial and kidney cancer.* Nat Rev Urol, 2015. **12**(10): p. 585-92.
55. Batai, K., et al., *Racial and Ethnic Disparities in Renal Cell Carcinoma: An Analysis of Clinical Characteristics.* Clin Genitourin Cancer, 2019. **17**(1): p. e195-e202.
56. Olshan, A.F., et al., *Racial difference in histologic subtype of renal cell carcinoma.* Cancer Med, 2013. **2**(5): p. 744-9.
57. Rashidkhani, B., P. Lindblad, and A. Wolk, *Fruits, vegetables and risk of renal cell carcinoma: a prospective study of Swedish women.* Int J Cancer, 2005. **113**(3): p. 451-5.
58. Zhang, S., et al., *Consumption of fruits and vegetables and risk of renal cell carcinoma: a meta-analysis of observational studies.* Oncotarget, 2017. **8**(17): p. 27892-27903.
59. Chen, Y., et al., *Vitamin D receptor suppresses proliferation and metastasis in renal cell carcinoma cell lines via regulating the expression of the epithelial Ca2+ channel TRPV5.* PloS one, 2018. **13**(4): p. e0195844-e0195844.
60. Krajewski, W., et al., *Vitamin D and urological cancers.* Central European journal of urology, 2016. **69**(2): p. 139-147.
61. Kim, J.K., et al., *Differentiation of Subtypes of Renal Cell Carcinoma on Helical CT Scans.* American Journal of Roentgenology, 2002. **178**(6): p. 1499-1506.
62. Cheville, J.C., et al., *Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma.* Am J Surg Pathol, 2003. **27**(5): p. 612-24.
63. Oliva, M.R., et al., *Renal cell carcinoma: t1 and t2 signal intensity characteristics of papillary and clear cell types correlated with pathology.* AJR Am J Roentgenol, 2009. **192**(6): p. 1524-30.
64. Zisman, A., et al., *Improved prognostication of renal cell carcinoma using an integrated staging system.* J Clin Oncol, 2001. **19**(6): p. 1649-57.

65. Whaley, J.M., et al., *Germ-line mutations in the von Hippel-Lindau tumor-suppressor gene are similar to somatic von Hippel-Lindau aberrations in sporadic renal cell carcinoma.* American journal of human genetics, 1994. **55**(6): p. 1092-1102.
66. Thoenes, W., S. Storkel, and H.J. Rumpelt, *Human chromophobe cell renal carcinoma.* Virchows Arch B Cell Pathol Incl Mol Pathol, 1985. **48**(3): p. 207-17.
67. Muglia, V.F. and A. Prando, *Renal cell carcinoma: histological classification and correlation with imaging findings.* Radiologia brasileira, 2015. **48**(3): p. 166-174.
68. Linehan, W.M., M.I. Lerman, and B. Zbar, *Identification of the von Hippel-Lindau (VHL) gene. Its role in renal cancer.* Jama, 1995. **273**(7): p. 564-70.
69. Knudson, A.G., Jr., *Mutation and cancer: statistical study of retinoblastoma.* Proc Natl Acad Sci U S A, 1971. **68**(4): p. 820-3.
70. Binkovitz, L.A., C.D. Johnson, and D.H. Stephens, *Islet cell tumors in von Hippel-Lindau disease: increased prevalence and relationship to the multiple endocrine neoplasias.* AJR Am J Roentgenol, 1990. **155**(3): p. 501-5.
71. Melmon, K.L. and S.W. Rosen, *Linda's Disease. Review of the Literature and Study of a Large Kindred.* Am J Med, 1964. **36**: p. 595-617.
72. Huson, S.M., et al., *Cerebellar Haemangioblastoma and von Hippel-Lindau disease.* Brain, 1986. **109**(6): p. 1297-1310.
73. Horton, W.A., V. Wong, and R. Eldridge, *Von Hippel-Lindau disease: clinical and pathological manifestations in nine families with 50 affected members.* Arch Intern Med, 1976. **136**(7): p. 769-77.
74. Lamiell, J.M., F.G. Salazar, and Y.E. Hsia, *von Hippel-Lindau disease affecting 43 members of a single kindred.* Medicine (Baltimore), 1989. **68**(1): p. 1-29.
75. Maher, E.R., J.R. Yates, and M.A. Ferguson-Smith, *Statistical analysis of the two stage mutation model in von Hippel-Lindau disease, and in sporadic cerebellar haemangioblastoma and renal cell carcinoma.* J Med Genet, 1990. **27**(5): p. 311-4.
76. Seizinger, B.R., et al., *Von Hippel-Lindau disease maps to the region of chromosome 3 associated with renal cell carcinoma.* Nature, 1988. **332**(6161): p. 268-9.

77. Tory, K., et al., *Specific Genetic Change in Tumors Associated With von Hippel-Lindau Disease.* JNCI: Journal of the National Cancer Institute, 1989. **81**(14): p. 1097-1101.
78. Lubensky, I.A., et al., *Allelic deletions of the VHL gene detected in multiple microscopic clear cell renal lesions in von Hippel-Lindau disease patients.* Am J Pathol, 1996. **149**(6): p. 2089-94.
79. Gnarra, J.R., et al., *Mutations of the VHL tumour suppressor gene in renal carcinoma.* Nat Genet, 1994. **7**(1): p. 85-90.
80. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
81. Latif, F., et al., *Identification of the von Hippel-Lindau disease tumor suppressor gene.* Science, 1993. **260**(5112): p. 1317-20.
82. Shuin, T., et al., *Frequent Somatic Mutations and Loss of Heterozygosity of the von Hippel-Lindau Tumor Suppressor Gene in Primary Human Renal Cell Carcinomas.* Cancer Research, 1994. **54**(11): p. 2852-2855.
83. Foster, K., et al., *Somatic mutations of the von Hippel-Lindau disease tumour suppressor gene in non-familial clear cell renal carcinoma.* Hum Mol Genet, 1994. **3**(12): p. 2169-73.
84. Gordan, J.D., et al., *HIF-alpha effects on c-Myc distinguish two subtypes of sporadic VHL-deficient clear cell renal carcinoma.* Cancer Cell, 2008. **14**(6): p. 435-46.
85. Kaelin, W.G., Jr., *The von Hippel-Lindau tumor suppressor protein and clear cell renal carcinoma.* Clin Cancer Res, 2007. **13**(2 Pt 2): p. 680s-684s.
86. Tyers, M. and R. Rottapel, *VHL: a very hip ligase.* Proc Natl Acad Sci U S A, 1999. **96**(22): p. 12230-2.
87. Ohh, M., et al., *Ubiquitination of hypoxia-inducible factor requires direct binding to the beta-domain of the von Hippel-Lindau protein.* Nat Cell Biol, 2000. **2**(7): p. 423-7.
88. Schmidt, L., et al., *Two North American families with hereditary papillary renal carcinoma and identical novel mutations in the MET proto-oncogene.* Cancer Res, 1998. **58**(8): p. 1719-22.

89. Schmidt, L., et al., *Germline and somatic mutations in the tyrosine kinase domain of the MET proto-oncogene in papillary renal carcinomas.* Nat Genet, 1997. **16**(1): p. 68-73.
90. Lubensky, I.A., et al., *Hereditary and sporadic papillary renal carcinomas with c-met mutations share a distinct morphological phenotype.* Am J Pathol, 1999. **155**(2): p. 517-26.
91. Schmidt, L., et al., *Novel mutations of the MET proto-oncogene in papillary renal carcinomas.* Oncogene, 1999. **18**(14): p. 2343-50.
92. Zhuang, Z., et al., *Trisomy 7-harbouring non-random duplication of the mutant MET allele in hereditary papillary renal carcinomas.* Nat Genet, 1998. **20**(1): p. 66-9.
93. Nickerson, M.L., et al., *Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt-Hogg-Dube syndrome.* Cancer Cell, 2002. **2**(2): p. 157-64.
94. Khoo, S.K., et al., *Inactivation of BHD in sporadic renal tumors.* Cancer Res, 2003. **63**(15): p. 4583-7.
95. Pavlovich, C.P., et al., *Renal tumors in the Birt-Hogg-Dube syndrome.* Am J Surg Pathol, 2002. **26**(12): p. 1542-52.
96. Bethesda, M.N.C.I., *PDQ® Adult Treatment Editorial Board. PDQ Renal Cell Cancer Treatment. Updated 02/06/2019. Available at: https://www.cancer.gov/types/kidney/hp/kidney-treatment-pdq. Accessed 05/20/2019*
97. Homma, Y., et al., *Increased incidental detection and reduced mortality in renal cancer--recent retrospective analysis at eight institutions.* Int J Urol, 1995. **2**(2): p. 77-80.
98. Bassil, B., D.E. Dosoretz, and G.R. Prout, Jr., *Validation of the tumor, nodes and metastasis classification of renal cell carcinoma.* J Urol, 1985. **134**(3): p. 450-4.
99. Golimbu, M., et al., *Renal cell carcinoma: survival and prognostic factors.* Urology, 1986. **27**(4): p. 291-301.
100. Robson, C.J., B.M. Churchill, and W. Anderson, *The results of radical nephrectomy for renal cell carcinoma.* J Urol, 1969. **101**(3): p. 297-301.

101. Maldazys, J.D. and J.B. deKernion, *Prognostic factors in metastatic renal carcinoma.* J Urol, 1986. **136**(2): p. 376-9.
102. Medeiros, L.J., A.B. Gelb, and L.M. Weiss, *Renal cell carcinoma. Prognostic significance of morphologic parameters in 121 cases.* Cancer, 1988. **61**(8): p. 1639-51.
103. Dall'Oglio, M.F., et al., *Impact of clinicopathological parameters in patients treated for renal cell carcinoma.* J Urol, 2007. **177**(5): p. 1687-91.
104. Lee, C.T., et al., *Surgical management of renal tumors 4 cm. or less in a contemporary cohort.* J Urol, 2000. **163**(3): p. 730-6.
105. Amin, M.B., Edge, S., Greene, F., Byrd, D.R., Brookland, R.K., Washington, M.K., Gershenwald, J.E., Compton, C.C., Hess, K.R., Sullivan, D.C., Jessup, J.M., Brierley, J.D., Gaspar, L.E., Schilsky, R.L., Balch, C.M., Winchester, D.P., Asare, E.A., Madera, M., Gress, D.M., Meyer, L.R., *AJCC Cancer Staging Manual*. 8th ed. 2017, New York.
106. Escudier, B., et al., *Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* Ann Oncol, 2019.
107. Delahunt, B., et al., *Grading of renal cell carcinoma.* Histopathology, 2019. **74**(1): p. 4-17.
108. Capogrosso, P., et al., *Follow-up After Treatment for Renal Cell Carcinoma: The Evidence Beyond the Guidelines.* European Urology Focus, 2016. **1**(3): p. 272-281.
109. Sene, A.P., et al., *Renal carcinoma in patients undergoing nephrectomy: analysis of survival and prognostic factors.* Br J Urol, 1992. **70**(2): p. 125-34.
110. Snow, R.M. and P.F. Schellhammer, *Spontaneous regression of metastatic renal cell carcinoma.* Urology, 1982. **20**(2): p. 177-181.
111. Lekanidi, K., et al., *Spontaneous regression of metastatic renal cell carcinoma: case report.* Journal of medical case reports, 2007. **1**: p. 89-89.
112. Amitabh-Ray and S. Ghosh, *Long natural history and spontaneous regression: A rare outcome of metastatic renal cell carcinoma.* Clinical Cancer Investigation Journal, 2014. **3**(5): p. 447-449.

113. Abdel-Rahman, O., *Impact of histological subtype on outcomes of renal cell carcinoma patients.* Journal of Drug Assessment, 2018. **7**(1): p. 14-20.
114. Feltrin, L.T., et al., *Sinais de hipervascularização em imagens de ressonância magnética em metástases ósseas de carcinoma de células renais.* Radiologia Brasileira, 2009. **42**: p. 155-157.
115. Keegan, K.A., et al., *Histopathology of surgically treated renal cell carcinoma: survival differences by subtype and stage.* J Urol, 2012. **188**(2): p. 391-7.
116. Lopez-Beltran, A., et al., *2004 WHO classification of the renal tumors of the adults.* Eur Urol, 2006. **49**(5): p. 798-805.
117. Srigley, J.R., et al., *The International Society of Urological Pathology (ISUP) Vancouver Classification of Renal Neoplasia.* Am J Surg Pathol, 2013. **37**(10): p. 1469-89.
118. Mancilla-Jimenez, R., R.J. Stanley, and R.A. Blath, *Papillary renal cell carcinoma: a clinical, radiologic, and pathologic study of 34 cases.* Cancer, 1976. **38**(6): p. 2469-80.
119. Beck, S.D., et al., *Effect of papillary and chromophobe cell type on disease-free survival after nephrectomy for renal cell carcinoma.* Ann Surg Oncol, 2004. **11**(1): p. 71-7.
120. Lopez-Beltran, A., et al., *2009 update on the classification of renal epithelial tumors in adults.* Int J Urol, 2009. **16**(5): p. 432-43.
121. Dabestani, S., et al., *Long-term Outcomes of Follow-up for Initially Localised Clear Cell Renal Cell Carcinoma: RECUR Database Analysis.* Eur Urol Focus, 2018.
122. Flanigan, R.C., et al., *Nephrectomy followed by interferon alfa-2b compared with interferon alfa-2b alone for metastatic renal-cell cancer.* N Engl J Med, 2001. **345**(23): p. 1655-9.
123. Mathieu, R., et al., *Nephrectomy improves overall survival in patients with metastatic renal cell carcinoma in cases of favorable MSKCC or ECOG prognostic features.* Urol Oncol, 2015. **33**(8): p. 339.e9-15.
124. Mejean, A., et al., *Sunitinib Alone or after Nephrectomy in Metastatic Renal-Cell Carcinoma.* N Engl J Med, 2018. **379**(5): p. 417-427.

125. Mickisch, G.H., et al., *Radical nephrectomy plus interferon-alfa-based immunotherapy compared with interferon alfa alone in metastatic renal-cell carcinoma: a randomised trial.* Lancet, 2001. **358**(9286): p. 966-70.
126. Choueiri, T.K., et al., *Cabozantinib Versus Sunitinib As Initial Targeted Therapy for Patients With Metastatic Renal Cell Carcinoma of Poor or Intermediate Risk: The Alliance A031203 CABOSUN Trial.* J Clin Oncol, 2017. **35**(6): p. 591-597.
127. Choueiri, T.K., et al., *Cabozantinib versus sunitinib as initial therapy for metastatic renal cell carcinoma of intermediate or poor risk (Alliance A031203 CABOSUN randomised trial): Progression-free survival by independent review and overall survival update.* Eur J Cancer, 2018. **94**: p. 115-125.
128. Escudier, B., et al., *Randomized, controlled, double-blind, cross-over trial assessing treatment preference for pazopanib versus sunitinib in patients with metastatic renal cell carcinoma: PISCES Study.* J Clin Oncol, 2014. **32**(14): p. 1412-8.
129. Motzer, R.J., et al., *Pazopanib versus sunitinib in metastatic renal-cell carcinoma.* N Engl J Med, 2013. **369**(8): p. 722-31.
130. Sternberg, C.N., et al., *Pazopanib in locally advanced or metastatic renal cell carcinoma: results of a randomized phase III trial.* J Clin Oncol, 2010. **28**(6): p. 1061-8.
131. Coppin, C., et al., *Immunotherapy for advanced renal cell cancer.* Cochrane Database Syst Rev, 2005(1): p. Cd001425.
132. Negrier, S., et al., *Medroxyprogesterone, interferon alfa-2a, interleukin 2, or combination of both cytokines in patients with metastatic renal carcinoma of intermediate prognosis: results of a randomized controlled trial.* Cancer, 2007. **110**(11): p. 2468-77.
133. Hudes, G., et al., *Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma.* N Engl J Med, 2007. **356**(22): p. 2271-81.
134. Escudier, B., et al., *Phase III trial of bevacizumab plus interferon alfa-2a in patients with metastatic renal cell carcinoma (AVOREN): final analysis of overall survival.* J Clin Oncol, 2010. **28**(13): p. 2144-50.

135. Escudier, B., et al., *Bevacizumab plus interferon alfa-2a for treatment of metastatic renal cell carcinoma: a randomised, double-blind phase III trial.* Lancet, 2007. **370**(9605): p. 2103-11.
136. Rini, B.I., et al., *Bevacizumab plus interferon alfa compared with interferon alfa monotherapy in patients with metastatic renal cell carcinoma: CALGB 90206.* J Clin Oncol, 2008. **26**(33): p. 5422-8.
137. Yang, J.C., et al., *A randomized trial of bevacizumab, an anti-vascular endothelial growth factor antibody, for metastatic renal cancer.* N Engl J Med, 2003. **349**(5): p. 427-34.
138. Motzer, R.J., et al., *Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma.* J Clin Oncol, 2009. **27**(22): p. 3584-90.
139. Motzer, R.J., et al., *Sunitinib versus interferon alfa in metastatic renal-cell carcinoma.* N Engl J Med, 2007. **356**(2): p. 115-24.
140. McDermott, D.F., et al., *Randomized phase III trial of high-dose interleukin-2 versus subcutaneous interleukin-2 and interferon in patients with metastatic renal cell carcinoma.* J Clin Oncol, 2005. **23**(1): p. 133-41.
141. Yang, J.C., et al., *Randomized comparison of high-dose and low-dose intravenous interleukin-2 for the therapy of metastatic renal cell carcinoma: an interim report.* J Clin Oncol, 1994. **12**(8): p. 1572-6.
142. Motzer, R.J., et al., *Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma.* N Engl J Med, 2018. **378**(14): p. 1277-1290.
143. Motzer, R.J., et al., *Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial.* Lancet, 2008. **372**(9637): p. 449-56.
144. Motzer, R.J., et al., *Axitinib versus sorafenib as second-line treatment for advanced renal cell carcinoma: overall survival analysis and updated results from a randomised phase 3 trial.* Lancet Oncol, 2013. **14**(6): p. 552-62.
145. Choueiri, T.K., et al., *Cabozantinib versus Everolimus in Advanced Renal-Cell Carcinoma.* N Engl J Med, 2015. **373**(19): p. 1814-23.

146. Escudier, B., et al., *Randomized phase II trial of first-line treatment with sorafenib versus interferon Alfa-2a in patients with metastatic renal cell carcinoma.* J Clin Oncol, 2009. **27**(8): p. 1280-9.
147. Motzer, R.J., et al., *Nivolumab for Metastatic Renal Cell Carcinoma: Results of a Randomized Phase II Trial.* J Clin Oncol, 2015. **33**(13): p. 1430-7.
148. Motzer, R.J., et al., *Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma.* N Engl J Med, 2015. **373**(19): p. 1803-13.
149. Weiss, G.R., et al., *A randomized phase II trial of continuous infusion interleukin-2 or bolus injection interleukin-2 plus lymphokine-activated killer cells for advanced renal cell carcinoma.* J Clin Oncol, 1992. **10**(2): p. 275-81.
150. Atkins, M.B., et al., *Randomized phase II trial of high-dose interleukin-2 either alone or in combination with interferon alfa-2b in advanced renal cell carcinoma.* J Clin Oncol, 1993. **11**(4): p. 661-70.
151. Sleijfer, D.T., et al., *Phase II study of subcutaneous interleukin-2 in unselected patients with advanced renal cell cancer on an outpatient basis.* J Clin Oncol, 1992. **10**(7): p. 1119-23.
152. Fisher, R.I., et al., *Metastatic renal cancer treated with interleukin-2 and lymphokine-activated killer cells. A phase II clinical trial.* Ann Intern Med, 1988. **108**(4): p. 518-23.
153. Fyfe, G., et al., *Results of treatment of 255 patients with metastatic renal cell carcinoma who received high-dose recombinant interleukin-2 therapy.* J Clin Oncol, 1995. **13**(3): p. 688-96.
154. Choueiri, T.K., et al., *Cabozantinib versus everolimus in advanced renal cell carcinoma (METEOR): final results from a randomised, open-label, phase 3 trial.* Lancet Oncol, 2016. **17**(7): p. 917-927.
155. Motzer, R.J., et al., *Long-term follow-up of overall survival for cabozantinib versus everolimus in advanced renal cell carcinoma.* Br J Cancer, 2018. **118**(9): p. 1176-1178.
156. Rini, B.I., et al., *Comparative effectiveness of axitinib versus sorafenib in advanced renal cell carcinoma (AXIS): a randomised phase 3 trial.* Lancet, 2011. **378**(9807): p. 1931-9.

157. deKernion, J.B. and D. Berry, *The diagnosis and treatment of renal cell carcinoma.* Cancer, 1980. **45**(7 Suppl): p. 1947-56.

158. Novick, A.C., et al., *Conservative surgery for renal cell carcinoma: a single-center experience with 100 patients.* J Urol, 1989. **141**(4): p. 835-9.

159. Swanson, D.A., S. Wallace, and D.E. Johnson, *The role of embolization and nephrectomy in the treatment of metastatic renal carcinoma.* Urol Clin North Am, 1980. **7**(3): p. 719-30.

160. Hatcher, P.A., et al., *Surgical management and prognosis of renal cell carcinoma invading the vena cava.* J Urol, 1991. **145**(1): p. 20-3; discussion 23-4.

161. Oliver, S.G., *Functional genomics: lessons from yeast.* Philos Trans R Soc Lond B Biol Sci, 2002. **357**(1417): p. 17-23.

162. Griffiths, W.J., et al., *Targeted metabolomics for biomarker discovery.* Angew Chem Int Ed Engl, 2010. **49**(32): p. 5426-45.

163. Ganti, S. and R.H. Weiss, *Urine metabolomics for kidney cancer detection and biomarker discovery.* Urol Oncol, 2011. **29**(5): p. 551-7.

164. Nicholson, J.K. and I.D. Wilson, *Opinion: understanding 'global' systems biology: metabonomics and the continuum of metabolism.* Nat Rev Drug Discov, 2003. **2**(8): p. 668-76.

165. Nicholson, J.K., J.C. Lindon, and E. Holmes, *'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data.* Xenobiotica, 1999. **29**(11): p. 1181-9.

166. Tweeddale, H., L. Notley-McRobb, and T. Ferenci, *Effect of slow growth on metabolism of Escherichia coli, as revealed by global metabolite pool ("Metabolome") analysis.* Journal of Bacteriology, 1998. **180**(19): p. 5109-5116.

167. ter Kuile, B.H. and H.V. Westerhoff, *Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway.* FEBS Lett, 2001. **500**(3): p. 169-71.

168. Griffin, J.L. and J.P. Shockcor, *Metabolic profiles of cancer cells.* Nat Rev Cancer, 2004. **4**(7): p. 551-61.

169. Mendes, P., D.B. Kell, and H.V. Westerhoff, *Why and when channelling can decrease pool size at constant net flux in a simple dynamic channel.* Biochimica Et Biophysica Acta-General Subjects, 1996. **1289**(2): p. 175-186.
170. Kell, D.B. and H.V. Westerhoff, *Towards a Rational Approach to the Optimization of Flux in Microbial Biotransformations.* Trends in Biotechnology, 1986. **4**(6): p. 137-142.
171. Fiehn, O., *Metabolomics – the link between genotypes and phenotypes.* Plant Molecular Biology, 2002. **48**(1): p. 155-171.
172. Fiehn, O., *Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks.* Comp Funct Genomics, 2001. **2**(3): p. 155-68.
173. Nicholson, J.K., et al., *Metabonomics: a platform for studying drug toxicity and gene function.* Nat Rev Drug Discov, 2002. **1**(2): p. 153-61.
174. Schmidt, C., *Metabolomics Takes Its Place as Latest Up-and-Coming "Omic" Science.* JNCI: Journal of the National Cancer Institute, 2004. **96**(10): p. 732-734.
175. Kind, T., et al., *A comprehensive urinary metabolomic approach for identifying kidney cancer.* Analytical Biochemistry, 2007. **363**(2): p. 185-195.
176. Katt, M.E., et al., *In Vitro Tumor Models: Advantages, Disadvantages, Variables, and Selecting the Right Platform.* Frontiers in Bioengineering and Biotechnology, 2016. **4**(12).
177. Louzada, S., F. Adega, and R. Chaves, *Defining the sister rat mammary tumor cell lines HH-16 cl.2/1 and HH-16.cl.4 as an in vitro cell model for Erbb2.* PLoS One, 2012. **7**(1): p. e29923.
178. Vargo-Gogola, T. and J.M. Rosen, *Modelling breast cancer: one size does not fit all.* Nat Rev Cancer, 2007. **7**(9): p. 659-72.
179. van Staveren, W.C., et al., *Human cancer cell lines: Experimental models for cancer cells in situ? For cancer stem cells?* Biochim Biophys Acta, 2009. **1795**(2): p. 92-103.

180. Kao, J., et al., *Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery.* PLoS One, 2009. **4**(7): p. e6146.
181. Gazdar, A.F., et al., *Lung cancer cell lines as tools for biomedical discovery and research.* J Natl Cancer Inst, 2010. **102**(17): p. 1310-21.
182. Ruhe, J.E., et al., *Genetic alterations in the tyrosine kinase transcriptome of human cancer cell lines.* Cancer Res, 2007. **67**(23): p. 11368-76.
183. Leon, Z., et al., *Mammalian cell metabolomics: experimental design and sample preparation.* Electrophoresis, 2013. **34**(19): p. 2762-75.
184. Burdall, S.E., et al., *Breast cancer cell lines: friend or foe?* Breast Cancer Res, 2003. **5**(2): p. 89-95.
185. Namekawa, T., et al., *Application of Prostate Cancer Models for Preclinical Study: Advantages and Limitations of Cell Lines, Patient-Derived Xenografts, and Three-Dimensional Culture of Patient-Derived Cells.* Cells, 2019. **8**(1).
186. Ferreira, D., F. Adega, and R. Chaves, *The Importance of Cancer Cell Lines as in vitro Models in Cancer Methylome Analysis and Anticancer Drugs Testing*, in *Oncogenomics and Cancer Proteomics - Novel Approaches in Biomarkers Discovery and Therapeutic Targets in Cancer*. 2013.
187. Schmidt, K. and I. Podmore, *Current Challenges in Volatile Organic Compounds Analysis as Potential Biomarkers of Cancer.* Journal of Biomarkers, 2015. **2015**: p. 16.
188. Kell, D.B., et al., *Metabolic footprinting and systems biology: the medium is the message.* Nature Reviews Microbiology, 2005. **3**(7): p. 557-565.
189. Mapelli, V., L. Olsson, and J. Nielsen, *Metabolic footprinting in microbiology: methods and applications in functional genomics and biotechnology.* Trends Biotechnol, 2008. **26**(9): p. 490-7.
190. Rowan, D.D., *Volatile Metabolites.* Metabolites, 2011. **1**(1): p. 41-63.
191. *<2014_Combined proteomics and metabolomics analysis reveals grade-dependent metabolism pathways in kidney cancer.pdf>.*
192. Holopainen, J.K. and J. Gershenzon, *Multiple stress factors and the emission of plant VOCs.* Trends Plant Sci, 2010. **15**(3): p. 176-84.

193. Jansen, R.M., et al., *Detection of diseased plants by analysis of volatile organic compound emission.* Annu Rev Phytopathol, 2011. **49**: p. 157-74.

194. Dourtoglou, V., et al., *Discrimination of varietal wines according to their volatiles.* Food chemistry, 2014. **159**: p. 181-187.

195. Lim, D.K., et al., *Optimized Mass Spectrometry-Based Metabolite Extraction and Analysis for the Geographical Discrimination of White Rice (Oryza sativa L.): A Method Comparison Study.* J AOAC Int, 2018. **101**(2): p. 498-506.

196. Di Lena, M., F. Porcelli, and D.F. Altomare, *Volatile organic compounds as new biomarkers for colorectal cancer: a review.* Colorectal Dis, 2016. **18**(7): p. 654-63.

197. Khalid, T., et al., *Urinary Volatile Organic Compounds for the Detection of Prostate Cancer.* PLoS One, 2015. **10**(11): p. e0143283.

198. Phillips, M., et al., *Volatile biomarkers in the breath of women with breast cancer.* J Breath Res, 2010. **4**(2): p. 026003.

199. Garner, C.E., et al., *Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease.* Faseb j, 2007. **21**(8): p. 1675-88.

200. Kimball, B.A., D.A. Wilson, and D.W. Wesson, *Alterations of the volatile metabolome in mouse models of Alzheimer's disease.* Scientific Reports, 2016. **6**: p. 19495.

201. Filipiak, W., et al., *A Compendium of Volatile Organic Compounds (VOCs) Released By Human Cell Lines.* Current medicinal chemistry, 2016. **23**(20): p. 2112-2131.

202. Batty, C.A., et al., *Use of the Analysis of the Volatile Faecal Metabolome in Screening for Colorectal Cancer.* PLoS One, 2015. **10**(6): p. e0130301.

203. Abaffy, T., et al., *Differential volatile signatures from skin, naevi and melanoma: a novel approach to detect a pathological process.* PLoS One, 2010. **5**(11): p. e13813.

204. Deng, C. and X. Zhang, *A simple, rapid and sensitive method for determination of aldehydes in human blood by gas chromatography/mass spectrometry and solid-phase microextraction with on-fiber derivatization.* Rapid Commun Mass Spectrom, 2004. **18**(15): p. 1715-20.

205. Silva, C.L., M. Passos, and J.S. Camara, *Investigation of urinary volatile organic metabolites as potential cancer biomarkers by solid-phase microextraction in combination with gas chromatography-mass spectrometry.* Br J Cancer, 2011. **105**(12): p. 1894-904.
206. Filipiak, W., et al., *Comparative analyses of volatile organic compounds (VOCs) from patients, tumors and transformed cell lines for the validation of lung cancer-derived breath markers.* J Breath Res, 2014. **8**(2): p. 027111.
207. Willis, C.M., et al., *Olfactory detection of human bladder cancer by dogs: proof of principle study.* Bmj, 2004. **329**(7468): p. 712.
208. Peng, G., et al., *Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors.* Br J Cancer, 2010. **103**(4): p. 542-51.
209. Pickel, D., et al., *Evidence for canine olfactory detection of melanoma.* Applied Animal Behaviour Science, 2004. **89**(1-2): p. 107-116.
210. McCulloch, M., et al., *Diagnostic accuracy of canine scent detection in early- and late-stage lung and breast cancers.* Integr Cancer Ther, 2006. **5**(1): p. 30-9.
211. Chen, X., et al., *A study of the volatile organic compounds exhaled by lung cancer cells in vitro for breath diagnosis.* Cancer, 2007. **110**(4): p. 835-44.
212. Shirasu, M. and K. Touhara, *The scent of disease: volatile organic compounds of the human body related to disease and disorder.* J Biochem, 2011. **150**(3): p. 257-66.
213. Ulanowska, A., et al., *The application of statistical methods using VOCs to identify patients with lung cancer.* J Breath Res, 2011. **5**(4): p. 046008.
214. Silva, C.L., M. Passos, and J.S. Camara, *Solid phase microextraction, mass spectrometry and metabolomic approaches for detection of potential urinary cancer biomarkers--a powerful strategy for breast cancer diagnosis.* Talanta, 2012. **89**: p. 360-8.
215. Krönert, K., et al., *Urinary excretion patterns of endogenously produced alcohols in type 1 (IDDM) and type 2 (NIDDM) diabetes mellitus compared with healthy control subjects.* Diabetes Research and Clinical Practice, 1990. **10**(2): p. 161-165.

216. Vaupel, P., *Tumor microenvironmental physiology and its implications for radiation oncology.* Semin Radiat Oncol, 2004. **14**(3): p. 198-206.
217. Nunes de Paiva, M.J., H.C. Menezes, and Z. de Lourdes Cardeal, *Sampling and analysis of metabolomes in biological fluids.* Analyst, 2014. **139**(15): p. 3683-94.
218. Risticevic, S., et al., *Recent developments in solid-phase microextraction.* Anal Bioanal Chem, 2009. **393**(3): p. 781-95.
219. Vuckovic, D., S. Risticevic, and J. Pawliszyn, *In vivo solid-phase microextraction in metabolomics: opportunities for the direct investigation of biological systems.* Angew Chem Int Ed Engl, 2011. **50**(25): p. 5618-28.
220. Zhang, Z. and J. Pawliszyn, *Headspace solid-phase microextraction.* Analytical Chemistry, 1993. **65**(14): p. 1843-1852.
221. Monteiro, M., et al., *Analysis of volatile human urinary metabolome by solid-phase microextraction in combination with gas chromatography-mass spectrometry for biomarker discovery: application in a pilot study to discriminate patients with renal cell carcinoma.* Eur J Cancer, 2014. **50**(11): p. 1993-2002.
222. Kwak, J., et al., *Volatile biomarkers from human melanoma cells.* Journal of chromatography. B, Analytical technologies in the biomedical and life sciences, 2013. **931**: p. 90-96.
223. Gupta-Elera, G., et al., *The role of oxidative stress in prostate cancer.* Eur J Cancer Prev, 2012. **21**(2): p. 155-62.
224. Miyajima, A., et al., *Role of reactive oxygen species in cis-dichlorodiammineplatinum-induced cytotoxicity on bladder cancer cells.* British Journal of Cancer, 1997. **76**(2): p. 206-210.
225. Szatrowski, T.P. and C.F. Nathan, *Production of large amounts of hydrogen peroxide by human tumor cells.* Cancer Res, 1991. **51**(3): p. 794-8.
226. Guéraud, F., et al., *Chemistry and biochemistry of lipid peroxidation products.* Free Radical Research, 2010. **44**(10): p. 1098-1124.
227. Jareño-Esteban, J.J., et al., *Volatile Organic Compounds in Exhaled Breath in a Healthy Population: Effect of Tobacco Smoking.* Archivos de Bronconeumología (English Edition), 2013. **49**(11): p. 457-461.

228. Kodama, M., et al., *Free radical chemistry of cigarette smoke and its implication in human cancer.* Anticancer Res, 1997. **17**(1a): p. 433-7.
229. Draper, H.H., A.S. Csallany, and M. Hadley, *Urinary aldehydes as indicators of lipid peroxidation in vivo.* Free Radical Biology and Medicine, 2000. **29**(11): p. 1071-1077.
230. Barrera, G., *Oxidative stress and lipid peroxidation products in cancer progression and therapy.* ISRN Oncol, 2012. **2012**: p. 137289.
231. Esterbauer, H., *Cytotoxicity and genotoxicity of lipid-oxidation products.* Am J Clin Nutr, 1993. **57**(5 Suppl): p. 779S-785S; discussion 785S-786S.
232. Esterbauer, H., R.J. Schaur, and H. Zollner, *Chemistry and biochemistry of 4-hydroxynonenal, malonaldehyde and related aldehydes.* Free Radical Biology and Medicine, 1991. **11**(1): p. 81-128.
233. Spiteller, G., *Linoleic acid peroxidation—the dominant lipid peroxidation process in low density lipoprotein—and its relationship to chronic diseases.* Chemistry and Physics of Lipids, 1998. **95**(2): p. 105-162.
234. Thornalley, P.J., A. Langborg, and H.S. Minhas, *Formation of glyoxal, methylglyoxal and 3-deoxyglucosone in the glycation of proteins by glucose.* Biochemical Journal, 1999. **344**(1): p. 109.
235. Laffel, L., *Ketone bodies: a review of physiology, pathophysiology and application of monitoring to diabetes.* Diabetes Metab Res Rev, 1999. **15**(6): p. 412-26.
236. Negre-Salvayre, A., et al., *Advanced lipid peroxidation end products in oxidative damage to proteins. Potential role in diseases and therapeutic prospects for the inhibitors.* Br J Pharmacol, 2008. **153**(1): p. 6-20.
237. Repetto, M., J. Semprine, and A. Boveris, *Lipid Peroxidation: Chemical Mechanism, Biological Implications and Analytical Determination*, in *Lipid Peroxidation*. 2012.
238. Anton, A.P., et al., *Headspace generation coupled to gas chromatography-mass spectrometry for the automated determination and quantification of endogenous compounds in urine. Aldehydes as possible markers of oxidative stress.* J Chromatogr A, 2014. **1367**: p. 9-15.

239. Guadagni, R., et al., *Solid-phase microextraction-gas chromatography-mass spectrometry method validation for the determination of endogenous substances: urinary hexanal and heptanal as lung tumor biomarkers.* Anal Chim Acta, 2011. **701**(1): p. 29-36.
240. Kuiper, H.C., et al., *Quantitation of mercapturic acid conjugates of 4-hydroxy-2-nonenal and 4-oxo-2-nonenal metabolites in a smoking cessation study.* Free radical biology & medicine, 2010. **48**(1): p. 65-72.
241. Spanel, P., et al., *Analysis of formaldehyde in the headspace of urine from bladder and prostate cancer patients using selected ion flow tube mass spectrometry.* Rapid Commun Mass Spectrom, 1999. **13**(14): p. 1354-9.
242. Yazdanpanah, M., et al., *Cytotoxic aldehydes as possible markers for childhood cancer.* Free Radic Biol Med, 1997. **23**(6): p. 870-8.
243. Lu, Y., D. Yao, and C. Chen, *2-Hydrazinoquinoline as a Derivatization Agent for LC-MS-Based Metabolomic Investigation of Diabetic Ketoacidosis.* Metabolites, 2013. **3**(4).
244. Des Rosiers, C., et al., *Gas chromatographic-mass spectrometric assay of tissue malondialdehyde, 4-hydroxynonenal, and other aldehydes after their reduction to stable alcohols.* Anal Biochem, 1993. **208**(1): p. 161-70.
245. De Zwart, L.L., et al., *Simultaneous determination of eight lipid peroxidation degradation products in urine of rats treated with carbon tetrachloride using gas chromatography with electron-capture detection.* Journal of Chromatography B: Biomedical Sciences and Applications, 1997. **694**(2): p. 277-287.
246. Luo, X.P., et al., *Determination of Aldehydes and Other Lipid Peroxidation Products in Biological Samples by Gas Chromatography-Mass Spectrometry.* Analytical Biochemistry, 1995. **228**(2): p. 294-298.
247. W. A. English, W.E.S., *Emerging Trends in Gas Chromatography and Mass Spectrometry Instrumentation for Analytical & Bioanalytical Techniques.* Journal of Analytical & Bioanalytical Techniques, 2015. **06**(01).
248. Ryan, D., et al., *Recent and potential developments in the analysis of urine: a review.* Anal Chim Acta, 2011. **684**(1-2): p. 8-20.

249. Zhang, Q., et al., *Solid-phase microextraction technology for in vitro and in vivo metabolite analysis.* Trends in analytical chemistry : TRAC, 2016. **80**: p. 57-65.
250. Gomase, V.S., et al., *Metabolomics.* Curr Drug Metab, 2008. **9**(1): p. 89-98.
251. Monteiro, M.S., et al., *Metabolomics analysis for biomarker discovery: advances and challenges.* Curr Med Chem, 2013. **20**(2): p. 257-71.
252. Chan, E.C., et al., *Metabonomic profiling of bladder cancer.* J Proteome Res, 2015. **14**(2): p. 587-602.
253. Wu, Y. and L. Li, *Sample normalization methods in quantitative metabolomics.* J Chromatogr A, 2016. **1430**: p. 80-95.
254. Halama, A., *Metabolomics in cell culture--a strategy to study crucial metabolic pathways in cancer development and the response to treatment.* Arch Biochem Biophys, 2014. **564**: p. 100-9.
255. Trygg, J., E. Holmes, and T. Lundstedt, *Chemometrics in Metabonomics.* Journal of Proteome Research, 2007. **6**(2): p. 469-479.
256. Davis, V.W., et al., *Metabolomics and surgical oncology: Potential role for small molecule biomarkers.* J Surg Oncol, 2011. **103**(5): p. 451-9.
257. Spratlin, J.L., N.J. Serkova, and S.G. Eckhardt, *Clinical applications of metabolomics in oncology: a review.* Clin Cancer Res, 2009. **15**(2): p. 431-40.
258. Ding, X.F., et al., *The tumor suppressor pVHL down-regulates never-in-mitosis A-related kinase 8 via hypoxia-inducible factors to maintain cilia in human renal cancer cells.* J Biol Chem, 2015. **290**(3): p. 1389-94.
259. Iliopoulos, O., et al., *Tumour suppression by the human von Hippel-Lindau gene product.* Nat Med, 1995. **1**(8): p. 822-6.
260. Shinojima, T., et al., *Renal cancer cells lacking hypoxia inducible factor (HIF)-1alpha expression maintain vascular endothelial growth factor expression through HIF-2alpha.* Carcinogenesis, 2007. **28**(3): p. 529-36.
261. Veliceasa, D., et al., *Transient potential receptor channel 4 controls thrombospondin-1 secretion and angiogenesis in renal cell carcinoma.* Febs j, 2007. **274**(24): p. 6365-77.

262. Lichner, Z., et al., *miR-17 inhibition enhances the formation of kidney cancer spheres with stem cell/ tumor initiating cell properties.* Oncotarget, 2015. **6**(8): p. 5567-81.

263. Liu, Y.H., et al., *Up-regulation of vascular endothelial growth factor-D expression in clear cell renal cell carcinoma by CD74: a critical role in cancer cell tumorigenesis.* J Immunol, 2008. **181**(9): p. 6584-94.

264. Miyake, M., et al., *Erythropoietin is a JAK2 and ERK1/2 effector that can promote renal tumor cell proliferation under hypoxic conditions.* J Hematol Oncol, 2013. **6**: p. 65.

265. Korhonen, M., et al., *Integrins and laminins in human renal carcinoma cells and tumors grown in nude mice.* Cancer Res, 1994. **54**(16): p. 4532-8.

266. Hakimi, A.A., et al., *Mp23-11 Genomic Comparison of Renal Cell Carcinoma Cell Lines to Human Tumors.* Journal of Urology, 2014. **191**(4S): p. e247-e247.

267. Furge, K.A., et al., *Detection of DNA copy number changes and oncogenic signaling abnormalities from gene expression data reveals MYC activation in high-grade papillary renal cell carcinoma.* Cancer Res, 2007. **67**(7): p. 3171-6.

268. Pulkkanen, K.J., et al., *Characterization of a new animal model for human renal cell carcinoma.* In Vivo, 2000. **14**(3): p. 393-400.

269. Pulkkanen, K.J., et al., *HSV-tk gene therapy for human renal cell carcinoma in nude mice.* Cancer Gene Ther, 2001. **8**(7): p. 529-36.

270. Brodaczewska, K.K., et al., *Choosing the right cell line for renal cell cancer research.* Mol Cancer, 2016. **15**(1): p. 83.

271. Looyenga, B.D., et al., *Chromosomal amplification of leucine-rich repeat kinase-2 (LRRK2) is required for oncogenic MET signaling in papillary renal and thyroid carcinomas.* Proc Natl Acad Sci U S A, 2011. **108**(4): p. 1439-44.

272. Furge, K.A., et al., *Combining differential expression, chromosomal and pathway analyses for the molecular characterization of renal cell carcinoma.* Can Urol Assoc J, 2007. **1**(2 Suppl): p. S21-7.

273. Ashida, S., et al., *Effects of von Hippel-Lindau gene mutation and methylation status on expression of transmembrane carbonic anhydrases in renal cell carcinoma.* J Cancer Res Clin Oncol, 2002. **128**(10): p. 561-8.

274. Kucejova, B., et al., *Interplay between pVHL and mTORC1 pathways in clear-cell renal cell carcinoma.* Mol Cancer Res, 2011. **9**(9): p. 1255-65.

275. Araújo, A.M., et al., *Analysis of extracellular metabolome by HS-SPME/GC–MS: Optimization and application in a pilot study to evaluate galactosamine-induced hepatotoxicity.* Toxicology Letters, 2018. **295**: p. 22-31.

276. Pluskal, T., et al., *MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data.* BMC Bioinformatics, 2010. **11**: p. 395.

277. Pasikanti, K.K., et al., *Metabolic footprinting of tumorigenic and nontumorigenic uroepithelial cells using two-dimensional gas chromatography time-of-flight mass spectrometry.* Analytical and bioanalytical chemistry, 2010. **398**(3): p. 1285-1293.

278. Worley, B. and R. Powers, *Multivariate Analysis in Metabolomics.* Curr Metabolomics, 2013. **1**(1): p. 92-107.

279. Wheelock, Å.M. and C.E. Wheelock, *Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine.* Molecular BioSystems, 2013. **9**(11): p. 2589-2596.

280. Berben, L., S.M. Sereika, and S. Engberg, *Effect size estimation: methods and examples.* Int J Nurs Stud, 2012. **49**(8): p. 1039-47.

281. Thriumani, R., et al., *A study on volatile organic compounds emitted by in-vitro lung cancer cultured cells using gas sensor array and SPME-GCMS.* BMC Cancer, 2018. **18**(1): p. 362.

282. Qin, T., et al., *The screening of volatile markers for hepatocellular carcinoma.* Cancer Epidemiol Biomarkers Prev, 2010. **19**(9): p. 2247-53.

283. Phillips, M., et al., *Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study.* Lancet, 1999. **353**(9168): p. 1930-3.

284. Abaffy, T., et al., *A case report - Volatile metabolomic signature of malignant melanoma using matching skin as a control.* J Cancer Sci Ther, 2011. **3**(6): p. 140-144.

285. Serasanambati, M., et al., *Profiling Single Cancer Cells with Volatolomics Approach.* iScience, 2019. **11**: p. 178-188.

286. Handa, H., et al., *Exhaled Breath Analysis for Lung Cancer Detection Using Ion Mobility Spectrometry.* PLOS ONE, 2014. **9**(12): p. e114555.
287. Wang, C., et al., *Noninvasive detection of colorectal cancer by analysis of exhaled breath.* Anal Bioanal Chem, 2014. **406**(19): p. 4757-63.
288. Hadi, N.I., et al., *Serum Metabolomic Profiles for Breast Cancer Diagnosis, Grading and Staging by Gas Chromatography-Mass Spectrometry.* Sci Rep, 2017. **7**(1): p. 1715.
289. Rodrigues, D., et al., *Volatile metabolomic signature of bladder cancer cell lines based on gas chromatography–mass spectrometry.* Metabolomics, 2018. **14**(5).
290. Tang, H., et al., *Determination of volatile organic compounds exhaled by cell lines derived from hematological malignancies.* Biosci Rep, 2017. **37**(3).
291. Sever, A., et al., *Biomarkers for Detection and Monitoring of B16 Melanoma in Mouse Urine and Feces.* Journal of Biomarkers, 2015. **2015**: p. 9.
292. Filipiak, W., et al., *TD-GC-MS analysis of volatile metabolites of human lung cancer and normal cells in vitro.* Cancer Epidemiol Biomarkers Prev, 2010. **19**(1): p. 182-95.
293. Wang, C., et al., *Blood volatile compounds as biomarkers for colorectal cancer.* Cancer Biology & Therapy, 2014. **15**(2): p. 200-206.
294. Jobu, K., et al., *Metabolomics study on the biochemical profiles of odor elements in urine of human with bladder cancer.* Biol Pharm Bull, 2012. **35**(4): p. 639-42.
295. Silva, C.L., et al., *Volatile metabolomic signature of human breast cancer cell lines.* Scientific Reports, 2017. **7**: p. 43969.
296. Jia, Z., et al., *Detection of Lung Cancer: Concomitant Volatile Organic Compounds and Metabolomic Profiling of Six Cancer Cell Lines of Different Histological Origins.* ACS Omega, 2018. **3**(5): p. 5131-5140.
297. Ayala, A., M.F. Munoz, and S. Arguelles, *Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal.* Oxid Med Cell Longev, 2014. **2014**: p. 360438.
298. Hakim, M., et al., *Volatile Organic Compounds of Lung Cancer and Possible Biochemical Pathways.* Chemical Reviews, 2012. **112**(11): p. 5949-5966.

299. Hanai, Y., et al., *Urinary volatile compounds as biomarkers for lung cancer.* Biosci Biotechnol Biochem, 2012. **76**(4): p. 679-84.

300. Jiménez-Pacheco, A., et al., *Furan and p-xylene as candidate biomarkers for prostate cancer.* Urologic Oncology: Seminars and Original Investigations, 2018. **36**(5): p. 243.e21-243.e27.

301. Wang, D., et al., *Urinary volatile organic compounds as potential biomarkers for renal cell carcinoma.* Biomed Rep, 2016. **5**(1): p. 68-72.

302. Peng, G., et al., *Diagnosing lung cancer in exhaled breath using gold nanoparticles.* Nat Nanotechnol, 2009. **4**(10): p. 669-73.

303. Liu, H., et al., *Investigation of volatile organic metabolites in lung cancer pleural effusions by solid-phase microextraction and gas chromatography/mass spectrometry.* Journal of Chromatography B, 2014. **945-946**: p. 53-59.

304. Sponring, A., et al., *Release of volatile organic compounds from the lung cancer cell line NCI-H2087 in vitro.* Anticancer Res, 2009. **29**(1): p. 419-26.

305. O'Neill, H.J., et al., *A computerized classification technique for screening for the presence of breath biomarkers in lung cancer.* Clin Chem, 1988. **34**(8): p. 1613-8.

306. Woollam, M., et al., *Detection of Volatile Organic Compounds (VOCs) in Urine via Gas Chromatography-Mass Spectrometry QTOF to Differentiate Between Localized and Metastatic Models of Breast Cancer.* Scientific Reports, 2019. **9**(1): p. 2526.

307. Droge, W., *Free radicals in the physiological control of cell function.* Physiol Rev, 2002. **82**(1): p. 47-95.

308. Valko, M., et al., *Free radicals and antioxidants in normal physiological functions and human disease.* Int J Biochem Cell Biol, 2007. **39**(1): p. 44-84.

309. Traverso, N., et al., *Role of Glutathione in Cancer Progression and Chemoresistance.* Oxidative Medicine and Cellular Longevity, 2013. **2013**: p. 10.

310. Minotti, G., et al., *Paradoxical inhibition of cardiac lipid peroxidation in cancer patients treated with doxorubicin. Pharmacologic and molecular reappraisal of anthracycline cardiotoxicity.* J Clin Invest, 1996. **98**(3): p. 650-61.

311. Phaniendra, A., D.B. Jestadi, and L. Periyasamy, *Free radicals: properties, sources, targets, and their implication in various diseases.* Indian J Clin Biochem, 2015. **30**(1): p. 11-26.
312. He, S.M. and B. Lambert, *Acetaldehyde-induced mutation at the hprt locus in human lymphocytes in vitro.* Environ Mol Mutagen, 1990. **16**(2): p. 57-63.
313. Chen, C.H., et al., *Targeting aldehyde dehydrogenase 2: new therapeutic opportunities.* Physiol Rev, 2014. **94**(1): p. 1-34.
314. Shibamoto, T., *Analytical methods for trace levels of reactive carbonyl compounds formed in lipid peroxidation systems.* Journal of Pharmaceutical and Biomedical Analysis, 2006. **41**(1): p. 12-25.
315. Salaspuro, M., *Acetaldehyde as a common denominator and cumulative carcinogen in digestive tract cancers.* Scandinavian Journal of Gastroenterology, 2009. **44**(8): p. 912-925.
316. Lorenti Garcia, C., et al., *Relationship between DNA lesions, DNA repair and chromosomal damage induced by acetaldehyde.* Mutat Res, 2009. **662**(1-2): p. 3-9.
317. Grafström, R.C., et al., *Pathobiological effects of acetaldehyde in cultured human epithelial cells and fibroblasts.* Carcinogenesis, 1994. **15**(5): p. 985-990.
318. Cui, C., et al., *Targeting calcium signaling in cancer therapy.* Acta Pharmaceutica Sinica B, 2017. **7**(1): p. 3-17.
319. Grafstrom, R.C., *In vitro studies of aldehyde effects related to human respiratory carcinogenesis.* Mutat Res, 1990. **238**(3): p. 175-84.
320. Grafström, R.C., *Carcinogenesis studies in human epithelial tissues and cells in vitro: emphasis on serum-free culture conditions and transformation studies.* Acta physiologica Scandinavica. Supplementum, 1990. **592**: p. 93-133.
321. Harris, C.C., *Chemical and physical carcinogenesis: advances and perspectives for the 1990s.* Cancer Res, 1991. **51**(18 Suppl): p. 5023s-5044s.
322. Dellarco, V.L., *A mutagenicity assessment of acetaldehyde.* Mutat Res, 1988. **195**(1): p. 1-20.

323. Lambert, B. and S.M. He, *DNA and chromosome damage induced by acetaldehyde in human lymphocytes in vitro.* Ann N Y Acad Sci, 1988. **534**: p. 369-76.
324. Hanouneh, I.A., et al., *The breathprints in patients with liver disease identify novel breath biomarkers in alcoholic hepatitis.* Clin Gastroenterol Hepatol, 2014. **12**(3): p. 516-23.
325. Koppenol, W.H., P.L. Bounds, and C.V. Dang, *Otto Warburg's contributions to current concepts of cancer metabolism.* Nat Rev Cancer, 2011. **11**(5): p. 325-37.
326. Tanda, N., et al., *Exhaled acetone and isoprene in perioperative lung cancer patients under intensive oral care: possible indicators of inflammatory responses and metabolic changes.* Biomed Res, 2019. **40**(1): p. 29-36.
327. Richard, J.P., *Mechanism for the formation of methylglyoxal from triosephosphates.* Biochem Soc Trans, 1993. **21**(2): p. 549-53.
328. Chiavarina, B., et al., *Methylglyoxal-Mediated Stress Correlates with High Metabolic Activity and Promotes Tumor Growth in Colorectal Cancer.* Int J Mol Sci, 2017. **18**(1).
329. Thornalley, P.J., *Dicarbonyl intermediates in the maillard reaction.* Ann N Y Acad Sci, 2005. **1043**: p. 111-7.
330. Nokin, M.-J., et al., *Methylglyoxal, a glycolysis side-product, induces Hsp90 glycation and YAP-mediated tumor growth and metastasis.* eLife, 2016. **5**: p. e19375.
331. Rabbani, N., et al., *Multiple roles of glyoxalase 1-mediated suppression of methylglyoxal glycation in cancer biology—Involvement in tumour suppression, tumour growth, multidrug resistance and target for chemotherapy.* Seminars in Cancer Biology, 2018. **49**: p. 83-93.
332. Chakraborty, S., K. Karmakar, and D. Chakravortty, *Cells producing their own nemesis: understanding methylglyoxal metabolism.* IUBMB Life, 2014. **66**(10): p. 667-78.

333. Monteiro, M., et al., *GC-MS metabolomics-based approach for the identification of a potential VOC-biomarker panel in the urine of renal cell carcinoma patients.* Journal of Cellular and Molecular Medicine, 2017. **21**(9): p. 2092-2105.

334. Takahashi, M., et al., *Effects of glyoxal and methylglyoxal administration on gastric carcinogenesis in Wistar rats after initiation with N-methyl-N'-nitro-N-nitrosoguanidine.* Carcinogenesis, 1989. **10**(10): p. 1925-7.

335. Svendsen, C., et al., *The food processing contaminant glyoxal promotes tumour growth in the multiple intestinal neoplasia (Min) mouse model.* Food and Chemical Toxicology, 2016. **94**: p. 197-202.

336. Schaeffeler, E., et al., *Metabolic and Lipidomic Reprogramming in Renal Cell Carcinoma Subtypes Reflects Regions of Tumor Origin.* European Urology Focus, 2018.

337. Viant, M.R., et al., *How close are we to complete annotation of metabolomes?* Curr Opin Chem Biol, 2017. **36**: p. 64-69.

9 Appendix

Supplementary Table 1 List of VOCs selected as important in discriminating the cancer cell lines from the normal cell line HK-2. The identification of the metabolites is based on the NIST (2019). They are characterized by their name, characteristic ions (m/z), Retention Time, average retention index according to literature (RIlit.), retention index calculated based on the n-alkanes series (RIcalc.) and Reverse Match (R. Match).

VOC	Identification	Retention Time (min)	Characteristic ions (m/z)	RIlit.	RIcalc.	R. Match
1,3-Di-tert-butylbenzene	L_2	17.94	175+190	1247	1247	909
2,4-Dimethyl-1-heptene	L_2	6.07	55+70	836	839	940
2,6-Di-tert-butylquinone	L_2	23.56	67+135	1461	1459	873
2-Ethyl-1-hexanol	L_2	11.48	57+70	1030	1027	902
Acetophenone	L_2	12.61	77+105	1065	1064	889
Cyclohexanol	L_2	7.29	57+67	880	886	909
Decane	L_2	10.66	57+71	1000	1000	865
Dodecane	L_2	16.62	57+71	1200	1200	816
Ethylbenzene	L_2	6.60	91+106	855	859	855
p-xylene/1,3-Dimethyl-benzene	L_2	6.87	91+106	865/866	869	895/871
Unknown 1	L_4	7.48	55+98+104	-	893	-
Unknown 2	L_4	8.58	55+69+71	-	930	-
Unknown 3	L_4	9.49	57+71+87	-	961	-
Unknown 4	L_4	9.65	57+71+83	-	966	-
Unknown 5	L_4	9.98	59+77+94	-	977	-
Unknown 6	L_4	10.85	57+71+85	-	1007	-
Unknown 7	L_4	10.86	57+69+93	-	1008	-
Unknown 8	L_4	10.96	57+71+85	-	1011	-
Unknown 9	L_4	11.23	56+70+85	-	1019	-
Unknown 10	L_4	12.76	59+77+105	-	1071	-
Unknown 11	L_4	12.93	55+69+83	-	1075	-
Unknown 12	L_4	13.21	51+77+91	-	1084	-
Unknown 13	L_4	17.04	55+57+71	-	1215	-

Note:

L1 - Identified metabolites (GC-MS analysis of the metabolite of interest and a chemical reference standard of suspected structural equivalence, with all analyses performed under identical analytical conditions within the same laboratory) [337]; **L2** - Putatively annotated compounds (spectral (MS) similarity with NIST database), when standards were not commercially available [337]; **L4 –** Unidentified [337].

Supplementary Table 2 List of VCCs selected as important in discriminating the cancer cell lines from the normal cell line HK-2. The identification of the metabolites is based on the NIST (2019). They are characterized by their name, characteristic ions (m/z), Retention Time, average retention index according to literature (RIlit.), retention index calculated (RIcalc.) and Reverse Match (R. Match).

VCC	Identification	Retention Time (min)	Characteristic ions (m/z)	$RI_{lit.}$	$RI_{calc.}$	R. Match
Acetaldehyde	**L_1**	13.24	209+239	1112	1104	906
		13.41			1110	926
Acetone	**L_1**	14.94	253	1167	1161	901
Cyclohexanone	**L_1**	24.61	276+293	1635	1523	826
Glyoxal	**L_1**	33.29	235+448	1935	1921	885
		33.45			1930	899
Methlyglyoxal	**L_1**	33.56	265	2174	1935	903
		34.06			1961	912
Unknown 1	**L_4**	10.62	99+117+195	-	1017	-
Unknown 2	**L_4**	19.78	69+93+362	-	1334	-
Unknown 3	**L_4**	21.48	55+58+240	-	1397	-
Unknown 4	**L_4**	24.00	93+117+161	-	1498	-
Unknown 5	**L_4**	26.07	93+117+161	-	1582	-
Unknown 6	**L_4**	26.51	55+69+112	-	1604	-
Unknown 7	**L_4**	30.42	77+79+91	-	1778	-

Note:

L1 - Identified metabolites (GC-MS analysis of the metabolite of interest and a chemical reference standard of suspected structural equivalence, with all analyses performed under identical analytical conditions within the same laboratory) [337]; **L2** - Putatively annotated compounds (spectral (MS) similarity with NIST database), when standards were not commercially available [337]; **L4 –** Unidentified [337].