



**HAL**  
open science

# Étude de l'hétérogénéité intra-tumorale et de l'impact des thérapies ciblées sur le carcinome rénal à cellules claires par une approche transcriptomique sur cellules uniques

Judikael Saout

► **To cite this version:**

Judikael Saout. Étude de l'hétérogénéité intra-tumorale et de l'impact des thérapies ciblées sur le carcinome rénal à cellules claires par une approche transcriptomique sur cellules uniques. Médecine humaine et pathologie. Université de Rennes, 2022. Français. NNT : 2022REN1B056 . tel-04893585

**HAL Id: tel-04893585**

**<https://theses.hal.science/tel-04893585v1>**

Submitted on 17 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THESE DE DOCTORAT DE

L'UNIVERSITE DE RENNES 1

ECOLE DOCTORALE N° 637

*Sciences de la vie et de la santé*

Spécialité : « *Génétique, génomique et bioinformatique* »

Par

**Judikael SAOUT**

**Étude de l'hétérogénéité intra-tumorale et de l'impact de thérapies ciblées sur le carcinome rénal à cellules claires par une approche transcriptomique sur cellules uniques**

Thèse présentée et soutenue à Rennes, le 1<sup>er</sup> décembre 2022

Unité de recherche : Irset – Inserm UMR\_S 1085

Équipe : « Physiologie et physiopathologie du tractus urogénital (URGENT) »

## Rapporteurs avant soutenance :

Thierry MASSFELDER

Joël LEMAULT

Directeur de Recherche, Inserm U1260, Université de Strasbourg

Directeur de Recherche, CEA-Service de Recherches en Hémato-Immunologie, INSERM U976, Hôpital Saint-Louis, Paris

## Composition du Jury :

Président

Joël LEMAULT

Directeur de Recherche, CEA-Service de Recherches en Hémato-Immunologie, INSERM U976, Hôpital Saint-Louis, Paris

Examineur

Thierry MASSFELDER

Directeur de Recherche, Inserm U1260, Université de Strasbourg

Directeur de thèse

Frédéric CHALMEL

Directeur de Recherche, Irset – Inserm UMR\_S 1085, Rennes

Co-directrice de thèse

Nathalie RIOUX-LECLERCQ

Professeur des Universités – Praticien Hospitalier, CHU Pontchaillou, Rennes  
Irset – Inserm UMR\_S 1085, Rennes

Invitée

Aurélien LARDENOIS

Maître de Conférences de l'Université de Rennes 1  
Inserm UMR\_S 1085, Irset – Inserm UMR\_S 1085, Rennes



**CONFIDENTIEL**





## REMERCIEMENTS

En premier lieu, je tiens à remercier **Bernard Jégou** et **Michel Samson**, directeurs successifs de l'Irset, pour leur accueil et leur bienveillance. En particulier, merci pour les marques de reconnaissance et d'encouragements qu'ils m'ont témoignés.

Merci au Docteur **Nathalie Dejuq-Rainsford**, et au Docteur **Frédéric Chalmel**, co-directeurs de l'équipe de recherche « Physiologie et physiopathologie du tractus urogénital », pour m'avoir accueillie au sein de leur équipe.

J'adresse mes sincères remerciements au Docteur **Joël Le Maoult**, directeur de recherche au CEA à l'hôpital Saint-Louis, et au Docteur **Thierry Massfelder**, directeur de recherche à l'INSERM à Strasbourg, pour avoir accepté d'être rapporteurs de cette thèse.

Pour le très bon encadrement dont j'ai bénéficié, je tiens à remercier le Docteur **Frédéric Chalmel**, le Docteur **Nathalie-Rioux Leclercq** et le Docteur **Aurélien Lardenois**, qui ont co-dirigé ces travaux. Je veux également remercier les Docteurs **Thomas Derrien**, **Céline Delaloy** et **Aurélien Serandour**, membres de mon comité de thèse, pour leur regard critique et leurs conseils avisés au long de cette thèse.

J'aimerais remercier également nos collaboratrices, le Docteur **Odile Filhol-Cochet** et **Chatherine Pillet** pour leur aide précieuse dans le développement de mes projets de thèse.

A tous les membres de **l'équipe 8**, j'adresse mes chaleureux remerciements, pour le cadre de travail dont ils contribuent à maintenir la haute qualité, à la fois professionnelle et humaine.

Une thèse ne se réalise pas seul. Bien que de nombreuses personnes y aient contribué, j'aimerais en remercier quelques unes en particulier, pour leur aide inestimable :

**Bertrand**. Sans toi, sans Le GoldFinger ; rien n'était possible. Et avec toi, tout était plus drôle (fou-rires mémorables en ce qui me concerne). Au laboratoire jusqu'à 3h du matin, à la piscine, en footing, aux palets, sur 40km de rando à travers champs... je ne compte plus les

moments passés ensemble et les mots me manquent. Merci pour tout, et davantage encore pour ton amitié.

**Gwendoline.** Je te suis très reconnaissant, pour l'aide concrète que tu m'as apportée et pour l'émulation autour de nos thématiques. Je suis heureux et vraiment satisfait que tu aies rejoint le groupe ; j'espère que toi aussi. Je te souhaite le meilleur pour la fin de ta thèse.

**Aurélie.** Merci. Tu as joué un rôle unique dans mon encadrement et j'en mesure la portée. Je te dois en bonne partie ce manuscrit de thèse, je voulais te dire ma gratitude.

**Fred.** Je comprends pourquoi mes illustres prédécesseurs voyaient en toi un mentor. Je ne saurais pas t'exprimer ma reconnaissance à sa juste valeur ici. Merci de m'avoir fait confiance et soutenu. Merci pour ton investissement qui confine à l'abnégation, parfois. Merci pour ton amitié. Grâce à toi (à la thèse, certes, mais aussi à toi dans cette thèse), j'ai pu en apprendre plus sur moi, pour ma plus grande satisfaction et pour mon bien ; je sais que cela me servira.

À tous les autres belles rencontres de cette thèse et dont les amitiés ont dépassé le cadre ; à celles qui composent l'équipe 8 et à la famille des doctorantes (et affiliées, Maryne ;) , à mes voisines de Choc, aux mousquetaires, au frère, aux partenaires sportifs, au grand blond avec une chaussure noire, au beau brun tatoué, aux gamers, aux multiples parisiennes, aux pékinois express, aux Barcelonais, aux acteurs studios, aux compagnons de labo, aux amateurs de bière, aux Doc7teurs, aux gourmets et chefs pâtisseries (miam) ...

A tous, Merci.

# TABLE DES MATIERES

REMERCIEMENTS.....	I
LISTE DES TABLEAUX.....	IX
LISTE DES FIGURES.....	XI
LISTE DES ABREVIATIONS.....	XIII
<b>1. INTRODUCTION.....</b>	<b>1</b>
1.1. DES CARCINOMES RÉNAUX À CELLULES CLAIRES.....	3
1.1.1. Première description de cancer du rein.....	3
1.1.2. À l'origine, le rein.....	3
1.1.2.1. Morphologie.....	4
1.1.2.2. Micro-anatomie.....	4
1.1.2.3. Vascularisation.....	4
1.1.3. Classification des carcinomes rénaux.....	5
1.1.4. Épidémiologie.....	9
1.1.4.1. Incidence et mortalité du cancer du rein.....	9
1.1.4.2. Facteurs de risque.....	10
1.1.4.2.1. Tabagisme.....	10
1.1.4.2.2. Indice de masse corporelle.....	11
1.1.4.2.3. Hypertension artérielle.....	11
1.1.4.2.4. Insuffisance rénale chronique.....	12
1.1.4.2.5. Prédisposition génétique.....	12
1.1.5. Facteurs pronostiques du ccRCC.....	12
1.1.5.1. Facteurs anatomiques.....	12
1.1.5.2. Facteurs histologiques.....	15
1.1.5.3. Facteurs cliniques.....	17
1.1.5.4. Facteurs moléculaires.....	17
1.1.5.5. Les modèles pronostics.....	17
1.1.5.6. Analyses de survie.....	19
1.1.5.7. Écueils au pronostic anatomo-pathologique.....	20
1.1.6. Mécanismes et pathophysiologie.....	21
1.1.6.1. Altérations génétiques.....	21
1.1.6.2. Étapes et processus métastatique.....	27
1.1.7. Microenvironnement.....	29
1.1.7.1. Microenvironnement vasculaire.....	30
1.1.7.2. Microenvironnement stromal.....	30
1.1.7.3. Microenvironnement immunitaire.....	31
1.1.8. Prise en charge médicale.....	33

1.1.8.1. Chirurgie.....	33
1.1.8.2. Surveillance active et thérapies ablatives .....	33
1.1.8.3. Traitement thérapeutique systémique .....	34
1.1.8.3.1. Thérapies ciblées anti-angiogéniques.....	35
1.1.8.3.2. Immunothérapies .....	36
1.1.8.3.3. Nouvelles cibles thérapeutiques.....	38
1.1.8.4. Échappement aux thérapies.....	38
1.2. DE L'ÉVOLUTION ET DE L'HÉTÉROGÉNÉITÉ TUMORALE .....	39
1.2.1. <i>Généralités</i> .....	39
1.2.1.1. Concepts historiques.....	39
1.2.1.2. Modèles d'évolution tumorale.....	42
1.2.1.2.1. Modèle d'évolution linéaire.....	42
1.2.1.2.2. Modèle d'évolution ramifiée .....	43
1.2.1.2.3. Modèle d'évolution neutre.....	44
1.2.1.2.4. Modèle d'évolution ponctuée .....	44
1.2.1.3. Implication des cellules souches cancéreuses.....	45
1.2.2. <i>Implication dans l'échappement thérapeutique</i> .....	46
1.2.2.1. Échappement aux inhibiteurs de tyrosine kinase .....	46
1.2.2.2. Échappement aux immunothérapies .....	47
1.2.3. <i>Hétérogénéité intra-tumorale dans le ccRCC</i> .....	48
1.3. DES MODELES D'ÉTUDE DES CCRCCS .....	51
1.3.1. <i>Systèmes 2D : lignées cellulaires de ccRCC</i> .....	51
1.3.2. <i>Système 3D</i> .....	53
1.3.2.1. Organoïdes et sphéroïdes .....	53
1.3.2.2. Modèles de xénogreffe murin.....	56
1.3.2.3. Culture organotypique ex-vivo.....	57
1.3.2.4. Les tumeurs sur puce : avenir des modèles de culture tumoraux ?.....	59
1.4. DE L'ÉVOLUTION DE LA TRANSCRIPTOMIQUE .....	61
1.4.1. <i>Le dogme central de la biologie moléculaire et les disciplines –omiques</i> .....	61
1.4.2. <i>La transcriptomique « bulk » et ses technologies</i> .....	63
1.4.2.1. Les débuts de la transcriptomique : EST & SAGE .....	63
1.4.2.2. Les puces à ADN .....	65
1.4.2.3. Le séquençage transcriptomique haut-débit .....	67
1.4.2.3.1. Le RNA-sequencing .....	67
1.4.2.3.2. Le séquençage « long-read » .....	68
1.4.2.3.3. BRB-seq.....	69
1.4.3. <i>La transcriptomique à l'échelle de la cellule unique ou single-cell RNA-seq</i> .....	70
1.4.3.1. Les technologies.....	70
1.4.3.1.1. Les premières méthodes.....	70
1.4.3.1.2. Le Chromium de 10X Genomics .....	72
1.4.3.1.2.1. Capture cellulaire et préparation de librairie de séquençage .....	72

1.4.3.1.2.2. Identifiant moléculaire unique ou « UMI ».....	77
1.4.3.1.3. Technologies émergentes et élargissement des protocoles .....	77
1.4.3.1.3.1. Single-nuclei RNA-sequencing .....	78
1.4.3.1.3.2. CITE-seq .....	78
1.4.3.1.3.3. Sci-RNA-seq .....	81
1.4.3.1.3.4. La transcriptomique spatiale .....	81
1.4.3.1.3.5. LIVE-seq .....	83
1.4.3.2. L'analyse de données de scRNA-seq .....	84
1.4.3.2.1. Contrôle qualité .....	85
1.4.3.2.2. Normalisation .....	86
1.4.3.2.3. Correction des biais techniques .....	86
1.4.3.2.4. Sélection des gènes variables et réduction de dimensions.....	87
1.4.3.2.5. Classification non supervisée .....	89
1.4.3.2.6. Expression différentielle .....	89
1.4.3.2.7. Inférence de CNVs .....	90
1.4.3.2.8. Déconvolution.....	91
1.4.3.2.9. Lignage cellulaire et reconstruction de pseudotemps .....	92
1.5. DE LA TRANSCRIPTOMIQUE DANS L'ÉTUDE DES CCRCCS.....	95
1.5.1. États des lieux et enjeux expérimentaux.....	95
1.5.2. Apports des approches bulk.....	96
1.5.2.1. RT-qPCR.....	96
1.5.2.2. Pucés à ADN .....	97
1.5.2.3. Séquençage ARN massif .....	98
1.5.3. Single-cell RNA-seq en oncologie et dans l'étude des ccRCCs.....	99
1.5.3.1. scRNA-seq en oncologie .....	99
1.5.3.1.1. Caractérisation des tumeurs à l'échelle de la cellule unique .....	99
1.5.3.1.2. Apports du scRNA-seq en oncologie .....	100
1.5.3.1.2.1. Lignage cellulaire, différenciation et plasticité .....	100
1.5.3.1.2.2. Classification tumorale .....	101
1.5.3.1.2.3. Microenvironnement tumoral .....	101
1.5.3.1.2.4. Processus métastatique.....	102
1.5.3.1.2.5. Réponse aux thérapies .....	103
1.5.3.2. scRNA-seq dans le ccRCC .....	104
1.5.3.2.1. Caractérisation du compartiment immunitaire .....	105
1.5.3.2.2. Prédiction du pronostic clinique .....	106
1.5.3.3. Approches émergentes en oncologie.....	108
<b>2. CONTEXTE ET OBJECTIFS .....</b>	<b>111</b>
2.1. CONTEXTE DU LABORATOIRE .....	113
2.2. CONTEXTE SCIENTIFIQUE .....	113
2.3. HYPOTHESE DE TRAVAIL .....	114

2.4. OBJECTIFS DE MES TRAVAUX DE THESE DE DOCTORAT .....	114
<b>3. MATÉRIEL ET METHODES .....</b>	<b>117</b>
3.1. PATIENTS .....	119
3.2. MODELE DE CULTURE ORGANOTYPIQUE DE COUPES DE CCRCC .....	119
3.3. TRAITEMENTS .....	120
3.4. BIOLOGIE CELLULAIRE .....	121
3.4.1. <i>Dissociation tissulaire</i> .....	121
3.4.2. <i>Congélation et décongélation de cellules</i> .....	121
3.4.3. <i>Multiplexage par cell-hashing</i> .....	121
3.4.4. <i>Cytométrie</i> .....	122
3.4.5. <i>Capture de cellules uniques</i> .....	122
3.5. BIOLOGIE MOLÉCULAIRE .....	123
3.5.1. <i>Marquage immuno-histochimique</i> .....	123
3.5.2. <i>Librairies de séquençage</i> .....	123
3.6. TRAITEMENT BIO-INFORMATIQUE .....	123
3.6.1. <i>Prétraitement, contrôles de qualité, intégration et normalisation</i> .....	124
3.6.2. <i>Réduction de la dimensionnalité et classification</i> .....	124
3.6.3. <i>Annotation des clusters et identification des gènes différentiellement exprimés</i> .....	125
3.6.4. <i>Analyses</i> .....	125
3.6.4.1. <i>Analyse pseudo-bulk</i> .....	125
3.6.4.2. <i>Analyse des variations du nombre de copies</i> .....	125
3.6.4.3. <i>Inférence de trajectoire de cellule unique et analyse de pseudo temps</i> .....	126
3.6.4.4. <i>Analyses fonctionnelles</i> .....	126
3.6.4.5. <i>Traitement des données du TCGA-KIRC et analyses de survie</i> .....	127
3.6.4.5.1. <i>Données d'ARN bulk</i> .....	127
3.6.4.5.2. <i>Analyse de survie</i> .....	127
3.6.4.5.3. <i>Sélection d'ensembles de gènes favorables et défavorables</i> .....	127
3.6.5. <i>Dépôt des données</i> .....	127
<b>4. RÉSULTATS.....</b>	<b>129</b>
4.1. ATLAS DE L'HETEROGENEITE DANS LE CCRCC ET VALEUR PRONOSTIQUE .....	131
4.1.1. <i>Introduction</i> .....	131
4.1.2. <i>Manuscrit soumis</i> .....	133
4.1.3. <i>Discussion et perspectives</i> .....	191
4.2. ÉTUDE DES IMPACTS D'INHIBITEURS DE KINASES SUR L'HETEROGENEITE INTRA-TUMORALE DU CCRCC PAR UNE APPROCHE COMBINANT UN MODÈLE DE CULTURE 3D ET SINGLE-CELL RNA-SEQ.....	195
4.2.1. <i>Introduction</i> .....	195
4.2.2. <i>Résultats supplémentaires - mises au point expérimentales</i> .....	197

4.2.2.1. Conception et contraintes expérimentales .....	197
4.2.2.2. Résultats de mises au point .....	199
4.2.3. <i>Manuscrit en préparation</i> .....	201
4.2.4. <i>Discussion et perspectives</i> .....	237
4.3. UNCOVER : UNE RESSOUCE WEB DE DONNEES GENOMIQUES DEDIEE AUX CANCERS UROLOGIQUES.....	241
4.3.1. <i>Introduction</i> .....	241
4.3.2. <i>Manuscrit en préparation</i> .....	243
4.3.3. <i>Discussion et perspectives</i> .....	269
<b>5. DISCUSSION</b> .....	<b>271</b>
<b>BIBLIOGRAPHIE</b> .....	<b>287</b>
<b>ANNEXES</b> .....	<b>319</b>
5.1. REVUE DE LA LITTERATURE - 1 .....	321
5.2. REVUE DE LA LITTERATURE -2 .....	343





## **LISTE DES TABLEAUX**

<b>TABLEAU 1.</b>	<b>CLASSIFICATION OMS 2016 DES TUMEURS RENALES .....</b>	<b>7</b>
<b>TABLEAU 2.</b>	<b>DESCRIPTION DES STADES DE LA CLASSIFICATION TNM.....</b>	<b>14</b>
<b>TABLEAU 3.</b>	<b>ÉCHELLE DE STATUT DE PERFORMANCE ECOG-PS .....</b>	<b>16</b>
<b>TABLEAU 4.</b>	<b>EXEMPLE DE MODELES PRONOSTIQUES DANS LE CCRCC ET METHODES DE CALCUL..</b>	<b>19</b>
<b>TABLEAU 5.</b>	<b>CNV LES PLUS FREQUENTS DANS LE CCRCC .....</b>	<b>24</b>
<b>TABLEAU 6.</b>	<b>MOLECULES POSSEDANT UNE AMM EN EUROPE EN 2021 .....</b>	<b>36</b>
<b>TABLEAU 7.</b>	<b>COMPARAISON DES SYSTEMES DE CULTURE EN ONCOLOGIE .....</b>	<b>55</b>
<b>TABLEAU 8.</b>	<b>CARACTERISTIQUES DES MOLECULES UTILISEES DANS LE PROJET DE THESE LIE AUX IMPACTS DES TRAITEMENTS.....</b>	<b>120</b>



# LISTE DES FIGURES

FIGURE 1.	ANATOMIE DU REIN.....	3
FIGURE 2.	LE NEPHRON .....	5
FIGURE 3.	ASPECT MACROSCOPIQUE D'UN CARCINOME RENAL A CELLULES CLAIRES.....	8
FIGURE 4.	HISTOLOGIE DU CCRCC.....	8
FIGURE 5.	INCIDENCE ET MORTALITE LIEES AUX PRINCIPAUX CANCERS, TOUS AGES ET SEXES CONFONDUS	9
FIGURE 6.	STADES DE DEVELOPPEMENT DU CCRCC.....	13
FIGURE 7.	CLASSIFICATION DE L'ISUP POUR LE CCRCC.....	15
FIGURE 8.	COURBE DE SURVIE KAPLAN-MEIER REPRESENTATIVE.....	20
FIGURE 9.	VOIES MOLECULAIRES DANS LE CCRCC .....	21
FIGURE 10.	GENES MUTES DANS LE CCRCC.....	23
FIGURE 11.	LES « HALLMARKS », OU CARACTERISTIQUES FONDAMENTALES DU CANCER.....	26
FIGURE 12.	IMPLICATION DE L'EMT DANS LE PROCESSUS METASTATIQUE .....	26
FIGURE 13.	ILLUSTRATION DU MICROENVIRONNEMENT TUMORAL .....	29
FIGURE 14.	ARBRE DE DECISION DANS LA PRISE EN CHARGE MEDICALE DU CARCINOME RENALE A CELLULES CLAIRES .....	32
FIGURE 15.	GUIDELINES DANS LE TRAITEMENT DU CCRCC. ....	34
FIGURE 16.	MECANISMES D' ACTIONS DES INHIBITEURS DE CHECKPOINT IMMUNITAIRE.....	37
FIGURE 17.	NOTION D'HETEROGENEITE TUMORALE .....	39
FIGURE 18.	EVOLUTION DU CONCEPT D'HETEROGENEITE INTRA-TUMORALE. ....	40
FIGURE 19.	MODELES D'EVOLUTION TUMORALE .....	40
FIGURE 20.	MODELES DE TUMORIGENESE .....	46
FIGURE 21.	HETEROGENEITE INTRA-TUMORALE A L'ECHELLE MACRO- ET MICROSCOPIQUE.....	48
FIGURE 22.	ARBRE PHYLOGENETIQUE DANS LE CCRCC.....	49
FIGURE 23.	ÉVOLUTION ET HETEROGENEITE TUMORALE DANS LE CCRCC .....	50
FIGURE 24.	DIFFERENTS SYSTEMES DE CULTURE TUMORALE .....	54
FIGURE 25.	CULTURE ORGANOTYPIQUE DE COUPES. ....	57
FIGURE 26.	SCHEMA SIMPLIFIE DU FLUX DE L'INFORMATION GENETIQUE.....	60

FIGURE 27.	ORGANISATION ET TRANSFERT DE L'INFORMATION ENTRE LE GENOME, L'EPIGENOME, LE TRANSCRIPTOME ET LE PROTEOME .....	60
FIGURE 28.	ÉVOLUTION DE L'UTILISATION DES PRINCIPALES TECHNIQUES TRANSCRIPTOMIQUES 62	
FIGURE 29.	FONCTIONNEMENT DES PUCES A ADN.....	64
FIGURE 30.	SCHEMA DU PROTOCOLE DE SEQUENÇAGE ARN SUR UNE PLATEFORME ILLUMINA..	66
FIGURE 31.	ILLUSTRATION DU MULTIPLEXAGE REALISE EN BRB-SEQ .....	69
FIGURE 32.	ÉVOLUTION DES TECHNOLOGIES ET DEBIT DE SINGLE-CELL RNA-SEQ .....	71
FIGURE 33.	PROCESSUS DE CAPTURE DU CHROMIUM 10X.....	73
FIGURE 34.	ÉTAPES DE PREPARATION DE LA LIBRAIRIE DE SCRNA-SEQ .....	74
FIGURE 35.	STRUCTURE DES ADNC A L'ISSUE DU PROTOCOLE CHROMIUM DE 10X GENOMICS ...	74
FIGURE 36.	L'IDENTIFIANT MOLECULAIRE UNIQUE AMELIORE LA QUANTIFICATION.....	76
FIGURE 37.	RESUME GRAPHIQUE DE DIVERSES TECHNOLOGIES MULTI-OMIQUES « <i>SINGLE-CELL</i> » 76	
FIGURE 38.	SCHEMA DU PROTOCOLE DE CITE-SEQ.....	79
FIGURE 39.	SCHEMA DU PROTOCOLE DE SCI-RNA-SEQ.....	80
FIGURE 40.	ILLUSTRATION DE L'APPROCHE DE TRANSCRIPTOMIQUE SPATIALE.....	82
FIGURE 41.	SCHEMA DE L'ANALYSE EN SCRNA-SEQ.....	84
FIGURE 42.	REPRESENTATION DE TYPE TSNE ET UMAP D'UN ENSEMBLE DE CELLULES .....	88
FIGURE 43.	ILLUSTRATION DE LA DECONVOLUTION TRANSCRIPTOMIQUE.....	91
FIGURE 44.	ILLUSTRATION DE L'INFERENCE DE TRAJECTOIRES UTILISEE EN SINGLE-CELL RNA-SEQ 93	
FIGURE 45.	APPROCHES SINGLE-CELL RECENTES ET QUESTIONS BIOLOGIQUES ET CLINIQUES EMERGENTES EN ONCOLOGIE.....	108
FIGURE 46.	MISES AU POINT DU PROTOCOLE DE CULTURE ORGANOTYPIQUE DE COUPES DE CCRCC	196
FIGURE 47.	ÉVALUATION DE LA QUALITE DE DISSOCIATION DE COUPES DE CCRCC APRES 48H DE CULTURE	198

## LISTE DES ABREVIATIONS

**ACP** : Analyse en composante principale  
**ADN** : Acide DésoxyriboNucléique  
**ADNc** : ADN complémentaire  
**AED** : Analyse d'expression différentielle  
**AJCC** : American Joint Commitee on Cancer  
**AMM** : Autorisation de mise sur le marché  
**ARN** : Acide RiboNucléique  
**ARNlc** : ARN long codant  
**ARNlnc** : ARN long non codant  
**ARNm** : ARN messenger  
**ARNnc** : ARN non codant  
**ARNr** : ARN ribosomique  
**BRB-seq** : Bulk RNA Barcoding and sequencing  
**CAF** : Cancer Associated Fibroblasts  
**CCA** : Canonical correlation analysis  
**CCAFU** : Comité de cancérologie de l'Association française d'urologie  
**ccRCC** : Clear cell Renal Cell Carcinoma (Carcinome rénal à cellules claires)  
**CDX** : Cell-line Derived Xenografts  
**chRCC** : Chromophobe renal cell carcinoma (Carcinome Rénal Chromophobe)  
**CHU** : Centre hospitalier universitaire  
**CITE-seq** :  
**CNV** : Copy number variation (Variation du nombre de copies)  
**CR** : Carcinome rénal  
**CRB** : Centre de Ressources Biologiques  
**CSC** : Cellule souches cancéreuses  
**CTC** : Cellules tumorales circulantes  
**CTL** : Lymphocytes T cytotoxiques  
**CytoF** : Cytometry by Time of Flight  
**DGE** : Digital Gene Expression  
**DMEM** : Dulbecco's Modified Eagle Medium  
**DMSO** : Diméthylsulfoxyde  
**EAU** : European Association of Urology  
**EBM** : Endothelial Cell Growth Basal Medium  
**ECOG-PS** : Eastern Cooperative Oncology Groupe – Performance Status  
**EGA** : European Genome-Phenome Archive  
**EGF** : Epidermal Growth Factor

**EGM** : Endothelial cell growth factor  
**EMA** : European Medicines Agency  
**EMT** : Epithelio Mesenchymal Transition  
**EST** : Expressed Sequenced Tag  
**FFPE** : Formalin fixed and paraffin embedded  
**FGF** : Fibroblasts Growth Factor (Facteur de croissance fibroblastique)  
**FluidFM** : Fluidic force microscopy  
**G&T-seq** : Genome & Transcriptome sequencing  
**HCA** : Human Cell Atlas  
**HIF** : Hypoxia Inducible Factor  
**HIT** : Hétérogénéité intra-tumorale  
**HNSCC** : Head and neck squamous cell carcinoma  
**HR** : Hazard Ratio  
**IC** : Intervalle de Confiance  
**ICI** : Immune Checkpoint Inhibitor  
**IKCS** : International Kidney Cancer Symposium  
**IL** : Interleukine  
**ILK** : Integrin-linked kinase  
**IMDC** : International Metastatic Database Consortium  
**INCa** : Institut National du Cancer  
**IRM** : Imagerie par résonance magnétique  
**IRSET** : Institut de Recherche en Santé, Environnement et Travail  
**ISUP** : International Society of Urological Pathology  
**ITK** : Inhibiteurs de tyrosine kinase  
**KIRC** : Kidney Renal clear Carcinoma  
**KO** : Knock-Out  
**LCR** : Liquide céphalo-rachidien  
**LOH** : Loss of heterozygosity  
**LTC** : Lymphocytes T cytotoxiques  
**MAST** : Model-based Analysis of Single-cell Transcriptomics  
**MET** : Microenvironnement tumoral  
miARN : micro-ARN  
**MRCA** : Most recent common ancestor  
mRCC : ccRCC métastatique  
**MuSiC** : Multi-subject Single-cell Deconvolution  
**NCI** : National Cancer Institute  
**NHGRI** : National Human Genome Research Institute  
**NP** : Néphrectomie partielle  
**NT** : Néphrectomie totale

**OMS** : Organisation mondiale de la santé  
**PCR** : Polymerase Chain Reaction  
**PDGF** : Platelet-Derived Growth Factor  
**PDTSC** : Patient Derived Tumor Slice Culture  
**PDX** : Patient Derived Xenografts  
**PLCO** : Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial  
**pRCC** : papillary renal cell carcinoma (Carcinome Rénal Papillaire)  
**RCC** : Renal Cell Carcinoma  
**RNA** : Ribosomal Ribonucleic Acid  
**RNA – seq** : RNA-sequencing  
**ROS** : Reactive Oxygen Species  
**RPCA** : Reciprocal canonical correlation analysis  
**RPMI** : Roswell Park Memorial Institute  
**RT** : Radiothérapie  
**RT-PCR** : Reverse transcription polymerase chain reaction  
**RT-qPCR** : Reverse transcription quantitative polymerase chain reaction  
**SAGE** : Serial Analysis of Gene Expression  
**scATAC-seq** : Single-cell Assay for Transposase-Accessible Chromatin  
**sci-RNA-seq** : Single-cell Combinatorial Indexing RNA-sequencing  
**scNMT-seq** : single-cell Nucleosome, Methylation and Transcription sequencing)  
**scRNA-seq** : single-cell RNA-seq ; séquençage des ARNs à l'échelle de la cellule unique  
**SEER** : Surveillance, Epidemiology and End Results  
**SNP** : Single nucleotid polymorphism  
**snRNA-seq** : Single nuclei RNA-seq ; séquençage ARN sur noyaux uniques  
**TAM** : Tumor Associated Macrophages  
**TCGA** : The Cancer Genome Atlas  
**TCGA-KIRC** : The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma  
**TCR** : T Cell Receptors  
**TGF** : Tumor Growth Factor (Facteur de croissance tumorale)  
**TKI** : Inhibiteurs de tyrosine kinase  
**TMB** : Tumor Mutational Burden  
**TME** : Tumor Micro Environment (Microenvironnement Tumoral)  
**TNM** : Tumor-Node-Metastasis

**TOC** : Tumor-on-chip  
**TRACERx** :  
 t-SNE : t-distributed stochastic neighbor embedding  
**UICC** : Union for International Cancer Control  
**UMAP** : Uniform Manifold Approximation and Projection for Dimension Reduction  
**UMI** : Unique molecular identifier  
**UncoVer** : Uro-oncogenomics Viewer  
**UTR** : Untranslated Region  
**VEGF** : Vascular Endothelial Growth Factor (Facteur de croissance des cellules endothéliales)  
**VHL** : Von-Hippel Lindau  
**VITAL** : Vitamin and lifestyle  
**WHO/ISUP** : World Health Organization / International Society of Urological Pathology

# **1.INTRODUCTION**





## 1.1. DES CARCINOMES RÉNAUX À CELLULES CLAIRES

### 1.1.1. PREMIERE DESCRIPTION DE CANCER DU REIN

Dans son texte « *Practicae Medicinae* » publié pour la première fois en 1613, le médecin allemand Daniel Sennert est le premier à suggérer une tumeur survenant dans le rein. Cette observation sera remise en cause en 1841 par le pathologiste Pierre Rayet dans l'introduction de sa série d'observations de tumeurs rénales (Delahunt and Eble, 2005). Il suggère que les commentaires de Sennert font référence à une inflammation chronique sclérosante du rein plutôt qu'à une tumeur maligne. Le premier cas confirmé de description de carcinome rénal a été rapporté par Miriel en 1810 et concernait Françoise Levelly, une femme de 35 ans reçue à l'hôpital de Brest. C'est dans le même temps, au cours du 19<sup>ème</sup> siècle, que commencent à émerger les premières connaissances liées à la physiologie rénale.

### 1.1.2. À L'ORIGINE, LE REIN

Les reins sont des organes vitaux appartenant, avec les uretères et la vessie, au système urinaire (Figure 1). Ils ont pour fonction principale d'exfiltrer les déchets métaboliques du sang vers l'urine. Situés dans l'espace retro-péritonéal, de part et d'autre de la colonne vertébrale ; ils ont la forme d'un haricot et ont une dimension moyenne de 12cm x 6cm x 3cm (Wallace, 1998).

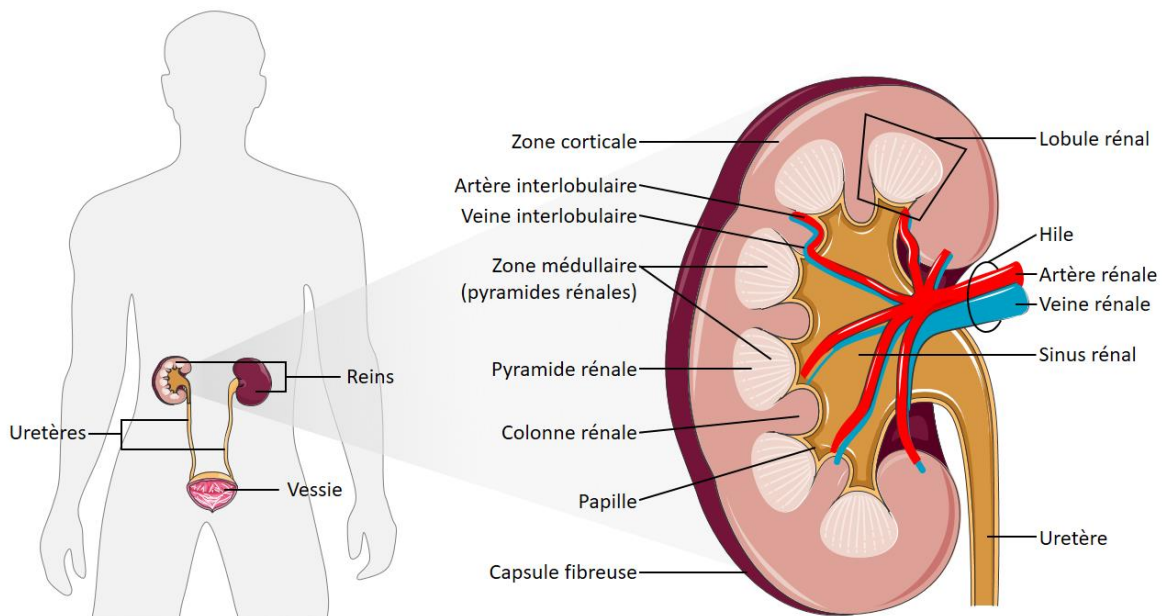


Figure 1. Anatomie du rein

Le schéma présente à gauche la situation anatomique des reins et, à droite, une coupe longitudinale de rein sain avec les principaux constituants assurant la fonction rénale.

### **1.1.2.1. Morphologie**

Le rein est composé du parenchyme dans lequel s'effectue la filtration, et du sinus rénal qui collecte les urines via les papilles rénales et les redirige vers l'uretère (Wallace, 1998) (Figure 2). L'ensemble est recouvert et maintenu par une capsule fibreuse, aussi appelée capsule rénale. Le parenchyme comprend deux parties ; 1 – la zone médullaire vers l'intérieur, constituée des pyramides rénales (ou pyramides de Malpighi) et, 2 – la zone corticale, en périphérie, qui constitue la zone de filtration glomérulaire. La partie corticale qui s'insinue entre les pyramides rénales est appelée colonne rénale. L'ensemble formé par une pyramide rénale et le cortex environnant est appelé lobule rénal.

### **1.1.2.2. Micro-anatomie**

C'est à la jonction entre cortex et médulla que l'on trouve les néphrons ; les unités fonctionnelles du rein (Mahadevan, 2019) (Figure 2). Le néphron comprend deux parties ; le corpuscule rénal et le tubule rénal. Le corpuscule rénal est situé dans le cortex et comprend le glomérule et la capsule de Bowman. Le tubule rénal est divisé en trois segments ; 1 – le tubule contourné proximal dans le cortex, proche du corpuscule rénal, 2 – les tubules droits ascendant et descendant qui forment l'anse de Henlé, dans la médulla et, 3 – le tube contourné distal, qui rejoint le tube collecteur. Un tube collecteur est ainsi relié à plusieurs néphrons et évacue l'urine vers les papilles rénales, en direction du sinus rénal. Un lobule rénal compte entre cinquante mille et cent mille néphrons, tandis qu'un rein en compte environ un million.

### **1.1.2.3. Vascularisation**

Le rein possède une vascularisation parmi les plus importantes et présente des caractéristiques spécifiques (Gomez, Norwood and Tufro-McCreddie, 1997). L'artère rénale, née de l'aorte abdominale, entre dans le rein par le hile (Figure 1B) et se ramifie une première fois en artères interlobulaires. Ces dernières se divisent ensuite en artérioles afférentes qui mènent chacune à un groupe de capillaires glomérulaires (Figure 2). À la sortie du glomérule, les capillaires glomérulaires se rejoignent à nouveau pour former les artérioles efférentes. Enfin, ces dernières se mêlent aux capillaires péri-tubulaires. Les réseaux capillaires glomérulaires et péri-tubulaires dans le néphron sont une caractéristique unique de la physiologie rénale. Ils permettent, pour les premiers, la filtration de grandes quantités de liquides et de solutés et, pour

les seconds, la filtration, sécrétion et réabsorption de minéraux ainsi que l'élimination de déchets métaboliques du sang vers l'urine. Les capillaires péri-tubulaires sont suivis par le système veineux qui se prolonge en parallèle des vaisseaux artériels et forme progressivement les veines interlobulaires puis la veine rénale. Celle-ci se déverse enfin dans la veine cave inférieure.

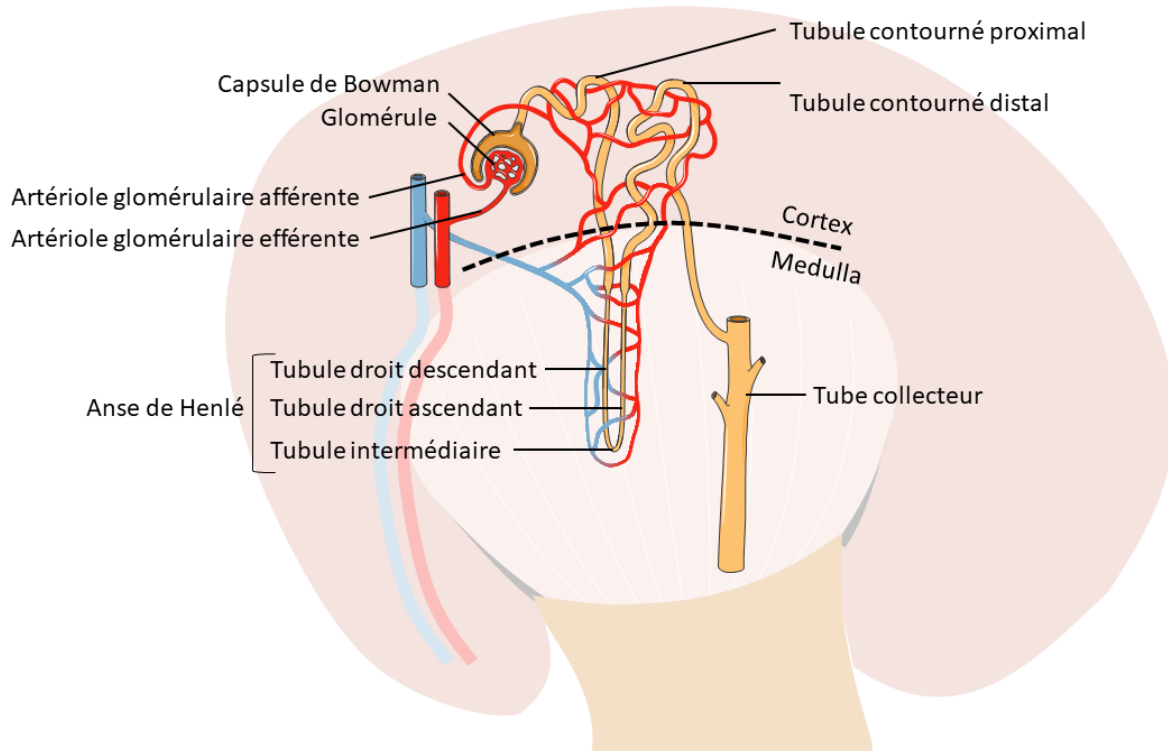


Figure 2. Le néphron

Ce schéma représente une coupe longitudinale de lobule rénal dans lequel est représenté un néphron. Le néphron n'est pas à l'échelle. Le glomérule et les tubules contournés se trouvent dans la partie corticale tandis que l'anse de Henlé se trouve dans la médulla.

### 1.1.3. CLASSIFICATION DES CARCINOMES RENAUX

Le cancer du rein désigne toute formation tumorale maligne prenant naissance dans le rein à partir du système tubulaire. Il survient quand les cellules de l'organe prolifèrent de façon incontrôlée. En fonction du compartiment anatomique, de particularités moléculaires, ou encore de l'âge auquel il survient, le cancer du rein se subdivise en de nombreux sous-types, répertoriés dans la classification de l'Organisation mondiale de la santé (OMS) en 2016 (Moch *et al.*, 2016) (Tableau 1). La majorité (90%) des cancers du rein sont des carcinomes rénaux, un sous-groupe hétérogène de tumeurs cancéreuses issues des cellules épithéliales du rein, généralement des

cellules des tubules proximaux ou distaux (Hsieh *et al.*, 2017). Les 10 % restant comprennent, 1 – les cancers urothéliaux, prenant naissance à partir de l'urothélium (5-10% des cancers du rein) ; 2 – les tumeurs de Wilms ou néphroblastomes, survenant chez l'enfant, et ; 3 – les sarcomes rénaux qui se forment à partir des cellules endothéliales ou du tissu conjonctif (< 1% des cancers du rein). Le carcinome rénal se subdivise à son tour en trois groupes majeurs (Chen *et al.*, 2016) ; le carcinome rénal à cellules claires (75%), le carcinome papillaire (15%) et le carcinome rénal chromophile (5%). Les 5% restant constituent des tumeurs rares (< 1% d'incidence) (Moch *et al.*, 2016).

Le carcinome rénal à cellules claires (ccRCC) provient des cellules épithéliales des tubules contournés proximaux (Muglia and Prando, 2015) (Figure 2). Macroscopiquement, le ccRCC est d'apparence jaune chamois souvent hémorragique (Figure 3). Il constitue une masse tumorale d'en moyenne sept centimètres prenant naissance en périphérie du cortex rénal. La tumeur est souvent distincte du rein sain et bien circonscrite par une pseudo-capsule. Elle peut présenter des zones nécrotiques ou des remaniements hémorragiques ou œdémateux. Le ccRCC est caractérisé par un cytoplasme riche en lipides et en glycogène, qui apparaît translucide sur les coupes histologiques, d'où la dénomination « cellules claires » (Muglia and Prando, 2015) (Figure 4). Son architecture peut varier entre des modèles solides, alvéolaires et/ou acinaires, et contient fréquemment un réseau dense de vaisseaux à parois fines. Les tumeurs de haut grade peuvent contenir des cellules au cytoplasme éosinophile entrecoupées de zones de nécrose ou d'hémorragie. Le ccRCC n'est pas seulement le plus fréquent mais également le plus létal parmi les carcinomes rénaux (Rabjerg, 2017; Shingarev and Jaimes, 2017). Il constitue le modèle d'étude dans ce travail de thèse.

<b>Tumeurs malignes à cellules rénales</b>	<b>Tumeurs mésenchymateuses de l'adulte</b>
Carcinome à cellules claires du rein	Léiomyosarcome
Néoplasie kystique multiloculaire rénale	Angiosarcome
Carcinome rénal papillaire	Rhabdomyosarcome
Carcinome rénal associé à la HLRCC	Ostéosarcome
Carcinome à cellules chromophores	Sarcome d'Ewing
Carcinome des tubes collecteurs	Angiomyolipome
Carcinome du rein médullaire	Angiomyolipome épithélioïde
Carcinome associé aux translocations MITF	Léiomyome
Carcinome lié à un déficit en succinate déshydrogénase	Hémangiome
Carcinome tubuleux et mucineux à cellules fusiformes	Lymphangiome
Carcinome tubulo-kystique	Hémangioblastome
Carcinome associé à la maladie kystique acquise	<b>Tumeur à cellules juxtaglomérulaires</b>
Carcinome rénal papillaire à cellules claires	Tumeur interstitielle rénale médullaire
Adénome papillaire	Schwannome
Carcinome non classés	Tumeur fibreuse solitaire
<b>Tumeurs métanéphriques</b>	<b>Tumeurs mésenchymateuses de l'enfant</b>
Adénome métanéphrique	Sarcome à cellules claires
Adénofibrome métanéphrique	Tumeur rhabdoïde
<b>Tumeurs stromales métanéphriques</b>	Néphrome mésoblastique congénital
Tumeur néphroblastique	Tumeur rénale ossifiante de l'enfant
Reste néphrogénique	<b>Tumeurs neuroendocrines</b>
Néphroblastome	Tumeur neuroendocrine différenciée
Néphroblastome kystique partiellement différencié	Carcinome neuroendocrine à grandes cellules
Néphrome kystique pédiatrique	Phéochromocytome
<b>Tumeurs mixtes épithéliales et mésenchymateuses</b>	<b>Tumeurs diverses</b>
Néphrome kystique	Néoplasme rénal hématopoïétique
Tumeur mixte épithéliale et stromale	Tumeur des cellules germinales
<b>Tumeur métastatiques</b>	<b>Tumeurs mésenchymateuses</b>

Tableau 1. Classification OMS 2016 des tumeurs rénales

Il existe de nombreux types et sous-types de cancers, parmi lesquels les carcinomes sont les plus fréquents et nombreux. Inspiré de Moch et al., 2016.



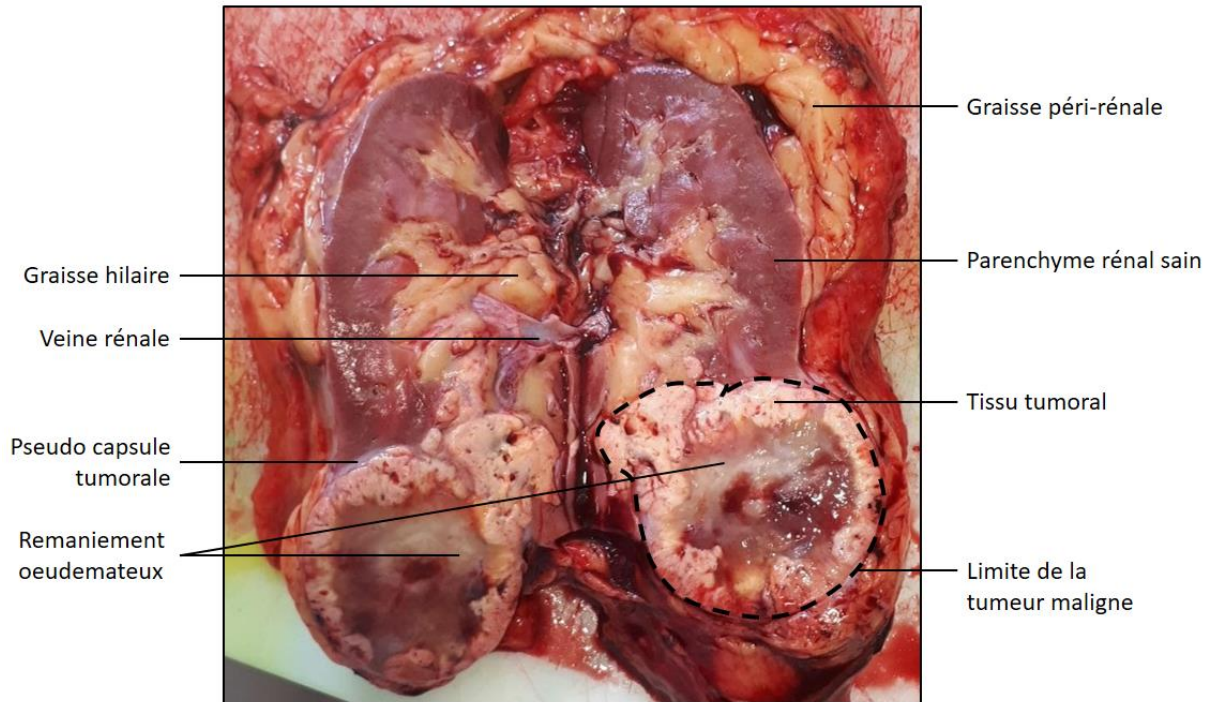


Figure 3. Aspect macroscopique d'un carcinome rénal à cellules claires

Sur cette photo, la tumeur est de stade II et ne dépasse pas le parenchyme rénal. On observe par ailleurs de la graisse péri-rénale infiltrant le hile, au centre de l'image, ainsi que des remaniements œdémateux au sein de la tumeur. Le tissu tumoral est d'apparence jaune chamois/beige, comme le sont typiquement les carcinomes rénaux.

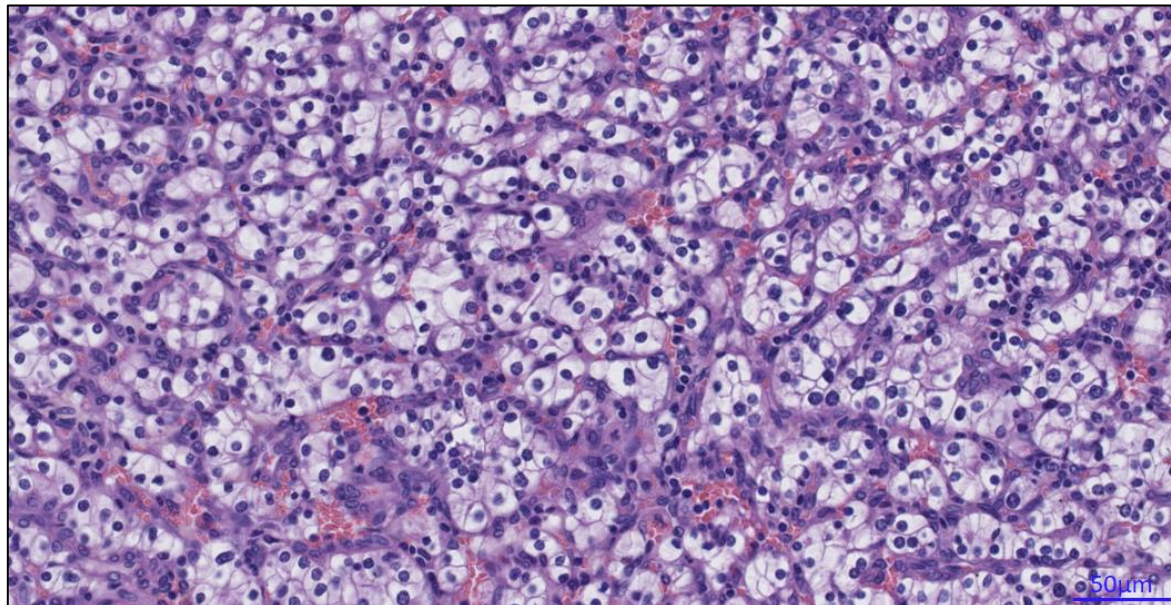


Figure 4. Histologie du ccRCC

Sur cette coupe histologique colorée à l'hématoxyline et à l'éosine, on distingue de nombreux îlots ovoïdes correspondant à des tubules coupés transversalement et entourés de stroma et de vaisseaux sanguins. Le stroma des cellules néoplasiques est clair en raison d'une accumulation en glycogène, d'où l'appellation du carcinome rénal à « cellules claires ».

### 1.1.4. ÉPIDÉMIOLOGIE

La majorité des données épidémiologiques disponibles concernent généralement le cancer du rein ou les carcinomes rénaux sans distinction de sous-types histologiques. Les chiffres d'incidence et de mortalité varient selon les pays et selon les méthodes d'analyse dans les études. Dans cette section sont présentés les chiffres de 2020 issues de l'OMS.

#### 1.1.4.1. Incidence et mortalité du cancer du rein

Le cancer du rein occupe la 14<sup>e</sup> place en terme d'incidence et la 15<sup>e</sup> place en terme de mortalité à travers le monde, tous âges et sexes confondus, avec respectivement 431 288 nouveaux cas et 179 368 décès en 2020 (H. Sung *et al.*, 2021) (Figure 5). En France, le cancer du rein est le 6<sup>e</sup> cancer le plus diagnostiqué, avec 14 705 nouveaux cas et 4 960 décès en 2020. Les données de l'Observatoire mondial des cancers (« *Global Cancer Observatory* » ou GLOBOCAN) de l'OMS montrent une légère augmentation des taux d'incidence et de mortalité entre 2012 et 2020 (Ferlay *et al.*, 2015; W. W. Sung *et al.*, 2021). En France de façon similaire selon les chiffres de l'INCa (Institut National du Cancer), l'incidence du cancer du rein a augmenté d'approximativement 1,6% entre 2010 et 2018 (Defossez *et al.*, 2019). Cette augmentation est en partie attribuée à l'amélioration de la performance des examens d'imagerie abdominale qui permettent d'augmenter le nombre de diagnostics précoces ou fortuits de formes asymptomatiques. Des données de la base SEER (Surveillance, Epidemiology, and End Results) recueillies entre 1973 et 2014 aux États-Unis permettent d'estimer un taux d'incidence du ccRCC chez les patients de plus de 40 ans d'environ 8.01 pour 100 000 (Feng *et al.*, 2019).

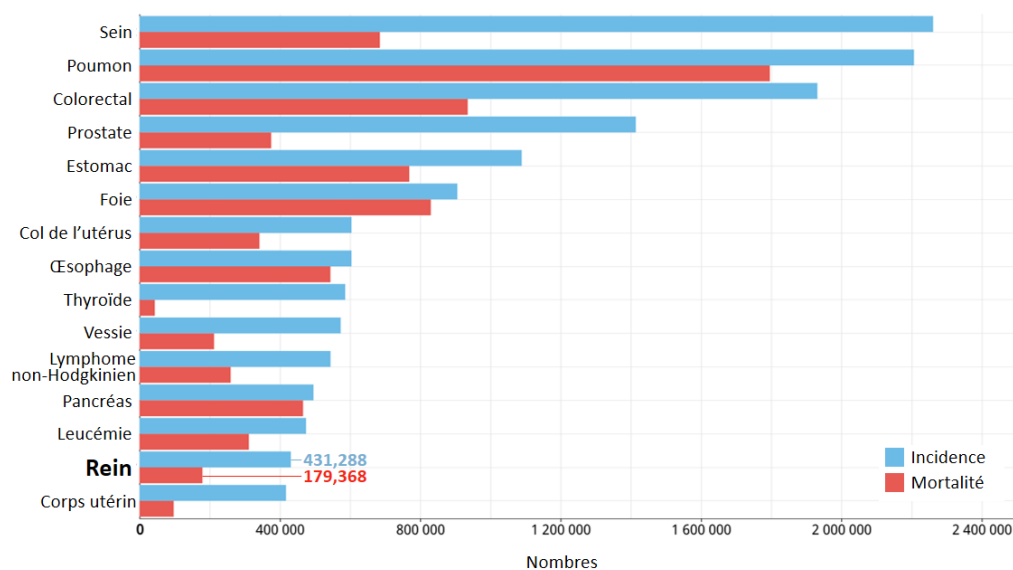


Figure 5. Incidence et mortalité liées aux principaux cancers, tous âges et sexes confondus

Données issues du GLOBOCAN 2020 (H. Sung *et al.*, 2021).



#### **1.1.4.2. Facteurs de risque**

L'incidence du cancer du rein est particulièrement corrélée avec le genre et l'âge. Le ratio standardisé en fonction de l'âge est de 0,5 cas pour 10 000 en dessous de 40 ans et augmente progressivement jusqu'à 35 pour 10 000 à partir de 75 ans (Capitanio *et al.*, 2019), tandis que le pic d'incidence se situe entre 60 et 70 ans. Cette tendance à l'augmentation de l'incidence avec l'âge s'observe d'ailleurs dans la majorité des cancers de l'adulte. L'incidence est plus élevée chez l'homme que chez la femme, avec un ratio de 2 pour 1 (Peired *et al.*, 2021). Ensuite, les facteurs de risque les mieux établis sont l'obésité, le tabagisme et l'hypertension artérielle, également en cause dans d'autres types de cancers (Lewandowska *et al.*, 2019). De nombreux autres facteurs ont été associés à un risque accru de cancer du rein mais sont moins bien établis, en raison de biais de sélection des patients ou d'inconsistance dans les données disponibles (Capitanio *et al.*, 2019). Deux études importantes comprenant les cohortes VITAL (Vitamin and Lifestyle) (White *et al.*, 2004) et PLCO (Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial) nous renseignent bien sur ces risques et leurs associations à la maladie (MacLeod *et al.*, 2013; Liao *et al.*, 2017).

##### **1.1.4.2.1. Tabagisme**

Le tabagisme est le facteur de risque le mieux établi dans le développement du carcinome rénal. La fumée de cigarette contient une variété de carcinogènes reconnus comme impliqués dans son étiologie. Dans l'étude VITAL, le tabagisme est associé de manière indépendante au carcinome rénal (>37,5 paquets par année contre 0 ; HR : 1,58, IC 95 % : 1,09-2,29) (Liao *et al.*, 2017). De même, dans la cohorte PLCO, l'intensité du tabagisme était significativement associée à un risque plus élevé de développer un carcinome rénal ainsi qu'une forme de haut grade (Lotan *et al.*, 2016). De plus, le risque relatif est directement associé à la durée de la consommation de tabac et diminue avec le temps après son arrêt. A cet égard, une méta-analyse de plus de 24 articles a démontré un risque relatif global d'incidence du carcinome rénal de 1,31 (1,22-1,40) pour tous les fumeurs, 1,36 (1,19-1,56) pour les fumeurs actuels et 1,16 (1,08-1,25) pour les anciens fumeurs. Les risques correspondants de mortalité spécifique au carcinome rénal sont respectivement de 1,23 (1,08-1,40), 1,37 (1,19-1,59) et 1,02 (0,90-1,15) (Cumberbatch *et al.*, 2016). Pour ces raisons, le Comité de cancérologie de l'association française d'urologie (CCAFU) recommande fortement l'arrêt du tabac (Bensalah *et al.*, 2020). Là encore, les mécanismes biologiques n'ont pas été clairement établis.

#### **1.1.4.2.2. Indice de masse corporelle**

L'obésité est communément associée à plusieurs cancers, y compris le carcinome rénal (Bergström *et al.*, 2001). Une étude sur la souris a démontré que le carcinome rénal pouvait être induit par un régime alimentaire riche en graisse et a pu confirmer l'induction de ces modifications sur coupes histologiques (Yang *et al.*, 2018). Dans l'étude VITAL, les réponses d'une cohorte américaine de 77 260 personnes âgées de 50 à 76 ans ont confirmé une association significative de l'obésité avec le carcinome rénal (MacLeod *et al.*, 2013). Le risque relatif correspondant à 5kg de poids corporel est estimé à 1.25 chez l'homme et 1.35 chez la femme, soit un sur-risque de respectivement 25% et 35% (MacLeod *et al.*, 2013; Gild, Ehdaie and Kluth, 2017; Liao *et al.*, 2017). Les mécanismes biologiques à l'origine de cette association ne sont pas bien établis, bien que des données récentes suggèrent l'implication des facteurs de croissance de type insulinique ou encore des adipokines, des hormones au potentiel pro-inflammatoire (Liao *et al.*, 2017). Les adipocytes, à travers la sécrétion d'adipokines, participeraient au remodelage du microenvironnement, favorisant la prolifération des cellules tumorales et l'angiogenèse (formation des vaisseaux sanguins), et réduisant l'immunité anti-tumorale (Gati *et al.*, 2014). Enfin, l'inflammation chronique et les hormones stéroïdiennes sexuelles sont d'autres facteurs causaux potentiels (Gild, Ehdaie and Kluth, 2017). La CCAFU dans son dernier rapport préconise l'activité physique et la perte de poids pour les patients obèses (Bensalah *et al.*, 2020). De façon intéressante, l'obésité est un facteur de risque établi dans le ccRCC, mais également un facteur de bon pronostic chez les personnes atteintes de ccRCC. Bien que des hypothèses avancent l'implication d'adiponectines comme facteurs protecteurs (Ito *et al.*, 2017), ce phénomène dit de « paradoxe de l'obésité » reste encore peu compris.

#### **1.1.4.2.3. Hypertension artérielle**

L'hypertension artérielle a également été identifiée comme pouvant favoriser le carcinome rénal (MacLeod *et al.*, 2013; Chien *et al.*, 2017; Mazzucotelli *et al.*, 2017), ce en quoi le contrôle de la tension des patients hypertendus semble être bénéfique (Weikert *et al.*, 2008). Dans l'étude VITAL, l'hypertension est significativement associée au risque de carcinome rénal (HR : 1,70, IC 95 % : 1,30-2,22). Une méta-analyse récente confirme cette association positive entre l'hypertension et le risque de cancer du rein (Hidayat *et al.*, 2017). Concernant les mécanismes biologiques, certains chercheurs ont émis l'hypothèse que les personnes hypertendues souffrent d'une hypoxie rénale chronique qui favoriserait la

prolifération des cellules tumorales et l'angiogenèse (Gago-Dominguez *et al.*, 2002; Sharifi and Farrar, 2006).

#### **1.1.4.2.4. Insuffisance rénale chronique**

L'insuffisance rénale augmente également le risque de développer un cancer du rein (Stengel, 2010). Des études ont démontré que le pronostic de ces patients était meilleur en comparaison des patients non-insuffisants rénaux (Shrewsberry *et al.*, 2014). Ils présentent des carcinomes de plus petite taille généralement asymptomatiques et de bas grade (Neuzillet *et al.*, 2011).

#### **1.1.4.2.5. Prédisposition génétique**

Enfin, la prédisposition génétique est un autre facteur de risque important puisque près de 3 % des carcinomes rénaux sont liés à un fond génétique familiale, principalement en lien avec une mutation sur le gène VHL, qui tire son nom de la maladie de Von Hippel-Lindau (Woodward *et al.*, 2000). D'autres syndromes héréditaires liés à des mutations des gènes FLCN, MET, FH ou aux sous-unités de la succinate déshydrogénase sont aussi associés au développement de carcinomes rénaux (Menko and Maher, 2016).

### **1.1.5. FACTEURS PRONOSTIQUES DU CCRCC**

Certaines caractéristiques de la tumeur permettent de prédire avec une relative efficacité la durée de survie d'un patient ; il s'agit de facteurs pronostiques. Ceux-ci peuvent être classés en quatre catégories principales ; les facteurs anatomiques, histologiques, cliniques et moléculaires. Il existe aussi des modèles, ou scores, intégrant plusieurs facteurs et permettant d'affiner la prédiction. Si l'enjeu est principalement d'informer le patient, ces facteurs pronostiques peuvent néanmoins aider à la décision des médecins, notamment au regard de la balance bénéfice/risque.

#### **1.1.5.1. Facteurs anatomiques**

Les principaux facteurs anatomiques sont la taille de la tumeur, son degré d'extension à des zones ou compartiments comprenant ; dans la veine, le fascia de Gerota, le système collecteur, la graisse péri-rénale ou encore le hile (Figure 6). La colonisation de la glande adrénaire et des ganglions lymphatiques, ainsi que la présence de métastases distantes sont

également considérées comme des facteurs anatomiques. Selon ces paramètres, la tumeur est classée du stade I (propagation faible) au stade IV (propagation importante). Le système de classification TNM (« *Tumor-Node-Metastasis* »), plus fin, est cependant le plus couramment utilisé (Tableau 2). Émis par l'AJCC (American Joint Committee on Cancer) et l'UICC (Union for International Cancer Control), ce système est régulièrement mis à jour et en est à sa 8<sup>e</sup> édition (Amin *et al.*, 2017).

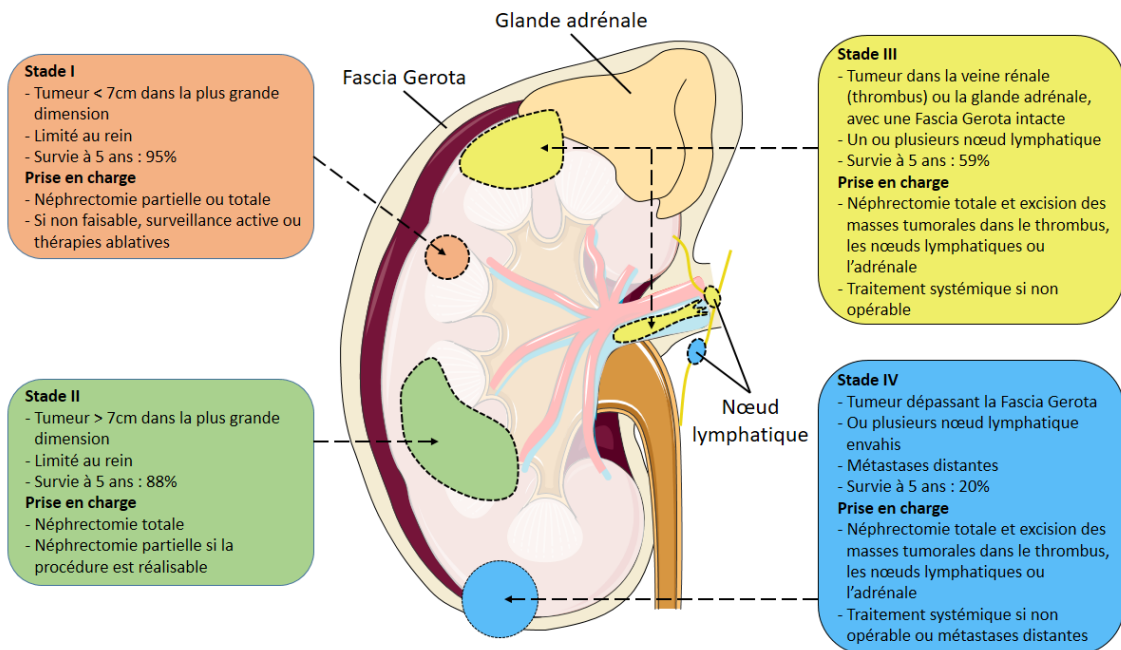


Figure 6. Stades de développement du ccRCC

Le schéma représente une coupe de rein sur laquelle sont superposées des zones illustrant le système de détermination du stade de développement. Celui-ci s'appuie sur la taille et la position de la tumeur, ainsi que sur l'atteinte des ganglions lymphatiques. Par exemple, une tumeur de stade I ou II est entièrement enfermée dans le rein tandis que les tumeurs de stade III peuvent s'étendre dans les veines principales ou les glandes surrénales, à l'intérieur du fascia de Gerota. Inspiré de Chen *et al.*, 2016.

Tumeur primitive (T)	
Tx	Impossible d'évaluer la tumeur primitive.
T0	Aucun signe de tumeur primitive.
T1	Tumeur limitée au rein $\leq 7$ cm de grand axe T1a : Tumeur $\leq 4$ cm T1b : Tumeur $> 4$ cm et $\leq 7$ cm
T2	Tumeur limitée au rein $> 7$ cm T2a : Tumeur $> 7$ cm et $\leq 10$ cm T2b : Tumeur $> 10$ cm
T3	T3a : Envahissement macroscopique de la veine rénale ou de ses branches, de la graisse péri-rénale ou hilare T3b : Envahissement de la veine cave inférieure sous diaphragmatique. T3c : Envahissement de la veine cave inférieure sus diaphragmatique ou de la paroi de la veine cave.
T4	Franchissement du fascia de Gerota.
Ganglions lymphatiques régionaux (N)	
Nx	Non déterminé.
T0	Pas de métastase ganglionnaire.
N1	Métastase au niveau d'un ganglion unique.
N2	Métastases au niveau de plusieurs ganglions.
Métastases à distance (M)	
Mx	Non déterminé
M0	Aucun signe de métastases à distance.
M1	Présence de métastases à distance.
Correspondance des stades de progression et des stades TNM	
Stage I	T1 - N0 - M0
Stage II	T2 - N0 - M0
Stage III	T3 - N0 - M0 T1/T2/T3 - N1 - M0
Stage IV	T4 – n'importe quel N - M0 N'importe quel T – n'importe quel N – M1

Tableau 2. Description des stades de la classification TNM

### 1.1.5.2. Facteurs histologiques

Ces facteurs incluent le grade de la tumeur, l'invasion lymphovasculaire, la nécrose et l'invasion du système collecteur (Sun *et al.*, 2011; L. Zhang *et al.*, 2018). Cependant, le grade est considéré comme le plus important. Le grade de Furhman s'appuie sur la taille et la forme du noyau, et sur la proéminence des nucléoles (Erdogan, Demirel and Polat, 2004). Très utilisé ces dernières décennies, il a maintenant été remplacé par le grade WHO/ISUP (*World Health Organization / International Society of Urological Pathology*) (Delahunt *et al.*, 2013). Cette classification établie en 2012 est recommandée par l'OMS (Delahunt *et al.*, 2013; Moch *et al.*, 2016). De façon similaire au grade de Furhman, l'évaluation repose sur l'aspect des nucléoles pour les grades 1 à 3, et sur le pléomorphisme des noyaux ou le type de différenciation (sarcomatoïde ou rhabdoïde) pour le grade 4 (Figure 7). Le grade 1 présente des nucléoles basophiles peu visibles ou absents à fort grossissement (x400) ; le grade 2 présente des nucléoles éosinophiles visibles à fort grossissement (x400) mais ne sont pas proéminents. Le grade 3 présente des nucléoles proéminents et facilement distinguables à faible grossissement (x100). Enfin, le grade 4 présente des cellules tumorales géantes et/ou un pléomorphisme nucléaire marqué. On parle également de grade 4 lorsque l'on observe une différenciation sarcomatoïde ou rhabdoïde. Il a été démontré que le système ISUP s'est révélé plus efficace que le grade de Furhman, en particulier pour les tumeurs de grade 2 et 3 (Dagher *et al.*, 2017).

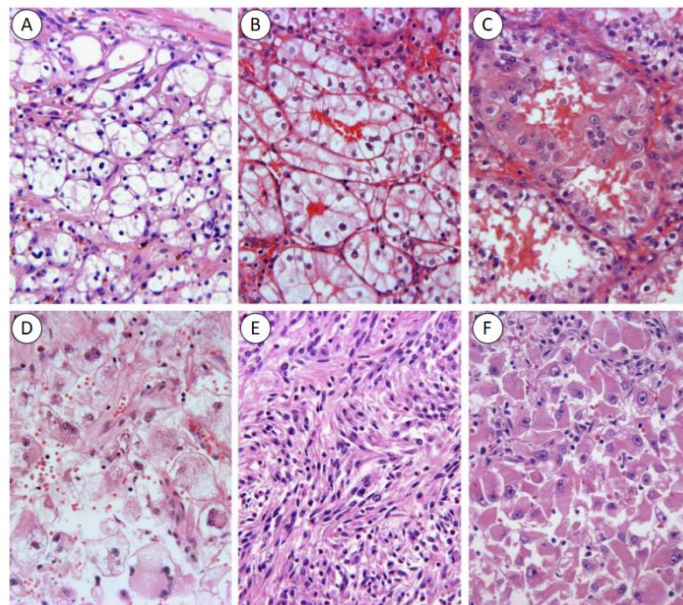


Figure 7. **Classification de l'ISUP pour le ccRCC**

Le système de classification de l'ISUP repose sur l'examen histologique des tumeurs. Sur ces coupes colorées à l'hématoxyline et à l'éosine, l'aspect des noyaux renseigne sur le grade et sur l'agressivité de la tumeur. A, Grade 1 ; B, Grade 2. C ; Grade 3 ; D, grade 4 ; E, Grade 4 : différenciation sarcomatoïde ; F, grade 4 : différenciation rhabdoïde.

GRADE	ECOG PERFORMANCE STATUS
<b>0</b>	Patient entièrement actif, capable d'effectuer les mêmes activités pré-morbides sans restriction
<b>1</b>	Patient restreint dans ses activités physiques, mais capable de se déplacer et d'effectuer des activités légères ou sédentaires, par ex. : travaux ménagers légers ou tâches administratives
<b>2</b>	Patient ambulatoire et capable de s'occuper de lui, mais incapable d'effectuer des activités. Debout > 50% de la journée
<b>3</b>	Patient capable de soins limités, alité ou au fauteuil > 50% de la journée
<b>4</b>	Patient complètement handicapé, ne pouvant s'occuper de lui. Totalement alité ou confiné au fauteuil

Tableau 3. Échelle de statut de performance ECOG-PS

### **1.1.5.3. Facteurs cliniques**

Les facteurs cliniques incluent le « statut de performance » et des indicateurs variés tels que l'anémie, la cachexie, le nombre de neutrophiles, lymphocytes et de plaquettes, le taux d'albumine (Kim *et al.*, 2003, 2004; Patard *et al.*, 2004; Bensalah *et al.*, 2006; Cho *et al.*, 2016; Lee *et al.*, 2020; Patel *et al.*, 2020). Le statut de performance est un score notant l'état général du patient et ses capacités motrices. Citons en particulier l'échelle ECOG-PS (*Eastern Cooperative Oncology Group - Performance Status*) (Tableau 4).

### **1.1.5.4. Facteurs moléculaires**

Il existe de nombreux facteurs moléculaires, parmi lesquels CA9, VEGF, MKI67 (Ki67), TP53 (p53), CDKN1A (p21), PTEN, CD44, CXCR4, CD274 (PD-L1), ou encore des miRNAs, SNPs et mutations (Sabatino *et al.*, 2009; Fan *et al.*, 2020; European Association of Urology, 2021) (Jeong *et al.*, 2012; Hakimi, Pham and Hsieh, 2013; Gayed *et al.*, 2014; Kumar *et al.*, 2018). Bien que ces marqueurs permettent d'améliorer certains modèles pronostiques, ils ne sont pas ou très peu validés par des études indépendantes. Par ailleurs, il n'existe pas de preuves concluantes sur leur valeur pour la sélection de traitement pour le ccRCC métastatique (Raimondi *et al.*, 2020; Ueda *et al.*, 2020; Fukuda *et al.*, 2021). Leur usage en pratique clinique courante n'est donc pas recommandé. Récemment, une signature de seize gènes a été déterminée par une équipe permettant de mieux prédire le pronostic des patients, notamment des patients à risque apparemment faible (Rini *et al.*, 2015). Cependant, la signature n'a pas encore été validée par d'autres études. De façon très intéressante, le consortium TRACERx a proposé une classification génique s'appuyant sur les modèles d'évolution (ponctuée, en branche ou linéaire ; voir la section **1.2.1.2. Modèles d'évolution tumorale**) qui semble être corrélée à l'agressivité de la tumeur et la survie du patient.

### **1.1.5.5. Les modèles pronostics**

Les modèles pronostics sont des scores combinant des facteurs pronostiques, développés et validés indépendamment (Karakiewicz *et al.*, 2007; Zigeuner *et al.*, 2010). Ils se révèlent plus précis que la classification TNM ou le grade nucléolaire seuls. Ils sont systématiquement testés et comparés aux modèles préexistants avant d'être adoptés en pratique clinique. Les scores MSKCC et IMDC concernent le ccRCC métastatique tandis que d'autres comme les scores UISS, VENUSS, GRANT ou encore Leibovich concernent le ccRCC localisé (Tableau 4). Le score de Leibovich, par exemple, est un algorithme conçu par le chirurgien urologue



américain Bradley Leibovich et formalisé en 2005 (Leibovich *et al.*, 2005). Il est aussi utilisé pour prédire le risque de progression de la tumeur après chirurgie, qu'il s'agisse d'une rechute locale chez les patients favorables ou du développement d'une métastase chez les patients ayant subi une néphrectomie élargie. (Blackmur *et al.*, 2021). Ce score est utilisé en clinique pour aider à la décision des oncologues. L'algorithme prend en compte la taille de la tumeur, le stade pTNM, le grade ISUP, le statut des ganglions et la présence de nécrose et détermine un score selon lequel le patient sera classé dans trois catégories ; risque faible (score compris entre 0 et 2 inclus), intermédiaire (score compris entre 3 et 5 inclus) ou élevé (score  $\geq 6$ ) (Tableau 4).

Modèle pronostic	Facteurs de risques
<b>UISS</b>	1. ECOG-PS
	2. Classification T
	3. Classification N (N+ = métastatique)
	4. Grade
	T1N0M0G1–2, ECOG PS 0 : risque faible T3N0M0G2–4, ECOG PS $\geq 1$ OR T4N0M0 : risque élevé Autre N0M0: risque intermédiaire
<b>Leibovich</b>	1. Classification T (pT1a : 0, pT1b : 1, pT2:3, pT3-4: 4 points)
	2. Classification N (pNx/N0 : 0, pN+: 2 points)
	3. Taille de la tumeur (< 10 cm : 0, $\geq 10$ cm : 1 point)
	4. Grade (G1-2 : 0, G3 : 1, G4 : 3 points)
	5. Nécrose (absente : 0, présente : 1 point)
	0–2 points : risque faible 3–5 points : risque intermédiaire 6 or more points : risque élevé
<b>VENUSS</b>	1. Classification T (pT1 : 0, pT2 : 1, pT3–4 : 2 points)
	2. Classification N (pNx/pN0 : 0, pN1 : 3 points)
	3. Taille de la tumeur ( $\leq 4$ cm : 0, $> 4$ cm : 2 points)
	4. Grade (G1/2 : 0, G3/4 : 2 points)
	5. Thrombus tumoral (absent : 0, présent : 2 points)
	0–2 points: risque faible 3–5 points: risque intermédiaire 6 or more points: risque élevé
<b>GRANT</b>	1. Age > 60 ans
	2. Classification T = T3b, pT3c or pT4
	3. Classification N = pN1
	4. Grade (Furhman = G3 or G4)
	0–1 facteurs : Groupe favorable 2 facteurs ou plus : Groupe non-favorable

Tableau 4. Exemple de modèles pronostiques dans le ccRCC et méthodes de calcul

### 1.1.5.6. Analyses de survie

Les analyses de survie sont des méthodes statistiques utilisées pour examiner l'évolution dans le temps d'un événement donné. Les analyses de survie de type Kaplan-Meier (K-M) sont les plus fréquemment utilisées, par exemple dans les phases III et certaines phases II des essais cliniques médicaux randomisés (N. Dudley *et al.*, 2016). Elles permettent dans ce cas de

mesurer la proportion de patients vivant pendant un certain temps après un traitement. Plus généralement, l'analyse K-M mesure les différences de probabilités de survie des patients dans différents groupes (contrôle versus traitement, expression génique forte ou faible, etc.) (Figure 8). L'analyse K-M nécessite, entre autres critères, que les patients soient affectés de manière aléatoire aux différentes conditions testées et qu'ils n'entrent pas dans l'étude au même moment. Elle permet ainsi de calculer combien de temps après le début d'un traitement particulier l'événement étudié (par exemple, le décès, la progression de la maladie, etc.) s'est produit pour les individus surveillés pendant l'étude (Peto *et al.*, 1977; Rich *et al.*, 2010).

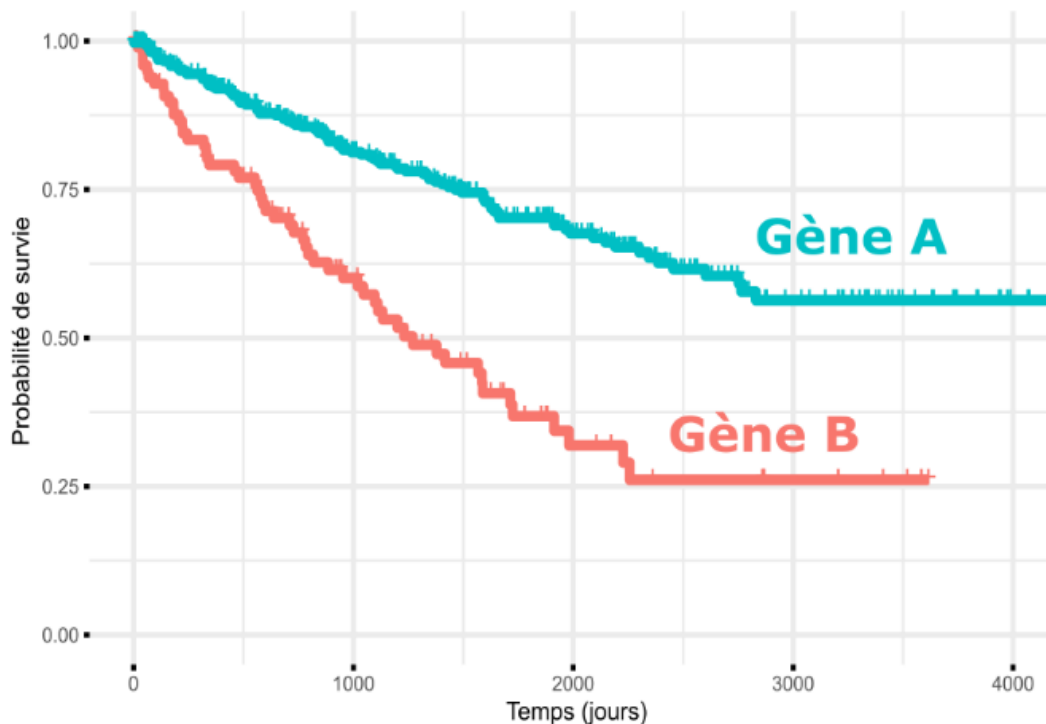


Figure 8. **Courbe de survie Kaplan-Meier représentative**

Dans cet exemple, les patients possédant la signature du gène A ont un pourcentage de chance de survie plus élevé que les patients possédant la signature du gène B (>50% à 3000 jours contre 0,25%, respectivement).

#### **1.1.5.7. Écueils au pronostic anatomo-pathologique**

Malgré l'établissement de ces nombreux critères pronostiques, la prise en charge du ccRCC reste compliquée puisque près de 30% des patients atteints de ccRCC localisé (stade I à II ; traités généralement uniquement par chirurgie ablative), finissent par avoir une rechute tandis qu'aucun traitement adjuvant ne leur était prescrit. (Rini, Campbell and Escudier, 2009). Les systèmes pronostiques s'avèrent donc insuffisants, parfois en raison d'une variabilité inter-observateur dans l'interprétation des critères ou bien vraisemblablement parce que ces facteurs

ne rendent pas compte de toute la complexité et la biologie de la tumeur (Ficarra *et al.*, 2005; Delahunt *et al.*, 2011). La recherche dans ce domaine s'intéresse en particulier aux facteurs moléculaires et doit encore parvenir à réduire l'écart entre prédiction et situation réelle.

## 1.1.6. MECANISMES ET PATHOPHYSIOLOGIE

### 1.1.6.1. Altérations génétiques

La perte d'activité du gène suppresseur de tumeur VHL (Von Hippel Lindau) est un événement majeur et quasi universel dans le carcinome rénal à cellules claires (Gnarra *et al.*, 1994). Elle intervient très précocement dans sa carcinogénèse, résultant de la délétion du locus de VHL (délétion de la région chromosomique 3p25) (60 à 90%), de mutation dans ses exons ou zones d'épissage (80%) et de la méthylation de son promoteur (10%). Le gène VHL code pour la protéine pVHL, qui, associée au complexe E3 ligase, participe à l'ubiquitination et la dégradation des facteurs HIF-1 $\alpha$  (Hypoxia Inducible Factor) et HIF-2 $\alpha$  via le protéasome (Figure 9) (Semenza, 2013; Masson and Ratcliffe, 2014).

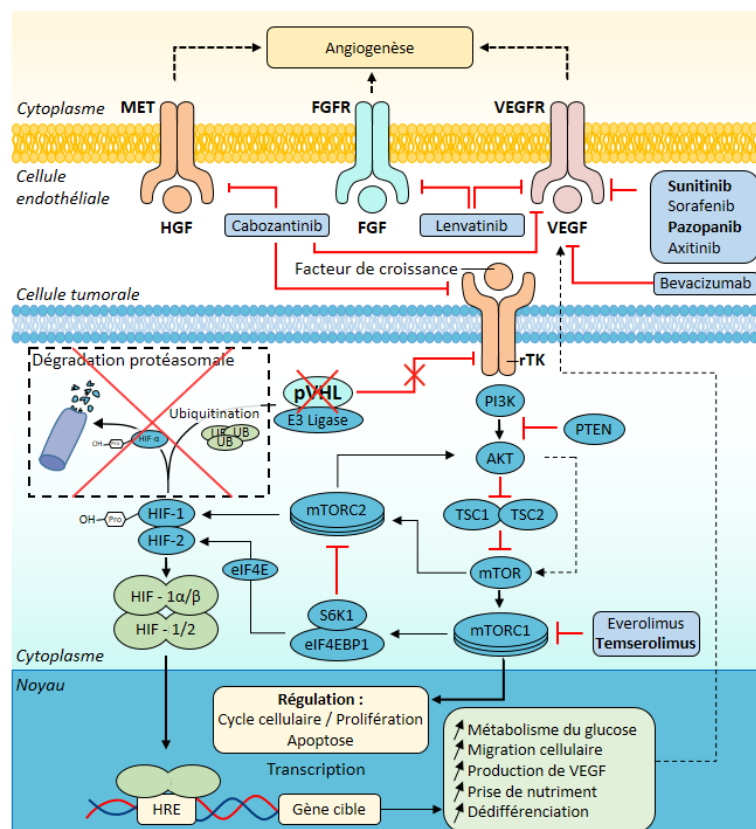


Figure 9. Voies moléculaires dans le ccRCC

Ce schéma représente les principales voies moléculaires connues pour être impliquées dans la carcinogénèse du ccRCC. Ces voies, souvent surexprimées, conduisent à un contrôle anormal de l'apoptose ou à l'augmentation du métabolisme et de la croissance cellulaire. Les thérapies indiquées en gras sont celles utilisées dans la partie appliquée de ce travail de thèse. Figure inspirée de Hsieh *et al.*, 2017.

En condition normale d'apport en oxygène (normoxie), les sous-unités  $\alpha$  de HIF-1 et HIF-2 sont hydroxylées sur un ou sur les deux résidus proline par PHD1/2/3 (prolyl hydroxylases 1, 2 et 3). La reconnaissance de cette forme hydroxylée par pVHL conduit à la formation d'un complexe responsable de sa poly-ubiquitinylation et de sa dégradation protéasomale (Gossage et al., 2015). En condition d'hypoxie, les PHDs ne sont pas actives, HIF- $\alpha$  s'accumule, s'hétérodimérise avec son partenaire HIF-1b (ou ANRT, aryl hydrocarbon receptor nuclear translocator) permettant la fixation du complexe sur des éléments de réponse à l'hypoxie et la transcription de gènes impliqués dans l'adaptation à l'hypoxie. La perte d'activité de VHL implique donc l'arrêt de la dégradation des protéines HIF et induit leur accumulation aberrante ainsi qu'un surcroît d'expression de leurs gènes cible. Ces gènes comprennent notamment GLUT1, HK1 et HK2, impliqués dans le métabolisme du glucose et VEGFA, EPO et PDGF $\beta$ , impliqués dans l'angiogenèse, ce qui explique que les ccRCCs ont les particularités d'être hautement vascularisés et riches en lipides et glycogènes (Hakimi *et al.*, 2016). CAIX, un marqueur connu du ccRCC, est également surexprimé par cette voie (Ivanov *et al.*, 2001). Par ailleurs, les facteurs HIF-1 et HIF-2 ont des rôles distincts et complémentaires, le premier régulant essentiellement l'expression des gènes des voies apoptotique et glycolytique, et le second régulant des gènes impliqués dans la prolifération et l'angiogenèse (Keith, Johnson and Simon, 2012).

Si la perte de VHL semble être un événement initial majeur dans la carcinogenèse du ccRCC, deux observations laissent penser qu'elle n'y suffit pas : 1 – Les individus porteurs de mutations germinales de VHL peuvent ne développer un ccRCC que tardivement, parfois avec une latence supérieure à 30 ans (Kaelin, 2007) ; 2 – la perte de VHL chez des souris KO ne suffit pas à l'induction du ccRCC (Kapitsinou and Haase, 2008). Cela indique que d'autres acteurs moléculaires sont probablement impliqués dans le développement du ccRCC (Wei and Hsieh, 2015). Des études de génomique ont par la suite révélé dans ce sens un ensemble de mutations prévalentes dans le ccRCC, bien que significativement moins fréquentes que la déplétion de VHL. Elles affectent notamment les gènes PBRM1 (dans 29% à 41% des ccRCC), SETD2 (8-12%), BAP1 (6-10%), KDM5C (4-7%) et mTOR (5-6%) (Peña-Llopis *et al.*, 2012; Hakimi, Chen, *et al.*, 2013; Sato *et al.*, 2013; Xu *et al.*, 2016; Liu *et al.*, 2019). De façon remarquable, VHL, PBRM1, SETD2 et BAP1 ont tous un rôle suppresseur de tumeurs d'une part, et se situent sur le bras court du chromosome 3 (3p), d'autre part. Ainsi, la perte d'une seule copie du 3p entraînerait l'haplo-insuffisance de ces quatre gènes, favorisant de façon nette la carcinogenèse. Ceci corrobore le constat de la perte quasi-universelle du 3p et conforte son

statut d'événement précoce dans la pathophysiologie du ccRCC (Hakimi, Pham and Hsieh, 2013; Wei and Hsieh, 2015). Une étude de référence parue dans le journal *Nature* intègre des analyses moléculaires variées de 385 échantillons de la cohortes du TCGA-KIRC (« *The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma* »), pour dresser un état des lieux des altérations somatiques du ccRCC (Figure 10) (Creighton *et al.*, 2013). Elle renforce les précédentes observations quant aux gènes fréquemment mutés et au rôle central de VHL, et permet de découvrir des nouvelles altérations avec une puissance statistique importante. De façon intéressante, la méthylation du promoteur de VHL et ses mutations apparaissent comme deux événements exclusifs entraînant la perte de fonction de VHL.

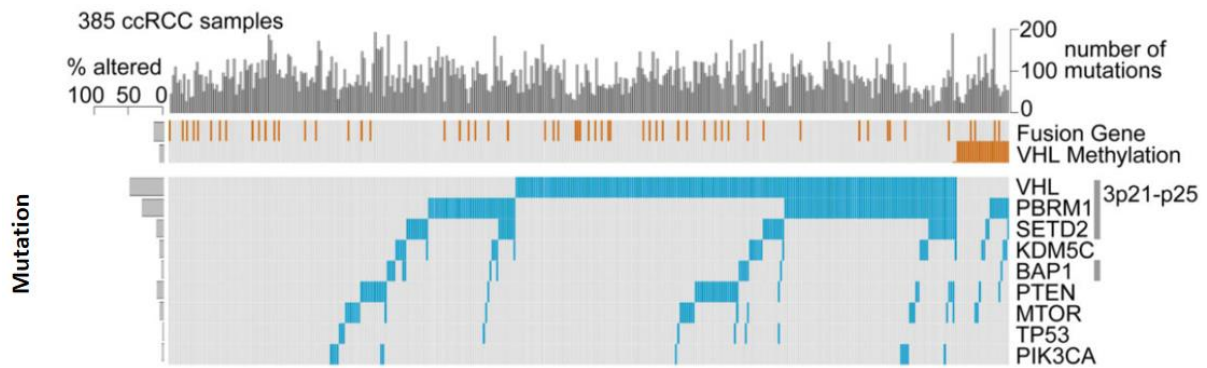


Figure 10. **Gènes mutés dans le ccRCC**

Les données sont issues de la base de données TCGA-KIRC. Les graphiques montrent le nombre et le type d'altérations dans 385 échantillons de ccRCC. Figure extraite de (Creighton *et al.*, 2013).

Type de CNV	#	Bande	Fréquence	#	Bande	Fréquence
<b>Perte</b>	1	<b>3p21.1</b>	92%	21	<b>8p11.21</b>	25%
	2	<b>3p26.1</b>	91%	22	<b>6q15</b>	24%
	3	<b>3p21.1</b>	89%	23	<b>10q23.31</b>	18%
	4	<b>3p14.2</b>	80%	24	<b>10q26.2</b>	18%
	5	<b>3p14.2</b>	78%	25	<b>1p36.13</b>	16%
	6	<b>3p14.1</b>	71%	26	<b>4q35.2</b>	16%
	7	<b>3p12.3</b>	65%	27	<b>4q24</b>	16%
	8	<b>3p12.3</b>	60%	28	<b>1p36.22</b>	16%
	9	<b>3p12.1</b>	51%	29	<b>1p35.3</b>	14%
	10	<b>3p11.2</b>	47%	30	<b>13q33.3</b>	14%
	11	<b>14q24.3</b>	45%	31	<b>13q12.13</b>	13%
	12	<b>14q24.3</b>	45%	32	<b>1p31.1</b>	12%
	13	<b>9p21.3</b>	32%	33	<b>10p15.3</b>	12%
	14	<b>3q11.2</b>	32%	34	<b>Xq28</b>	10%
	15	<b>9p24.1</b>	31%	35	<b>2q37.1</b>	8%
	16	<b>3q12.1</b>	31%	36	<b>15q21.1</b>	8%
	17	<b>8p23.2</b>	31%	37	<b>11q23.3</b>	6%
	18	<b>8p21.1</b>	30%	38	<b>1q42.3</b>	5%
	19	<b>6q26</b>	28%	39	<b>1q41</b>	5%
	20	<b>6p27</b>	28%	40	<b>7q31.1</b>	1%
<b>Gain</b>	1	<b>5q35.2</b>	67%	21	<b>17q25.3</b>	8%
	2	<b>5q21.1</b>	50%	22	<b>Xq28</b>	7%
	3	<b>5q15</b>	44%	23	<b>17q12</b>	6%
	4	<b>7q36.2</b>	34%	24	<b>15q26.1</b>	5%
	5	<b>7q11.22</b>	34%	25	<b>Xq12</b>	5%
	6	<b>2q33.1</b>	16%	26	<b>9p24.1</b>	3%
	7	<b>3q26.2</b>	15%	27	<b>14q13.1</b>	3%
	8	<b>8q24.22</b>	15%	28	<b>4q32.1</b>	3%
	9	<b>1q32.1</b>	14%	29	<b>3p26.3</b>	2%
	10	<b>1q21.3</b>	14%			

Tableau 5. CNV les plus fréquents dans le ccRCC

Les données sont issues de la base de données TCGA-KIRC. Tableau inspiré de Creighton *et al.*, 2013.

Les aberrations chromosomiques ou CNV (« *Copy Number Variation* ») les plus fréquemment retrouvées dans cette étude sont présentées dans le Tableau 5. Ces résultats concordent avec ceux de nombreuses études et classent notamment aux premiers rangs des CNV la perte du 3p21.1 (> 90%) et du 14q24 (45%) et le gain du 5q35 (67%) (Wilhelm *et al.*, 2002; Arai *et al.*, 2008; Toma *et al.*, 2008; Beroukhim *et al.*, 2009; Chen *et al.*, 2009; Shuib *et al.*, 2011; Moore *et al.*, 2012; Karekla *et al.*, 2017).

Les gènes PTEN et mTOR sont également fréquemment mutés. Tous deux font partie de la voie PI3K/AKT/mTOR, une voie de signalisation intracellulaire activée en réponse à la fixation d'un facteur de croissance sur son récepteur à activité tyrosine kinase (rTK) (Figure 9). Cette voie régule entre autres le cycle cellulaire, la prolifération et l'apoptose. Elle régule également l'angiogenèse en influant sur les niveaux de HIF.

L'activation du rTK entraîne le recrutement à la membrane de la PI3K (phosphatidyl inositol 3-kinase), la phosphorylation d'AKT et l'activation de mTOR soit par voie directe, soit par l'inhibition des inhibiteurs TSC1 et TSC2 (tuberous sclerosis complex 1 et 2) (Figure 9). La protéine mTOR s'associe à d'autres protéines pour former les complexes mTORC1 et mTORC2 (Dreyer *et al.*, 2009). Le complexe mTORC1 permet l'activation du complexe eIF4EBP/S6K1 et la libération du facteur d'initiation de la traduction eIF4E qui promeut la traduction de la cycline D1 (prolifération), de HIF-1a et HIF-2a, et de GLUT-1 (transport du glucose). Le complexe mTORC2 régule notamment la formation du cytosquelette et l'activité de la PKC (proteïn kinase C). La phosphatase PTEN, quant à elle, régule la voie par déphosphorylation de PIP3 et inhibition de l'activation d'AKT.

En raison de l'implication des gènes PTEN et MTOR dans la signalisation intracellulaire du ccRCC, leurs mutations ont un impact conséquent sur sa physiologie globale. Les mutations de PTEN sont généralement inactivatrices tandis que les mutations de mTOR sont généralement activatrices, favorisant, dans les deux cas, l'activité de la voie PI3K/AKT/mTOR et sa fonction pro-tumorale (Xu *et al.*, 2016).



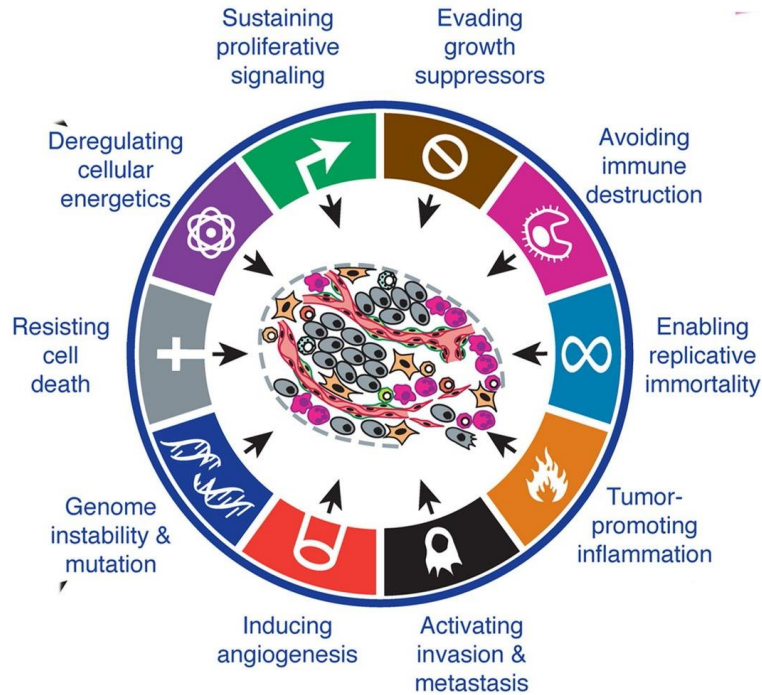


Figure 11. Les « hallmarks », ou caractéristiques fondamentales du cancer

Le schéma représente la nouvelle génération des *hallmarks*, que l'on peut traduire par « caractéristiques fondamentales », publiées en 2011. Ces caractéristiques participent à la protection ou à la promotion du cancer. Figure adaptée de Hanahan and Weinberg, 2011.

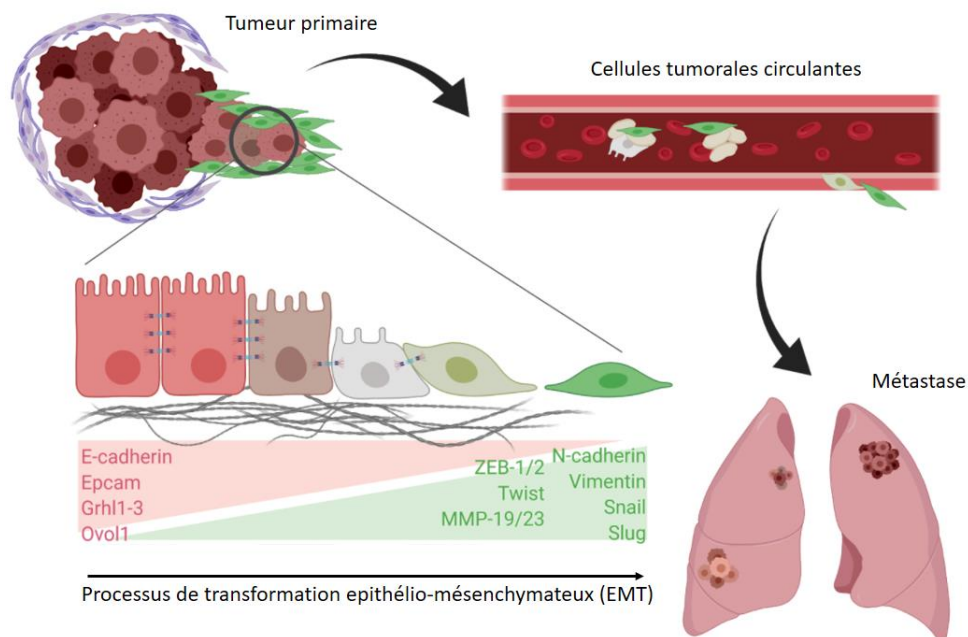


Figure 12. Implication de l'EMT dans le processus métastatique

Les cellules malignes d'origines épithéliales perdent progressivement leur polarité (pôles basal et apicale indistincts) ainsi que certains marqueurs tels que Epcam, et acquièrent un phénotype mésenchymateux en exprimant notamment la Vimentin et la N-cadherin. Elles peuvent ensuite passer dans la circulation sanguine et coloniser des organes distants comme les poumons. Figure tirée de <https://oncobites.blog/2019/11/01/the-other-emt-exploring-the-controversial-driver-of-metastasis/>.

### 1.1.6.2. *Étapes et processus métastatique*

La pathogénèse moléculaire du ccRCC reste peu comprise aujourd'hui. Toutefois, sur la base de la perte de fonction de VHL qui semble être universelle et précoce, la littérature fait état d'une séquence partielle d'événements qui peut être décrite en six étapes (Shenoy and Pagliaro, 2016) : (i) perte de l'activité de VHL, par mutation germinale ou somatique, et inactivation de l'allèle sauvage ; (ii) activation constitutive de la voie HIF en raison de la perte de l'activité de VHL et transcription de gènes impliqués dans l'angiogenèse, la transition épithélio-mésenchymateuse, le processus métastatique, la survie, la glycolyse anaérobie et la voie du pentose phosphate ; (iii) interactions de la voie HIF avec d'autres voies oncogéniques ; (iv) modifications épigénétiques à l'échelle du génome (potentiellement dues à l'hyperactivité de la voie HIF) ; (v) influence de l'épigénétique sur diverses voies oncogéniques, apoptotiques, régulatrices du cycle cellulaire et réparatrices de mésappariements (inhibition de plusieurs gènes suppresseurs de tumeurs) ; (vi) évasion immunitaire, au moins partiellement causée par des modifications de l'épigénome. Ces mécanismes interagiraient tout au long de la carcinogenèse et de la progression de la maladie, et confèreraient également la chimiorésistance et la radiorésistance connues dans le ccRCC.

Dans le ccRCC, près de 30 % des patients ont des métastases au moment du diagnostic, et environ 60 % ont des métastases dans les deux ou trois premières années après le diagnostic (Casuscelli *et al.*, 2019). Les sites métastatiques les plus fréquents du cancer du rein sont le poumon, l'os, les ganglions, le foie et les glandes surrénaliennes (Bianchi *et al.*, 2012). Parmi les autres sites secondaires se trouvent la thyroïde (Beutner *et al.*, 2015), l'estomac (Yamamoto *et al.*, 2009), le cœur (Zustovich *et al.*, 2008), la peau (Mandal, Littler and Libertiny, 2012), le pancréas (Singla *et al.*, 2020) ou encore la cavité nasale (N. Zhang *et al.*, 2018). Dans les « *hallmarks* » du cancer décrites par Weinberg et Hannahan, la capacité des cellules tumorales à envahir le tissu et métastaser est une des propriétés fondamentales retrouvées dans le cancer (Hanahan and Weinberg, 2011) (Figure 11). Cette capacité est acquise progressivement à travers plusieurs étapes, notamment la transition épithélio-mésenchymateuse (ou « EMT » pour *Epithelio Mesenchymal Transition*). L'EMT est un événement du développement embryonnaire lors duquel une cellule épithéliale acquiert un phénotype mésenchymateux, modifiant ses propriétés (Kim *et al.*, 2014). Dans le cadre du cancer, ce processus est détourné et permet aux cellules épithéliales de perdre leur polarité et leur attache au site primaire afin de gagner les tissus environnants, le système vasculaire et, éventuellement, de produire une métastase distante (Brabletz *et al.*, 2018) (Figure 12). De nombreuses études ont démontré le rôle crucial

de l'EMT dans la progression du ccRCC (Harada *et al.*, 2012; Dumanskiy *et al.*, 2013; Fang *et al.*, 2013; Weygant *et al.*, 2015; Landolt *et al.*, 2017) ainsi que l'implication du facteur HIF dans son initiation (He and Magi-Galluzzi, 2014). En dehors de HIF, des cytokines comme IL-6, IL-8, IL-15 et TNF- $\alpha$  promeuvent également l'EMT via la voie de signalisation Akt/GSK-3 $\beta$  / $\beta$ -catenin (Khawam *et al.*, 2009; Ho *et al.*, 2012; Zhou *et al.*, 2016; Q. Chen *et al.*, 2017), rappelant le rôle du microenvironnement tumoral dans la carcinogenèse rénale.

### 1.1.7. MICROENVIRONNEMENT

Le microenvironnement tumoral (ou « *TME* » pour « *Tumor Micro Environment* ») est l'environnement direct de la tumeur. Il constitue un réseau hétérogène et dynamique formé de plusieurs types cellulaires et de la matrice extracellulaire. Les interactions entre cellules malignes, myofibroblastes, fibroblastes, cellules adipeuses, cellules immunitaires et cellules endothéliales conditionnent certaines caractéristiques de la tumeur telles que sa croissance, sa dormance ou son potentiel invasif et métastatique (Figure 13) (Wang *et al.*, 2017). Plus précisément, les cellules tumorales agissent soit directement (en libérant, par exemple, des facteurs de croissance ou des cytokines), soit indirectement (en induisant l'hypoxie ou la nécrose), recrutant ou activant les cellules du TME. De plus, les cellules tumorales sont capables de modifier la matrice extracellulaire, ce qui conduit le plus souvent à la progression de la tumeur. Dans l'autre sens, les événements du microenvironnement tumoral favorisent la progression de la tumeur en stimulant la croissance et la survie des cellules malignes, ce qui promeut notamment le processus métastatique (Lorusso and Rüegg, 2008).

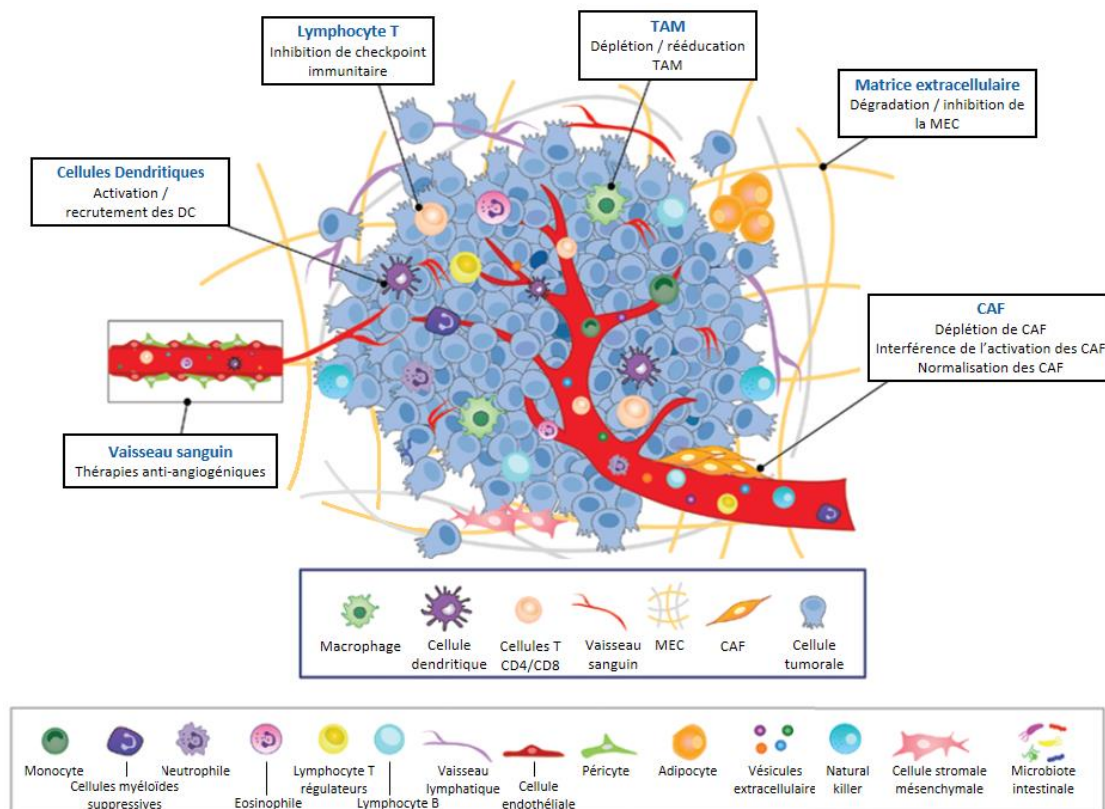


Figure 13. Illustration du microenvironnement tumoral

Les cellules tumorales sont entourées par le microenvironnement tumoral, constitué de nombreux types cellulaires distincts. Celui-ci participe au métabolisme tumoral. TAM : « *Tumor Associated Macrophages* », CAF : « *Cancer Associated Fibroblasts* ». Figure adaptée de Bejarano, Jordão and Joyce, 2021. Figure adaptée de García-Figueiras *et al.*, 2019.

### **1.1.7.1. Microenvironnement vasculaire**

Les cellules endothéliales forment la barrière entre le sang ou la lymphe et la paroi vasculaire. Les cellules endothéliales vasculaires sont au contact avec les cellules sanguines tandis que les cellules endothéliales lymphatiques sont au contact avec la lymphe. Les cellules endothéliales sont très actives du point de vue métabolique et jouent un rôle essentiel dans plusieurs processus physiologiques, comme l'oxygénation des tissus, le maintien de la fluidité du sang et la pression artérielle ou l'immunité innée et adaptative (Rohlenova *et al.*, 2018). Les cellules endothéliales sont également impliquées dans des processus pathologiques, y compris l'initiation et la progression des tumeurs, principalement en contribuant à l'angiogenèse (Hida *et al.*, 2018). Dans le ccRCC, la perte de VHL mène à l'accumulation de HIF, entraînant une réponse hypoxique et une angiogenèse via la sécrétion de VEGF. Ceci confère au ccRCC une hyper-vascularisation caractéristique (Hakimi *et al.*, 2016) et généralement supérieure aux autres types de carcinomes (Yoshihara *et al.*, 2013; Şenbabaoğlu *et al.*, 2016). Il a été démontré que le recrutement de cellules endothéliales vasculaires dans le TME joue un rôle essentiel dans la progression du ccRCC (Albiges *et al.*, 2011; Pircher *et al.*, 2011, 2017). Par ailleurs, les cellules endothéliales circulantes ainsi que les progéniteurs endothéliaux reflètent l'activité des agents anti-angiogéniques sur la néovasculature tumorale (Kraan *et al.*, 2012; Zhou *et al.*, 2017), démontrant leur potentiel prédictif dans le traitement du ccRCC. Dans une étude sur une cohorte de 55 patients atteints de ccRCC métastatique traités par des agents anti-angiogéniques, les auteurs ont démontré, par exemple, que des niveaux plus élevés de cellules progénitrices circulantes positives pour les marqueurs CD34 et VEGFR2 sont corrélés avec de moins bons résultats (Farace *et al.*, 2011). Les cellules endothéliales dans le TME semblent cependant avoir un profil différent des cellules endothéliales normales (Lambrechts *et al.*, 2018). Dans le ccRCC, Akino et ses collègues sont les premiers à avoir étudié les aberrations chromosomiques dans des cellules endothéliales cancéreuses fraîchement isolées et ont constaté que 33 % d'entre elles étaient aneuploïdes, alors que les cellules endothéliales normales sont diploïdes (Akino *et al.*, 2009). Or, le fait que les cellules du TME puissent également être génétiquement ou morphologiquement perturbées pourrait offrir de nouvelles opportunités thérapeutiques (Heidegger, Pircher and Pichler, 2019).

### **1.1.7.2. Microenvironnement stromal**

Les fibroblastes associés au cancer (ou CAF pour « *Cancer Associated Fibroblasts* ») constituent un type cellulaire à part dans le TME. Ils promeuvent des caractéristiques

carcinogènes en participant au remodelage de la matrice extracellulaire ou via la sécrétion de cytokines (Errarte, Larrinaga and López, 2020). Les CAF sont une sous-population de fibroblastes possédant un phénotype myofibroblastique principalement exprimé dans les lésions tumorales. Leurs principales fonctions incluent la stimulation de l'angiogenèse et la croissance métastatique via la sécrétion de facteurs de croissance tels que le VEGF, le facteur de croissance dérivé des plaquettes (PDGF), le facteur de croissance transformant- $\beta$  (TGF- $\beta$ ), le facteur de croissance épidermique des plaquettes (EGF) ou le facteur de croissance des fibroblastes (FGF) (Heidegger, Pircher and Pichler, 2019). Dans le ccRCC, de nombreuses études ont démontré l'implication des CAFs dans la prolifération et la migration des cellules malignes, établissant ainsi leur caractère pro-tumoral (Motzer *et al.*, 2007; Bakhtyar *et al.*, 2013; H. Wang *et al.*, 2014; Chuanyu *et al.*, 2017; Vickman *et al.*, 2020). Alors que leur implication dans la carcinogenèse a été démontrée dans plusieurs sites tumoraux, les CAFs sont cependant moins abondants dans le ccRCC (Öhlund, Elyada and Tuveson, 2014; Yu *et al.*, 2017). Leurs rôles dans la carcinogenèse en font toutefois des cibles thérapeutiques potentielles dans les thérapies innovantes à venir (Errarte, Larrinaga and López, 2020).

### **1.1.7.3. Microenvironnement immunitaire**

Le ccRCC est connu pour être l'un des cancers les plus immunogènes et infiltrés en cellules immunitaires, en lien avec l'angiogenèse qui le caractérise (Vuong *et al.*, 2019). Des signatures transcriptomiques montrent en effet que le ccRCC est très fortement infiltré comparé à dix-huit autres types de cancers épithéliaux (Yoshihara *et al.*, 2013; Şenbabaoğlu *et al.*, 2016). De façon générale, les cellules immunitaires majoritaires infiltrant le TME dans le ccRCC sont les lymphocytes T, les macrophages et les cellules suppressives dérivées des myéloïdes (Heidegger, Pircher and Pichler, 2019; Díaz-Montero, Rini and Finke, 2020). Plus précisément, le ccRCC présente les scores les plus élevés de cellules T CD8, Th1, de cellules dendritiques (DC), de neutrophiles et de cellules cytotoxiques. A l'inverse, il présente des scores relativement faibles de cellules Th2 et de cellules T régulatrices (Treg), ce qui suggère un profil pro-inflammatoire global (Pircher *et al.*, 2011). Cependant, d'autres données continuent de mettre en évidence l'hétérogénéité du ccRCC, avec des sous-groupes distincts qui diffèrent dans l'expression des programmes d'angiogenèse, les modèles d'infiltration immunitaire et les résultats cliniques (Pircher *et al.*, 2011; Kraan *et al.*, 2012; Zhou *et al.*, 2017; Azizi *et al.*, 2018). En 2017, une étude est parue décrivant un atlas immunitaire du ccRCC, établi avec la cytométrie de masse à partir de 73 tumeurs précédemment non traitées (Chevrier *et al.*, 2017). Les auteurs

ont mis au jour 17 phénotypes de macrophages associés à la tumeur (TAM), 22 phénotypes de cellules T ainsi qu'une composition immunitaire distincte, corrélés avec le pronostic des patients. Ils ont aussi identifié des marqueurs pronostiques potentiels. Par exemple, un score CA-2 élevé (reflétant des fréquences élevées de macrophages M-11 ou M-13 et faibles de macrophages M-5) est associé à un mauvais pronostic. Ces sous-types de macrophages sont donc proposés comme cibles potentielles pour le traitement du RCC (Chevrier *et al.*, 2017). Conformément à ces résultats, des données encore plus conséquentes provenant de la base de données TCGA-KIRC décrivent des signatures immunitaires associées à des taux de survie réduits, présentant des cellules T, des cellules B, des macrophages, des cellules dendritiques et des cellules NK (Ricketts *et al.*, 2018). Une autre étude dévoile 22 profils cellulaires distinguant à la fois le type de cellule et son état actif ou inactif, retrouvés dans 40 ccRCCs (Pan *et al.*, 2020). Les auteurs déterminent que les cellules dendritiques activées et au repos, les mastocytes et les éosinophiles sont associés à un pronostic favorable tandis que les cellules B mémoire, les cellules T folliculaires auxiliaires et les cellules T régulatrices (Tregs) sont corrélées à un résultat moins favorable. Ainsi, le profil immunitaire du ccRCC est complexe et très lié à la survie du patient. Contrairement à de nombreux autres types de cancers, une importante infiltration immunitaire dans le ccRCC est généralement associée à un mauvais pronostic (Yuhao Wang *et al.*, 2019).

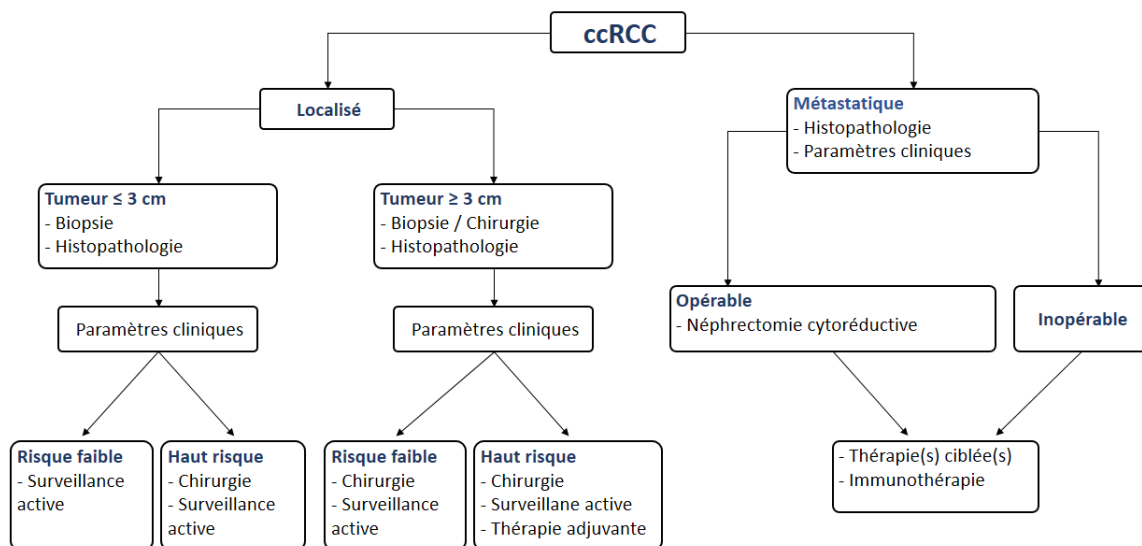


Figure 14. Arbre de décision dans la prise en charge médicale du carcinome rénal à cellules claires

La prise en charge médicale du ccRCC suit un arbre de décisions qui détermine notamment le caractère localisé et la taille de la tumeur.

### **1.1.8. PRISE EN CHARGE MEDICALE**

La décision de la stratégie médicale à adopter pour contrer le ccRCC dépend de nombreux facteurs cliniques (EAU guidelines, 2021) Figure 14. De façon générale pour les patients atteints d'un ccRCC opérable, le traitement standard est l'excision chirurgicale. Dans les cas de ccRCC inopérables ou métastatiques, la prise en charge inclut des traitements par thérapies ciblées et/ou des inhibiteurs de points de contrôle immunitaires (« *Immune Checkpoint Inhibitor* » ou ICI) (Ljungberg *et al.*, 2019). Le traitement par radiothérapie (RT) est globalement inefficace dans le ccRCC qui développe une résistance. La RT peut être utilisée cependant pour certains cas de palliation de métastases ou de contrôle de croissance localisée (Li *et al.*, 2017).

#### **1.1.8.1. Chirurgie**

Le choix de l'ablation chirurgicale du ccRCC dépend du stade clinique de la maladie et de l'état général du patient (Figure 14) (Ljungberg *et al.*, 2019). Les tumeurs localisées sont généralement retirées chirurgicalement, par néphrectomie totale (NT) ou partielle (NP). La néphrectomie peut aussi être utilisée en intention cytoréductrice dans le cas où le patient présente un bon état général, une petite métastase et un site primaire conséquent (Bex *et al.*, 2016). Le but de la NP est de retirer la masse tumorale complète en préservant au mieux le parenchyme rénal sain. Elle est indiquée en première ligne pour les ccRCC de stade T1 et T2. Elle est aussi recommandée dans les cas où le patient n'a qu'un rein, un carcinome rénal bilatéral, ou bien des conditions de santé qui atténuent sa fonction rénale, comme des calculs rénaux, de l'hypertension ou du diabète. La NT consiste à retirer le rein atteint dans son entièreté, ainsi en général que le tissu adipeux périrénal, avec ou sans la glande surrénale et les nœuds lymphatiques environnants. Elle est recommandée pour les ccRCC localement avancés (stades non-T1, non-T2) et non métastatiques (M0), ainsi que dans les cas de multiples petites tumeurs dans le rein ou quand la tumeur a envahi les vaisseaux sanguins. La glande surrénale peut être conservée dans le cas où la masse tumorale est inférieure à 5 cm.

#### **1.1.8.2. Surveillance active et thérapies ablatives**

Dans les cas où le patient a un âge avancé ou souffre de comorbidités rendant la chirurgie délicate, la surveillance active ou les thérapies ablatives telles que la cryothérapie et les radiofréquences peuvent être utilisées (Motzer *et al.*, 2015; Ljungberg *et al.*, 2020). L'approche de surveillance la plus commune consiste en des scanners ou examens IRM tous les trois mois



la première année, tous les 6 mois la seconde année, puis tous les ans ensuite. Les techniques ablatives, quant à elles, sont recommandées pour traiter les petites tumeurs pT1. Elles ont pour but de détruire entièrement le tissu tumoral. Cependant, des études suggèrent que le taux de récurrence locale est supérieur dans le cas des thérapies ablatives, comparé à celui des néphrectomies partielles (Campbell *et al.*, 2009).

### 1.1.8.3. Traitement thérapeutique systémique

À l'heure actuelle, aucune recommandation dans le traitement du ccRCC métastatique (mRCC) ne repose sur l'utilisation de chimiothérapies conventionnelles, en raison de leur inefficacité. Cependant, le paysage thérapeutique du mRCC a été profondément transformé ces 15 dernières années avec l'avènement des thérapies ciblées (Falzone *et al.*, 2018). Celles-ci comprennent les inhibiteurs de tyrosine kinase (TKI) anti-angiogéniques et, plus récemment, les ICI. Leur efficacité repose sur le fait que le ccRCC est un cancer particulièrement immunogène et vascularisé. (Escudier *et al.*, 2014; Motzer *et al.*, 2015; Lee *et al.*, 2016; Hsieh *et al.*, 2017; Powles *et al.*, 2017).

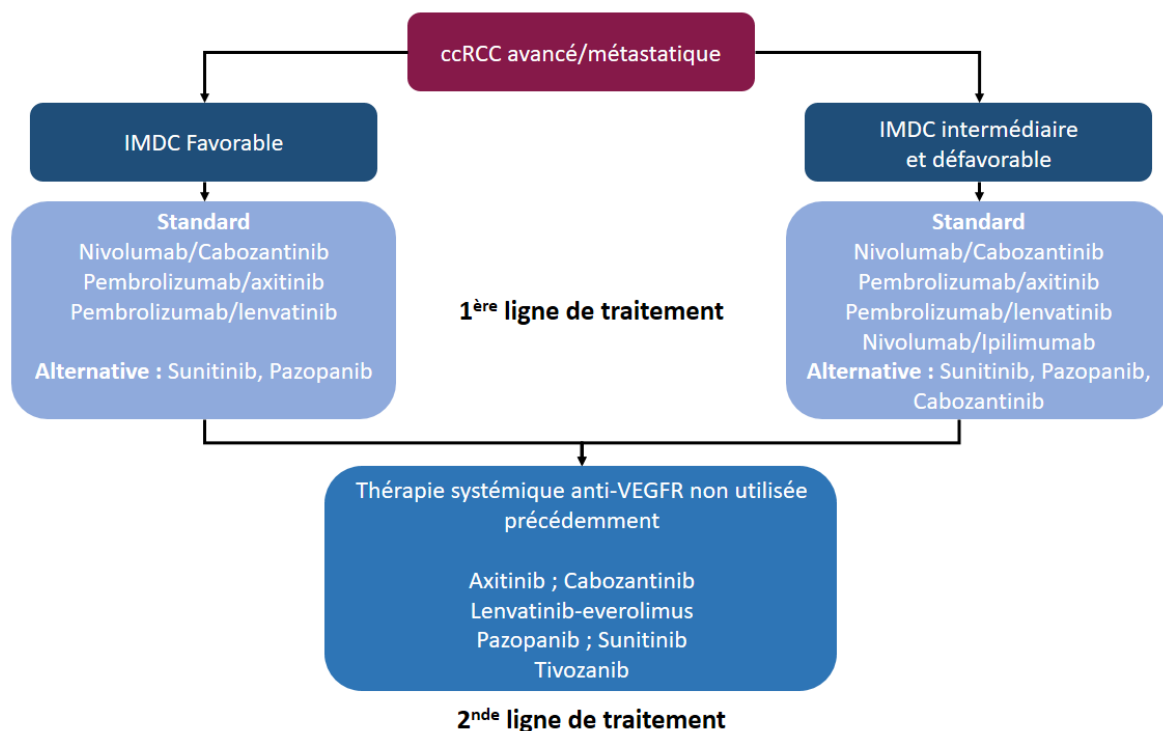


Figure 15. **Guidelines dans le traitement du ccRCC.**

Ces indications sont issues de l'Association européenne d'urologie. Les groupes de pronostics favorable, intermédiaire et défavorable sont déterminés en fonction du score IMDC (International Metastatic Database Consortium) (Ko *et al.*, 2015).

Les recommandations les plus récentes sont les « *Guidelines* » émises par l'EAU (European Association of Urology, 2021), prenant en compte les résultats des derniers essais cliniques (Figure 15). Le recours à des thérapies différentes appliquées séquentiellement est validé dans la prise en charge du mRCC et permet d'améliorer la survie globale des patients. Parmi les changements récents majeurs, les mRCC de bon pronostic peuvent désormais être traités par ICI dès la première ligne de traitement (Figure 15), alors que ceux-ci n'étaient recommandés qu'en seconde ligne dans le cas d'une rechute après traitement par TKI en 2019 (Ljungberg *et al.*, 2019). L'immunothérapie devient ainsi le standard de traitement pour tous les groupes pronostiques, notamment en combinaison soit avec un autre ICI, soit avec un TKI. Les TKI sont utilisés préférentiellement en alternative, dans le cas de toxicité trop importante due à l'immunothérapie, ou bien en seconde ligne de traitement après immunothérapie (Figure 15).

#### **1.1.8.3.1. Thérapies ciblées anti-angiogéniques**

De nombreuses stratégies de thérapies ciblées dans le mRCC exploitent son haut degré de vascularisation et ciblent l'axe de signalisation du VEGF. En particulier, les TKI approuvés en première et deuxième lignes pour le traitement du mRCC dans l'Union européenne sont le sorafenib, le sunitinib, le pazopanib, l'axitinib, le lenvatinib et le cabozantinib (Tableau 6) (Hemminki *et al.*, 2020). Une étude de phase 3 rapportée en 2007 fait date en matière de traitement systémique de première ligne du mRCC en révélant la supériorité du sunitinib sur l'interféron- $\alpha$  en termes de taux de réponse, de survie sans progression et de survie globale (Motzer *et al.*, 2007). Le sunitinib est devenu par la suite une norme de soins et, bien qu'il ait été rejoint, voire remplacé, par de nombreux TKI et ICI (Dutcher *et al.*, 2020), il reste une référence pour la comparaison des nouvelles molécules en phase 3. Dans les recommandations actuelles, le Sunitinib et le Pazopanib sont des alternatives possibles en première ou seconde ligne de traitement. L'Axitinib, le Lenvatinib et le Cabozantinib sont eux recommandés en combinaison avec un ICI en première ligne ou bien en seconde ligne de traitement, notamment après une première ligne d'ICI (Figure 15).

Le Temsirolimus et l'Everolimus, deux molécules inhibitrices ciblant mTOR, ont aussi un potentiel anti-angiogénique. Le Temsirolimus était encore recommandé en 2019 en alternative aux ICI en première ligne de traitement dans le mRCC de mauvais pronostic ; aujourd'hui, l'Everolimus reste recommandé en seconde ligne de traitement.

Molécule	Approbation de l'EMA	Cibles	Type de molécule
Interféron-alpha	-	Augmente les Natural Killer et l'activité cytotoxique des lymphocytes T	Cytokine
Aldesleukin (IL-2)	1992	Facteur de croissance des lymphocytes T	Cytokine
Sorafenib	2005	RAF kinase, VEGFR-2, VEGFR-3, PDGFR, KIT, FLT-3	TKI
Sunitinib	2006	PDGFRs, c-KIT, RET, CD114, CD135	TKI
Temsirolimus	2007	mTOR	Molécule inhibitrice
Bevacizumab	2009	VEGF-A	TKI
Everolimus	2009	mTOR	Molécule inhibitrice
Pazopanib	2009	VEGFR, PDGFR, c-KIT, FGFR	TKI
Axitinib	2012	VEGFR1/2/3	TKI
Nivolumab	2015	PD-1	Anticorps
Cabozantinib	2016	VEGFR2, MET, AXL	TKI
Lenvatinib	2016	VEGFR1/3, FGFR1/4, PDGFR $\alpha$ , RET, KIT	TKI
Ipilimumab	2018	CTLA-4 + PD-L1	Anticorps
Pembrolizumab	2019	PD-1 + VEGFR1, VEGFR2/3	Anticorps
Avelumab	2019	PD-L1 + VEGFR1, VEGFR2/3	Anticorps

Tableau 6. **Molécules possédant une AMM en Europe en 2021**

Ce tableau est inspiré de la revue de Hemminki *et al.*, 2020. EMA = European Medicines Agency. AMM = Autorisation de mise sur le marché.

### 1.1.8.3.2. Immunothérapies

Avant l'introduction du Sunitinib, les cytokines telles que l'interféron- $\alpha$  et l'IL-2 à haute dose qui renforcent l'activité immunitaire anti-tumorale constituaient la norme de soin du mRCC depuis les années 1990. Ces deux médicaments, associés à une toxicité importante ; en particulier dans le cas de l'IL-2 à forte dose, ne profitaient généralement qu'à un petit sous-groupe de patients. Bien qu'ils induisent également certains degrés de toxicité selon le profil du patient (Ramos-Casals *et al.*, 2020), les immunothérapies de nouvelle génération, introduites avec le Nivolumab en 2015 ont progressivement remplacé les cytokines et ont révolutionné le paysage thérapeutique du mRCC (Hemminki *et al.*, 2020). Elles se présentent sous la forme d'inhibiteurs de point de contrôle immunitaire (ICI) (Figure 16).

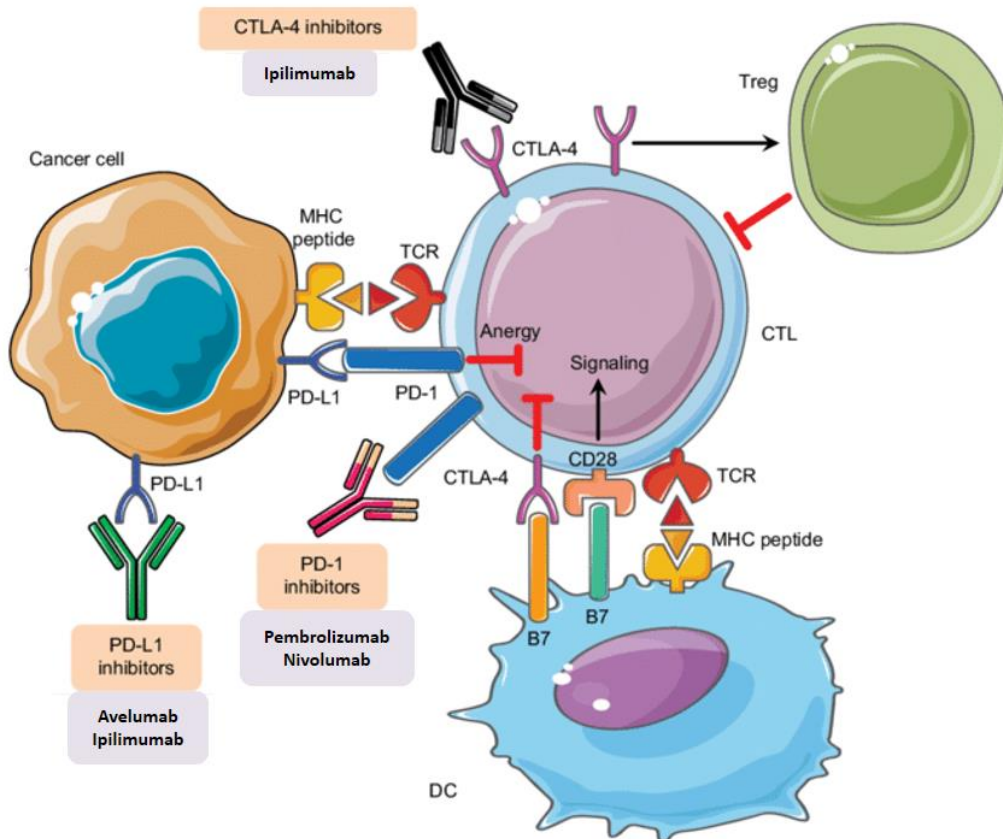


Figure 16. Mécanismes d'actions des inhibiteurs de checkpoint immunitaire.

L'inhibition des lymphocytes T cytotoxiques (CTL) dans le microenvironnement tumoral par les Tregs ou les points de contrôle immunitaires permet aux cellules cancéreuses d'échapper à l'attaque immunitaire et de se développer. Les ligands B7 exprimés sur les cellules présentatrices d'antigènes se lient au récepteur CD28 des CTL, ce qui entraîne une amplification des cellules T et une réponse immunitaire. Par ailleurs, la liaison des ligands B7 au CTLA-4 exprimé sur les cellules T supprime leur activité. CTLA-4 renforce également l'activité des Tregs, ce qui entraîne une activité immunosuppressive. PD-1 est exprimé sur les cellules T activées. PD-1 se lie à son ligand PD-L1, ce qui entraîne l'anergie des CTL et favorise les signaux inhibiteurs. L'inhibition pharmacologique des points de contrôle immunitaire à l'aide d'anticorps monoclonaux restaure l'activité antitumorale des CTL et lève l'immunosuppression. Abréviations : CTLA-4, antigène 4 des lymphocytes T cytotoxiques ; CTLs, lymphocytes T cytotoxiques ; DC, cellule dendritique ; CMH, complexe majeur d'histocompatibilité ; PD-1, mort cellulaire programmée-1 ; PD-L1, ligand de la mort cellulaire programmée-1 ; TCR, récepteur des cellules T ; Tregs, cellules T régulatrices. Figure tirée de Ayoub, Al-Shami and Yaghan, 2019.

L'Ipilimumab est un anticorps inhibiteur de CTLA-4, qui joue un rôle dans l'activation de la réponse immunitaire. Le Nivolumab et le Pembrolizumab sont des anticorps dirigés contre la protéine de mort cellulaire programmée 1 (PD1). L'Avelumab est un anticorps dirigé contre le ligand 1 de la protéine de mort cellulaire programmée (PDL1). PD1 régule négativement la fonction des cellules T et son ligand PDL1 s'exprime de façon aberrante à la surface des cellules cancéreuses. Par conséquent, le blocage de l'axe PD1-PDL1 ou de CTLA-4 favorise l'activation des cellules T et la destruction immunitaire du cancer. Ces ICI sont actuellement le standard de traitement, utilisés en combinaison avec des TKI ou d'autres ICI (Figure 15).

### 1.1.8.3.3. Nouvelles cibles thérapeutiques

Les protéines kinases sont des molécules que l'on peut aisément cibler avec des molécules inhibitrices. De nombreux essais en cours tentent de mettre à jour de nouvelles cibles pertinentes dans le traitement du ccRCC métastatique. Des études ont par exemple démontré l'intérêt de cibler CK2 dans le ccRCC (Rabjerg *et al.*, 2017), ou bien CK2 et la protéine ATM en combinaison thérapeutique (Giacosa *et al.*, 2021). D'autres molécules ciblant la voie PI3K/AKT (Pictilisib ou GDC-0941) et la molécule SRC (Saracatinib ou AZD0530) sont également potentiellement intéressantes étant donnés les rôles de leur cible dans la biologie du ccRCC (Lee *et al.*, 2014; Guo *et al.*, 2015; Roelants *et al.*, 2018a).

### 1.1.8.4. Échappement aux thérapies

Comme dans d'autres maladies cancéreuses, il y a un échappement thérapeutique important dans le ccRCC (Kamli, Li and Gobe, 2019; Jian *et al.*, 2021). En effet, près de 30% des patients sont d'emblée insensibles aux TKI anti-angiogéniques et 70% de ceux qui répondent finissent par développer une résistance (Sharma *et al.*, 2021). Des études récentes ont montré l'établissement d'une voie angiogénique alternative et sa conséquence dans l'acquisition de résistances thérapeutiques (X. Wang *et al.*, 2014; Rezapourbehnagh *et al.*, 2019). Concernant les immunothérapies développées ces cinq dernières années, une importante proportion de patients y est insensible ou développe une résistance après quelques mois de traitement (Sharma *et al.*, 2021). Ce champ d'étude est vaste et, pour l'essentiel, peu exploré. Le rôle du TME dans la résistance acquise a été mise en évidence dans le ccRCC (Finke *et al.*, 2013). Notamment, l'hypoxie résultant de la perte d'activité de VHL induit des modifications d'expression génique dans la tumeur et les cellules associées au TME, accélérant la carcinogenèse et induisant des phénotypes malins plus résistants (Jing *et al.*, 2019). Par ailleurs, l'hypoxie diminue également le potentiel de pénétration des molécules thérapeutiques et augmente l'expression de gènes impliqués dans l'efflux de molécules, à la fois dans les cellules malines et endothéliales. Enfin, dans le cancer de l'ovaire, une équipe a démontré qu'un environnement hypoxique favorisant la glycolyse – comme c'est le cas dans le ccRCC, augmentait la production d'acides lactiques, ayant pour conséquence un pH bas qui supprimerait les fonctions prolifératives et cytotoxiques des cellules immunitaires (Ahmed *et al.*, 2018). Cette chaîne mécanistique pourrait se retrouver dans le ccRCC et expliquer en partie les acquisitions de résistance observées.

## 1.2. DE L'ÉVOLUTION ET DE L'HÉTÉROGÉNÉITÉ TUMORALE

### 1.2.1. GENERALITES

Bien que le cancer ait des grandes caractéristiques fondamentales ou « *hallmarks* », il se manifeste de façon très variée selon les individus, y compris quand il naît dans un même organe. L'hétérogénéité inter-tumorale désigne l'hétérogénéité génétique et phénotypique des tumeurs malignes d'un patient à l'autre (Figure 17). À une autre échelle, l'hétérogénéité intra-tumorale (HIT) quant à elle se réfère à la variété de cellules malignes et non-malignes qui peuvent co-exister au sein d'une tumeur et éventuellement ses métastases, chez un même individu. Cette HIT résulte de processus évolutifs complexes au sein de la tumeur.

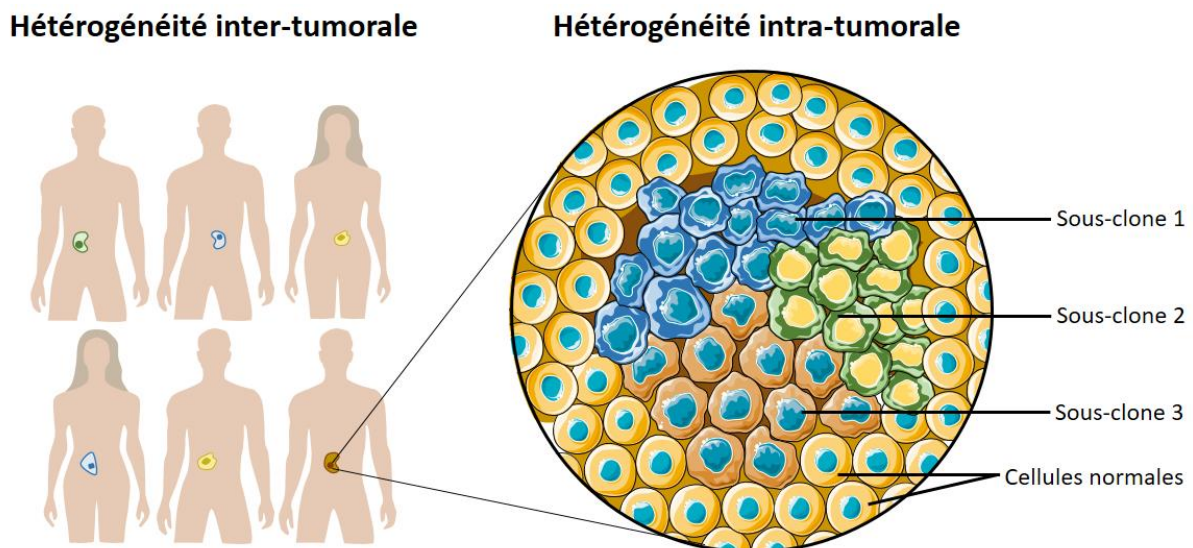


Figure 17. **Notion d'hétérogénéité tumorale**

L'hétérogénéité tumorale désigne à la fois l'hétérogénéité inter-tumorale et intra-tumorale, bien que le terme soit employé essentiellement pour décrire l'HIT.

#### 1.2.1.1. Concepts historiques

En 1833, le physiologiste Johannes Muller est le premier à utiliser le microscope, inventé près de 200 ans auparavant ('Invention of the Microscope', 1858), pour étudier et décrire les tissus humains cancéreux, employant des techniques alors utilisées par les botanistes et physiologistes (Hajdu, 2004) (Figure 18). Avant cela, les observations étaient réalisées à l'œil nu et ne permettaient pas d'apprécier les subtilités histologiques des tissus. Il a découvert au sein d'une même tumeur des sous-types histologiques différents et a noté des variations entre les zones de tissu conjonctif adjacentes à la tumeur.

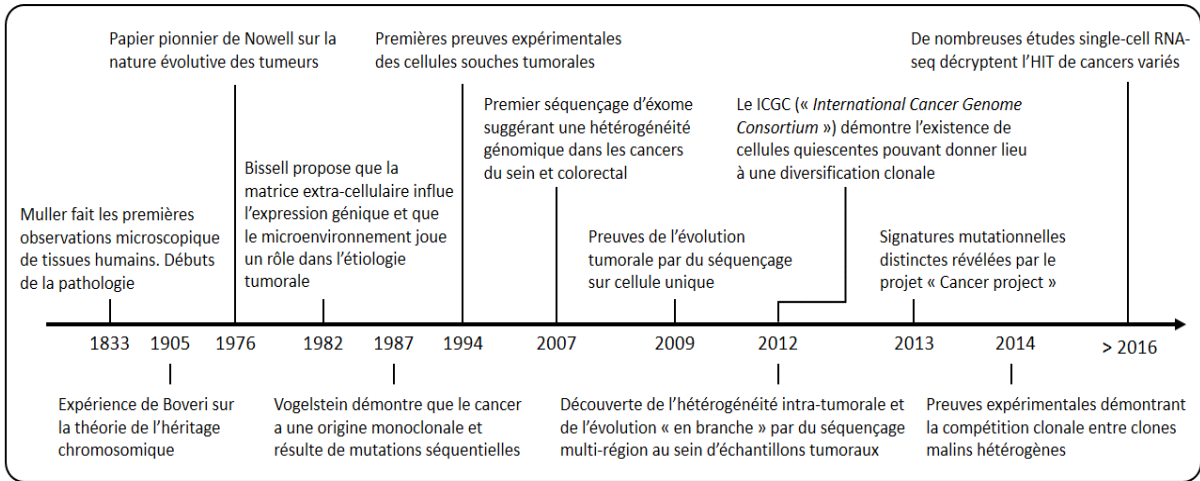


Figure 18. **Evolution du concept d'hétérogénéité intra-tumorale.**

Les prémices du concept d'hétérogénéité intra-tumorale tel qu'il est connu aujourd'hui, remontent au début du 19<sup>e</sup> siècle lorsque les premières observations microscopiques des tissus humains laissent voir différentes populations de cellules.

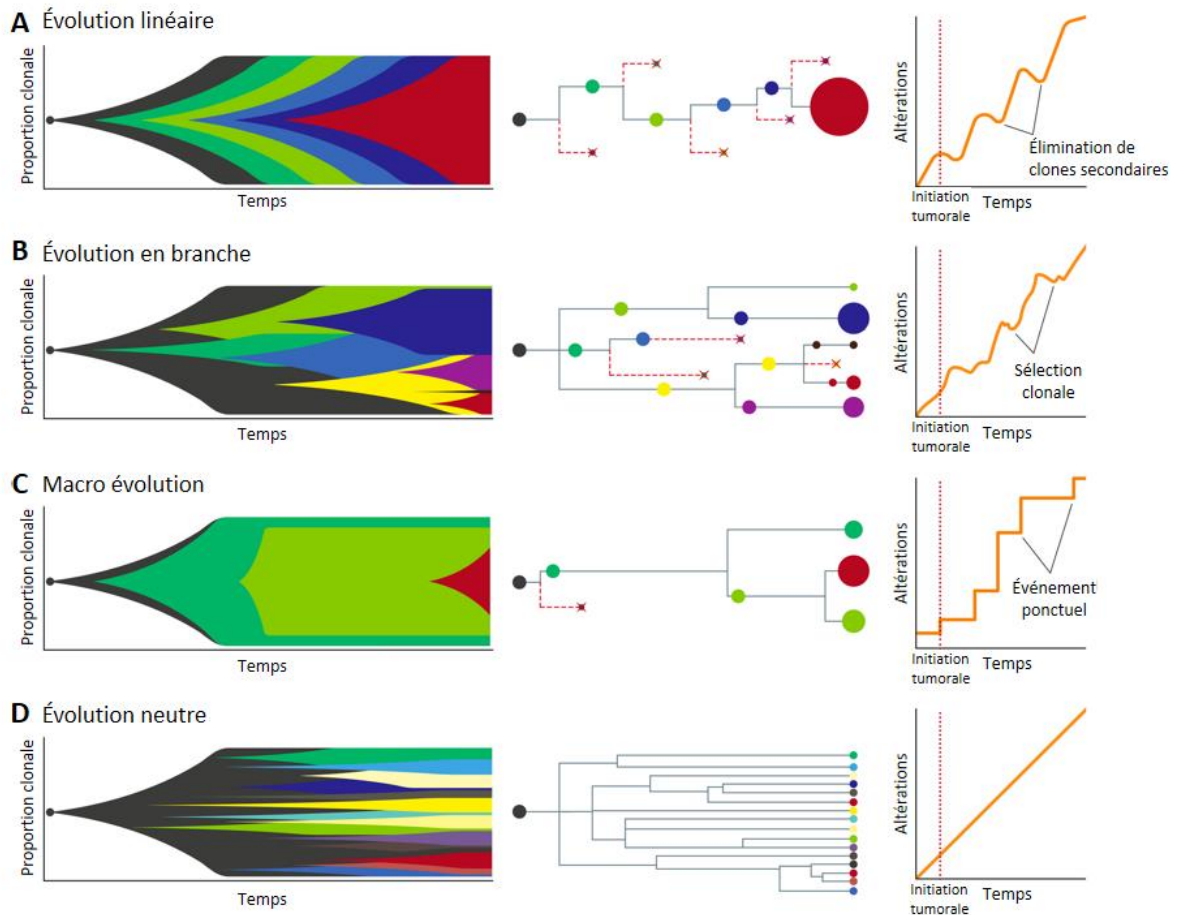


Figure 19. **Modèles d'évolution tumorale**

Figure adaptée de Vendramin, Litchfield and Swanton, 2021.



Ses travaux l'ont mené à formuler l'hypothèse que les tumeurs sont composées de nouvelles cellules au sein d'un organe malade. C'est l'élève de Muller, le pathologiste Rudolf Virchow, qui déterminera plus tard que toutes les tumeurs dérivent en fait de la transformation de cellules normales. Muller et Virchow ont transformé la médecine moderne non seulement en inventant le domaine de la pathologie, mais aussi en apportant certaines des premières preuves de l'hétérogénéité tumorale (Balkwill and Mantovani, 2001).

Plus tard, au début des années 1900, les travaux du biologiste allemand Theodor Boveri ont posé les fondations d'une grande partie de notre compréhension actuelle des origines et de l'hétérogénéité tumorales (Boveri, 2008). Il a observé l'association entre les mitoses aberrantes et les tumeurs malignes, constaté que les cellules ayant des combinaisons chromosomiques différentes étaient phénotypiquement dissemblables. Cela l'a mené à deux conclusions principales ; 1- les chromosomes transmettent des facteurs d'hérédité différents et, 2- la distribution inégale des chromosomes est préjudiciable au développement normal des cellules (Brokaw, 2000). Il a ainsi compris que c'est à la suite d'altérations génétiques que les cancers naissent de cellules normales, et a postulé que la plupart des tumeurs et leurs métastases proviennent d'une seule cellule. Des décennies plus tard, ses hypothèses ont été confirmées par des études clefs, faisant des travaux de Boveri l'un des jalons les plus importants de la recherche sur le cancer (Rowley, 1973; Nowell, 2007).

Aujourd'hui, le modèle d'évolution tumorale est généralement assimilé à la théorie de l'évolution par la sélection naturelle élaborée par Darwin au 19<sup>e</sup> siècle (Darwin, 1860). En 1976, Peter Nowell est le premier à appliquer concrètement les concepts évolutionnistes à l'oncologie et à formuler une hypothèse de l'évolution des tumeurs (Nowell, 1976) (Figure 18). Son modèle stipule que la plupart des cancers provient d'une seule cellule néoplasique et évolue selon un processus de sélection des altérations somatiques, conduisant à la prolifération et à la survie des clones les plus agressifs. Par la suite, en dépit de cette avancée conceptuelle, le point de vue de Peter Nowell a été globalement négligé, et l'évolution des tumeurs fut traditionnellement considérée comme une succession linéaire de divisions cellulaires clonales. Dans le cadre du modèle linéaire, les altérations s'accumulent dans les cellules progénitrices de manière progressive et confèrent aux cellules un avantage sélectif important, leur permettant de surpasser les clones précédents (Figure 19). Par conséquent, les tumeurs seraient composées de clones identiques résultant de l'élimination continue des clones devenus secondaires. Cependant, ce point de vue a été remis en cause par les travaux de Gloria Heppner, qui a démontré que les tumeurs sont composées de sous-clones génétiquement et phénotypiquement



distincts (Dexter *et al.*, 1978). En s'inspirant de la génétique des populations, elle assimile les tumeurs à des "sociétés hautement adaptées à la survie" et postule que les tumeurs survivent à la sélection naturelle et artificielle (thérapeutique) grâce à l'hétérogénéité, en produisant de nouveaux clones susceptibles d'être sélectionné (Heppner, 1984). Ces observations dévoilent un modèle selon lequel les tumeurs se développent de manière non pas linéaire mais ramifiée, avec de multiples sous-clones dérivés du même ancêtre commun, qui finissent par diverger et acquérir des aptitudes différentes (Greaves and Maley, 2012; Swanton, 2012) (Figure 19). Dans les années 1980, Bissell initie l'idée désormais bien acceptée que le microenvironnement tumoral (MET) est intimement lié à l'étiologie tumorale et participe à son hétérogénéité cellulaire (Bissell, Hall and Parry, 1982; Bissell and Hines, 2011). Il avance également le rôle de la matrice extracellulaire dans la régulation de l'expression génique, ce qui sera validé par la suite (Meacham and Morrison, 2013; C. C. Wang *et al.*, 2014). Les expériences et avancées technologiques à la suite de ces travaux pionniers ont permis de déchiffrer plus avant l'hétérogénéité tumorale dans de nombreux sites et de façonner le concept tel qu'on le connaît aujourd'hui (Figure 18).

### **1.2.1.2. Modèles d'évolution tumorale**

L'HIT est largement reconnue dans les cancers humains, mais les raisons de son émergence pendant le développement de la maladie restent peu connues (Davis, Gao and Navin, 2017). La principale raison à cela est la difficulté de collecter chez un patient des échantillons tumoraux provenant de la même tumeur à des temps différents. L'absence de ces jalons chronologiques a poussé à inférer l'évolution tumorale à partir d'échantillons ne représentant qu'une fraction temporelle donnée. Cela a mené à l'émergence de différents modèles d'évolution tumorale en compétition, comprenant essentiellement les modèles linéaire, ramifié, neutre, et ponctué (Vendramin, Litchfield and Swanton, 2021) (Figure 19). Chaque modèle suppose sa propre logique d'apparition des mutations et de la sélection des clones, et en conséquence a des implications différentes concernant le diagnostic et le traitement des patients.

#### **1.2.1.2.1. Modèle d'évolution linéaire**

Ce modèle postule que les mutations sont acquises de manière linéaire, étapes après étapes, conduisant à des stades de plus en plus agressifs (Fearon R. Eric and Bert, 1990). Les nouvelles mutations offrent un avantage sélectif si important que les clones qui les portent supplantent tous les clones précédents (Figure 17A). L'arbre phylogénétique qui en résulte

montre un clone dominant avec les restes de quelques clones qui ont persisté sans pour autant croître après la pression de sélection. Du point de vue conceptuel, le modèle linéaire a été un cadre utile pour comprendre que l'acquisition séquentielle de mutations menait à un "potentiel agressif" plus important. Cependant, en dépit de certaines études supportant le modèle linéaire (Linder and Gartler, 1965; Fialkow, 1974; Fearon R. Eric and Bert, 1990; Sawada *et al.*, 1994), les interprétations expérimentales ne constituent pas de preuves tangibles car ces études ne considèrent les statuts mutationnels que d'un ou de quelques gènes, laissant de côté d'autres marqueurs potentiellement hétérogènes et aptes à définir des clones distincts.

#### 1.2.1.2.2. Modèle d'évolution ramifiée

Dans le modèle d'évolution ramifiée, les clones malins divergent à partir d'un ancêtre commun et évoluent en parallèle dans la tumeur pour donner naissance à de multiples lignées clonales (Figure 17B). Contrairement au modèle linéaire, les balayages sélectifs sont rares et plusieurs clones peuvent s'y adapter et se développer simultanément. Le degré d'HIT fluctue au cours de la progression de la tumeur. Le modèle prévoit que de multiples clones soient présents lors d'un échantillonnage clinique. Les arbres phylogénétiques correspondant comportent des clones dominant en raison de l'avantage sélectif conféré par des mutations « *drivers* », et des clones intermédiaires (Figure 17B). Ce modèle est supporté par de nombreuses études génomiques réalisées sur différents cancers ; à l'échelle de la cellule unique (Navin *et al.*, 2011; Hou *et al.*, 2012; Li *et al.*, 2012; Gawad, Koh and Quake, 2014; Y. Wang *et al.*, 2014; Yu *et al.*, 2014; Gao *et al.*, 2016) , par séquençage multi-régional (Yachida *et al.*, 2010; Gerlinger *et al.*, 2012, 2014; De Bruin *et al.*, 2014; Zhang *et al.*, 2014; Aubry *et al.*, 2015; Boutros *et al.*, 2015; Kim *et al.*, 2015; Yates *et al.*, 2015; Hoffman *et al.*, 2016) ou par séquençage profond (Campbell *et al.*, 2008; Ding *et al.*, 2012; Nik-Zainal *et al.*, 2012; Shah *et al.*, 2012). De manière cohérente, ces études identifient des mutations appartenant soit à un tronc commun d'évolution, soit à quelques sous-clones, soit exclusivement à un clone tumoral. Elles peuvent différer cependant par le nombre de clones rapportés et la forme des arbres évolutifs. Dans certaines tumeurs, les arbres évolutifs ont des troncs courts et de nombreuses branches, indiquant des sous-clones avec des degrés de divergence importants comme c'est le cas dans certains cancers du sein et du rein (Gerlinger *et al.*, 2014; Y. Wang *et al.*, 2014; Yates *et al.*, 2015). D'autres arbres ont des troncs très longs et peu de branches, révélateurs de sous-clones présentant plus de similitudes que de différences (De Bruin *et al.*, 2014; Zhang *et al.*, 2014).

Le nombre de sous-populations clonales identifiées varie également de manière significative non seulement d'un type de cancer à l'autre mais également entre les patients atteints du même cancer. Dans le cancer du sein triple négatif par exemple, une étude de séquençage profond sur une cohorte de 104 patientes révèle entre 1 et 19 sous-clones par patiente (Shah *et al.*, 2012) tandis qu'une autre étude recourant au séquençage multirégional identifie 1 à 4 sous-clones majeurs par patiente (Yates *et al.*, 2015). Ces données suggèrent que le nombre de sous-clones présents dans les tumeurs humaines peut varier grandement. Cependant, il est possible que ces écarts soient dus à des différences dans la profondeur de séquençage ou le nombre de régions échantillonnées. Le modèle d'évolution ramifiée est considéré comme le plus pertinent pour décrire l'évolution tumorale vis-à-vis des mutations ponctuelle (Davis, Gao and Navin, 2017).

#### **1.2.1.2.3. Modèle d'évolution neutre**

L'évolution neutre est un cas extrême d'évolution ramifiée qui postule qu'il n'y a pas de sélection ou d'adaptation pendant la majeure partie de la durée de vie de la tumeur (Figure 17C). Ce modèle suppose que les mutations aléatoires s'accumulent au fil du temps, entraînant une dérive génétique et une HIT étendue. Dans ce cas, l'hétérogénéité ne participe pas à la progression tumorale mais n'en est qu'une conséquence. L'arbre phylogénétique correspond à une structure très ramifiée, constituée de nombreux nœuds intermédiaires. Le modèle d'évolution neutre était proposé à l'origine dans le cadre de l'évolution des espèces et remettait en question l'idée de sélection naturelle au cœur de la théorie de Darwin (Freese, 1962; Kimura, 1983). Dans les tumeurs, la présence de sous-clones et le principe d'évolution convergente de sous-clones distincts sont la marque d'une sélection naturelle. Cependant toutes les tumeurs ne présentent pas ces caractéristiques, ce qui laisse la possibilité d'une évolution neutre dans certains cas. Quelques études apportent des preuves que l'évolution neutre peut se produire dans certaines tumeurs, bien qu'elles ne permettent pas de démontrer que c'est le seul modèle à l'œuvre de la progression tumorale, notamment en raison de la faible profondeur de séquençage (Williams *et al.*, 2016) ou de faible nombre de réplica ( $n=1$ ) (Ling *et al.*, 2015).

#### **1.2.1.2.4. Modèle d'évolution ponctuée**

Les modèles linéaires ramifié et neutre supposent généralement que les mutations sont acquises de manière séquentielle et progressive au fil du temps. Cependant, des études récentes

ont suggéré que dans certains cancers, un grand nombre d'aberrations génomiques peut se produire dans des périodes très courtes, généralement aux stades les plus précoces de la progression tumorale (Figure 17D). Ces observations sont les fondements du modèle d'évolution ponctuée ; des événements dramatiques et soudains qui conditionnent la progression tumorale. Dans ce modèle, l'HIT est très importante aux premiers stades de l'initiation de la tumeur, à la suite de quoi un ou quelques clones dominants se développent de manière stable pour former la masse tumorale (Figure 17D). L'arbre phylogénétique correspondant est caractérisé par l'absence de clones intermédiaires pendant l'évolution de la tumeur, résultant en un nœud à la racine et un ou quelques clones dominants. Le terme "évolution ponctuée" est inspiré de "l'équilibre ponctué" (« *punctuated equilibrium* »), imaginé par Gould et Eldredge en 1970 dans le domaine de l'évolution des espèces pour remettre en question le paradigme l'évolution darwinienne progressive (Eldredge and Gould, 1972; Williams *et al.*, 2016).

### **1.2.1.3. Implication des cellules souches cancéreuses**

Les cellules souches cancéreuses (CSC) ont été identifiées pour la première fois en 1994 par John Dick et ses collègues (Lapidot *et al.*, 1994). Elles suscitent un très vif intérêt dans la recherche en oncologie depuis. En général faiblement représentées dans le tissu tumoral, elles sont considérées comme un facteur déterminant de l'HIT (Prasetyanti and Medema, 2017). Comme les cellules souches normales lors du développement, elles possèdent les capacités d'auto renouvellement et de différenciation (Prasetyanti and Medema, 2017). Dans leur genèse, les modèles d'évolution comprenant des CSC, autrement appelés modèles hiérarchiques, étaient opposés aux modèles d'évolution clonale ou « stochastique » qui prennent essentiellement en compte l'accumulation de mutations au sein de cellules différenciées (voir la section **1.2.1.2. Modèles d'évolution tumorale**) (Shackleton *et al.*, 2009). Le modèle hiérarchique propose que la tumorigenèse soit régulée par une population phénotypiquement distincte capable de donner naissance aux cellules qui formeront ensuite la tumeur. Ce modèle a des implications toutes autres car il suppose que l'élimination de cette seule population permettrait d'éradiquer totalement la tumeur. Cependant, des études ont démontré que les CSC et cellules non-souches au sein d'une tumeur pouvaient présenter une plasticité importante, et que la cellule originelle ne dérivait pas toujours d'une population souche (Kreso and Dick, 2014). Pour ces raisons, un troisième modèle dit « unifié » a été décrit et propose que les CSC pourraient acquérir des mutations générant des populations de CSC hétérogènes à l'origine de l'HIT et que les cellules

non-souches seraient capables d'acquérir des caractéristiques de CSC, notamment par la transition épithélio-mésenchymateuse (Figure 20).

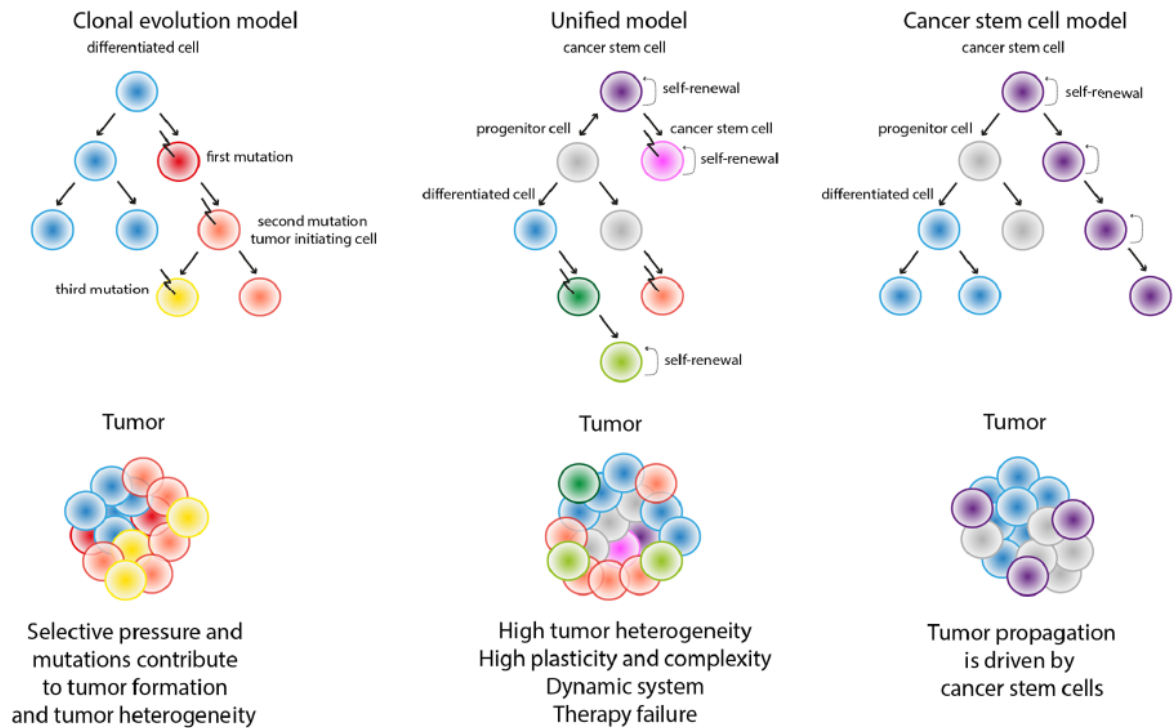


Figure 20. **Modèles de tumorigénèse**

Le modèle d'évolution clonale correspond à une évolution branchée. Tiré de Corrà and Moch, 2018.

### 1.2.2. IMPLICATION DANS L'ÉCHAPPEMENT THERAPEUTIQUE

La majorité des cancers avancés et métastatiques reste incurable, même dans les cas où les thérapies disponibles sont capables d'éliminer la grande majorité des cellules tumorales. Dans une tumeur donnée, un haut degré d'HIT augmente les chances qu'au moins certaines cellules tumorales survivent à l'élimination induite par le traitement, tandis que la diversification continue des phénotypes des cellules tumorales pendant le traitement permet aux cellules tumorales de s'adapter aux pressions sélectives imposées par le traitement, ce qui entraîne une résistance *de novo* et une éventuelle rechute (Marusyk, Janiszewska and Polyak, 2020).

#### 1.2.2.1. Échappement aux inhibiteurs de tyrosine kinase

Malgré les réponses initiales aux TKI parfois spectaculaires et le développement continu d'inhibiteurs plus efficaces, la résistance émerge de manière quasi inévitable dans les cancers avancés et métastatiques. Dans de nombreux cas de rechute, la résistance est associée à des

mutations génétiques spécifiques, telles que l'amplification génomique de la cible thérapeutique, des mutations ponctuelles réduisant l'impact inhibiteur, ou les amplifications et mutation d'autres gènes permettant aux cellules tumorales de maintenir leur potentiel oncogène (Lovly and Shaw, 2014). Cependant, ces changements mutationnels ne suffisent pas pour conférer une résistance totale expliquant la rechute clinique (Shaffer *et al.*, 2017). En réalité, le développement de la résistance est multifactoriel (Hong *et al.*, 2019; Vander Velde *et al.*, 2020) et repose en partie sur la diversité préexistante et la diversification en cours (Merlo *et al.*, 2006). Bien que son rôle soit communément admis, la pertinence fonctionnelle de l'HIT dans la résistance aux traitements reste mal comprise (Marusyk, Janiszewska and Polyak, 2020).

### **1.2.2.2. Échappement aux immunothérapies**

Des données récentes démontrent que l'HIT influence le succès de l'immunothérapie. Des résistances acquises à l'immunothérapie via l'hétérogénéité des néo antigènes, de la présentation de l'antigène et de la signalisation de l'interféron ont notamment été rapportées (Havel, Chowell and Chan, 2019; Kalbasi and Ribas, 2020). La charge mutationnelle tumorale (ou TMB pour « *Tumor Mutational Burden* ») définie comme le nombre de mutations somatiques au sein des tumeurs, est l'un des plus forts prédicteurs de la réponse aux inhibiteurs de checkpoint immunitaire (ICI), un TMB élevée étant en général associée à une meilleure réponse (Ribas and Wolchok, 2018). Cependant, un TMB élevé peut également augmenter l'HIT et des données récentes, tant dans des modèles expérimentaux que chez des patients, montrent que les tumeurs présentant une HIT élevée sont moins susceptibles de répondre aux ICI que les tumeurs plus homogènes (McGranahan *et al.*, 2016; Anagnostou *et al.*, 2017; Gejman *et al.*, 2018; Milo *et al.*, 2018; Y. Wolf *et al.*, 2019). La résistance à l'immunothérapie via l'HIT pourrait être due à différents mécanismes. D'une part, une HIT plus élevée est associée à un risque plus élevé d'avoir un clone résistant. Par ailleurs, certains sous-clones peuvent ne pas déclencher de réponse immunitaire efficace.

### 1.2.3. HÉTÉROGÉNÉITÉ INTRA-TUMORALE DANS LE CCRCC

L'hétérogénéité inter- et intra-tumorale dans le ccRCC est connue des pathologistes depuis longtemps et a un enjeu diagnostique important. Le ccRCC est notamment hétérogène des points de vues clinique, macroscopique et microscopique (López and Angulo, 2018) (Figure 21).

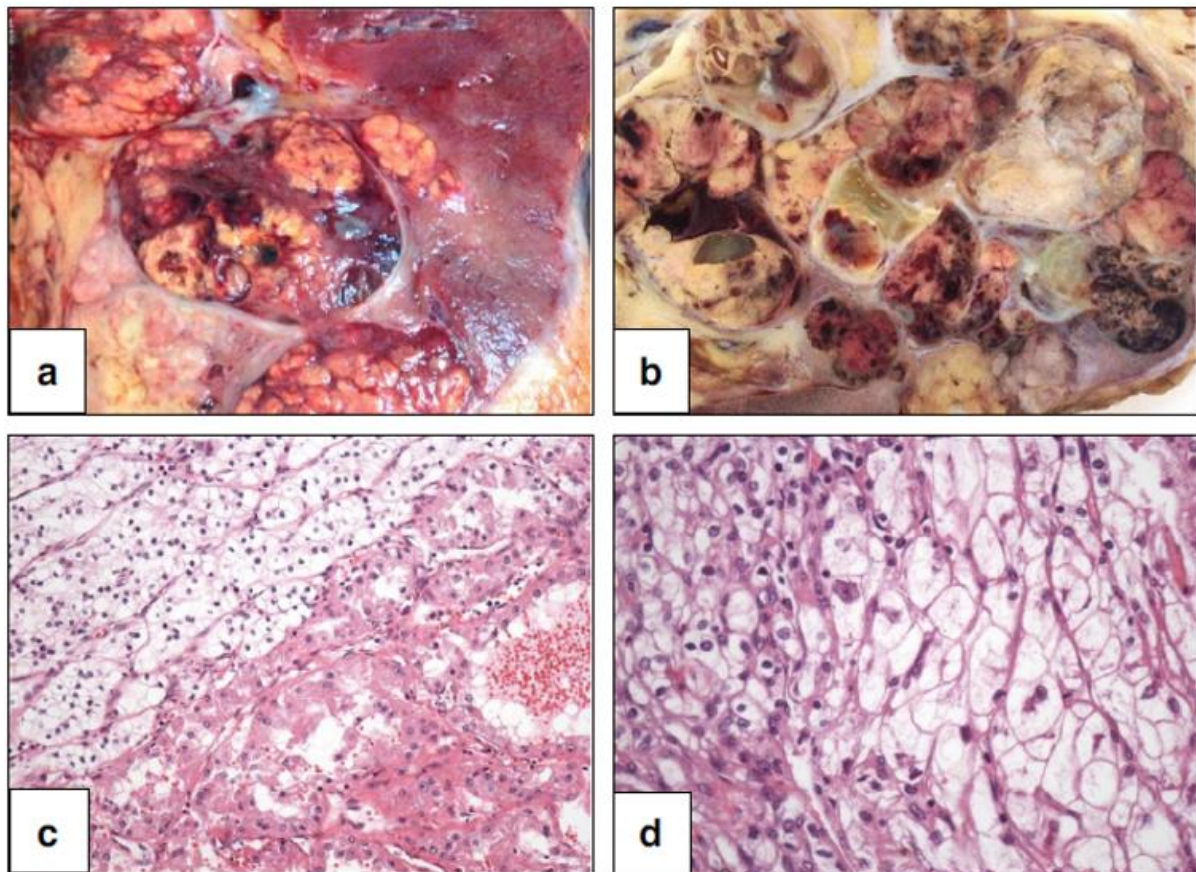


Figure 21. **Hétérogénéité intra-tumorale à l'échelle macro- et microscopique.**

Exemples frais (A) et fixés au formol (B) de carcinome rénal à cellules claires présentant des hétérogénéités macroscopiques évidentes, avec différents nodules tumoraux ainsi que des différences de couleurs et de textures. C-D. Vues microscopiques de l'hétérogénéité intra-tumorale sur des coupes à l'hématoxyline-éosine de deux cas de carcinome rénal à cellules claires. Tiré de (López and Angulo, 2018).

Une étude a également décrit une HIT au niveau métabolique dans le ccRCC, avec des variations intra-tumeur des voies du pyruvate et des niveaux de métaboliques issus de la glycolyse (Okegawa *et al.*, 2017). À l'échelle génomique, des stratégies d'échantillonnage multiple ont été employées pour démontrer l'HIT (Gerlinger *et al.*, 2012; Dagher *et al.*, 2021). Dans l'étude de Dagher et ses collaborateurs, les auteurs ont étudié par génomique et transcriptomique une tumeur rénale et trois sites distants, dont un thrombus veineux et deux métastases surrénaliennes avec plusieurs prélèvements par site. Ils ont démontré que les sites



étaient hétérogènes et polyclonaux, chaque site tumoral contenant deux populations clonales agressives différentes, avec des gènes exprimés de manière différentielle et impliqués dans des fonctions biologiques distinctes. Dans une étude portant sur un plus grand nombre d'individus, l'équipe de Charles Swanton a réalisé les séquençages multirégionaux de quatre patients ccRCC avant et après traitement dans les sites primaires ou métastatiques (Gerlinger *et al.*, 2012), mettant en évidence des mutations communes à tous les échantillons analysés chez chaque patient. Ces mutations dites « drivers », dans le cas du carcinome rénal, consistent essentiellement en des altérations du gène VHL et la perte d'hétérozygotie (« LOH », *loss of heterozygosity*) du bras 3p du chromosome 3. D'autres mutations bien connues telles que celles affectant SETD2, PBRM1, MTOR, PI3KCA, PTEN et KDM5C ont en revanche été retrouvées de manière hétérogène dans les échantillons de la tumeur primaire ou des territoires métastatiques (Figure 22).

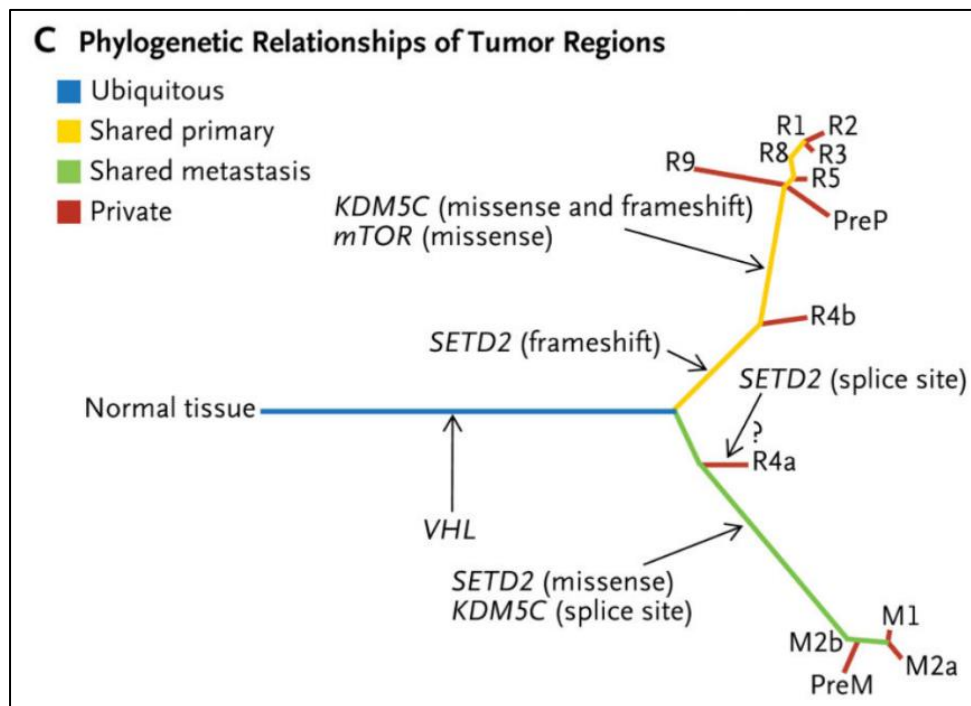


Figure 22. **Arbre phylogénétique dans le ccRCC**

L'hétérogénéité génomique peut résulter de l'accumulation séquentielle et parallèle de mutations, contribuant à l'évolution du ccRCC. Dans cet exemple, "R" représente les caractéristiques génomiques de la tumeur primaire et "M" celles des sites métastatiques, numérotés en conséquence. Les principales lésions génétiques acquises après la mutation de VHL se retrouvent dans différents échantillons et sont indiquées sur les branches. La perte de VHL est donc l'événement « troncale », commun à tous les échantillons. Les autres mutations sont retrouvées de façon partagée ou unique selon les échantillons, impliquant une évolution branchée. Figure tirée de Gerlinger *et al.*, 2012



Ces caractéristiques génétiques permettent de construire des phylogénies de tumeurs, dans lesquelles le "tronc" de l'arbre évolutif représente les mutations trouvées dans l'ancêtre commun le plus récent (MRCA) qui sont présentes dans chaque cellule tumorale (Figure 23). Ces résultats démontrent que c'est le modèle d'évolution ramifiée qui est à l'œuvre dans la progression tumorale du ccRCC (Figure 19).

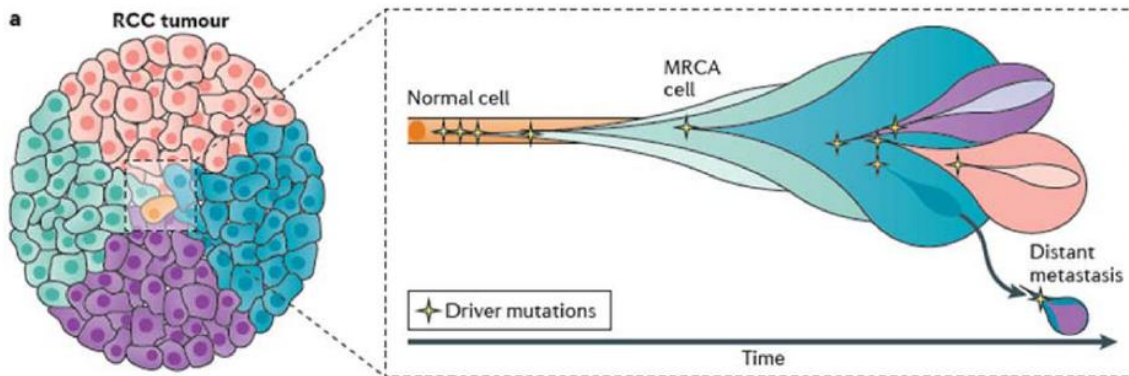


Figure 23. Évolution et hétérogénéité tumorale dans le ccRCC

Les sous-clones tumoraux sont issus de l'ancêtre commun le plus récent (« Most recent common ancestor » ou MRCA), dans lequel une cellule normale a acquis toutes les capacités nécessaires à la transformation maligne. Tiré de Hsieh et al., 2017.

### 1.3. DES MODELES D'ETUDE DES CCRCCS

Les modèles biologiques permettent de reproduire et étudier les caractéristiques des tumeurs. Ils ont ouvert la voie à l'innovation thérapeutique en cancérologie. Les « *hallmarks* » du cancer définies par Hanahan et Weinberg décrivent un ensemble dynamique de caractéristiques destinées à synthétiser et simplifier la complexité de la biologie du cancer et de la progression tumorale et à préparer le terrain pour des approches plus approfondies et des questions précises (Hanahan and Weinberg, 2011) (Figure 11). Cependant, toutes ces caractéristiques ne sont pas reproductibles dans un même système d'étude. Différents modèles de culture existent en oncologie et parmi eux, les lignées cellulaires primaires ou immortalisées ainsi que les organoïdes, modèles murins et modèles de culture *ex-vivo* ont été exploités dans la recherche sur le ccRCC (Roelants *et al.*, 2020; Wolf, Kimryn Rathmell and Beckermann, 2020).

#### 1.3.1. SYSTEMES 2D : LIGNEES CELLULAIRES DE CCRCC

Ross Granville Harrison a réalisé les premières cultures cellulaires en 1907 dans ses recherches sur l'origine des fibres nerveuses (Harrison *et al.*, 1907). Depuis, la culture cellulaire a été améliorée et utilisée pour observer en dehors du corps de nombreux mécanismes cellulaires telles que la croissance et la différenciation des cellules (Harrison, 1910; Scudiero *et al.*, 1988). Aujourd'hui, les expériences peuvent être menées à l'aide de cellules primaires isolées directement à partir de tissus humains ou à l'aide de cultures immortalisées, établies et déposées dans des banques de cellules (Paul, 1980). Les lignées immortalisées sont le système biologique le plus fréquemment utilisé dans la recherche en oncologie et ont considérablement contribué à notre compréhension moléculaire et génétique actuelle de la biologie du ccRCC (Kondo *et al.*, 2001). De par son rapport coût-efficacité, sa facilité de maintenance et de manipulation, et sa capacité de réplication infinie, ce système respecte la loi des trois "R" en recherche : remplacement, réduction et raffinement des modèles animaux (Flecknell, 2002; Baumans, 2004). Dans les années 1980 aux Etats-Unis, le NCI (National Cancer Institute) débutait l'étude de nombreux médicaments anticancéreux en utilisant 60 lignées cellulaires comme modèle (NCI-60) (Shoemaker, 2006). L'étude a conduit à la caractérisation intensive et à la mise à disposition du public de données génétiques et moléculaires spécifiques de chacune des lignées. Parmi elles se trouvaient huit lignées de RCC ; 786-O, A-498, ACHN,

CAKI-1, RXF 393, SN12C, TK-10, and UO-31. Ces lignées ont dès lors été établies comme modèle de premier plan dans la recherche sur le ccRCC.

La lignée 786-O est la plus fréquemment citée aujourd'hui (Brodaczewska *et al.*, 2016). Classée dans le type ccRCC, elle comprend une délétion de VHL, et les cellules 786-O affichent systématiquement une expression accrue des protéines HIF-2 $\alpha$  et VEGF, ce qui en fait un modèle de premier ordre pour le ccRCC (Tate *et al.*, 2019). C'est dans cette lignée qu'a été démontré pour la première fois le rôle régulateur de l'axe pVHL-HIF. Lorsqu'elles sont transfectées avec la protéine VHL sauvage, la croissance tumorale *in vivo* est inhibée (Iliopoulos *et al.*, 1995, 1996). Les cellules 786-O ont également été utilisées pour montrer que la surexpression de HIF seule favorise la croissance tumorale, validant ainsi que l'activation de HIF se produit en aval de l'interaction pVHL (Kondo *et al.*, 2002). Plus récemment, les cellules 786-O ont été utilisées pour déchiffrer le rôle de la résistance induite par l'hypoxie à l'inhibiteur du VEGF, le sorafenib, via la voie HIF-2 $\alpha$ -Cox2, illustrant ainsi un mécanisme potentiel de résistance aux médicaments (Zhao, Luo and Wu, 2015). Enfin, des populations semblables à des cellules souches ont été observées dans les cellules 786-O, en faisant un modèle potentiel pour les études sur les CSC (Brodaczewska *et al.*, 2016).

La lignée A-498 est également classée comme ccRCC et est la deuxième plus citée (Wolf, Kimryn Rathmell and Beckermann, 2020). Elle a été utilisée avec la lignée 786-O pour démontrer que l'activation de mTOR activait la tumorigenèse dans le ccRCC (Robb *et al.*, 2007). Les mêmes lignées ont permis d'illustrer la redondance fonctionnelle et le rapport de dominance des sous-unités HIF-1 $\alpha$  et HIF-2 $\alpha$  dans la transcription de VEGF (facteur angiogénique) (Shinojima *et al.*, 2007).

Les cultures 2D ont les avantages d'être facilement accessibles, simples à maintenir et peu coûteuses. Cependant, elles présentent également de nombreux inconvénients. Premièrement, les cellules cultivées en 2D ne reproduisent pas les structures naturelles des tumeurs ni la physiologie tumorale. Les interactions cellule-cellule et cellule-environnement n'existent pas comme dans la masse tumorale. Or, ces interactions sont responsables de la différenciation, de la prolifération et de la vitalité des cellules, de l'expression des gènes et des protéines, de la réactivité aux stimuli, du métabolisme des médicaments et d'autres fonctions cellulaires. En raison de ces inconvénients dans les systèmes 2D, des modèles alternatifs ont été recherchés afin d'imiter une masse tumorale naturelle, tels que les systèmes de culture 3D.

### 1.3.2. SYSTEME 3D

#### 1.3.2.1. *Organoïdes et sphéroïdes*

Les organoïdes et sphéroïdes sont des structures en trois dimensions formées à partir de culture de cellules. Parfois confondus, les termes renvoient à l'origine cellulaire : un organoïde est la version simplifiée et miniaturisée d'un organe produit *in vitro* à partir de cellules souches ou primaires, tandis qu'un sphéroïde est formé à partir d'une lignée cellulaire uniforme. Dans le cas d'un organoïde tumoral, les cellules proviennent d'une tumeur et peuvent contenir plusieurs lignages cellulaires, ce qui le distingue comme le modèle le plus complexe des deux. Hamburg et Salmon sont les premiers à avoir rapporté l'établissement d'un système de culture tridimensionnel dans les années 1970 (Hamburger and Salmon, 1977). Depuis, de nombreuses similarités ont été décrites entre les organoïdes tumoraux et les masses tumorales naturelles correspondantes en termes de morphologie et de développement cellulaire (Pampaloni, Reynaud and Stelzer, 2007; Mazzoleni, Di Lorenzo and Steimberg, 2009).

Les organoïdes permettent d'obtenir une organisation structurelle *in vitro* comparable à l'architecture d'une tumeur *in vivo*, tout en conservant le milieu vasculaire, stromal et inflammatoire qui constitue le microenvironnement tumoral (Clevers, 2016; Meijer, Jager and Gent, 2017). Des organoïdes tumoraux ont été créés avec succès à partir de tumeurs primaires colorectales (Van De Wetering *et al.*, 2015), pancréatiques (Boj *et al.*, 2015), mammaires (Sachs *et al.*, 2018) ou pulmonaires (Finnberg *et al.*, 2017). Dans le domaine du ccRCC, Grassi et ses collègues ont été les premiers à établir et à caractériser avec succès des cultures d'organoïdes générés à partir de cellules souches pluripotentes issues à la fois de rein sain et du tissu tumoral correspondant (Grassi *et al.*, 2019). Les auteurs ont comparé par immunofluorescence les structures des organoïdes normaux et tumoraux et ont constaté des différences structurelles significatives. Ils ont pu identifier et mesurer dans les deux types de cultures des marqueurs de régions spécialisées du néphron comme les tubules contournés proximaux et distaux, la capsule de Bowman ou l'anse de Henlé. Un autre groupe a développé une méthode d'interface air-liquide pour cultiver des organoïdes à partir de biopsies humaines de ccRCC ; ils ont pu retenir dans une matrice de collagène les cellules malignes primaires aux côtés des compartiments immunitaires et non-immunitaires transformés du microenvironnement (Neal *et al.*, 2018).

De façon remarquable, ils ont appliqué des immunothérapies anti PD-1 et PD-L1 et observé l'expansion de lymphocytes et la production de lymphocytes T CD8 cytotoxiques qui ont eu pour conséquence l'élimination des cellules tumorales et la restauration des points de contrôle immunitaire.

Malgré ces avancées considérables dans le domaine de la culture cellulaire, la technique organoïde présente quelques limitations qui nécessiteront de futures développements (Wolf, Kimryn Rathmell and Beckermann, 2020). Pour des raisons qui restent à élucider, il a été observé dans plusieurs études un avantage prolifératif des cellules épithéliales normales présentes dans le microenvironnement en présence de facteurs de croissance dans le milieu (Karthaus *et al.*, 2014). Dans certains modèles, la privation initiale de facteurs de croissance, suivie d'une forte supplémentation en nutriments réduit la contamination des cellules épithéliales normales (Neal *et al.*, 2018). Ce phénomène révèle que les cellules répondent de façon particulière aux conditions de culture *in vitro*, qui sont très probablement différentes dans une tumeur *in vivo*. Ainsi, les améliorations des modèles organoïdes dans le ccRCC passeront par l'analyse approfondie du microenvironnement tumoral.

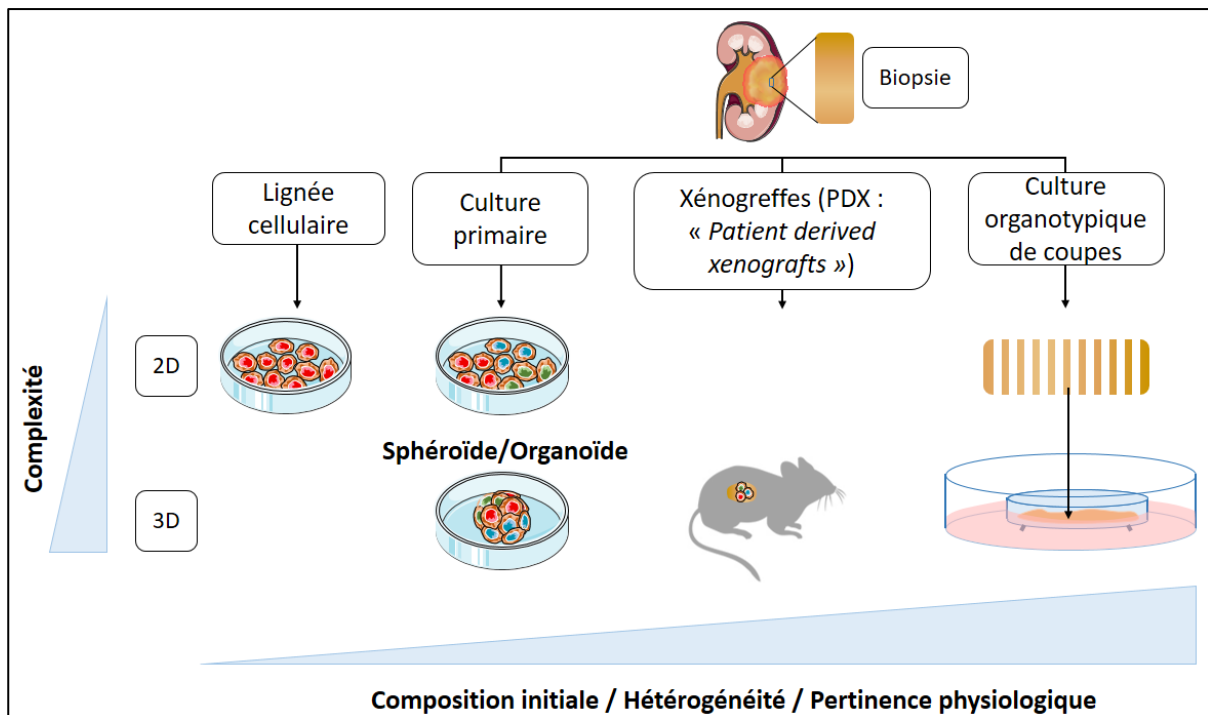


Figure 24. Différents systèmes de culture tumorale

Ce diagramme représente différents systèmes de culture. Un système de culture *ex vivo* doit être choisi en fonction de l'objectif de recherche. Chacun présente des avantages et inconvénients. Un échantillon de tumeur peut être dissocié à l'aide de méthodes enzymatiques et/ou mécaniques, puis cultivé sous la forme d'une monocouche en 2D ou d'une culture organoïde de tumeur en 3D. Les échantillons de tumeurs peuvent être implantés dans des souris immunodéficientes pour générer des modèles PDX. Des coupes de tissu tumoral peuvent être réalisées puis cultivées, permettant de conserver l'architecture générale de la tumeur.

Un autre désavantage dans ce modèle réside dans sa capacité limitée à restituer la composition et l'hétérogénéité initiale de la tumeur, notamment en raison de la faible proportion de tumeur échantillonnée et cultivée (Figure 24 ; Tableau 7) (Wolf, Kimryn Rathmell and Beckermann, 2020). Par ailleurs, la reconstitution en 3D du tissu diffère nécessairement de la structure tissulaire initiale de la tumeur. Dans le cas du ccRCC, certaines caractéristiques telles que les remaniements nécrotique et hémorragique, l'inflammation ou l'hypoxie, rendent difficile l'extraction et le maintien à la fois de cellules malignes et de cellules du microenvironnement tumorale (Majmundar, Wong and Simon, 2010; Hsieh *et al.*, 2017). Malgré ces limites, les avantages des organoïdes par rapport aux systèmes 2D devraient favoriser leur développement et leur utilisation dans l'étude du ccRCC ou des développements thérapeutiques (Wolf, Kimryn Rathmell and Beckermann, 2020).

	Culture primaire	Lignée cellulaire	Lignée cellulaires 3D	Organoïde 3D	PDX	Culture organotypique de coupes	Tumeur sur puce
<b>Facilité de maintenance</b>	+/-	+	-	-	-	+/-	-
<b>Préservation de la morphologie tumorale</b>	-	-	-	-	+/-	+	+
<b>Viabilité cellulaire à long-terme</b>	+/-	+	+	+	+	-	+/-
<b>Prolifération non-sélective</b>	+/-	-	-	-	-	+	+
<b>Préservation du microenvironnement tumoral</b>	-	-	-	-	+/-	+	+
<b>Criblage haut-débit de molécules</b>	+/-	+	+/-	+/-	-	-	+/-
<b>Taux de réussite de génération du modèle</b>	+/-	-	+/-	+/-	-	+	
<b>Temps de génération</b>	+	-	-	-	-	+	+
<b>Ressemblance avec la tumeur initiale</b>	-	-	+/-	+/-	+/-	+	+
<b>Coût</b>	+	+	+/-	+/-	-	+	-/+

Tableau 7. Comparaison des systèmes de culture en oncologie

Le tissu tumoral frais peut être cultivé *ex vivo* de plusieurs façons, chacune ayant ses avantages et inconvénients. + : Avantage ; - : désavantage. Tableau adapté de (Meijer, Jager and Gent, 2017).

### 1.3.2.2. Modèles de xénogreffe murin

La souris est un modèle très largement utilisé en oncologie. Souvent génétiquement modifiées, elles sont aussi utilisées comme organisme hôte en cancérologie. Les premiers modèles de CDX (« *Cell-line Derived Xenografts* ») et PDX (« *Patient Derived Xenografts* ») ont été générés dans les années 1970 et sont encore très utilisés aujourd'hui (Kameya, Ohsawa and Nomura, 1976). Ils sont établis par transplantation orthotopique de cellules (CDX) ou de fragments tumoraux (PDX), sous la capsule rénale ou par transplantation hétérotopique (région sous-cutanée ; intrapéritonéale ou intramusculaire) (Klinghammer, Walther and Hoffmann, 2017). Le recours à des souris immunodéficientes, diabétiques non-obèses ou gravement immunodéprimées permet d'éviter le rejet des cellules tumorales (Prochazka *et al.*, 1992; Chen *et al.*, 2019). Dans les transplantations réussies, plusieurs études ont démontré que les modèles de xénogreffes conservaient l'histologie des tumeurs des patients, l'expression des gènes, les altérations du nombre de copies d'ADN, les profils mutagènes et les profils de réponse aux médicaments (Grisanzio *et al.*, 2011; Pavía-Jiménez, Tcheuyap and Brugarolas, 2014), ce qui en fait un des modèles les plus proches des modèles *in vivo*. Les modèles PDX ont notamment été exploités pour le dépistage des médicaments, la découverte de biomarqueurs et l'identification des mécanismes de résistance (Hidalgo *et al.*, 2014).

Dans le ccRCC, les modèles CDX et PDX ont permis le développement et la compréhension des mécanismes de certaines molécules anti-angiogéniques pour le traitement des formes métastatiques (Lee and Motzer, 2017). Une étude a permis de détecter des signes de réponses au sunitinib et de prédire avec succès l'évolution du volume de la tumeur près d'une semaine à l'avance (Rojas *et al.*, 2019). Dans une autre étude, des modèles de CDX utilisant les lignées cellulaires 786-O, A498 et CAKI-1 ont mis en évidence des modifications des débits sanguins tumoraux corrélées à la résistance au traitement par sorafénib, une molécule anti-angiogénique (Schor-Bardach *et al.*, 2009). Ces résultats sont imputables à la capacité des modèles CDX/PDX à reproduire des caractéristiques importantes du microenvironnement tumoral, notamment du réseau vasculaire dense typique ccRCC (Rojas *et al.*, 2019).

Bien que les modèles PDX fournissent un modèle proche de l'homme, leur maintien par passage de la tumeur entre souris induit d'importantes pressions sélectives qui biaisent probablement la distribution des phénotypes tumoraux (Lai *et al.*, 2017). Les PDX ont par exemple tendance à favoriser la croissance des formes agressives de cancers ; biais par lequel a été découvert le gène suppresseur de tumeur BAP1 (Peña-Llopis *et al.*, 2012). Il a depuis été démontré que les tumeurs des patients présentant des mutations de BAP1 sont associées à une survie plus faible et présentent des caractéristiques tumorales agressives (Linehan, Srinivasan

and Schmidt, 2010; Peña-Llopis *et al.*, 2012; Hakimi, Ostrovnaya, *et al.*, 2013). Si ce biais de sélection des PDX est intéressant pour étudier les formes agressives, il réduit cependant la marge de manœuvre. Le temps de génération est un autre inconvénient du modèle, notamment en comparaison des modèles de lignées cellulaires (Tableau 7). Le délai entre l'implantation et la croissance progressive de la tumeur greffée peut varier de 2 à 12 mois (Dangles-Marie *et al.*, 2007; Bergamaschi *et al.*, 2009), rendant son utilisation caduque dans le domaine du diagnostic par exemple, car un patient métastatique a des risques importants de décéder dans un délai inférieur à ce temps de génération.

### 1.3.2.3. Culture organotypique *ex-vivo*

Les systèmes 3D évoqués précédemment ne reproduisent que partiellement la complexité de la tumeur (Meijer, Jager and Gent, 2017). Parmi les limitations, des études rapportent que les organoïdes ne reproduisent pas les interactions tumeur-stroma et induisent un biais de sélection (Lancaster *et al.*, 2013; Van De Wetering *et al.*, 2015), et que les PDX dérivent par rapport au tissu tumoral initial (Morgan *et al.*, 2017), conduisant à un taux d'échecs important dans les essais pré-cliniques sur de petites molécules inhibitrices (Maeda and Khatami, 2018; Ward *et al.*, 2021). Dans ce contexte, la culture organotypique de coupes de tumeur ou « PDTSC » (*Patient Derived Tumor Slice Culture*) représente un système attractif et complémentaire déjà validé dans plusieurs sites tumoraux (Figure 25) (Guyot *et al.*, 2007; Vaira *et al.*, 2010; Gerlach *et al.*, 2014; Roelants *et al.*, 2018a; Misra *et al.*, 2019).

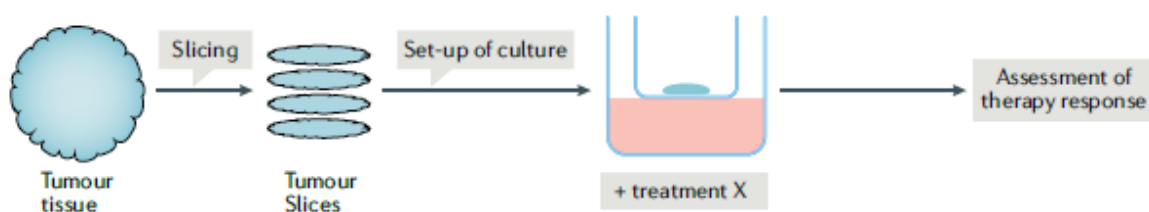


Figure 25. **Culture organotypique de coupes.**

Après résection, l'échantillon de tumeur frais est découpé en une série de tranches d'épaisseur égale qui sont cultivées sur inserts dans un milieu de culture adapté contenant le traitement à tester. D'après (van de Merbel *et al.*, 2021).

Elle consiste en la culture *ex vivo* de tranches de la tumeur, considérées comme organotypiques, soit maintenant la cohésion entre les cellules et les tissus. Elle est principalement employée pour évaluer la sensibilité aux médicaments (Centenera *et al.*, 2013; Davies *et al.*, 2015; Van De Merbel *et al.*, 2018) et pour déterminer leur accumulation, absorption (Fleck *et al.*, 2000) et toxicité (Vickers *et al.*, 2004). Des études notables ont permis



de prédire les réponses à des thérapies ciblées ou chimiothérapies ayant recours au PDTSCs (Vaira *et al.*, 2010; Gerlach *et al.*, 2014; Carranza-Torres *et al.*, 2015; Koerfer *et al.*, 2016; Jiang *et al.*, 2017). Globalement, cette méthode présente les avantages d'être rapide et peu coûteuse à mettre en place, de requérir de petits échantillons de tissus frais et de préserver l'architecture tumorale et l'interaction spatiale entre la tumeur et le stroma (Roelants *et al.*, 2020).

La technique a quelques limites importantes cependant. La durée de vie des tranches de tumeur est limitée et non renouvelable, ce qui empêche les expériences longues ou répétées. Par ailleurs, dédiées à certaines questions, elles ne permettent pas par exemple de réaliser des criblages haut-débit de molécules (Kenerson *et al.*, 2020). L'hypoxie est un autre problème éventuel, selon le design expérimental, puisque le système vasculaire est incomplet dans ces systèmes et que la quantité d'oxygène disponible pour le tissu dépend des échanges avec le milieu. Enfin, ces systèmes ont une autre limitation ; leur défaut de certaines caractéristiques systémiques telles que le système immunitaire ou vasculaire, ce qui entraîne des défaut d'oxygénation, seulement compensé par les échanges avec le milieu, ou bien l'incapacité à rendre compte d'une réponse immunitaire pertinente (Meijer, Jager and Gent, 2017). De nombreuses pistes d'amélioration sont en cours de développement dans ces modèles, telles que la mise en place d'écosystèmes personnalisés conservant le microenvironnement à travers le recours à des matrices et sérums spécifiques du grade de la tumeur. Les cellules immunitaires issues du sérum du patient peuvent par exemple infiltrer le tissu, étendant les champs d'investigation.

Dans le ccRCC, une équipe a utilisé un modèle de culture organotypique de coupes tumorales (Roelants *et al.*, 2018a, 2020). Les auteurs ont montré qu'une combinaison de deux inhibiteurs ciblant à la fois les kinases PI3K et Src impacte significativement la viabilité cellulaire (Roelants *et al.*, 2018a). Ils ont également démontré que le PDTSC conserve les caractéristiques histopathologiques d'origine ainsi que les composants du microenvironnement stromal et immunitaire. Dans une autre étude, la même équipe a démontré qu'une seconde combinaison d'inhibiteurs ciblant les protéines CK2 et ATM induisait une importante inhibition de la prolifération, une réduction de la migration et une apoptose passant par la production d'espèces réactives de l'oxygène (« ROS » ; *Reactive Oxygen Species*) (Giacosa *et al.*, 2021).

#### **1.3.2.4. Les tumeurs sur puce : avenir des modèles de culture tumoraux ?**

Les tumeurs sur puce, ou « *tumor-on-chip* » (TOC) sont un autre type de système de culture *ex-vivo* en trois dimensions faisant l'objet de recherches actives dans de nombreux types de cancers (Del Piccolo *et al.*, 2021). Également appelés « laboratoires sur puce » (« *lab-on-a-chip* »), il s'agit de dispositifs microphysiologiques divers qui intègrent la micro-ingénierie (impression 3D et conception de supports variés), la microfluidique (canaux micrométriques permettant des échanges de fluides) et la manipulation fine de tissus (tissu micro-disséqué) (Dorrigiv *et al.*, 2021). En plus des particularités des précédents systèmes de culture, les TOCs ont permis de recréer les dynamiques d'échanges, notamment en ce qui concerne l'apport en nutriments et en oxygène, entre la tumeur et son milieu (Bērziņa *et al.*, 2021). Des expériences ont déjà démontré leur capacité à maintenir des micro-tissus fonctionnels pendant des périodes relativement longues (Esch, Bahinski and Huh, 2015; Zheng *et al.*, 2016; Trujillo-de Santiago *et al.*, 2018). Permettant de recréer un microenvironnement tumoral représentatif, ils sont également prometteurs en tant que nouveau modèle pour le criblage rentable haut-débit de nouvelles molécules anticancéreuses (Tsai *et al.*, 2017; Shang *et al.*, 2019; Sontheimer-Phelps, Hassell and Ingber, 2019; W. Sun *et al.*, 2019) et support de la médecine de précision, de façon générale (Van Den Berg *et al.*, 2019). Comme pour chaque modèle existant, ces avantages ont des contreparties. Les TOCs sont encore coûteux et, les solutions commerciales étant encore peu nombreuses, leur mise en œuvre compliquée est à la charge du laboratoire. Comme toutes les technologies cependant, ces limitations seront amenées à être dépassées, ce qui pourrait bien donner une nouvelle impulsion au domaine des modèles de culture tumoraux, qui intégreront systématiquement, à long terme, un haut degré de complexité.

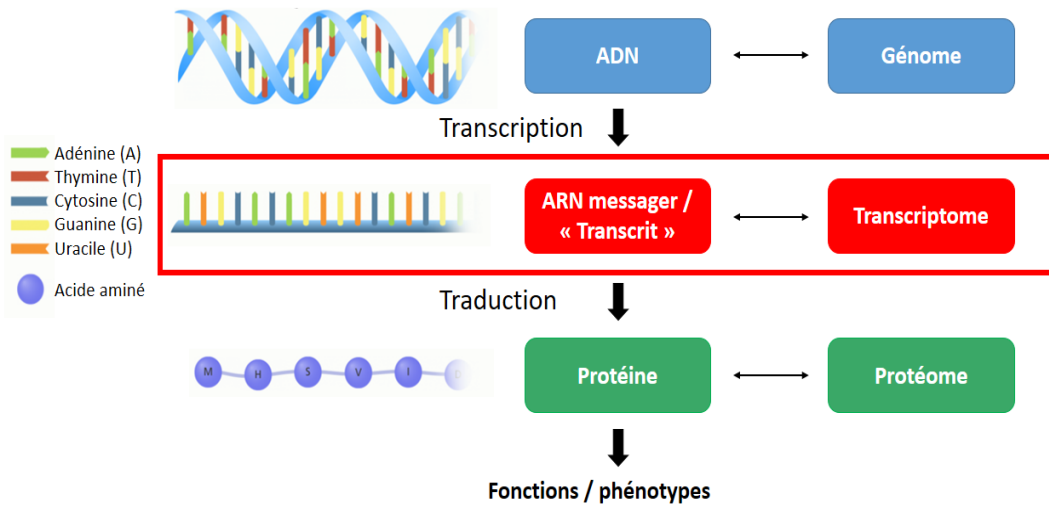


Figure 26. Schéma simplifié du flux de l'information génétique

Dans le cadre du dogme central tel qu'initialement proposé, le flux génétique de l'information passe de l'ADN à l'ARN, puis est traduite en protéines fonctionnelles.

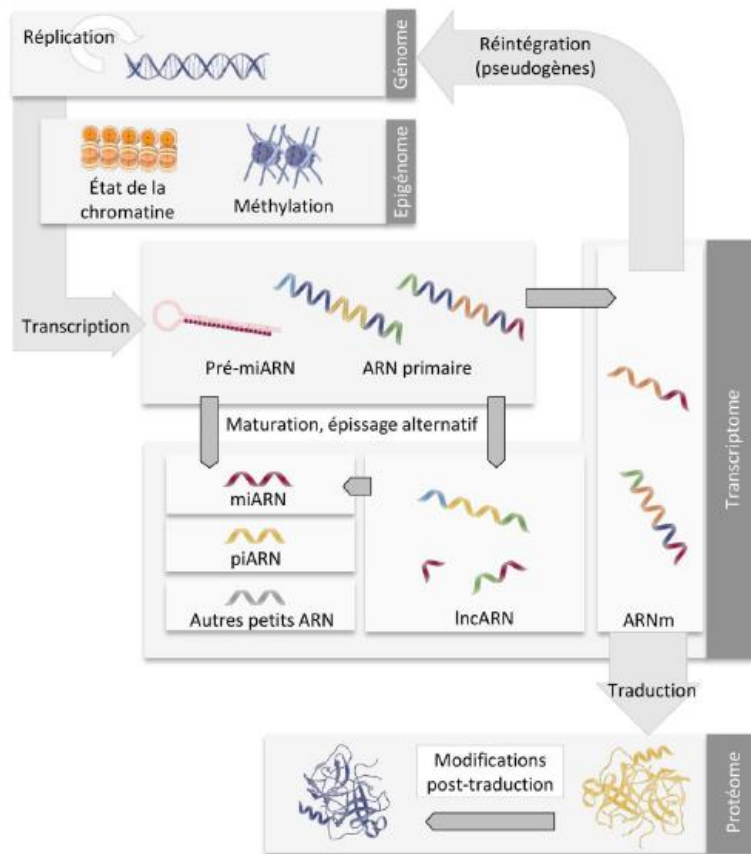


Figure 27. Organisation et transfert de l'information entre le génome, l'épigénome, le transcriptome et le protéome

L'organisation et les interactions entre les différents domaines -omiques dessinent une image complexe et en évolution du flux de l'information génétique, faisant intervenir de nombreux acteurs moléculaires tels que les longs ARNs non-codant (lncARN), les micro-ARNs (miARN) ou les « PIWI-interacting » ARNs (piARN) qui occupent notamment des fonctions de régulation de l'expression génique et de défenses immunitaires.

## 1.4. DE L'ÉVOLUTION DE LA TRANSCRIPTOMIQUE

### 1.4.1. LE DOGME CENTRAL DE LA BIOLOGIE MOLECULAIRE ET LES DISCIPLINES *-OMIQUES*

Le dogme central de la biologie moléculaire est un cadre de compréhension du flux de l'information génétique élaboré au milieu du 20<sup>e</sup> siècle. Dans le schéma le plus fréquent au sein de nos cellules, l'ADN contient l'information génétique nécessaire à la fabrication des protéines, et l'ARN est un messenger faisant transiter l'information depuis l'ADN vers les ribosomes au niveau desquels il est traduit en protéine. Ce processus s'appelle « l'expression génique » et comporte deux étapes clés : la transcription (passage de l'ADN à l'ARN) et la traduction (passage de l'ARN à la protéine) (Figure 26). Les ARNs sont également appelés « transcrits » pour cette raison. L'ensemble des ARNs forment le transcriptome, dont l'étude se nomme transcriptomique. De la même façon, les disciplines « -omiques » comme la génomique et la protéomique étudient le génome (ensemble des gènes du l'ADN) et le protéome (ensemble des protéines fonctionnelles).

Cependant, d'autres disciplines ont émergé avec les avancées scientifiques et technologiques et avec elles, certains aspects du dogme central sont apparus inexacts ou incomplets. En particulier, celui-ci ne prévoyait pas l'épigénétique, qui décrit les mécanismes réversibles, transmissibles et adaptatifs influençant l'expression génique sans affecter la séquence du gène. Ces mécanismes concernent notamment la méthylation de l'ADN et les modifications post-traductionnelles des histones, qui influent sur le degré de compaction de l'ADN et donc sur l'accessibilité des transcriptases et facteurs de transcription aux gènes ; ou encore les micro-ARNs dont les rôles dans la régulation de l'expression génique font l'objet de nombreuses études actuellement. Les longs ARN non-codants (« *long non-coding RNAs* » ou lncRNAs ) sont d'autres acteurs moléculaires ayant des fonctions très variées à l'interface dans la régulation de l'expression génique, de l'épigénétique (Mercer, Dinger and Mattick, 2009). En plus de ces grands mécanismes régulateurs, d'autres items particuliers dessinent aujourd'hui un schéma plus complexe et encore en évolution du flux de l'information génétique (Figure 27), telles que par exemple les modifications post-transcriptionnelles des ARNs (maturation, épissage alternatif) ou les modifications post-traductionnelles des protéines (addition de groupe fonctionnels ou de structure protéique), qui impactent leurs fonctions.

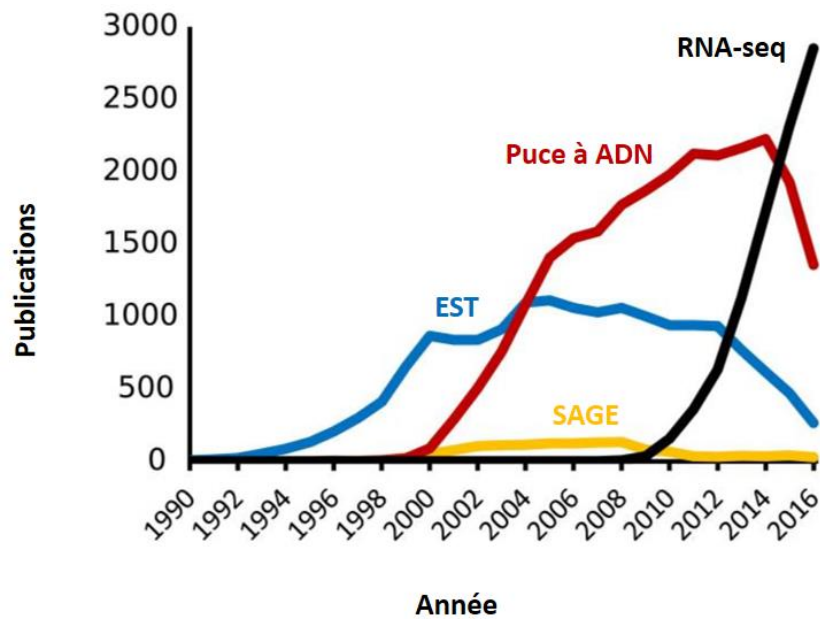


Figure 28. Évolution de l'utilisation des principales techniques transcriptomiques

Ce graphique décrit l'évolution du nombre de publications scientifiques citant les principales techniques de transcriptomiques des années 1990 à 2016. Figure tirée de Lowe *et al.*, 2017.

Comme vu ci-avant, la transcriptomique désigne l'étude du transcriptome, c'est-à-dire la somme de tous les ARNs, codants et non-codants, présents dans une cellule à un temps « t ». La caractérisation et la quantification du transcriptome dans un tissu et des conditions données sont indicateurs des principales tendances d'expression des gènes actifs et permettent déterminer les processus biologiques et voies métaboliques à l'œuvre. La transcriptomique regroupe un ensemble de techniques permettant notamment de déterminer la quantité relative des ARNs.

#### **1.4.2. LA TRANSCRIPTOMIQUE « BULK » ET SES TECHNOLOGIES**

Les premières méthodes séquençage de l'ADN furent inventées indépendamment l'une de l'autre dans la deuxième moitié des années 1970. Elles sont conçues pour l'une par l'équipe de Walter Gilbert, aux États-Unis (Maxam and Gilbert, 1977), et pour l'autre par celle de Frederick Sanger (Sanger, Nicklen and Coulson, 1977), au Royaume-Uni. Ces deux méthodes sont fondées sur des principes diamétralement opposés : l'approche de Sanger fonctionne par synthèse enzymatique, tandis que celle de Gilbert repose sur la dégradation chimique sélective. Pour cette découverte, Gilbert et Sanger seront co-réceptaires du prix Nobel de chimie en 1980.

##### **1.4.2.1. Les débuts de la transcriptomique : EST & SAGE**

Le séquençage EST (Expressed Sequenced Tag), développée dans les années 1990, est dérivé du séquençage Sanger et constitue la première technique ayant permis d'identifier les transcrits de gènes. Elle consiste à séquencer l'ADN complémentaire, ou ADNc, cloné dans des plasmides bactériens (Putney, Herlihy and Schimmel, 1983; Adams *et al.*, 1991). Cette méthode fut déterminante dans la découverte des gènes et de leur séquence. Plus tard, la méthode « *Serial Analysis of Gene Expression* » (SAGE), a permis non seulement de caractériser l'ADN à plus grande échelle mais aussi de quantifier les gènes et transcrits d'intérêt (Velculescu *et al.*, 1995). C'est à cette époque que la notion d'assemblage de transcrits est apparue, où les « tags » séquencés étaient réunis en contigs. Bien que ces technologies préfigurent la transcriptomique haut-débit, elles étaient particulièrement coûteuses en temps, en expériences et en main d'œuvre. Elles n'ont réellement été employées pour caractériser le transcriptome qu'au tournant des années 2000 (Figure 28).

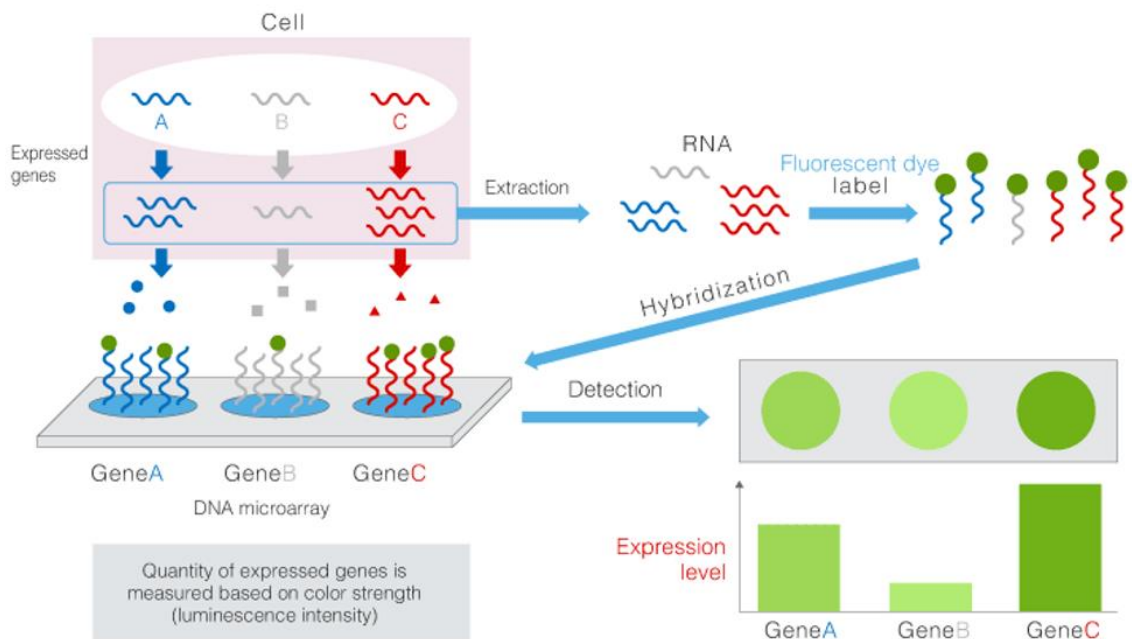


Figure 29. Fonctionnement des puces à ADN

Schéma représentant le principe général des puces à ADN. Les molécules d'ARN sont extraites à partir des cellules de l'échantillon, rétro-transcrites en ADN puis couplées à des billes fluorescentes excitables à différentes longueurs d'onde. Les molécules ADN complémentaires sont hybridées sur la puce à ADN sur les spots contenant les sondes nucléotidiques complémentaires et définissant chaque gène. En utilisant un appareil, les billes fluorescentes sont ensuite excitées et réémettent une fluorescence dont l'intensité est proportionnelle à la quantité d'ADNc hybridé.

#### 1.4.2.2. Les puces à ADN

Les puces à ADN sont apparues au début des années 2000. Elles permettent d'estimer simultanément la concentration de milliers de transcrits au sein d'un échantillon donné. En cela, la technologie de puce à ADN, ou « *microarrays* », est la première à avoir permis une étude haut-débit et quasi-exhaustive du transcriptome. Son principe repose sur l'hybridation spontanée entre l'ADN simple brin (dénaturé) d'intérêt et une sonde nucléotidique, en vertu du principe de complémentarité des bases. (Figure 29).

Bien qu'il existe plusieurs technologies, le principe général reste le même : des fragments d'ADN correspondant à des séquences complémentaires de transcrits, appelés sondes nucléotidiques, sont fixés sur une surface solide, composée de verre, de silicium ou de plastique (Figure 29). Un spot contient des milliers de sondes identiques spécifiques et complémentaires d'un transcrit (transcrits A, B ou C dans l'illustration). Le tissu d'intérêt est broyé et les ARNs sont extraits selon un protocole dédié puis amplifiés par PCR (« *polymerase chain reaction* ») afin d'obtenir une quantité de matériel génétique suffisante. L'ARN est rétro-transcrit en ADNc et marqué, le plus souvent avec un fluorochrome. Les cibles ainsi marquées sont mises en contact avec les sondes de la puce et s'y hybrident lorsque leurs séquences sont complémentaires. La puce est ensuite lavée pour éliminer les ADNc non hybridés. Le signal fluorescent est détecté dans un scanner à haute résolution en fonction de la longueur d'onde d'excitation du fluorochrome. L'analyse informatique permet enfin d'associer une valeur d'intensité à chaque sonde fixée sur la puce (Nelson, 2001; Pozhitkov, Tautz and Noble, 2007). Ainsi, la conception des sondes est cruciale pour la spécificité des hybridations, ce qui limite l'ensemble des gènes étudiés à ceux ayant déjà été caractérisés. Cette technologie ne permet donc pas d'identifier de nouveaux transcrits ou de nouveaux gènes. De plus, une hybridation non spécifique des sondes peut entraîner des erreurs dans la quantification (Pozhitkov, Tautz and Noble, 2007). D'autres applications de cette technologie permettent d'étudier la composition exonique des ARN (« *exon arrays* »), les polymorphismes nucléotidiques (« *single nucleotide polymorphisme* », ou SNP) (LaFramboise, 2009), les variations du nombre de copies d'un gène dans l'ADN (« *comparative genomic hybridization* », ou CGH array) (Shinawi and Cheung, 2008), ou les interactions protéines-ADN (ChIP-on-chip) (Bulyk, 2006). Le « *Tiling array* » est une technologie de puce permettant l'étude de nouveaux transcrits ou gènes. Elle diffère des puces à ADN classiques dans la conception des sondes. Celles-ci sont conçues pour couvrir l'ensemble d'un génome donné (Mockler and Ecker, 2005; Samanta, Tongprasit and Stolc, 2007).



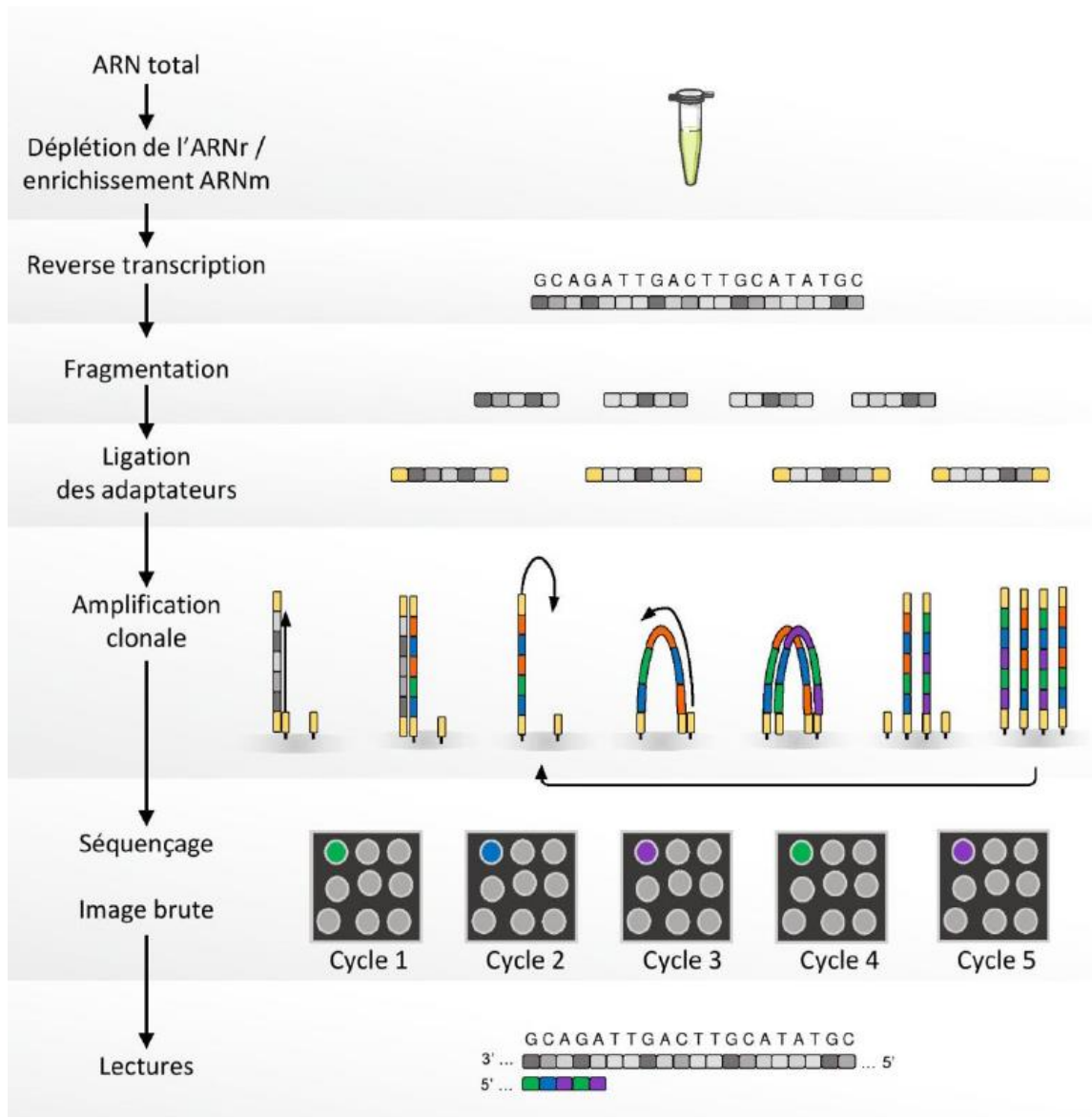


Figure 30. Schéma du protocole de séquençage ARN sur une plateforme Illumina

Dans le protocole illumina, l'un des plus courant, les ARN totaux sont extraits et raffinés pour enrichir et rétro-transcrire les ARN messagers. A la suite d'étapes de fragmentation, d'ajout d'adaptateurs de séquençage et d'amplification, les molécules d'ADNc sont séquençées par l'ajout et la détection fluorescente des bases nucléotidiques dans une molécule complémentaire néosynthétisée.

Cependant, une telle technologie est coûteuse et peu adaptée aux génomes eucaryotes complexes. Bien qu'elle reste une manière rapide d'analyser l'expression de gènes au sein d'un échantillon à un coût réduit, la technologie des puces à ADN a été progressivement supplantée par les technologies de séquençage haut-débit (Figure 28).

### **1.4.2.3. Le séquençage transcriptomique haut-débit**

#### **1.4.2.3.1. Le RNA-sequencing**

Depuis 2009, plusieurs générations de séquenceurs se sont succédées, chacune apportant des améliorations notables. Toutes sont développées autour de la méthode du séquençage Sanger (Sanger, Nicklen and Coulson, 1977). Cette section est illustrée avec le séquençage de type Illumina HiSeq (Figure 30) (Bentley *et al.*, 2008), très largement utilisé ces dernières années pour le séquençage ARN. Dans le protocole Illumina, les ARNs totaux sont extraits à partir d'un échantillon puis la fraction d'intérêt est isolée selon un protocole dédié. L'échantillon est par exemple souvent enrichi en ARN polyadénylés (protocole polyA) grâce à des billes poly-T, ce qui permet de capturer les ARNs messagers (ARNm). Il est aussi fréquent d'éliminer les ARNs ribosomique (ARNr), qui peuvent représenter près de 90 % des ARNs dans une cellule. Ce protocole permet aussi de conserver également les ARN non polyadénylés, utile pour l'étude des ARNs longs non-codants (ARNlnc). Les ARNs peuvent également être sélectionnés en fonction de leur taille pour étudier les petits ARN non-codants (ARNnc), ou les grands ARNs (>200 nucléotides), ARNm et ARNlnc.

Le protocole de RNA-seq passe par la fragmentation de l'ARN en morceaux de taille aléatoire, puis par leur rétrotranscription en ADNc. À cette étape, il est possible de conserver l'information sur le sens de transcription de l'ARN (soit le sens 3'-5', soit le sens 5'-3') en ajoutant d'adaptateurs permettant de reconnaître l'ADN à séquencer. Ces adaptateurs peuvent également servir de code-barres pour identifier les transcrits d'un échantillon donné et ainsi effectuer le séquençage de plusieurs échantillons en mélange. Enfin, l'ADNc est amplifié par PCR (« *Polymerase Chain Reaction* »). En raison de la taille variable des fragments d'ADNc et de la composition des séquences, cette dernière étape peut introduire un biais en augmentant artificiellement la proportion de certains fragments au sein de l'échantillon.

Les adaptateurs permettent la fixation aléatoire de chaque fragment sur le support de séquençage, appelé « *flowcell* ». Le nombre de copies de chaque fragment est amplifié pour créer des regroupements d'une même molécule, appelés « *clusters* », et optimiser la mesure du

signal de séquençage. La lecture de chaque fragment se fait base par base, de manière simultanée, par l'ajout de nouveaux nucléotides couplés à des fluorochromes. Ceux-ci s'incorporent au brin en cours de synthèse et émettent un signal fluorescent indiquant la nature du nucléotide. Les lectures de séquençage (appelées « *reads* ») ainsi générées ont une longueur fixe comprise entre 50 et 200 bases selon le protocole et de la technologie utilisée. Les *flowcells* permettent de générer trois milliards de lectures de séquençage dans le protocole HiSeq d'Illumina, mais une nouvelle génération de supports, dans le protocole Nova-seq, permet de générer 13 milliards de lectures en une expérience. Plus la taille des fragments séquencés est importante, plus les erreurs de séquençage sont fréquentes. Il est possible de séquencer un fragment d'ADNc à partir de l'une ou des deux extrémités. Cela permet d'obtenir deux séquences pairées pour chaque ADNc/ARN et ainsi améliore la qualité de l'alignement et de l'assemblage bio-informatique. Cela facilite aussi l'identification des différentes isoformes d'un même gène.

#### **1.4.2.3.2. Le séquençage « long-read »**

Avec le RNA-seq traditionnel, l'amplification des transcrits par PCR peut introduire un biais dans la quantification après le séquençage (Parekh *et al.*, 2016). De plus, la fragmentation nécessaire des transcrits pour la préparation des bibliothèques implique de les réassembler lors de l'analyse pour identifier les isoformes, ce qui peut aboutir à des erreurs et des transcrits chimériques (Phillippy, 2017). Le séquençage *long-read* a pour but de contourner ces limitations et de préciser la caractérisation et la quantification des isoformes. Ces méthodes ne nécessitent pas d'étapes d'amplification du matériel génétique à séquencer et permettent le séquençage de fragments significativement plus longs, voire de transcrits entiers contenant plusieurs kilobases (Laver *et al.*, 2015; Lu, Giordano and Ning, 2016; Garalde *et al.*, 2018). Deux technologies dominantes permettent ce type de séquençage : PacBio (Rhoads and Au, 2015) et Nanopore (Lu, Giordano and Ning, 2016). La première a recours à une ADN polymérase fixée au fond d'un puits, qui synthétise l'ADNc d'un brin d'ARN en incorporant successivement l'un des quatre nucléotides marqués par un fluorochrome spécifique. La seconde technologie détecte les propriétés biochimiques spécifiques de chaque nucléotide. L'ADNc est guidé à travers un pore qui détecte les changements électriques induits par chaque base. Le taux d'erreurs de séquençage reste important (~10 %), mais les optimisations actuelles tendent vers une meilleure spécificité. À l'avenir, les différents transcrits pourront être plus précisément caractérisés et quantifiés. Par ailleurs, ces modes de capture semblent également

permettre d'évaluer certaines modifications épigénétiques, dans le même temps que son identification. (Schatz, 2017). Ces techniques intègrent ainsi une composante multi-omique.

### 1.4.2.3.3. BRB-seq

Le BRB-seq (« *Bulk RNA Barcoding and sequencing* ») est une technologie de séquençage ARN dérivé du DGE-seq (« *Digital Gene Expression sequencing* ») (Rodríguez-Esteban *et al.*, 2015). Ces techniques reposent sur le marquage des ARNs par un code-barres (14 nucléotides) identifiant leur échantillon d'origine, ce qui permet leur mélange dans la librairie de séquençage (autrement appelé « multiplexage »), puis l'identification de l'échantillon de chaque ARN dans l'analyse. Elles incluent également l'utilisation d'identifiants moléculaires uniques (« *Unique molecular identifiers* » ou UMI ; 14 nucléotides aléatoires), propre à chaque molécule d'ARN (Figure 31) (Alpern *et al.* 2019). Ces principes sont apparus pour la première fois dans le domaine de la transcriptomique à l'échelle de la cellule unique (voir la section 1.4.3.1.2. **Le Chromium de 10X Genomics**) comme un moyen de réduire le coût et le temps de manipulation en générant une seule librairie de séquençage contenant plusieurs échantillons distincts (Ziegenhain *et al.*, 2017). Les UMIs présentent l'intérêt de permettre un comptage précis des transcrits séquencés et de s'affranchir des biais d'amplification par PCR. Un UMI équivaut à un transcrit ; deux UMIs identiques ou plus reflètent un biais d'amplification et correspondant à des duplicas de PCR. Ce principe avantage la technique par rapport aux méthodes classiques d'amplification d'ADNc.

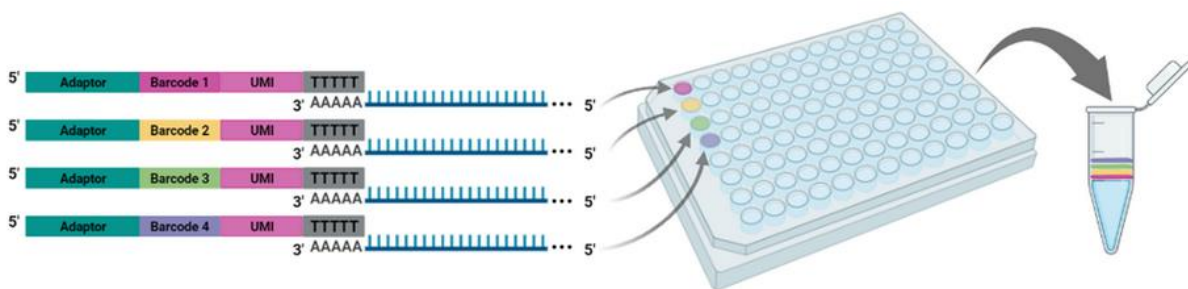


Figure 31. Illustration du multiplexage réalisé en BRB-seq

Chaque puits dans une plaque 96-puits contient un code-barres unique permettant de marquer des ARNs d'un échantillon. Après la transcription inverse et les marquages, les échantillons marqués sont regroupés pour le reste du protocole. Image issue de <https://alitheagenomics.com/>.

Le BRB-seq a le potentiel de révolutionner le domaine de la transcriptomique car il permet d'analyser de nombreux échantillons simultanément, de façon sensible et peu coûteuse (environ 30€ par échantillon). Des centaines d'échantillons peuvent être préparés en plaques 96-puits par un seul expérimentateur, en une ou deux journées. Cette approche a été utilisée

plusieurs fois récemment pour étudier les adipocytes bruns (Bast-Habersbrunner *et al.*, 2021), évaluer l'expression des haplotypes mitochondriaux chez la drosophile (Bever *et al.*, 2019), évaluer l'efficacité des thérapies ciblées contre le cancer dans les cellules de carcinome rénal à cellules claires en interrogeant plus de 2880 paires médicament-gène (Giacosa *et al.*, 2021), et fournir de nouvelles informations sur le rythme circadien (Litovchenko *et al.*, 2021) ainsi que sur l'évolution et le développement (Liu *et al.*, 2020).

### **1.4.3. LA TRANSCRIPTOMIQUE A L'ECHELLE DE LA CELLULE UNIQUE OU SINGLE-CELL RNA-SEQ**

Les analyses génomiques dites « *bulk* », ou globales, ont contribué à faire progresser rapidement des pans entiers de la biologie. L'étude transcriptomique d'échantillons cliniques a contribué en l'amélioration des connaissances sur les voies de signalisations des cancers. Cependant, les tumeurs humaines sont des écosystèmes complexes composés de diverses cellules, y compris dans le micro-environnement tumoral, dont la caractérisation transcriptomique précise n'est pas possible en séquençage global (à l'échelle du tissu entier). Dans ce contexte, le séquençage des ARNs à l'échelle de la cellule unique (« single-cell RNA-seq » ou scRNA-seq) a connu progression fulgurante ces dix dernières années (Chen, Ning and Shi, 2019). Cette innovation technique permet d'abolir la limite du séquençage global pour lequel un signal moyen est détecté pour le tissu entier ou une population cellulaire. En obtenant les profils transcriptomiques de chaque cellule après dissociation d'un tissu, le scRNA-seq permet en particulier d'identifier les principaux types cellulaires. Elle nécessite généralement l'obtention rapide de tissus frais en bon état pour faciliter la dissociation et obtenir des cellules viables. Cette technique est très utilisée dans les problématiques d'étude de l'hétérogénéité de tissus complexes, notamment de tumeurs cancéreuses.

#### **1.4.3.1. Les technologies**

##### **1.4.3.1.1. Les premières méthodes**

Le premier protocole scRNA-seq a été publié en 2009 (Tang *et al.*, 2009), un an seulement après la première publication sur le séquençage ARN massif (Mortazavi *et al.*, 2008). Bien que cette approche ait permis de mesurer le transcriptome de cellules individuelles, elle nécessitait une manipulation précise et était limitée à l'étude de quelques cellules. Depuis, de nombreux protocoles scRNA-seq ont été développés (notamment le CEL-Seq (Hashimshony *et al.*, 2012), le CEL-Seq2 (Hashimshony *et al.*, 2016), le Quartz-Seq (Sasagawa *et al.*, 2013), le Quartz-Seq2 (Sasagawa *et al.*, 2018) et le Smart-seq2 (Picelli *et al.*, 2013)) et le nombre de

cellules dans les expériences scRNA-seq a augmenté de façon exponentielle (Figure 32) (Svensson, Vento-Tormo and Teichmann, 2018).

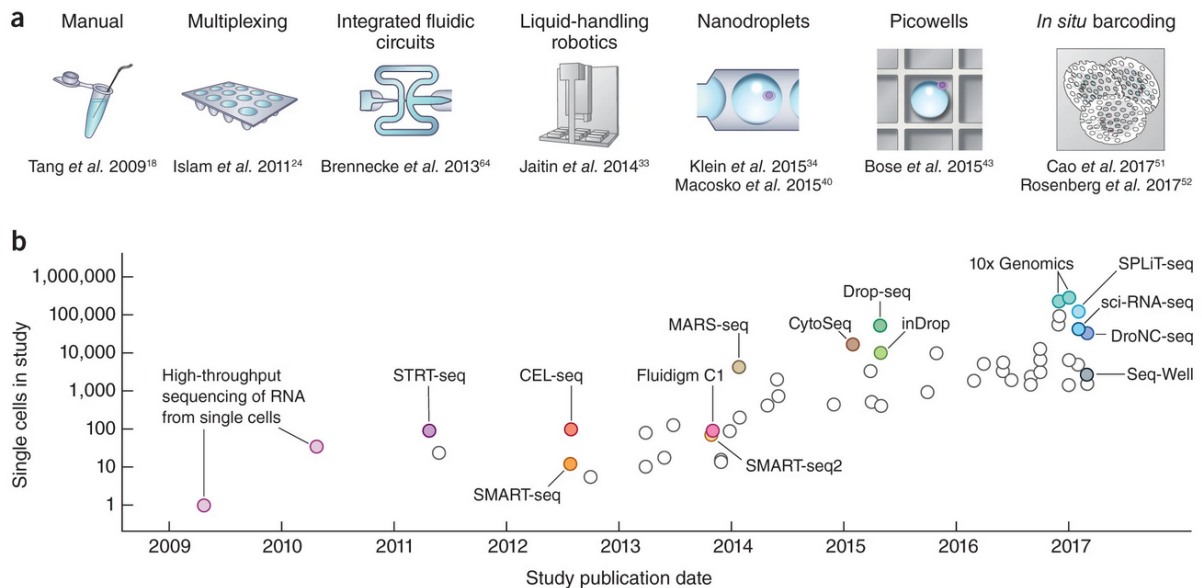


Figure 32. Évolution des technologies et débit de single-cell RNA-seq

(a) technique des capture de cellules. Le multiplexage des échantillons a permis de passer à ~100 cellules, puis à ~1 000 cellules grâce à des études utilisant la microfluidique. Les technologies de capture aléatoires utilisant des *nanodroplets* et des technologies *picowell* ont permis d'atteindre les dizaines de milliers de cellules. Enfin, des études ont utilisé le barcodage in situ pour atteindre à peu de frais l'ordre de grandeur suivant, soit des centaines de milliers de cellules. (b) Nombre de cellules rapporté par technologie et par année. Figure tirée de Svensson, Vento-Tormo and Teichmann, 2018.

Le premier système de capture commercialisé était le Fluidigm C1. Ce système utilise la technologie microfluidique pour séparer les cellules dans des puits individuels d'une plaque où elles sont ensuite lysées ; l'ARN rétro-transcrit en l'ADNc est collecté puis amplifié par PCR. Après cette étape, Les ADNc sont extraits de la plaque et les bibliothèques sont préparées en vue d'un séquençage Illumina. La plupart des données C1 ont été produites à partir de plaques de 96 puits mais des plaques 800 puits ont également été commercialisées, augmentant le nombre de cellules capturées simultanément. Les cellules étant capturées dans des puits individuels, elles peuvent être imagées avant la lyse, ce qui permet d'identifier les cellules endommagées, les puits vides ou les puits contenant plus d'une cellule. L'un des inconvénients cependant est que la technologie limite la taille maximale des cellules capturées. La capture de plusieurs cellules au lieu d'une seule en est un autre problème, assez fréquent. Macosko et ses collaborateurs ont constaté que lors de la préparation d'un mélange de cellules de souris et de cellules humaines, 30 % des bibliothèques obtenues contenaient des transcrits des deux espèces, mais que seul un tiers environ de ces doublets était visible sur les images de microscopie (Macosko *et al.*, 2015). Un autre système du même constructeur ; le Polaris intègre la possibilité

de sélectionner les cellules en fonction de leur coloration ou d'un marquage fluorescent, puis les conserver jusqu'à 24 heures tout en les soumettant à diverses conditions ou stimuli. Les cellules peuvent être imagées pendant cette période avant d'être lysées et préparées pour le séquençage de l'ARN. Cette plateforme offre des possibilités pour toute une série d'expériences qui ne sont pas possibles avec d'autres technologies de capture.

#### **1.4.3.1.2. Le Chromium de 10X Genomics**

Le protocole Chromium de l'entreprise 10X Genomics est actuellement la solution de single-cell RNA-seq la plus utilisée. Ce système sorti en 2017 permet de capturer avec efficacité des milliers de cellules par expérience et a permis de générer les premiers jeux de données conséquents (~100 000 cellules) (Figure 32). Cette section décrit le protocole de préparation d'échantillons avec le Chromium, et illustre certains principes communs à d'autres technologies tels que la microfluidique et la capture en microgouttelettes, ou le marquage des ARNs par des UMIs.

##### **1.4.3.1.2.1. Capture cellulaire et préparation de librairie de séquençage**

Le Chromium est un dispositif de capture microfluidique qui repose sur les technologies « *Drop-seq* », développées à partir de 2015 (Macosko *et al.*, 2015). Dans ce procédé, une suspension de cellules dissociées et d'enzymes de lyse d'une part, et des billes enrobées de sondes de capture nucléotidique d'autre part, rentrent dans un système microfluidique (Figure 33A). Ces billes (ou « *Gel-Beads* ») sont au cœur du système ; elles ont à leur surface des centaines de millions de sondes poly-T permettant de capturer les ARNm, et contenant deux séquences spéciales : un code-barres 10X, et une séquence UMI (Figure 33B). Ces deux séquences seront par la suite intégrées à l'ADNc par rétro-transcription (RT). La première séquence, le code-barres cellulaire 10X, mesure 16 nucléotides et est identique pour toutes les sondes d'une même bille. Il permettra d'identifier la cellule capturée dans la goutte, et donc la cellule d'origine des ARNs. La seconde séquence, l'UMI, est un identifiant qui marque de façon unique une molécule d'ARN, ce qui permettra, par un comptage absolu des UMIs, de quantifier les ARNs dans l'analyse en s'affranchissant des biais d'amplification induits par PCR pendant la préparation de la librairie de séquençage.

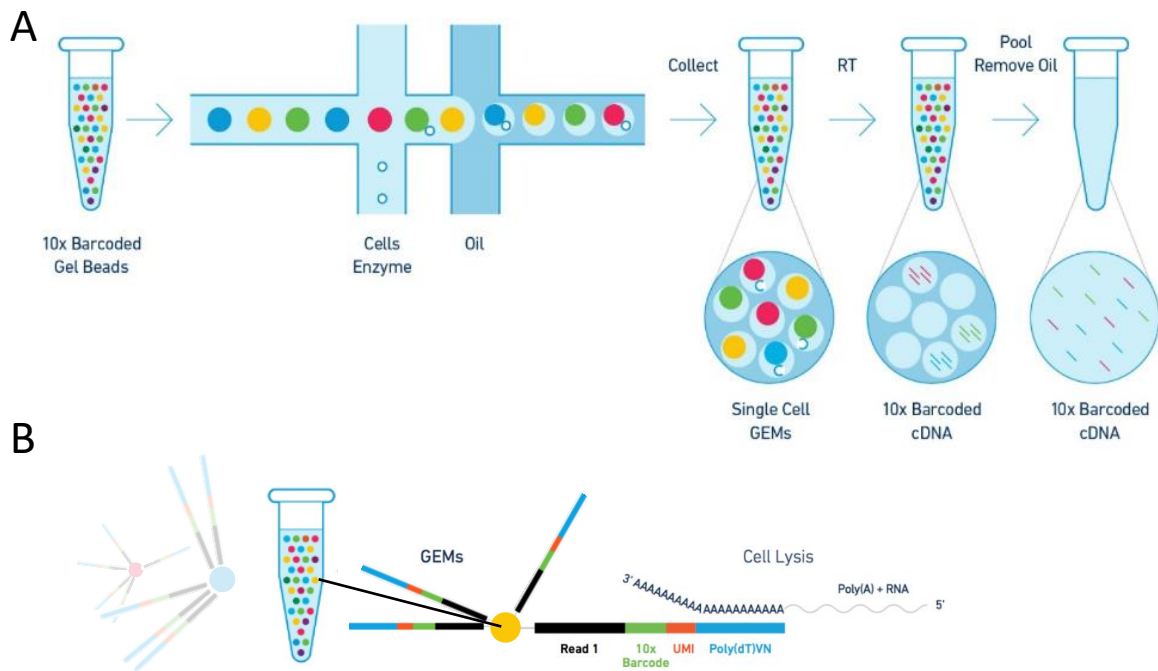


Figure 33. Processus de capture du Chromium 10X.

(A) Étapes principales. Des billes à codes-barres sont injectées dans un système microfluidique avec des cellules dissociées où elles sont capturées dans des gouttelettes aqueuses. Les cellules sont lysées à l'intérieur des gouttelettes et l'ARNm subit une rétro-transcription en ADNc à code-barres. Les gouttelettes sont ensuite cassées et l'ADNc est collecté pour le séquençage. (B) Illustration de la structure des sondes de capture au sein des microgouttelettes en suspensions (GEMs). Les sondes comprennent un adaptateur de séquençage, le code-barres cellulaire (« 10X Barcode »), l'UMI et la queue poly(T). Ils capturent les ARNs par leur queue poly-A, soit en 3'. Chaque bille possède des centaines de millions de sondes avec le même code-barres. Figure adaptée à partir d'images de 10x Genomics.

Le dispositif est conçu de sorte à former des gouttelettes aqueuses dans une suspension d'huile, appelées GEMs (« *Gel-bead-in-EMulsion* »), et de manière à ce qu'une cellule et une bille puissent être capturées simultanément dans l'une de ces gouttelettes. En leur sein, les réactifs transportés avec les cellules lysent la cellule, ce qui libère tout leur contenu en ARN. Les molécules d'ARN avec une queue poly(A), tels que les ARNm ou les ARNs longs non-codant, sont capturés par les sondes de capture poly-T sur la bille. Les ARNs capturés subissent une RT, puis les gouttelettes sont détruites (Figure 33A). A cette étape, il est à noter que c'est la partie 3' des ARNm (queue poly-A) qui s'hybride aux sondes. D'autres protocoles permettent de séquencer la partie 5' et comprennent des sondes de structures différentes, mais ils sont moins fréquents. Lors de la RT, des de courtes séquences mimant une extrémité 5' (« *switch oligo* ») sont ajoutées par hybridation aux nucléotides C (eux-mêmes ajoutés par les rétro-transcriptase pendant la RT).



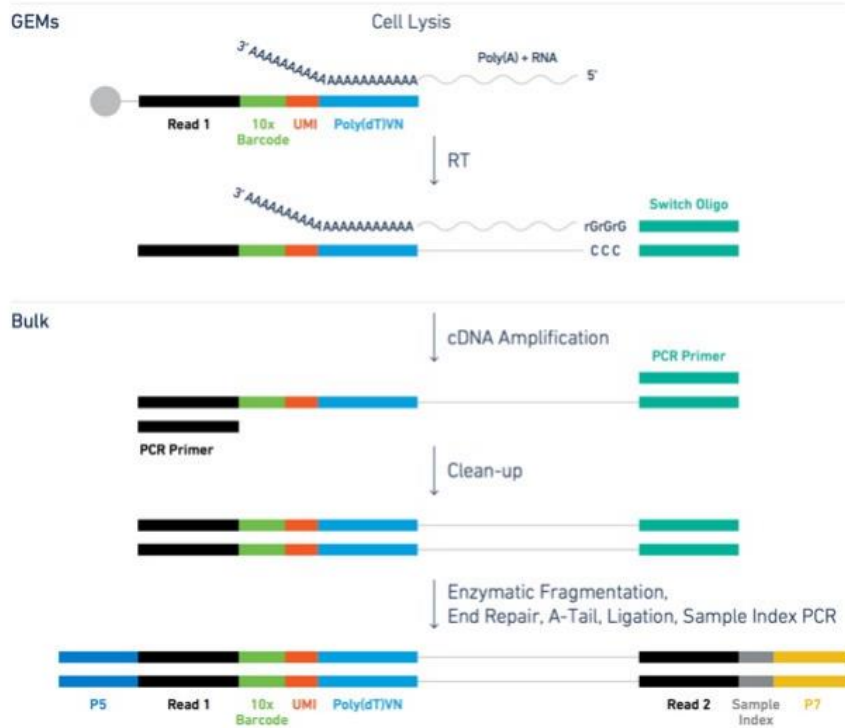


Figure 34. Étapes de préparation de la librairie de scRNA-seq

La première, la rétro-transcription, a lieu à l'intérieur des gouttelettes en suspension (GEMs) tandis que la suite de la préparation de la librairie se déroule en *bulk*, tous ARNs étant mélangés.

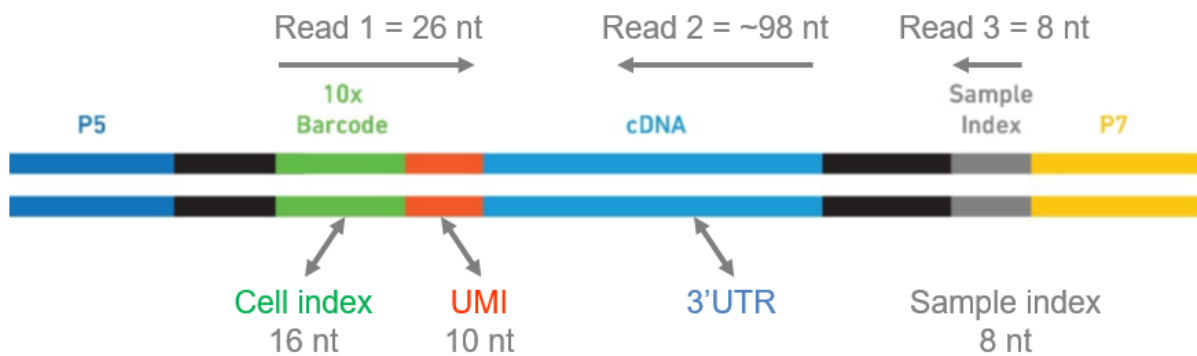


Figure 35. Structure des ADNc à l'issue du protocole Chromium de 10X Genomics

Les ADNc comprennent des séquences permettant l'hybridation au support de séquençage Illumina (P5 et P7) ; le code-barres 10X de 16 nucléotides et l'UMI de 10 nucléotides, compris dans l'amorce de séquençage du Read 1 ; la séquence 3' UTR de l'ADNc, correspondant à l'amorce du Read 2, et un index d'échantillon de 8 nucléotides (Read 3). L'ARN est partiellement séquencé.

Alors que toutes les étapes précédentes se déroulaient à l'intérieur du système microfluidique, la suite se déroule en dehors, avec du matériel conventionnel de biologie moléculaire. Toutes les billes et leurs ARNs capturés se retrouvent mélangées pour la suite du protocole, dans lequel les ADNc sont amplifiés par PCR, purifiés et fragmentés avant l'ajout d'index spécifiques pour le séquençage (Figure 34 ; Figure 35).

Dans la structure finale de l'ADNc qui sera séquencée se trouve le code-barres permettant d'identifier la cellule d'origine de l'ARN ainsi que l'UMI. Il est à noter que l'ARN n'est que partiellement séquencé, sur une longueur de 98 nucléotides dans sa partie 3'. Cette technologie donne donc accès à un transcriptome partiel. Le principal avantage des technologies de capture à base de gouttelettes est la possibilité de capturer beaucoup plus de cellules par échantillon, jusqu'à des dizaines de milliers. Ces approches sont également moins sélectives quant à la taille des cellules et produisent moins de doublets. Par conséquent, elles sont moins chères par cellule, les études utilisant la capture par gouttelettes séquençant généralement des cellules individuelles à une profondeur faible pour compenser leur grand nombre. La technologie Chromium de 10X Genomics permet diverses applications avec notamment le scRNA-seq, mais aussi le scATAC-seq (« *Single-cell Assay for Transposase-Accessible Chromatin* »), visant à étudier les états de compaction de la chromatine. Aujourd'hui, le Chromium 10X est l'une des technologies dominantes ; une étude suggère même que cette technique est la plus appropriée pour la majorité des expériences en dehors des contraintes de coût ou de protocole personnalisé (X. Zhang *et al.*, 2019).

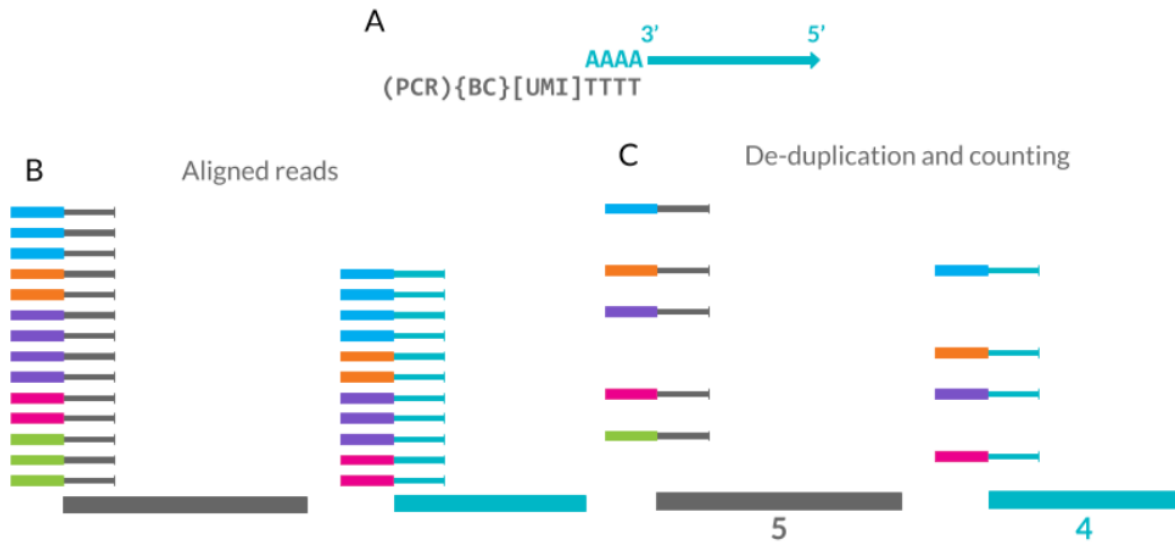


Figure 36. L'identifiant moléculaire unique améliore la quantification.

(A) Les UMI sont des séquences aléatoires de 8 à 10 paires de bases incluses dans la sonde de capture d'ARNm avec le code-barres cellulaire (BC) et l'amorce PCR. (B) La séquence de l'ARNm peut être alignée sur un génome de référence. (C) Chaque UMI n'est compté qu'une seule fois à chaque emplacement, ce qui élimine les doublons de PCR et améliore la quantification.

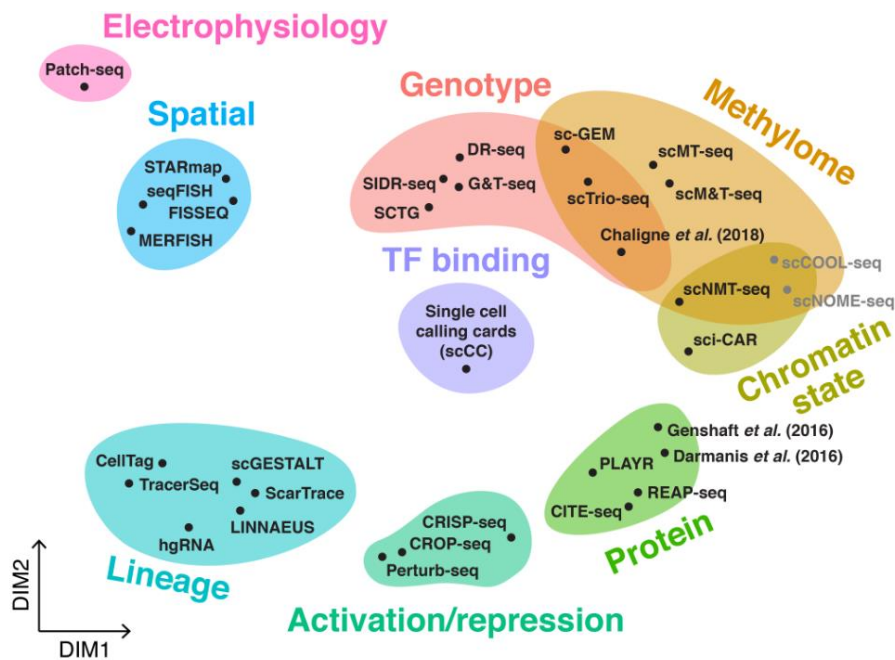


Figure 37. Résumé graphique de diverses technologies multi-omiques « *single-cell* »

Des protocoles ont été développés pour permettre des mesures multi-omiques à partir des mêmes cellules individuelles, dont une sélection est présentée ici. Ces autres méthodes mesurent le génotype, la méthylation, l'état de la chromatine, la liaison des facteurs de transcription, l'expression des protéines, l'activation ou la répression transcriptionnelle, le lignage cellulaire, la localisation spatiale ou encore l'électrophysiologie. Image et références provenant de <https://github.com/arnavm/multimodal-scRNA-seq>.

#### 1.4.3.1.2.2. Identifiant moléculaire unique ou « UMI »

La technologie 10X intègre les UMIs, une innovation moléculaire centrale pour les techniques basées sur la capture en microgouttelettes (Kivioja *et al.*, 2012). C'est à la suite de son développement pour le scRNA-seq que les UMIs ont inspiré les améliorations menant au protocole de BRB-seq présenté précédemment. Les cellules individuelles contiennent de très petites quantités d'ARN (de 10 à 30 pg, environ, dont moins de 5 % d'ARNm). Pour obtenir suffisamment d'ADNc pour le séquençage, une étape d'amplification par PCR est nécessaire. En fonction de leur séquence nucléotidique, les transcrits peuvent avoir des taux d'amplification différents, ce qui peut fausser leurs proportions relatives dans une librairie. Les UMI permettent d'améliorer la quantification de l'expression des gènes en permettant l'élimination des doublons de PCR produits pendant l'amplification (Figure 36) (Islam *et al.*, 2014). Les sondes nucléotidiques utilisées dans les protocoles de capture par gouttelettes, dont la technologie Chromium, comprennent une séquence UMI qui est spécifique à chaque sonde. Les séquences UMIs sont générées aléatoirement et sont suffisamment longues pour que la probabilité de capturer deux copies d'un même transcrit sur deux sondes ayant le même UMI soit quasi nulle. Après la rétro-transcription, l'amplification, le séquençage et l'alignement, la déduplication peut être effectuée en identifiant les lectures avec le même UMI qui s'alignent sur la même position et correspondent donc à des dupliquas PCR plutôt que des copies réellement exprimées d'un transcrit.

Pour que cette méthode soit efficace, chaque lecture doit être associée à un UMI, ce qui signifie que seule une petite section à l'extrémité 3' de chaque transcrit est séquencée. Cela a pour effet secondaire de réduire la quantité d'ADNc à séquencer et donc d'augmenter le nombre de cellules qui peuvent être séquencées en une fois. Cependant, si la quantification est améliorée, c'est au détriment de la couverture sur toute la longueur d'un transcrit, qui est nécessaire par exemple pour la détection de variants et l'assemblage de-novo. L'utilisation des UMIs est donc associée à un séquençage partiel du transcriptome, par exemple dans la technologie Chromium.

#### 1.4.3.1.3. Technologies émergentes et élargissement des protocoles

Des extensions aux protocoles standards ont également été proposées pour permettre toute une variété de mesures à l'échelle de la cellule unique (Figure 37). En particulier, des techniques permettent de réaliser des analyses multi-omiques, en combinant mesure d'expression génique et mesure du niveau de protéines, comme dans le cas du CITE-seq. Dans

cette section sont présentées une sélection de technologies adaptées des premières approches de scRNA-seq ou émergentes, illustrant partiellement la dynamique d'évolution des technologies *single-cell*.

#### **1.4.3.1.3.1. Single-nuclei RNA-sequencing**

Le séquençage ARN sur noyaux uniques, (ou « *single-nuclei RNA-sequencing* », scRNA-seq), est une alternative au scRNA-seq (Ding et al., 2020; Slyper et al., 2020). Selon un principe très similaire, il séquence le contenu en ARN des noyaux, à la place du contenu en ARN du cytoplasme pour le scRNA-seq. Cette approche a été développée pour étudier les noyaux des cellules difficiles à isoler vivantes et intègres. Alors que le scRNA-seq requiert des cellules viables pour la capture et le bon déroulement du protocole (viabilité cellulaire recommandée >70%), donc des tissus dans le meilleur état possible, généralement fraîchement retiré de leur contexte physiologique ; le snRNA-seq est plus facile à mettre en œuvre puisque les noyaux sont plus aisés à isolés à partir du broyat du tissu. La technique est particulièrement utilisée pour l'étude des tissus congelés, comme ceux accumulés dans les banques de tissus cancéreux. Il permet de façon général d'accéder à des populations cellulaires plus fragiles (Wu et al., 2019). Il est à noter que le transcriptome est légèrement différent puisque l'ARN des noyaux est immature et non épissé. Bien que les noyaux contiennent moins d'ARNm que les cellules, les analyses effectuées sur des cellules entières et des noyaux isolés à partir d'un tissu congelé du même échantillon ont démontré des profils transcriptomiques cohérents et des classifications similaires (Slyper et al., 2020). La possibilité d'analyser commodément des tumeurs congelées permet certaines applications telles que l'étude de paires d'échantillons prélevés chez un même individu à des moments différents. Cela peut s'avérer précieux pour mieux comprendre l'évolution d'une tumeur et les réponses aux traitements par exemple.

#### **1.4.3.1.3.2. CITE-seq**

L'un de ces protocoles, le CITE-seq, permet de mesurer les quantités de protéines à surface des cellules dans le même temps que le séquençage de leur transcriptome (Figure 38) (Stoeckius et al, 2017). Compatible avec la technologie 10X, cette technique repose sur l'utilisation d'anticorps spécifiques des protéines d'intérêt, conjugués à des séquences oligonucléotidiques appelées ADT (« *Antibody-Derived Tag* ») (Figure 38A). Ces séquences contiennent une amorce PCR et une séquence spécifique de chaque anticorps/protéine. Les ADTs sont commercialisés par l'entreprise Biolegend, notamment, et existent pour des

centaines de protéines de surface. Les analyses peuvent ainsi comprendre la caractérisation de plusieurs dizaines de protéines (Scheyltjens *et al.*, 2022). Les ADTs sont incubés avec les cellules dissociées afin de reconnaître et de se fixer spécifiquement sur leur protéine cible (Figure 38B). Ceux qui n'ont pas trouvé leur antigène sont éliminés par lavage avant la capture des cellules. Après la lyse des cellules dans les gouttelettes, les ARNm et les ADTs s'hybrident aux sondes des billes, ce qui permet de relier les codes-barres des cellules aux transcrits cellulaires et les codes-barres des ADTs à leurs protéines de surface. Ensuite, la librairie de séquençage est préparée selon un protocole adapté.

La technique du « *cell-hashing* » une technique de multiplexage (permettant de combiner les cellules de différents échantillons), utilise le même principe pour étiqueter avec un code-barres des protéines ubiquitaires sur toutes les cellules, mais dans des échantillons distincts. Les séquences ADTs comprennent cette fois un code-barres spécifique de chaque lot et permettent de construire une seule librairie de séquençage puis d'identifier les échantillons de provenance, dans l'analyse (Stoeckius *et al.*, 2018). En permettant l'utilisation commune d'un kit de réactifs pour plusieurs échantillons, la technique permet de réduire considérablement les coûts (compter 1500 euros par kit de réactif). Un avantage supplémentaire de cette approche est la détection simplifiée de doublets contenant des cellules provenant de différents échantillons.

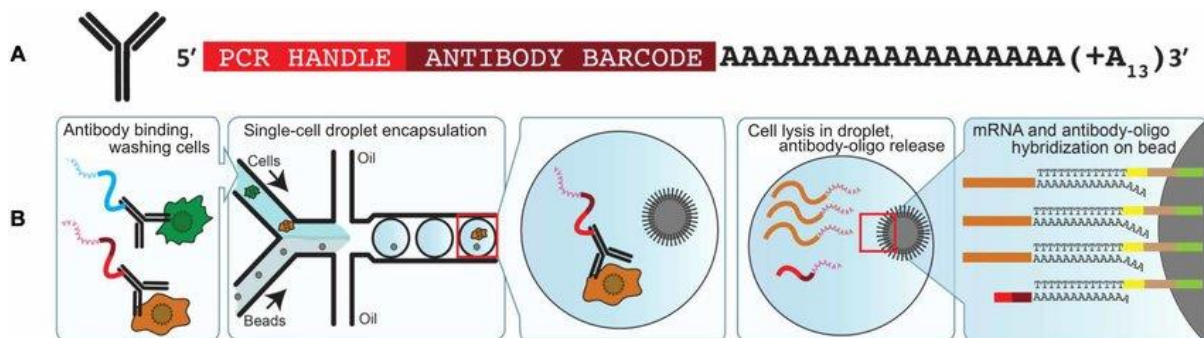


Figure 38. Schéma du protocole de CITE-seq

(A) Illustration des ADTs utilisés en CITE-seq. (B) Représentation schématique du CITE-seq dans un dispositif microfluidique de capture en gouttelettes. Les cellules sont incubées avec des anticorps, lavées et passées dans un système microfluidique où une seule cellule et une bille sont capturées dans la même gouttelette. Après la lyse des cellules, les ARNm et les ADTs s'hybrident aux sondes des billes, ce qui permet de relier les codes-barres des cellules aux transcriptions cellulaires et les codes-barres des ADTs à leurs protéines de surface.

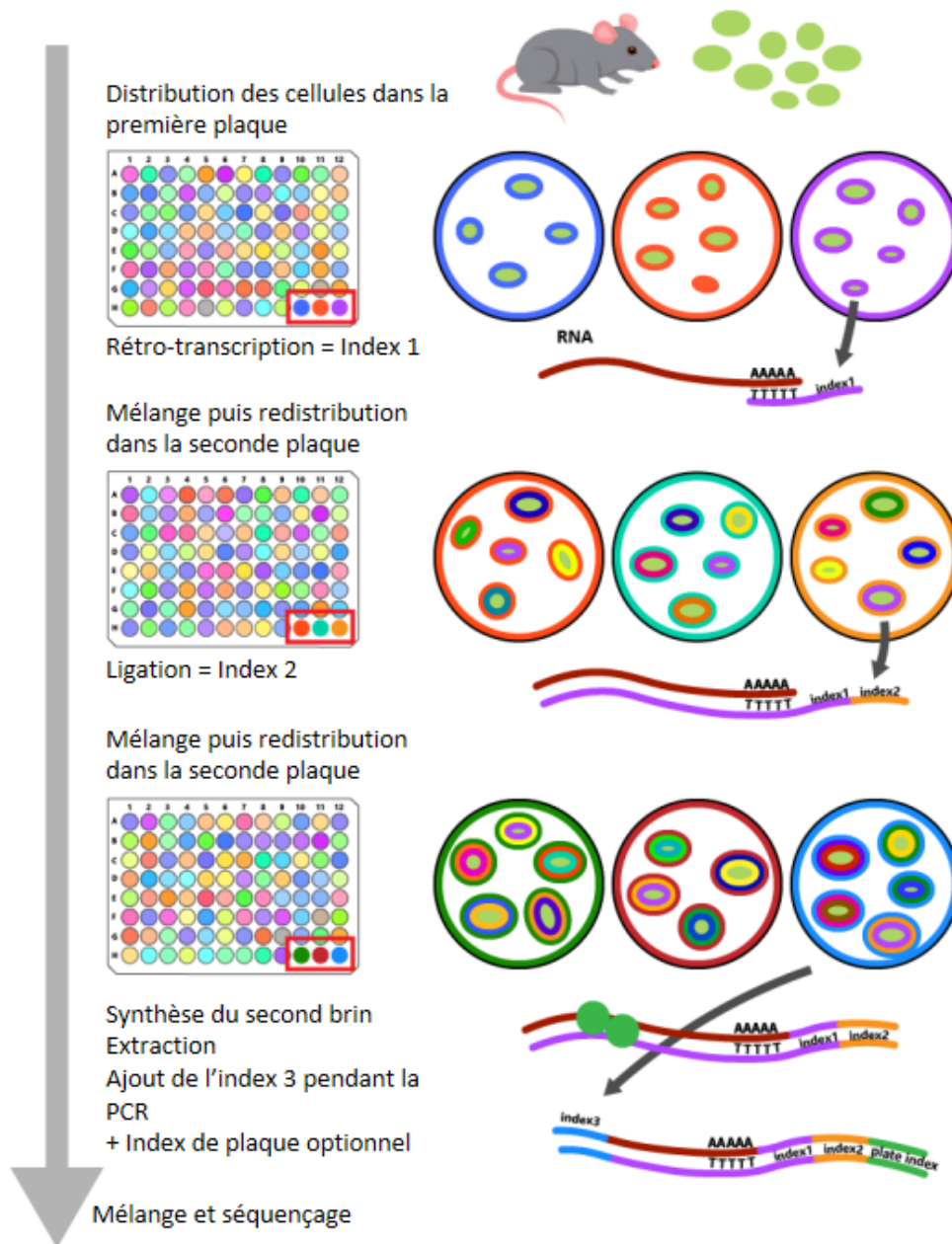


Figure 39. Schéma du protocole de sci-RNA-seq

Les noyaux/cellules sont isolés dans un tampon de lyse puis fixés. Les cellules sont ensuite réparties dans une plaque à 96 puits pour la transcription inverse, où le premier index est introduit. Si on le souhaite, les cellules de différents échantillons peuvent être déposées dans différents puits pendant ce premier tour d'indexation, ce qui facilite le traitement de plusieurs échantillons tout en minimisant les effets de lot. Après la transcription inverse et l'ajout du premier index, les cellules sont regroupées et réparties dans une nouvelle plaque pour ajouter un deuxième index par ligature, puis les cellules sont regroupées et réparties à nouveau. Après la synthèse du second brin et le marquage, le troisième index est ajouté par PCR. Enfin, la librairie est purifiée et séquencée. Dans cet exemple, près de deux millions de cellules ont été séquencées chez la souris et ont nécessité trois plaques, à titre d'exemple. De façon proportionnelle, il est possible de modifier l'échelle de l'approche pour séquencer davantage de cellules, moyennant des modifications de protocole (Martin *et al.*, 2021).

#### 1.4.3.1.3.3. Sci-RNA-seq

Bien que les techniques basées sur les gouttelettes soient actuellement les technologies de capture de cellules les plus couramment utilisées, d'autres approches ont été proposées qui promettent de capturer davantage de cellules à un coût moindre. Le sci-RNA-seq, ou « *Single-cell Combinatorial Indexing RNA-sequencing* » dont la première version de protocole parue en 2017 (J. Cao *et al.*, 2017) (Figure 39) a été récemment améliorée (Martin *et al.*, 2021), permet aujourd'hui de séquencer jusqu'à des millions de cellules ou noyaux à bas coût. Dans cette section, la technique du sci-RNA-seq est décrite en prenant pour exemple l'utilisation de cellules. Le principe repose sur l'indexation combinatoire, à savoir les ajouts successifs pour chaque cellule de plusieurs index (une séquence nucléotidique unique identifiant la cible à laquelle elle est ajoutée), dont la combinaison, au terme de suffisamment de passages, devient unique et spécifique de chaque cellule. Les cellules sont d'abord perméabilisées et fixées au méthanol puis sont réparties dans les puits d'une ou plusieurs plaques de 96 puits. Un premier index permettant d'identifier l'échantillon d'origine de chaque cellule est introduit dans chaque puits par une transcription inverse qui incorpore une amorce poly-T spécifique du puits, un code-barres, et des UMIs (Figure 39). Les cellules sont ensuite regroupées et réparties dans un nouvel ensemble d'une ou plusieurs plaques de 96 puits, à raison de 10 à 100 cellules par puits. Le deuxième index est ajouté à l'extrémité du premier index, puis les cellules ou les noyaux sont regroupés et divisés à nouveau. Dans le troisième ensemble de plaques, la synthèse du deuxième brin est réalisée et le produit double brin est ensuite marqué avec la transposase Tn5. L'amplification par PCR ajoute le troisième index et, enfin, la librairie de séquençage est purifiée et séquencée (Martin *et al.*, 2021). Entre autres applications, sci-RNA-seq3 a été utilisé pour générer les plus grands atlas scRNA-seq, chez la souris (~2 millions de cellules) (J. Cao *et al.*, 2017) et l'homme (~4 millions de cellules) (Cao *et al.*, 2020) à ce jour. La capture des cellules et la préparation des librairies de séquençage dans ces études ont d'ailleurs été réalisées dans un même laboratoire, en quelques semaines, et par un seul expérimentateur, illustrant l'intérêt de la technique.

#### 1.4.3.1.3.4. La transcriptomique spatiale

Le lien étroit entre l'organisation cellulaire au sein du tissu et les fonctions biologiques est bien connu, par exemple en biologie du développement, où des thèmes centraux tels que la perte de symétrie entre les cellules filles et leur devenir s'appuient en partie sur leurs relations spatiales (Gilbert and Barresi, 2017). En oncologie, similairement, comprendre la structure



spatiale des tumeurs est un enjeu important pour comprendre son fonctionnement en lien avec le TME. Beaucoup d'approches ont cherché depuis une dizaine d'années à capturer les profils transcriptomiques de ces structures spatiales. Cependant, ces méthodes limitent l'analyse à quelques gènes ou protéines à la fois, or jusqu'à récemment, les techniques de séquençage haut-débit ne pouvaient pas être appliquées *in situ*. Les premières méthodes permettant de contourner cette limitation, comme le Tomo-Seq, ont analysé par transcriptomique des séries de coupes pour reconstruire un axe spatial (Lacraz *et al.*, 2017). De même, d'autres approches utilisaient la microdissection laser pour isoler manuellement des régions spécifiques, allant parfois jusqu'à l'échelle de la cellule et reconstituant des informations distribuées spatialement (Nichterwitz *et al.*, 2016, 2018; J. Chen *et al.*, 2017; Moor *et al.*, 2018). Les approches les plus utilisées désormais s'appuient sur le séquençage haut-débit avec une haute résolution spatiale réelle et découlent des innovations liées au scRNA-seq, en s'appuyant sur l'ajout d'un code-barres spatial avant la préparation de la librairie (Larsson, Frisén and Lundeberg, 2021). En 2016, Stahl et ses collègues ont rapporté la première méthode de ce type leur ayant permis de capturer des transcriptomes entiers à partir de sections de tissus (Ståhl *et al.*, 2016). L'innovation centrale consiste à capturer l'ARN polyadénylé sur des lames adaptées, au niveau de points (ou « *spots* ») contenant des codes-barres spatiaux (Figure 40).

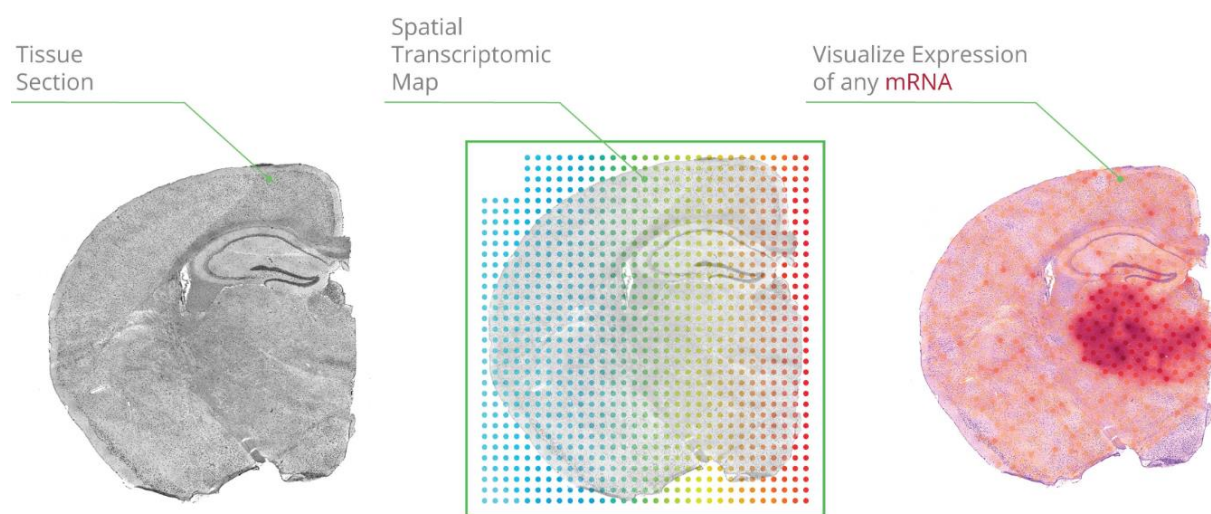


Figure 40. Illustration de l'approche de transcriptomique spatiale

Des sections de tissus réalisées à partir de tissus FFPE ou congelés sont analysées par points (« *spots* ») organisés sur des lames spéciales selon un quadrillage. Chaque point capture, dans ses coordonnées, le transcriptome du tissu. L'expression des gènes est ensuite mesurée informatiquement à l'échelle de ces points. La résolution des points et la taille des lames varient selon les technologies. Image tirée de <https://spatialtranscriptomics.com/technology/>.

Une librairie de séquençage d'ADNc incluant les barres-codes de ces points est ensuite préparée, à l'image des codes-barres identifiant les cellules en scRNA-seq. La méthode a été validée la première fois sur le bulbe olfactif de souris, et a depuis été employée par plusieurs

autres groupes (Thrane *et al.*, 2018; Moncada *et al.*, 2020). L'entreprise 10x Genomics a commercialisé technologie reposant sur cette technique, appelée *Visium*. Les points sur les lames dans la dernière version de cette technologie ont une résolution de 55µm, soit un diamètre de très peu supérieur à la taille moyenne des cellules. Ce diamètre a évolué depuis les premières versions à 100µm (~2 à 10 cellules), et atteindra très probablement bientôt la résolution de la cellule unique. Bien que la sensibilité soit globalement moins bonne qu'avec le scRNA-seq, elle permet d'obtenir plus de 10 000 transcrits par point (Lomakin *et al.*, 2021). Cette technologie a été utilisée dans de nombreux domaines par la suite, notamment en neurosciences (Ortiz *et al.*, 2020), cancérologie (Grauel *et al.*, 2020; Ji *et al.*, 2020) et en biologie du développement (Asp *et al.*, 2019). En oncologie où il y a une forte expertise en anatomopathologie et en histologie, cette dimension de l'analyse peut être très importante, suggérant que l'utilisation de cette technologie continuera d'y croître.

#### **1.4.3.1.3.5. LIVE-seq**

Le LIVE-seq est une technologie très récente, présentée dans une étude parue dans *Nature*, qui permet de séquencer le transcriptome individuel des cellules tout en les maintenant vivantes, là où les précédentes techniques nécessitent leur lyse (Chen *et al.*, 2022). Au cœur de cette innovation se trouve une technique de microscopie appelée "*fluidic force microscopy*", ou « FluidFM », qui utilise des canaux microscopiques - plus fins qu'un cheveu humain - pour manipuler sous microscope de minuscules volumes de fluides contenant les cellules. Les utilisateurs peuvent insérer des substances ou extraire le cytoplasme et l'ARNm de cellules individuelles sans avoir à les tuer. Le Live-seq permet d'analyser le transcriptome correspondant de façon plus immédiate à un phénotype cellulaire à un instant donné. En réutilisant les mêmes cellules, ils peuvent obtenir un transcriptome dit temporel. Les auteurs ont montré qu'ils pouvaient identifier divers types et états cellulaires sans introduire de perturbations majeures. Ils ont fourni une preuve de concept en cartographiant la trajectoire de différenciation de macrophages individuels avant et après leur activation, ainsi que des cellules stromales avant et après leur différenciation en cellules graisseuses.

### 1.4.3.2. L'analyse de données de scRNA-seq

L'élément central de l'analyse en scRNA-seq est la matrice d'expression, qui représente le nombre de transcrits uniques (« *counts* ») détectés pour chaque gène et dans chaque cellule (Figure 41).

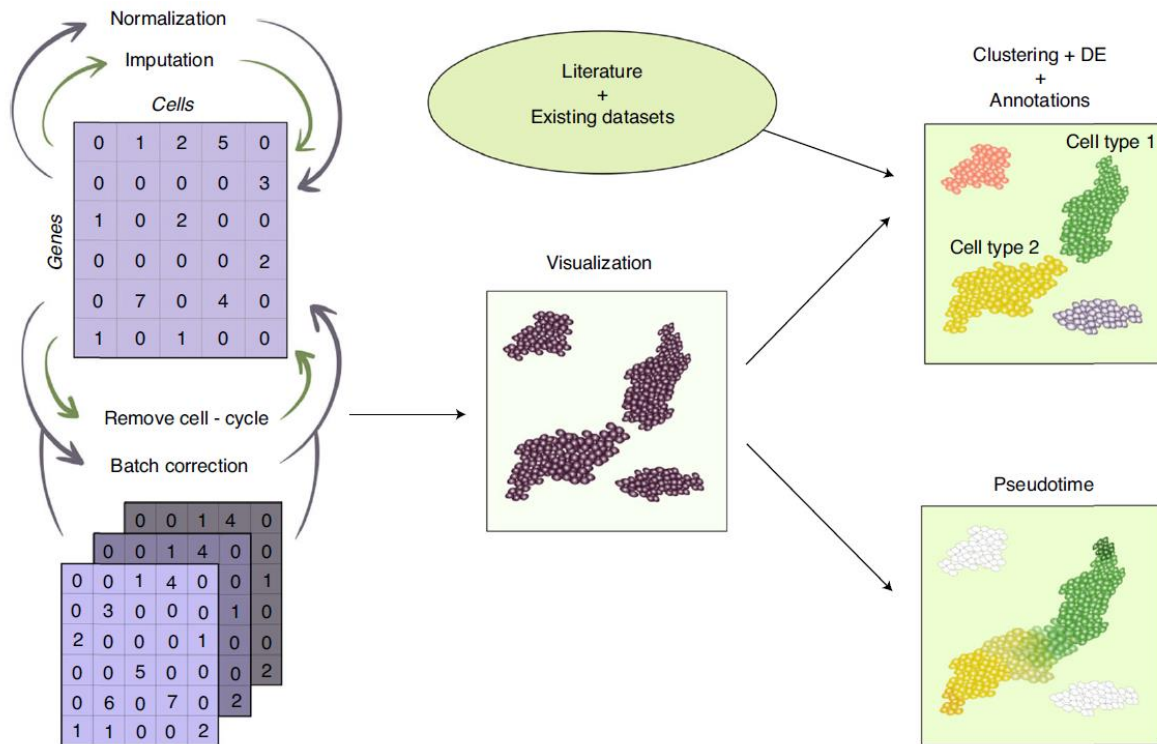


Figure 41. Schéma de l'analyse en scRNA-seq

Dans un scénario d'analyse typique, des matrices d'expression individuelles sont combinées en une matrice finale, qui est corrigée pour la profondeur de séquençage, le cycle cellulaire et d'autres biais potentiels. Les données sont ensuite visualisées et les groupes de cellules pertinents sont identifiés par classification (ou « *clustering* »), analyse d'expression différentielle et annotation à partir de gènes marqueurs. Souvent, des analyses de lignage établissent un « *pseudotime* » modélisant un temps de développement sont entreprises. Les données sont enfin comparées à la littérature existante. Figure tirée de Andrews *et al.*, 2021.

La plupart des gènes ne sont exprimés que dans un sous-ensemble de types cellulaires. En raison des faibles quantités de matériel de départ, de l'efficacité de capture et de la relativement faible profondeur de séquençage employée en single-cell RNA-seq, certains gènes ne seront pas détectés, même s'ils sont exprimés. En moyenne, sur les 20 à 30000 gènes du génome humain, seuls 1000 à 3000 gènes sont détectés dans chaque cellule. Il en résulte un grand nombre de valeurs nulles dans la matrice d'expression des gènes, ce qui est problématique car certains de ces zéros peuvent correspondre à plusieurs cas de figures : une expression nulle, une expression réelle mais faible, l'absence de détection de cette expression, ou encore variabilité dans la mesure (Sarkar and Stephens, 2021). Ce format de matrice a conduit au

développement nécessaire d'outils d'analyse dédiés aux données de scRNAseq, notamment du point de vue statistique. La difficulté de distinguer et de modéliser de façon appropriée ces sources de zéros est l'un des principaux enjeux dans l'analyse informatique. Certains outils informatiques intégrés tels que la suite Seurat dans l'environnement R (Butler *et al.*, 2018), ou Scanpy dans l'environnement Python (Wolf, Angerer and Theis, 2018), prennent en compte ces particularités et permettent d'inclure toutes les étapes de prétraitement et d'analyses standards des données.

#### 1.4.3.2.1. Contrôle qualité

La première étape de l'analyse de données issues de scRNA-seq consiste à exclure les codes-barres cellulaires qui ne correspondent pas à des cellules individuelles intactes. Ceux-là sont dus aux transcrits libres dans le milieu qui sont capturés par des billes dans des gouttelettes alors qu'aucune cellule ne s'y trouve. L'approche la plus simple consiste à estimer un seuil minimum d'UMIs requis pour considérer un code-barres comme correspondant à une cellule (G. X. Y. Zheng *et al.*, 2017). Dans une deuxième phase de contrôle qualité, certaines cellules sont considérées comme endommagées ou mourantes quand elles présentent des proportions trop élevées de gènes mitochondriaux (apoptose), trop faibles de gènes ribosomiaux, trop peu de gènes détectés ou des proportions élevées de lectures (« *read* ») non alignées ou multi-alignées. Ces cellules sont alors retirées de l'analyse ou mourantes (Ilicic *et al.*, 2016). Les seuils utilisés pour les paramètres précédents sont généralement déterminés par l'utilisateur, car le seuil optimal dépend du tissu, du protocole de dissociation et d'autres facteurs techniques. Définir les cellules aberrantes (en terme d'éloignement à la médiane) pour ces paramètres peut permettre d'établir des seuils spécifiques de chaque jeu de données, mais doit être fait avec précaution, en particulier quand les échantillons sont hétérogènes (Amezquita *et al.*, 2020). Il est également possible que les codes-barres cellulaires correspondent en fait à plus d'une cellule ; il s'agit de « doublets » ou de « multiplats ». Avec la technologie de microfluidique du Chromium (*10X Genomics*) par exemple, le pourcentage de doublets dépend du nombre de cellules analysées. Plus le nombre de cellules que l'on soumet au système est élevé, plus ce taux est élevé (Lareau *et al.*, 2020). Pour s'en affranchir, des outils tels que *scrublet* (Wolock, Lopez and Klein, 2019) et *DoubletFinder* (McGinnis, Murrow and Gartner, 2019) estiment un seuil maximal d'UMI correspondant à une cellule, ou, en deux temps, simulent des doublets possibles à partir du jeu de données puis calculent la similarité entre les mesures réelles et les doublets simulés pour définir un seuil.

#### 1.4.3.2.2. Normalisation

Le nombre de lectures de séquençage utiles varie d'une cellule à l'autre et doit être corrigé. Pour les données scRNA-seq, cet effet est prononcé, car la quantité d'ARN par cellule peut varier considérablement en raison du stade du cycle cellulaire et d'autres facteurs biologiques, même au sein d'un même type cellulaire (Marinov *et al.*, 2014). Par ailleurs, des facteurs techniques comme la taille des gouttelettes peuvent davantage accroître la variabilité dans la profondeur de séquençage (Macosko *et al.*, 2015). De façon générale, les différences de profondeur de séquençage peuvent être améliorées par une étape de normalisation. Pour les données de bulk RNA-seq, la normalisation consiste à calculer une quantité liée à la profondeur de séquençage de l'échantillon, et à diviser l'expression de tous les gènes par ce facteur. Une approche similaire peut être utilisée pour scRNA-seq en principe, mais le grand nombre de zéros présents dans la matrice d'expression nécessite de modifier la stratégie. Le programme *scran* sur R simule ce facteur en rassemblant un ensemble de cellules (Lun, Bach and Marioni, 2016). Il est également possible d'utiliser les ARN « *spike-in* » de référence, issus du « *External RNA Control Consortium* », ou des gènes de ménage pour estimer ces facteurs de normalisation (Lun *et al.*, 2017). En raison du grand nombre de zéros, la profondeur de séquençage peut impacter différemment les gènes faiblement exprimés de ceux fortement exprimés. Pour y remédier, il existe des méthodes de normalisation prenant en compte les niveaux d'expression de chaque gène. Par exemple, le programme *SCnorm* peut être utilisé pour les données bas-débit avec une grande profondeur de séquençage (Bacher *et al.*, 2017), et *SCTransform* (programme Seurat) peut être utilisé à l'inverse pour les données haut-débit et à faible profondeur de séquençage comme les données générées par le Chromium (Hafemeister and Satija, 2019). La commande par *SCTransform* réalise à la fois une normalisation impliquant par défaut un facteur de 10 000, une standardisation, et une détection des gènes variables entre les cellules.

#### 1.4.3.2.3. Correction des biais techniques

De façon similaire au biais potentiels induits par la profondeur de séquençage, les biais techniques (« *batch effet* » ou effet de lot), doivent être pris en compte afin de distinguer le signal biologique réel. Ces biais provenant de facteurs non-biologiques sont fréquents ; il s'agit par exemple de la durée de l'expérience, de variabilité de manipulation entre deux expérimentateurs ou de différences dans les réactifs utilisés. Généralement, ils peuvent être

évités par un design expérimental soigneux et adapté (Kim *et al.*, 2016). Cependant, cela reste compliqué, par exemple dans le cas du traitement simultané d'échantillons obtenus de façon asynchrone ou dans différents laboratoires. Il existe plusieurs outils de correction des effets batch tels que mnnCorrect (« *mutual nearest neighbors* ») (Haghverdi *et al.*, 2018) qui compare les conditions biologiques entre les cellules de différents lots, ainsi la « *canonical correlation analysis* » (CCA) et la « *Reciprocal canonical correlation analysis* » (RPCA) du programme Seurat (Butler *et al.*, 2018), adaptées de la première. Un autre programme, Harmony, sur l'environnement R, est fréquemment utilisé également pour ce type de correction (Korsunsky *et al.*, 2019).

#### 1.4.3.2.4. Sélection des gènes variables et réduction de dimensions

En scRNA-seq, chaque cellule est représentée dans un espace possédant autant de dimensions que de gènes exprimés. Chez l'humain ou la souris, un jeu de données est classiquement représenté dans un espace à environ 20 000 dimensions. Cependant, comme vus précédemment, de nombreux gènes ne sont pas détectés et le nombre de gènes détecté peut dépendre de la sensibilité du protocole ou du type cellulaire. Les méthodes de capture par gouttelettes identifient classiquement entre 1000 et 4000 gènes par cellule (Svensson *et al.*, 2017; Ziegenhain *et al.*, 2017). Le grand nombre de gènes peut rendre l'analyse difficile car ce sont autant de dimensions statistiques qu'il faut prendre en compte. Généralement, les gènes avec la plus grande variabilité d'expression entre les cellules sont donc sélectionnés (souvent, entre 2000 et 3000) (Yip, Sham and Wang, 2018). L'analyse par la suite, notamment la classification des cellules, est limitée à ce groupe de gènes pour plus de simplicité et pour réduire le bruit de fonds statistique induit par les milliers d'autres gènes. Dans l'outil Seurat, la commande SCTransform évoquée précédemment réalise ces étapes avec une approche non paramétrique permettant d'ajuster empiriquement la relation entre la variance et l'expression moyenne.

Même avec cette sélection de gènes variables, une cellule reste caractérisée par l'expression de milliers de gènes. Afin d'affiner le signal et réduire la caractérisation des cellules à certaines dimensions plus explicatives, le nombre de dimensions est réduit par des méthodes statistiques (S. Sun *et al.*, 2019) dont la plus courante est l'ACP (Analyse en composante principale). L'ACP permet une transformation linéaire qui préserve les distances euclidiennes entre les cellules dans un nouvel espace, celui de l'ACP. Cette méthode est particulièrement adaptée pour réduire les dimensions et résumer l'information dans les grands

jeux de données. Le nombre de composantes retenues pour la suite de l'analyse dépend de la complexité du jeu de données et peut être déterminé par divers algorithmes (Peres-Neto, Jackson and Somers, 2005). L'approche la plus courante consiste à afficher le pourcentage de variance expliquée par chaque composante, puis à identifier visuellement la composante à partir de laquelle cette variance chute significativement, afin de ne conserver que les composantes précédentes. Cependant, d'autres algorithmes sont utilisés dans l'analyse scRNA-seq pour réduire davantage la dimensionnalité afin de pouvoir représenter visuellement des cellules dans des espaces à deux ou trois dimensions. Ces algorithmes s'appuient généralement sur les 30 premières dimensions de l'ACP. L'une des méthodes les plus utilisées actuellement est la représentation UMAP (« *Uniform Manifold Approximation and Projection for Dimension Reduction* ») (McInnes *et al.*, 2018) (Figure 42).

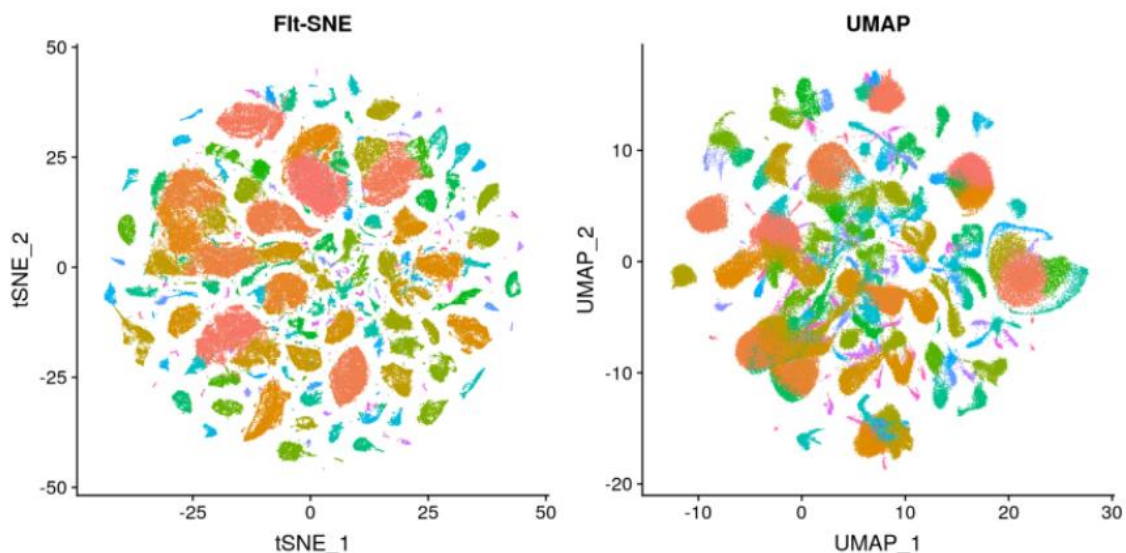


Figure 42. Représentation de type tSNE et UMAP d'un ensemble de cellules

Chaque point correspond à une cellule, elle-même caractérisée par l'expression de milliers de gènes, soit autant de dimensions. Ces dimensions ont été réduites à deux pour permettre l'affichage des cellules selon leurs coordonnées dans le nouvel espace défini par la méthode tSNE (gauche) ou UMAP (droite).

Cet algorithme fait l'approximation d'une structure des données qui s'appuie sur un réseau de plus proches voisins (« *nearest neighbors* ») entre les cellules, puis estime une manière de projeter les données en deux dimensions qui préserve au mieux cette structure. La représentation UMAP a largement remplacé la méthode t-SNE (« *t-distributed stochastic neighbor embedding* ») (Van Der Maaten and Hinton, 2008) en raison de sa capacité à mieux préserver les structures de données à grande échelle, notamment en ce qui concerne la pertinence des distances entre les cellules. Cependant, la représentation UMAP a tendance à favoriser les grands groupes (ou « *clusters* ») de cellules, donnant une représentation parfois moins résolutive que le t-SNE qui laisse apparaître des clusters discrets. Il est à noter que ces

deux méthodes requièrent des paramètres définis par l'utilisateur et susceptibles d'impacter significativement le résultat.

#### 1.4.3.2.5. Classification non supervisée

La classification, ou *clustering*, non supervisée des données scRNA-seq est indispensable car elle permet d'identifier des groupes de cellules présentant des profils transcriptomiques similaires – l'hypothèse étant que ces *clusters* sont composés de cellules appartenant au même sous-type cellulaire dans un état proche ou identique (par exemple une même phase du cycle cellulaire). Plus le nombre de clusters est important, plus la définition des sous-populations cellulaires est précise mais plus l'analyse devient complexe. Diverses méthodes de classification ont été développées bien avant l'avènement du scRNA-seq, et les outils existants sont des applications de méthodes classiques, tels que les algorithmes k-means (Hartigan and Wong, 1979) et de Louvain (Blondel *et al.*, 2008). Ce dernier est implémenté dans les programmes Seurat (Butler *et al.*, 2018) et Scanpy (Wolf, Angerer and Theis, 2018). Les méthodes basées sur l'algorithme de Louvain construisent un réseau de plus proches voisins, pour les cellules, puis identifient des communautés distinctes dans le réseau. La force de ces méthodes est leur rapidité, même pour de très grands ensembles de données. Chaque méthode de *clustering* possède ses propres paramètres qui peuvent affecter de manière significative les résultats et l'interprétation biologique. Par exemple, l'algorithme de Louvain requière un paramètre de résolution qui affecte le nombre de groupes ; plus ce paramètre est petit, plus le nombre de clusters identifiés est faible. De même, pour les méthodes basées sur les kmeans, le paramètre k détermine directement le nombre de clusters. Bien que certaines méthodes permettent d'estimer le nombre de clusters adéquat, tels que les packages R Mclust ou NbClust, l'utilisateur doit généralement prendre des décisions en fonction des particularités de son jeu de données et de la finesse biologique souhaitée. Ceci implique généralement de procéder à l'annotation des populations cellulaires à travers l'examen des gènes marqueurs exprimés dans les clusters de cellules.

#### 1.4.3.2.6. Expression différentielle

L'analyse d'expression différentielle (AED) est l'une des applications les plus importantes du bulk RNA-seq car elle permet d'obtenir une liste de gènes dont l'expression varie d'une condition biologique à l'autre. Dans le cas du scRNA-seq ou l'échelle se situe à la résolution de la cellule, l'enjeu immédiat après la classification des cellules est d'identifier leurs



gènes spécifiques, différenciellement exprimés d'un *cluster* à l'autre, afin d'identifier des marqueurs ou des voies de signalisation spécifiques. Pour le scRNA-seq, l'AED consiste à comparer des distributions de niveaux d'expression. Une particularité à noter est que les clusters de cellules à comparer ne sont identifiés comme types cellulaires qu'après leur classification, sur la base des mêmes gènes qui font l'objet de l'AED. Cette particularité invalide l'hypothèse des tests standards, la classification et l'AED n'étant plus indépendants. Il a été démontré que la classification non supervisée suivie d'une analyse d'expression différentielle peut donner lieu à des *pvalues* artificiellement basses (Zhang, Kamath and Tse, 2019). Une comparaison récente a conclu que le test non paramétrique de Wilcoxon quant à lui donne des résultats remarquablement bons (Soneson and Robinson, 2018). Les auteurs ont également conclu que les méthodes développées pour le bulk RNA-seq donnaient de bons résultats, en particulier lorsqu'elles attribuaient des poids à chaque élément de la matrice d'expression (Van den Berge *et al.*, 2018). Une autre étude comparative est parvenue à des conclusions similaires, ajoutant que la normalisation peut influencer le résultat (Vieth *et al.*, 2019). L'une des méthodes spécialement conçue pour le scRNA-seq, MAST (« *Model-based Analysis of Single-cell Transcriptomics* ») utilise un modèle gaussien pour combiner à la fois les différences de taux de détection et de moyenne d'expression en un test, et a été reportée comme donnant les meilleurs résultats (Soneson and Robinson, 2018; Luecken and Theis, 2019). À mesure que le scRNA-seq est appliqué sur des cohortes plus importantes et à des études comparatives, nous assistons aux développements de nouveaux modèles statistiques plus précis, tels que les outils TRADE-seq (Van den Berge *et al.*, 2020) ou IDEAS (Zhang *et al.*, 2022). Ces méthodes évoluent beaucoup, et sont souvent intégrées dans des outils d'analyse tels que Seurat, dont la fonction *FindAllMarkers* inclut nombres des méthodes évoquées précédemment.

#### 1.4.3.2.7. Inférence de CNVs

Plusieurs algorithmes ont été développés pour inférer à partir de données de scRNA-seq les variations du nombre de copies de gènes (« *copy number variation* » ou CNV) dans les échantillons tumoraux, reflétant le degré de réarrangements chromosomiques et le dérèglement des cellules tumorales (Lobo, 2008). Le plus connu et utilisé est InferCNV (inferCNV of the Trinity CTAT Project <https://github.com/broadinstitute/inferCNV>). Le programme explore les données de scRNA-seq de tumeurs afin d'identifier les gains et pertes à grande échelle (bras chromosomiques entiers ou large segments), ceci en comparant l'intensité relative d'expression des gènes à travers les positions du génome tumoral par rapport à des cellules "normales"

utilisées comme référence. Cette approche permet de déterminer CNV connus ou peu renseignés, et a notamment été utilisée avec succès dans des études récentes pour distinguer les cellules malignes, portant des aberrations chromosomiques, des cellules non-malignes (Filbin *et al.*, 2018; Hovestadt *et al.*, 2019; Neftel *et al.*, 2019).

#### 1.4.3.2.8. Déconvolution

La déconvolution consiste à déduire à partir d'un signal global la multitude de signaux plus petits le composant. En transcriptomique, cela revient à utiliser les profils d'expression génique de populations cellulaires connus pour déduire leur représentativité (signaux multiples, issus du scRNA-seq) dans un profil d'expression génique correspondant à un tissu complexe (signal global, issu du RNA-seq *bulk*) (Figure 43).

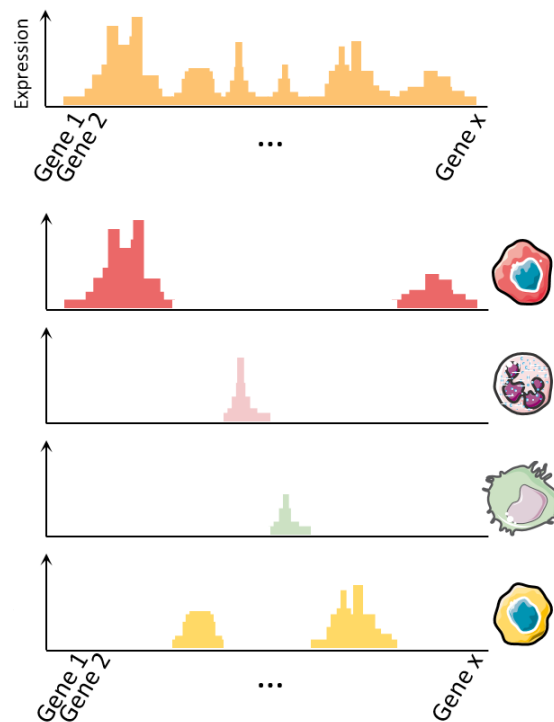


Figure 43. Illustration de la déconvolution transcriptomique

Un signal global, issu du séquençage ARN global, est représenté en haut. Le signal est déconvolué en plusieurs signaux composés de gènes marqueurs et permettent d'identifier des populations cellulaires.

Il existe déjà de nombreuses données de séquençage d'ARN de type « bulk » ou provenant de puces à ADN, issues de cohortes massives qui présentent l'intérêt d'associer données transcriptomiques et données cliniques. Cela peut servir par exemple à 1- déconvoluer les données RNA-seq *bulk* et, 2- associer les types cellulaires déduits (ou « déconvolués ») à des paramètres cliniques telles que le pronostic des patients. Depuis l'avènement du scRNA-

seq, plusieurs méthodes ont été développées pour déconvoluer les types cellulaires et leur proportion dans ce type de données, à partir de références à l'échelle de la cellule unique. (Avila Cobos et al., 2020), telles que CIBERSORT (Newman et al., 2015) et son extension BSEQ-sc (« *Deconvolution of Bulk Sequencing Experiments using Single Cell Data* ») (Baron et al., 2016) ou MuSiC (X. Wang et al., 2019). La composition cellulaire elle-même ou bien son évolution pendant la prise d'un traitement peuvent ainsi être évaluées par déconvolution informatique. Cela a permis dans certaines études de montrer que les patients présentant des niveaux élevés de lymphocytes infiltrant les tumeurs répondaient mieux aux immunothérapies (Darvin et al., 2018) ou que des changements dans certains types immunitaires survenaient au cours d'une infection grippale (Altboum et al., 2014).

#### **1.4.3.2.9. Lignage cellulaire et reconstruction de pseudotemps**

Les cellules de nombreux systèmes biologiques présentent une continuité et des transitions entre leurs différents états cellulaires. Ces processus dynamiques, au sein d'une partie des cellules peuvent être modélisés en reconstruisant à partir de données scRNA-seq la trajectoire « évolutive » ou transitionnelle de ces cellules et de les projeter ainsi dans sur une échelle de temps fictive, ou « pseudo-temps » (Figure 44). Le pseudo-temps est une mesure de la progression d'une cellule le long de la trajectoire. Il traduit les variations transcriptionnelles d'une cellule à l'autre et permet de les ordonner le long de cette trajectoire pour les situer dans le processus étudié (Griffiths, Scialdone and Marioni, 2018). Outre la révélation de la dynamique d'expression génique entre les transitions cellulaires, l'inférence de trajectoire peut également servir à l'identification des facteurs déclenchant les transitions d'état. Une multitude d'outils est disponible pour l'inférence de trajectoires, par exemple Monocle (Trapnell *et al.*, 2014; Qiu *et al.*, 2017; Cao *et al.*, 2019), Waterfall (Shin *et al.*, 2015), Wishbone (Setty *et al.*, 2016), TSCAN (Ji and Ji, 2016), Slingshot (Street *et al.*, 2018), CellRouter (Lummertz Da Rocha *et al.*, 2018), Velocity (La Manno *et al.*, 2018) et PAGA (F. A. Wolf *et al.*, 2019). Le format de la trajectoire peut varier en fonction de l'outil et être linéaire, bifurquée, ou en arborescente. Monocle3 utilise par exemple une approche comprenant des méthodes non supervisées axées sur les données et l'intégration de graphes inversés, qui améliore significativement la rapidité et la robustesse de l'analyse par rapport aux premières versions de l'outil. Slingshot s'appuie sur les groupes de cellules pour identifier avec différents niveaux de supervision de multiples trajectoires possibles. CellRouter utilise des réseaux de flux (« *flow networks* ») pour identifier les transitions d'états cellulaires. Une revue parue dans *Nature*

*Biotechnology* a évalué un certain nombre de ces approches et conclut que Slingshot, TSCAN et Monocle2 étaient les plus efficaces et pertinentes (Saelens *et al.*, 2019). A cet égard, la suite d'outils Dynverse permet d'utiliser et comparer les différentes méthodes afin de sélectionner la plus adaptée selon le modèle étudié (Cannoodt *et al.*, 2021).

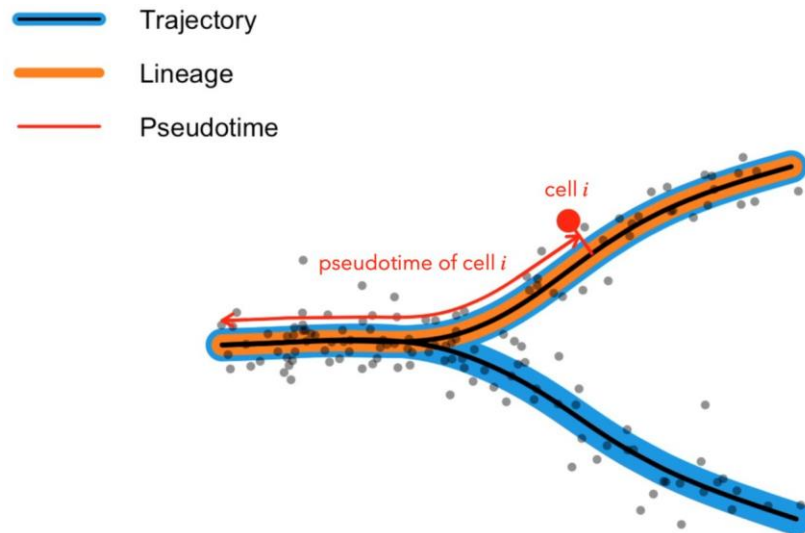


Figure 44. Illustration de l'inférence de trajectoires utilisée en single-cell RNA-seq

Une trajectoire est une collection de un ou plusieurs lignages cellulaires. Un lignage est un chemin de différenciation unique. Le pseudo-temps d'une cellule est la longueur de son lignage, depuis sa création jusqu'à sa projection.



## 1.5. DE LA TRANSCRIPTOMIQUE DANS L'ÉTUDE DES CCRCCS

### 1.5.1. ÉTATS DES LIEUX ET ENJEUX EXPERIMENTAUX

Ces dernières décennies, l'avènement des techniques « -omiques » a permis de mieux comprendre le ccRCC à l'échelle moléculaire et de développer des thérapies ciblées ou immunothérapies pour son traitement. Ces molécules ne sont cependant que partiellement efficaces. Il reste beaucoup à découvrir et à comprendre sur les bases moléculaires de cette résistance, ainsi que sur les causes des rechutes ou du développement de métastases. Il existe ainsi un besoin de caractériser le ccRCC du point de vue transcriptomique car sa carcinogenèse est peu connue à ce jour.

Du point de vue expérimental, l'un des problèmes liés à l'utilisation d'échantillons primaires, de façon immédiate après prélèvement ou à partir de biobanques, concerne les conditions d'obtention variables du tissu. En ce qui concerne l'analyse transcriptomique, certains gènes sont très sensibles et peuvent être rapidement impactés en fonction de la température ou de l'ischémie. L'impact transcriptionnel de l'hypoxie induite par le transport est par exemple observé dans la plupart des types de tumeurs (Park, 2017). Si un retard dans le circuit de la pièce opératoire a peu d'incidence sur sa gestion clinique, les analyses transcriptomiques peuvent en revanche s'en trouver affectées. Dans la pratique chirurgicale liée au cancer du rein, certains aspects peuvent également affecter les profils des ARNm. Souvent, lors d'une chirurgie laparoscopique ou robotique, la tumeur est prélevée et placée dans un sac au sein la cavité corporelle en attendant d'agrandir l'incision par laquelle elle sera retirée. Ce temps supplémentaire après résection à la température du corps peut entraîner une dégradation ou un changement du niveau d'ARNm. Le clampage de l'artère rénale avant la néphrectomie partielle peut également contribuer à l'ischémie, impactant de façon générale la qualité moléculaire de l'échantillon tumoral. Une étude a évalué l'impact des conditions de prélèvement sur la qualité et l'abondance de l'ARN (Liu *et al.*, 2013) et démontré que l'ARN du RCC se dégrade après quatre heures et que jusqu'à 4 000 gènes peuvent être significativement modifiés en fonction des conditions de stockage de la tumeur ou du temps ischémique. Pour optimiser la qualité des échantillons biologiques, la standardisation et la rapidité des conditions d'obtention de la tumeur sont ainsi décisives, dans le traitement du cancer du rein comme dans d'autres cancers. Ces enjeux sont importants à souligner étant donné leurs impacts éventuels sur le

transcriptome, et permettent d'expliquer une part de la variabilité observée dans les données de transcriptomiques.

### **1.5.2. APPORTS DES APPROCHES BULK**

Les techniques les plus utilisées dans l'étude transcriptomique *bulk* (globale) du ccRCC sont la RT-PCR (« *Reverse transcription polymerase chain reaction* »), les puces à ADN et le séquençage ARN. Cette section propose de décrire les résultats d'études notables par technologie de transcriptomique *bulk*.

#### **1.5.2.1. RT-qPCR**

En 1993, Kary Mullis a reçu le prix Nobel de chimie pour son invention de la PCR. Depuis ses prémices dans les années 1970, la RT-qPCR a révolutionné les sciences de la vie car elle permet de mesurer aisément l'expression de gènes spécifiques dans un échantillon biologique, à la condition de connaître leur séquence permettant de construire des amorces spécifiques pour l'amplification. Bien qu'elle ait permis des avancées conceptuelles importantes, comme par exemple la détection des diverses translocations du facteur de transcription E3 dans le cancer du rein (Lee *et al.*, 2017), la RT-qPCR est généralement utilisée pour la validation expérimentale de l'expression de certains gènes ; elle a permis de démontrer l'implication d'un très grand nombre d'entre eux dans la biologie du ccRCC (Syed *et al.*, 2020). Récemment, des études se sont par exemple intéressées à des candidats tels que les micro-ARN et ont pu démontrer leur rôles oncogènes (Jin *et al.*, 2017; Lai *et al.*, 2018; Quan *et al.*, 2019) ou suppresseurs de tumeur (P. Chen *et al.*, 2018; Jiang *et al.*, 2019) dans le cancer du rein, leur donnant une valeur prédictive potentielle. En dépit de son caractère « bas-débit », la RT-qPCR a également été employée pour cribler a priori d'importantes séries de gènes candidats permettant de prédire les résultats cliniques après opération. Dans une étude d'envergure en 2015, une équipe américaine a rassemblé une grande cohorte ccRCC de stade I à III. Ils ont réalisé des RT-qPCR sur des tissus FFPE (« formalin fixed and paraffin embedded ») issus de 942 ccRCC de stades I à III et ont criblé 732 gènes sélectionnés à partir de données de puce à ADN (Rini *et al.*, 2015). Ils ont ensuite sélectionné 11 gènes liés au ccRCC et 5 gènes « de ménage » (dont le niveau d'expression est constant dans les tissus) pour leur association statistique avec la rechute des tumeurs et la survie des patients, constituant une signature transcriptomique prédictive. Après validation dans une cohorte indépendante de 626 patients, leurs résultats démontrent la validité de la signature. Leur score a notamment identifié un

nombre significatif de patients à haut risque au stade I et à faible risque au stade II-III, laissant envisager une amélioration des systèmes diagnostics. Dans une autre étude évaluant le sunitinib en situation adjuvante, cette signature a également permis de stratifier efficacement en différentes catégories de risques 221 patients. En revanche, l'étude montre que la signature, si elle a une valeur pronostique, n'a pas de valeur prédictive du bénéfice du traitement, puisque qu'elle n'est pas corrélée avec les groupes de patients répondeurs non répondeurs. Une autre étude publiée en 2018 validait avec des RT-qPCR la qualité prédictive de ce score (Rini *et al.*, 2018), démontrant l'intérêt de l'usage de la RT-qPCR à des fins d'amélioration de la prise en charge clinique, en particulier s'agissant de d'évaluer l'expression de seulement 16 gènes.

### **1.5.2.2. Puces à ADN**

Les puces à ADN représentent un gain de débit conséquent par rapport à la RT-qPCR ; selon leur conception, elles permettent d'analyser l'expression de plusieurs milliers de gènes. L'une des façons d'exploiter ces données consiste à identifier les gènes exprimés de façon différentielle et statistiquement significative (« *differentially expressed genes* » ou DEGs) entre différentes conditions, par exemples des situations de traitements ou des stades pathophysiologiques distincts. Dans une étude visant à étudier la pathogenèse du ccRCC, les auteurs ont par exemple utilisé cette approche pour identifier les DEGs chez des patients ccRCC par rapport à des patients contrôle. Ils ont utilisé ces DEGs pour bâtir un réseau de régulation permettant d'identifier 126 facteurs de transcription qui, en lien avec leurs gènes cibles, pourraient jouer le rôle de biomarqueurs dans le ccRCC (Yao *et al.*, 2016). Ces facteurs incluent par exemple FOXC1, GATA3 et ESR1. Dans les mêmes DEGs, ils ont étudié l'enrichissement en gènes liés à des voies moléculaires et proposent les voies ErbB et MAPK comme acteurs de la pathogenèse du ccRCC. Avec une approche similaire, une autre étude a proposé sur la base de données de puces à ADN des gènes candidats impliqués dans la carcinogenèse du ccRCC (Wang *et al.*, 2018), comprenant MME, ALB et CDH1. Ils ne relient pas ces gènes à des spécificités du ccRCC cependant. De la même façon, ils proposent des voies « candidates », bien que leur étude d'enrichissement révèle des termes génériques tels que « voies métaboliques » et « processus métaboliques des petites molécules ». Dans une autre étude, Khan et ses collègues ont utilisé la même approche à nouveau pour déterminer les voies moléculaire impliquées dans le processus métastatique du ccRCC (Khan *et al.*, 2016). Ils montrent ou confirment l'implication des voies ILK (« *integrin-linked kinase* »), de l'extravasation de leucocytes, de la voie CXCR4 ou encore de la voie PI3K/AKT, suggérant



leur analyse plus approfondie pour élucider les mécanismes métastatiques dans le ccRCC. Dans autre étude, Hakimi et ses collaborateurs rapportent les résultats d'une des plus grandes analyses moléculaires de patients atteints de ccRCC métastatique et traités par TKI (Hakimi *et al.*, 2019). Après avoir dressé les profils d'expression génique de 409 patients (212 traités par sunitinib, 197 par pazopanib) à l'aide de puce à ADN, ils ont identifié quatre sous-groupes moléculaires distincts associés à la réponse et à la survie. Les analyses basées sur l'ARN ont révélé quatre sous-groupes moléculaires distincts associés à la réponse et à la survie. La caractérisation de ces sous-groupes leur a permis d'identifier les programmes d'angiogenèse et l'infiltration de macrophages comme étant de bons facteurs prédictifs de l'issue du traitement par TKI. Notamment, ces facteurs différaient selon le type de TKI reçu. Enfin, de façon notable dans plusieurs études, les puces à ADN ont aussi été utilisées pour établir les profils transcriptomiques de biopsies de carcinome rénaux et se sont notamment avérées utiles pour en distinguer les principaux sous-types du point de vue moléculaire. (Tan *et al.*, 2004, 2010; Lovisolo *et al.*, 2006; Petillo *et al.*, 2009).

### **1.5.2.3. Séquençage ARN massif**

Le TCGA-KIRC comprend des données de génomiques, de transcriptomique, de génotypage et de méthylome sur 537 ccRCCs (Syed *et al.*, 2020). Cet effort collectif est la plus vaste collection de données de transcriptomiques sur le ccRCC à ce jour (Tomczak, Czerwińska and Wiznerowicz, 2015). Elle a l'avantage de contenir également les données cliniques des patients, ce qui permet de les associer aux données transcriptomiques dans les études. De façon générale, une majorité des études traitant des données de RNA-seq dans le ccRCC utilisent les données du TCGA. Des études s'en sont par exemple fréquemment servi pour identifier des biomarqueurs de progression tumorale et de survie (Hu, Zeng and Liu, 2019; Yejinpeng Wang *et al.*, 2019; Zuo *et al.*, 2019; Z. Zhang *et al.*, 2020; Bi *et al.*, 2021). En dehors de cet effort international, quelques études ont eu recours au séquençage ARN sur des cohortes relativement importantes, comme une étude japonaise parue dans Nature Genetics en 2013, qui analyse les profils RNA-seq et Exome-seq de plus de 100 patients atteints de ccRCC (Sato *et al.*, 2013). Les auteurs révèlent une altération de la protéolyse normale dans les cellules causée par l'inactivation de VHL comme une caractéristique est commune dans tous les ccRCC. Ils identifient également d'autres voies et composants mutés de manière récurrente, comprenant la signalisation PI3K-AKT-mTOR, le complexe KEAP1-NRF2-CUL3, la méthylation de l'ADN et le métabolisme de p53. Cette étude donne de nouveaux éclairages sur les mécanismes

moléculaires à l'œuvre dans le ccRCC, qui reste une thématique à largement explorer. Dans une autre cohorte plus modeste, Parasramka et ses collaborateurs ont comparé par RNA-seq les tumeurs ccRCC à faible risque issues des blocs FFPE de 24 patients décédés et de 24 patients guéris, et ont déterminé une signature de dix gènes caractéristiques des premiers (Parasramka *et al.*, 2016). Cette signature rejoint de nombreuses autres évoquées précédemment, mais n'a pas été validée dans des cohortes externes. De façon similaire, une autre étude étudié les profils RNA-seq de 33 patients comprenant des tumeurs de haut grade et des métastases, afin de déterminer une signature associée aux ccRCC agressifs (Batai *et al.*, 2018). Ils constituent en deux groupes les patients avec des tumeurs ayant de hautes ou de faibles valeurs pour les paramètres suivants : la taille de la tumeur, le grade ISUP, le stade d'avancement et score de nécrose. Ils identifient ensuite 290 DEGs, dont quatre (G6PD, APLP1, GCNT3, et PLPP2) sont surexprimés dans les tumeurs de stades avancés (III et IV) et de hauts grades ISUP (3 et 4) ; et deux (APLP1 et PLPP2) sont significativement associés à une un mauvais pronostic dans les stades précoces. De façon intéressante, le RNA-seq comme les technologies précédentes a beaucoup servi à établir des signatures pronostiques et identifier des gènes marqueurs dans le ccRCC. D'autres approches ont été employées toutefois, comme la technologie de RBR-seq, pour étudier de l'impact de traitements anticancéreux sur des cultures organotypiques de ccRCC (Giacosa *et al.*, 2021). L'étude permet notamment de révéler un mécanisme potentiel par lequel une combinaison de thérapies ciblant les protéines CK2 et ATM semble efficacement tuer les cellules malignes.

### **1.5.3. SINGLE-CELL RNA-SEQ EN ONCOLOGIE ET DANS L'ETUDE DES CCRCCS**

Dans cette section, une première partie décrit les enjeux et les apports du scRNA-seq dans l'étude des cancers en général, puis une seconde partie décrit les études scRNA-seq dans le ccRCC, afin de donner un aperçu plus global et permettre d'envisager les approches de façon transversale.

#### **1.5.3.1. scRNA-seq en oncologie**

##### **1.5.3.1.1. Caractérisation des tumeurs à l'échelle de la cellule unique**

Comme décrit précédemment, l'HIT est observée dans de nombreux types de cancers. Elle désigne la coexistence au sein d'une même tumeur de clones tumoraux distincts et de cellules non-tumorales dans son microenvironnement. L'HIT, à travers cette diversité cellulaire, gouverne des phénotypes tumoraux tels que la progression tumorale, la rechute ou

l'échappement thérapeutique (Burrell *et al.*, 2013). L'un des enjeux majeur de la recherche en oncologie consiste à mieux comprendre et utiliser l'information de cette HIT (Suvà and Tirosh, 2019), ce qui nécessite d'accéder à la résolution de la cellule unique. En effet, la cellule est l'unité opérationnelle fondamentale du cancer (Baslan and Hicks, 2017). Le scRNA-seq, fournissant une information transcriptomique très précise et non-biaisée de chaque cellule, s'est récemment imposé comme la technique de référence pour parvenir à cet objectif. Idéalement, la technique requière la dissociation rapide de tissus fraîchement opérés afin qu'ils aient une bonne viabilité. Bien que la technique soit également possible sur tissu congelé, cela est plus compliqué car les cellules auront été fragilisées et la dissociation du tissu entraîne parfois de grosses pertes. Cependant, la caractérisation du transcriptome de tissus tumoraux congelés est également possible via le séquençage ARN sur noyaux uniques (snRNA-seq). Que ce soit sur cellules ou noyaux, ce profilage transcriptomique des tumeurs comprend quatre étapes principales : 1 – des échantillons frais ou congelés sont obtenus et dissociés via digestion mécanique et enzymatique pour obtenir des suspensions cellulaires ou nucléaires ; 2 – les cellules ou noyaux sont isolés ; 3 – leurs ARNs sont rétro-transcrits en ADNc (soit pour la longueur totale, soit pour les extrémités 3' ou 5') et éventuellement barre-codés selon la technique ; 4 – Le pool d'ADNc est séquencé et le démultiplexage permet de mesurer l'expression de chaque gène dans chaque cellule. Ces étapes doivent généralement être optimisées selon l'objectif et le cancer étudié.

### **1.5.3.1.2. Apports du scRNA-seq en oncologie**

#### **1.5.3.1.2.1. Lignage cellulaire, différenciation et plasticité**

Les programmes de développement jouent un rôle important dans de nombreuses tumeurs malignes (Shibue and Weinberg, 2017). Cette observation est à la base du modèle des cellules souches cancéreuses, qui suggère que les tumeurs contiennent des sous-ensembles de cellules de type souche auto-renouvelables à l'origine de la propagation tumorale (et dans certains cas de la résistance aux thérapies), ainsi qu'une majorité de cellules plus différenciées (Kreso and Dick, 2014; Lathia *et al.*, 2015). Des hiérarchies cellulaires et modèles de différenciation ont été décrits dans un grand nombre de cancers comprenant entre autres le gliome (Filbin *et al.*, 2018; Gojo *et al.*, 2020), le mélanome (Su *et al.*, 2017), le neuroblastome (Dong *et al.*, 2020), le médulloblastome (Hovestadt *et al.*, 2019; Jessa *et al.*, 2019; Dong *et al.*, 2020) et le cancer du poumon (Ireland *et al.*, 2020; Maynard *et al.*, 2020). Ces études démontrent globalement une similitude entre le développement normal et celui du cancer,

même si elles ouvrent des pistes pour expliquer en quoi ils pourraient différer. Le scRNA-seq peut également illustrer la plasticité cellulaire et les transitions vers des états antérieurs, sous la pression thérapeutique. Par exemple, dans des lignées cellulaires de mélanome mutant BRAF V600E dérivées de patients après traitement par le vemurafenib (inhibiteur de BRAF), il a permis de révéler une transition vers un phénotype primitif de type crête neurale, avec une forte expression du récepteur du facteur de croissance neurale (NGFR) et une expression réduite du facteur de transcription des mélanocytes, MITF, après seulement trois jours de traitement (Su *et al.*, 2017). Cette transition pourrait alors être la marque d'une résistance aux médicaments.

#### **1.5.3.1.2.2. Classification tumorale**

Le Bulk RNA-seq a joué un rôle majeur dans la classification tumorale de nombreux cancers, y compris le cancer du sein, le mélanome, les tumeurs cérébrales et autres (Verhaak *et al.*, 2010; Koboldt *et al.*, 2012; Muzny *et al.*, 2012; Brennan *et al.*, 2013; Akbani *et al.*, 2015). Le scRNA-seq permet d'aller plus loin en révélant l'hétérogénéité au sein des sous-classes et en précisant le rôle des sous-types cellulaires dans la classification. L'analyse du glioblastome a montré que des cellules distinctes d'une même tumeur ressemblaient à différents types identifiés en bulk (les glioblastomes dits proneuraux, classiques et mésenchymateux) (Gill *et al.*, 2014; Patel *et al.*, 2014). Plus récemment, l'analyse scRNA-seq d'un nombre beaucoup plus important de cellules et d'échantillons a permis de relier les sous-types aux abondances relatives des compartiments malin et immunitaire (Neftel *et al.*, 2019). Dans le cas du médulloblastome, la technique permet de distinguer les sous-types de tumeurs en fonction de l'état de différenciation des cellules malignes qui les composent (Hovestadt *et al.*, 2019; Gojo *et al.*, 2020). Dans le cas du cancer du poumon à petites cellules, cela a permis d'identifier la coexistence de plusieurs sous-types, préalablement définis en bulk RNA-seq, dans des cellules distinctes au sein de tumeurs individuelles (Ireland *et al.*, 2020). Dans l'ensemble, ces études et d'autres soulignent l'impact important de l'analyse scRNA-seq sur la définition et la compréhension de la classification des tumeurs (Castro, Tirosh and Suvà, 2021).

#### **1.5.3.1.2.3. Microenvironnement tumoral**

Le scRNA-seq permet également de disséquer les composants du microenvironnement tumoral (TME). Dans le cas du compartiment immunitaire, des études ont identifié des états distincts des cellules T CD8+ associés à la régression ou à la progression de la tumeur chez des patients atteints de mélanomes (Sade-Feldman *et al.*, 2018) ou encore une signature

transcriptomique d'exclusion des cellules T dans les cellules malignes, prédisant la réponse au traitement par les inhibiteurs de points de contrôle immunitaire (Jerby-Arnon *et al.*, 2018). Parmi d'autres exemples se trouvent le cancer du poumon non à petites cellules, dans lequel les études scRNA-seq ont identifié des populations de cellules T pré-fatiguées, associées à un meilleur pronostic (Guo *et al.*, 2018) et le cancer du foie, où l'expression accrue de LAYN est associée à l'épuisement des cellules T cytotoxiques et à un plus mauvais pronostic (C. Zheng *et al.*, 2017). Dans le cancer colorectal, scRNA-seq a mis en évidence des sous-ensembles de macrophages tumoraux avec des profils transcriptionnels inflammatoires et pro-angiogéniques distincts (L. Zhang *et al.*, 2020).

#### **1.5.3.1.2.4. Processus métastatique**

Le scRNA-seq permet au moins trois approches distinctes pour améliorer notre compréhension des métastases : 1 – la caractérisation des états cellulaires facilitant les métastases, 2 – la comparaison d'échantillons primaires et métastatiques appariés, et 3- l'analyse des cellules tumorales circulantes.

La transition épithéliale-mésenchymateuse (EMT) est le mécanisme le plus associé à la génération de métastases, bien que cela soit débattu car il est principalement étudié sur des modèles établis et difficile à caractériser dans les échantillons primaires. Le scRNA-seq permet désormais contourner cette difficulté. Dans le cancer de la tête et du cou a été identifié un programme de type EMT au niveau du bord invasif de la tumeur, qui était modulé par les CAF et enrichi chez les patients présentant des métastases (Puram *et al.*, 2017). De façon remarquable, ce programme ne présente pas certaines caractéristiques du programme EMT traditionnel telles que l'expression de facteurs de transcription types et la perte des marqueurs épithéliaux. Par ailleurs, il n'a été mesuré que dans deux des quatre sous-types de HNSCC. Ce principe d'une EMT contexte-spécifique est soutenu par d'autres études récentes et peut aider à comprendre certaines observations contradictoires. Cela souligne l'importance de définir les différents états cellulaires dans les tumeurs humaines.

La caractérisation scRNA-seq des cellules tumorales circulantes (CTCs) ouvre une fenêtre sur les premiers stades du processus métastatique. De façon cohérente avec la suspicion de leur rôle dans l'établissement des métastases, les CTCs expriment généralement des programmes EMT (Ting *et al.*, 2014; Cheng *et al.*, 2019). Dans un autre exemple, l'analyse scRNA-seq des CTCs du cancer du sein a montré une abondance de CTCs regroupées, avec un potentiel métastatique plus élevé que les CTCs individuels (Aceto *et al.*, 2014). Leur

caractérisation a révélé une expression accrue de la plakoglobine, une protéine de jonction cellulaire, qui maintient les grappes de cellules et semble ainsi augmenter le taux de métastases.

Enfin, en comparant par scRNA-seq des échantillons primaires et métastatiques appariés, il est possible d'identifier et caractériser les différences dans la composition et l'état des types cellulaires, permettant de mieux appréhender l'adaptation au site envahi (Navin *et al.*, 2011; Aceto *et al.*, 2014; Puram *et al.*, 2017; Su *et al.*, 2019; Chi *et al.*, 2020; Rao *et al.*, 2020). Par exemple, une étude sur les métastases leptoméningées a dressé le profil des cellules malignes dans le liquide céphalo-rachidien (LCR) et a identifié une régulation positive de LCN2 et d'autres gènes qui permettent de récupérer le fer des cellules inflammatoires, ce qui favorise leur survie dans l'environnement pauvre en fer du LCR (Chi *et al.*, 2020).

#### **1.5.3.1.2.5. Réponse aux thérapies**

Il est admis que la réponse au traitement reflète une combinaison de sélection clonale et d'adaptation cellulaire qui influence à la fois les cellules malignes et les cellules non malignes dans l'EMT. Lorsqu'ils sont étudiés en *bulk* après traitement, ces effets sont difficiles à distinguer, mais la résolution de la cellule unique permet de déchiffrer ces effets. Dans le cas du mélanome muté pour BRAF, cela par exemple permis d'identifier des cellules résistantes et d'établir qu'elles étaient déjà présentes les échantillons avant traitement (Tirosh *et al.*, 2016). D'autres études *in vitro* sur cellules uniques ont montré que ce programme de résistance est accru lors du traitement, non seulement en raison d'une sélection clonale, mais aussi en raison d'une reprogrammation épigénétique induite par le médicament, ce qui pourrait contribuer à expliquer la résistance généralisée à l'inhibition de BRAF dans le mélanome (Shaffer *et al.*, 2017). Dans le cancer du poumon non à petites cellules, il a été démontré que les inhibiteurs de tyrosine kinase induisent l'expression d'une signature cellulaire alvéolaire-régénérative dans les cellules tumorales résiduelles, soit un changement d'état cellulaire induit par le traitement (Maynard *et al.*, 2020). Dans le cancer du sein ER+/HER2-, l'analyse scRNA-seq a démontré une augmentation de l'expression de HER2 après plusieurs traitements, avec un phénotype moins sensible à la thérapie ciblée que dans le cancer du sein HER2+ primaire (Jordan *et al.*, 2016). Dans les CTCs du cancer de la prostate, le scRNA-seq a révélé le rôle de l'activation de la voie de signalisation Wnt non canonique dans la résistance à aux thérapies anti-androgéniques (Miyamoto *et al.*, 2015).

Dans l'ensemble, ces études et d'autres mettent en évidence le potentiel d'adaptation des cellules malignes sous pression thérapeutique. Cela laisse penser que les études scRNA-seq d'échantillons post-traitement sont amenées à se développer considérablement dans les prochaines années, voire qu'elles seront intégrées aux essais cliniques afin de comprendre les réponses cellulaires aux traitements et caractériser les cellules malignes résiduelles (Suvà and Tirosh, 2019). Cependant, dans de nombreux contextes, les cellules résiduelles issues de traitements partiellement réussis ne sont pas détectées. Ces cellules peuvent rester dormantes pendant un certain temps jusqu'à ce qu'elles reprennent leur croissance et donnent lieu à une rechute tumorale. Or, la tumeur résultante aura alors subi de nouvelles modifications génétiques et épigénétiques, qui masquent les mécanismes initiaux de résistance. Il est donc difficile de savoir dans quelle mesure les mécanismes de résistance peuvent être décelés à partir de l'analyse des tumeurs récurrentes. Cela incite à davantage caractériser les cellules résiduelles. Ceci peut être réalisé dans les hémopathies malignes (Ediriwickrema *et al.*, 2020) ou par l'isolement des CTCs après les traitements. C'est également possible avec dans les modèles animaux (Lawson *et al.*, 2015; Rambow *et al.*, 2018). Cette approche a été utilisée initialement pour évaluer les cellules résiduelles dans un modèle de mélanome après une réponse à l'inhibition de BRAF (Rambow *et al.*, 2018). La caractérisation approfondie de ces cellules après traitements dans les modèles animaux devrait en général faire progresser notre compréhension des mécanismes de résistance.

### **1.5.3.2. scRNA-seq dans le ccRCC**

Dans le ccRCC, plus d'une dizaine d'études de scRNA-seq a été publiée depuis 2016. Ces études ont permis d'aborder des questions variées telles que l'origine cellulaire du ccRCC (Young *et al.*, 2018), la caractérisation de cellules souches cancéreuses et d'organoïdes 3D (Fendler *et al.*, 2020), la caractérisation en profondeur des tumeurs de Wilm's chez l'enfant (Young *et al.*, 2018, 2021), ainsi que la détermination de nouveaux biomarqueurs plus résolutifs (Wu *et al.*, 2018; Su *et al.*, 2021). Elles ont aussi servi à valider des analyses de cytométrie de masse (CyTOF) dans le compartiment immunitaire, une autre technique *single-cell* haut-débit (Argüello *et al.*, 2020). Notamment, nombre d'études se sont attaché à établir des nouveaux marqueurs pronostics, et beaucoup ont généré d'énormes jeux de données sur le compartiment immunitaire du TME, permettant déjà de bien le décrire (Bi *et al.*, 2021; Braun *et al.*, 2021; Krishna *et al.*, 2021).

Kim et ses collaborateurs ont publié en 2016 une étude pionnière portant sur la tumeur primaire et la métastase pulmonaire d'un patient, qu'ils ont implantées chez des souris (modèle PDX) (Kim *et al.*, 2016). Avec le scRNA-seq, ils ont mis en évidence plusieurs clones et ont révélé des activations contrastées de diverses voies métaboliques selon les sites considérés. En s'appuyant sur ces résultats, ils ont pu déterminer une combinaison thérapeutique qui s'est avérée efficace pour stopper la prolifération d'une lignée primaire issue de la tumeur, illustrant ainsi l'application possible du scRNA-seq en médecine personnalisée. Cependant, leurs travaux laissent entrevoir quelques points d'amélioration. Par exemple, le recours à des souris immunodéprimées et les passages successifs de la tumeur pour sa croissance ont altéré sa composition cellulaire initiale, à la fois en ce qui concerne la tumeur et le microenvironnement. Or, ces composantes sont essentielles à la compréhension des mécanismes tumoraux, y compris l'échappement aux traitements. Malgré ces limitations et le faible nombre de cellules, ils ont cependant fourni la première preuve de l'HIT à l'échelle de la cellule unique dans le ccRCC. En 2018, Young et ses collaborateurs se sont intéressés à l'origine des tumeurs et ont publié dans *Science* la première étude scRNA-seq de grande ampleur (Young *et al.*, 2018). En utilisant la technologie 10X Genomics, ils ont séquencé les transcriptomes de plus de 72 000 cellules issues de reins normaux adultes et fœtaux ainsi que de tumeurs de Wilm's chez l'enfant ou de ccRCC chez l'adulte. En étudiant la similarité transcriptomique entre les types cellulaires, ils ont montré que les tumeurs de Wilm's sont issues de cellules fœtales aberrantes. En procédant de même dans le ccRCC, ils ont mis à jour trois populations de cellules tubulaires proximales dans le rein normal et avancent que l'une d'entre elles est à l'origine du ccRCC. Ils révèlent par ailleurs avec une résolution inédite les principales populations cellulaires présentes dans le ccRCC et de nouveaux marqueurs géniques pour chacune d'entre elles ; des cellules endothéliales, des fibroblastes et myofibroblastes, des cellules immunitaires et les cellules malignes.

#### **1.5.3.2.1. Caractérisation du compartiment immunitaire**

Plusieurs études scRNA-seq dans le ccRCC ont par la suite permis de caractériser le compartiment immunitaire (Bi *et al.*, 2021; Borcherdig *et al.*, 2021; Braun *et al.*, 2021; Krishna *et al.*, 2021). Les cellules immunitaires sont souvent majoritaires dans la suspension cellulaire après dissociation d'un ccRCC. Ceci est dû à un biais de sélection, ces cellules étant déjà individualisées et plus résistantes. Par ailleurs, le ccRCC est particulièrement vascularisé, ce qui augmente le contingent immunitaire, à la fois infiltrant la tumeur et dans le sang



périphérique. Ainsi le TME immunitaire avant ou après immunothérapies est déjà bien caractérisé. Borchering et ses collaborateurs ont séquencé l'ARN et les récepteurs de lymphocytes T (« *T Cell Receptors* » ou TCR) de plus de 25 000 cellules lymphoïdes et myéloïdes et démontré qu'il y a plus de lymphocytes T CD8+ et de macrophages parmi les cellules infiltrées dans le ccRCC par rapport au tissu rénal normal (Borchering *et al.*, 2021). Par ailleurs, l'étude démontre des différences transcriptomiques au sein de ces populations, indiquant une hétérogénéité du TME immunitaire. En s'appuyant sur la base de données du TCGA-KIRC, ils ont montré que les gènes spécifiques des populations TAM et CD8+ de leur études sont associés à un mauvais pronostic. Dans une étude parue peu après, une équipe a caractérisé le TME immunitaire avec ou sans immunothérapies chez des patients différents (Bi *et al.*, 2021). De façon cohérente, ils retrouvent majoritairement des lymphocytes T cytotoxiques (LTC) et des TAM. Ils démontrent que les répondeurs ont des LTC avec un niveau accru de récepteurs co-inhibiteurs et de molécules effectrices. Par ailleurs, les TAM évoluent vers un phénotype pro-inflammatoire dans un contexte riche en interféron à la suite du traitement. Une seconde étude a également caractérisé plus de 160 000 cellules du compartiment immunitaire avant et après immunothérapies (Krishna *et al.*, 2021). Les auteurs détectent une grande hétérogénéité inter- et intra-tumorale et un enrichissement des lymphocytes T CD8+ chez les patients répondeurs, tandis que les TAMs sont enrichis chez les patients résistants. Enfin, une autre étude incontournable a caractérisé plus de 160 000 cellules immunitaires du TME dans le ccRCC (Braun *et al.*, 2021), et démontré que les lymphocytes T CD8+ épuisés sont enrichis dans les métastases et présentent peu de diversité, en accord avec les résultats de l'étude de Borchering. Ils démontrent par ailleurs que les macrophages de type M2 augmentent dans les cancers avancés.

#### **1.5.3.2.2. Prédiction du pronostic clinique**

La transcriptomique permet depuis plusieurs années d'identifier des gènes dits biomarqueurs, associés à la survie des patients dans différentes pathologies cancéreuses. Le scRNA-seq porte ainsi la promesse d'obtenir des signatures plus précises car spécifiques de certaines populations cellulaires, immunitaires ou malines par exemple. Dans le ccRCC, la première étude du genre, parue en 2019, utilise les données de Kim *et al* pour établir une signature métastatique de 44 gènes en s'appuyant sur les profils transcriptomiques de 121 cellules (C. Zhang *et al.*, 2019). Les auteurs prouvent que la signature identifiée est associée à la survie globale et est corrélée à différentes caractéristiques cliniques. Ils établissent alors le

premier nomogramme prédictif basé sur des données de single-cell RNA-seq. Par la suite, d'autres études construisent des nomogrammes prédictifs incluant des gènes associés à la différenciation (Xia *et al.*, 2022) ou à certaines sous-populations cancéreuses (Zhang *et al.*, 2021), identifiés à partir de données scRNA-seq. Bien que les signatures soient susceptibles d'être plus spécifiques de populations d'intérêt pronostique grâce aux approches scRNA-seq, elles s'appuient toujours le gène comme facteur, et restent similaires à celles employées en bulk RNA-seq. Une seconde approche consiste à utiliser une technique de déconvolution dans des cohortes massives de données bulk RNA-seq et à analyser l'association des sous-populations identifiées avec la survie. Dans ce cas, l'échelle n'est plus génique mais populationnelle et prend en compte un ensemble de gènes. Dans leur étude, Hu et ses collaborateurs ont employé le scRNA-seq et caractérisé finement des populations de lymphocytes CD8 (Hu *et al.*, 2020). Les auteurs ont utilisé ensuite l'outil CIBERSORT (B. Chen *et al.*, 2018) pour déconvoluer les données du TCGA-KIRC à partir des profils transcriptomiques de leurs trois populations de lymphocytes T CD8, et ont déterminé trois groupes de patients parmi lesquels un groupe montrant une faible proportion de T CD8 activés et une importante proportion de T CD8 fatigués, significativement associé à un mauvais pronostic. Une autre étude a utilisé le même algorithme pour corréler les résultats de leur analyse de l'immunothérapie à l'échelle single-cell avec les types et proportions cellulaires déconvoluées dans la cohorte Checkmate 009, dans laquelle les patients ont été traités par immunothérapies (Bi *et al.*, 2021). Cependant, les auteurs estiment que la cohorte, comprenant sept ccRCC, n'est pas assez conséquente pour tirer des conclusions.

### 1.5.3.3. Approches émergentes en oncologie

Cette section évoque succinctement quelques voies émergentes d'études des tumeurs à l'échelle de la cellule unique. En effet, les développements technologiques récents, par exemple multi-omiques, offrent des voies nouvelles pour aborder des questions d'intérêt biologique et clinique en oncologie (Figure 45).

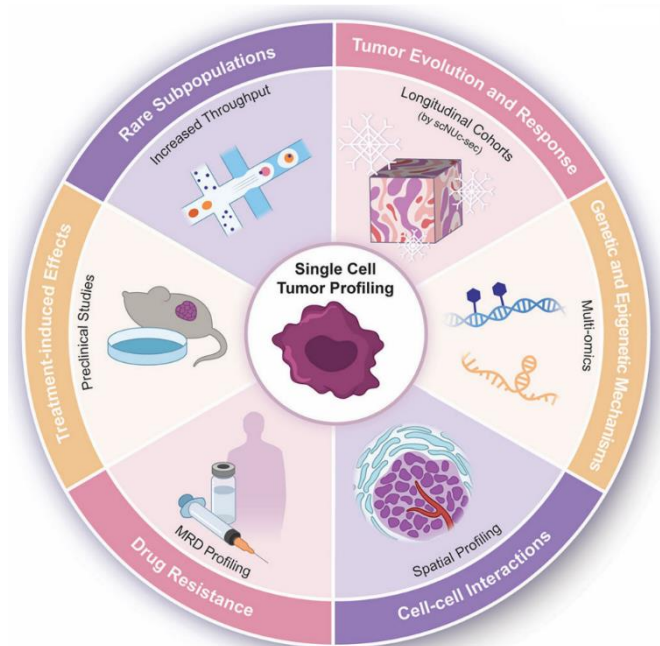


Figure 45. Approches single-cell récentes et questions biologiques et cliniques émergentes en oncologie

Les développements récents des approches –omiques à l'échelle de la cellule unique permettront d'aborder des questions variées en oncologie. Figure tirée de Castro, Tirosh and Suvà, 2021.

L'augmentation du débit des technologies de scRNA-seq permettra bientôt de caractériser des dizaines de milliers de cellules en routine (Picelli, 2017; Baran-Gale, Chandra and Kirschner, 2018; See *et al.*, 2018) et permettra la découverte de sous-populations tumorales rare mais cliniquement significatives, notamment celles qui présentent des adaptations liées à la résistance aux médicaments ou la progression métastatique. La possibilité de caractériser des échantillons congelés par snRNA-seq (Slyper *et al.*, 2020) permet d'ores et déjà de caractériser des échantillons congelés stockés dans des biobanques, ce qui facilite et appelle à très court terme des collaborations entre centres et la caractérisation de tumeurs rares ou de paies d'échantillons primaires et de rechutes tumorale, qui éclaireront les processus d'évolution tumorale. Les expériences qui évaluent simultanément les états transcriptionnels et épigénomiques (méthylation de l'ADN ou accessibilité de la chromatine) dans des cellules individuelles (Xu *et al.*, 2019; Nam, Chaligne and Landau, 2021) permettent de mieux

comprendre les mécanismes qui régissent la diversité des états et les transitions dans les tumeurs. La transcriptomique spatiale de coupes de tissus (Ståhl *et al.*, 2016; Rodriques *et al.*, 2019) (106,107), permet de mieux caractériser les interactions au sein du microenvironnement tumoral (Berglund *et al.*, 2018; Thrane *et al.*, 2018; Moncada *et al.*, 2020) et devrait largement se développer à l'avenir, en lien avec l'expertise histopathologie des cliniciens. Toutes ces études marquent le jeune âge de l'ère du profilage de cellules uniques dans les tumeurs humaines, dont la résolution, l'échelle et la portée progressent rapidement. Bien que la complexité nouvelle qu'elles laissent entrevoir représente un défi conséquent, la compréhension plus profonde acquise grâce à ces approches offre des opportunités sans précédent en vue d'améliorer de la prise en charge des patients atteints de cancer.



# **2.CONTEXTE ET OBJECTIFS**



## 2.1. CONTEXTE DU LABORATOIRE

Mon travail de thèse a été réalisé au sein de l'équipe « Physiologie et physiopathologie du tractus urogénital » (URGENT) à l'Institut de Recherche en Santé, Environnement et Travail (IRSET) de Rennes (UMR\_S 1085) ; qui comprend des volets virologique, toxicologique et physio-génomique. L'équipe a été renforcée en 2017 par l'arrivée de cliniciens du CHU de Pontchaillou, spécialisés dans les cancers urologiques et notamment le cancer du rein. C'est dans ce contexte qu'ont émergé des projets visant à répondre à plusieurs questions relatives aux carcinomes rénaux à cellules claires (ccRCC) par des approches de transcriptomique sur cellules uniques, développées dans l'équipe par Frédéric Chalmel, Nathalie Rioux-Leclercq et Aurélie Lardenois.

## 2.2. CONTEXTE SCIENTIFIQUE

L'efficacité des thérapies ciblées utilisées pour le traitement du ccRCC est mitigée, avec des patients d'emblée non répondeurs, des hétérogénéités de réponse chez un même patient, ou encore un échappement tumoral fréquent à ces traitements (Motzer *et al.*, 2007; Molina *et al.*, 2014; Crusz *et al.*, 2016). Ces écueils sont en partie imputables à l'hétérogénéité, non seulement inter, mais aussi intra-tumorale (HIT), de plus en plus décrite comme jouant un rôle crucial dans la biologie tumorale, la réponse aux thérapies anti-cancéreuses et la survie des patients (Gerlinger *et al.*, 2012; McGranahan and Swanton, 2015; López and Angulo, 2018). La technologie de scRNA-seq s'est récemment imposée comme la technique de choix pour déchiffrer l'HIT. Elle permet de mesurer l'expression génique à l'échelle de la cellule individuelle et donc d'identifier et caractériser fonctionnellement les différentes sous-populations d'un tissu biologique. Parallèlement, ces dernières années, des progrès significatifs ont été réalisés dans le domaine des modèles tumoraux. En particulier, la culture organotypique de coupes tumorales constitue un modèle très attractif pour étudier l'impact de traitements anti-cancéreux car il est peu coûteux, facile à mettre en œuvre, et permet de conserver à la fois la morphologie, la composition cellulaire tumorale et le microenvironnement d'origine de la tumeur.



## 2.3. HYPOTHESE DE TRAVAIL

Dans ce contexte, l'hypothèse principale de mon travail de thèse postule que l'hétérogénéité intra-tumorale (HIT) influence l'issue de la maladie et la réponse d'une tumeur aux thérapies ciblées. Ma thèse ambitionne d'apporter des éléments de réponses aux questions scientifiques suivantes : Comment la physiologie et les phénotypes d'une tumeur cancéreuse, telles que ses capacités de prolifération ou de résistance aux traitements anti-cancéreux, sont-ils impactés par l'HIT ? D'un point de vue appliqué, est-il possible de quantifier l'impact des traitements anti-cancéreux sur cette HIT ? Quelles sont les populations cellulaires sensibles ou résistantes aux traitements anti-cancéreux ? De façon plus transversale, peut-on tirer parti d'une meilleure connaissance de l'hétérogénéité tumorale pour améliorer la prise en charge thérapeutique des patients ?

## 2.4. OBJECTIFS DE MES TRAVAUX DE THESE DE DOCTORAT

L'objectif général consiste à développer des approches et donner des clefs de compréhension permettant tester les hypothèses de travail. La thèse, intitulée « Étude de l'hétérogénéité intra-tumorale et de l'impact de thérapies ciblées sur le carcinome rénal à cellules claires par une approche transcriptomique sur cellules uniques », dégage deux axes d'études principaux.

Le **premier axe**, fondamental et descriptif, consistait à établir un atlas décrivant l'hétérogénéité intra-tumorale du ccRCC à l'échelle de la cellule unique. Les objectifs étaient notamment de décrire les profils transcriptomiques des cellules malignes, étant donnée leur contribution à l'hétérogénéité intra-tumorale, et d'évaluer leur potentiel dans l'amélioration du pronostic des patients. Le **deuxième axe**, davantage appliqué et transversal, visait à développer une stratégie innovante permettant d'évaluer par scRNA-seq l'impact de traitements anti-cancéreux sur des modèles de culture 3D de ccRCC. Le premier objectif était d'apporter des preuves de faisabilité, de point de vue expérimental et analytique. Ensuite, les objectifs étaient de parvenir à identifier, caractériser et comparer les profils transcriptomiques de populations cellulaires résistantes ou sensibles aux traitements. Alors que ces deux premiers axes s'articulaient pour répondre à la forte demande de recherche translationnelle, un **troisième axe** a pris forme tôt dans cette thèse afin de valoriser nos recherches : l'élaboration d'un site web

interactif mettant à disposition des chercheurs et cliniciens les données et analyses de scRNA-seq dans le domaine de l'uro-oncologie. Son objectif est fonctionnel ; il s'agit de supporter l'effort de recherche translationnelle et de permettre à un public non-expert de visualiser et de manipuler ce type de données afin d'exploiter leur riche potentiel au bénéfice des patients.

Les trois axes de thèse décrits précédemment sont déclinés en trois chapitres de résultats dans ce manuscrit de thèse.



# **3. MATÉRIEL ET METHODES**



### **3.1. PATIENTS**

Les échantillons primaires traités dans ce travail de thèse ont été obtenus via le circuit de traitement des échantillons biologiques du Centre de Ressources Biologiques (CRB) de Rennes (BB-0033-00056) suite à une néphrectomie réalisée au service d'urologie du CHU de Rennes. Le protocole de recherche répond aux exigences du comité d'éthique local. Les patients subissant une néphrectomie ont donné leur consentement quant à l'utilisation de la pièce opératoire à des fins de recherche. Les caractéristiques et informations cliniques et histopathologiques ont été obtenues auprès du service d'anatomopathologie du CHU de Rennes.

### **3.2. MODELE DE CULTURE ORGANOTYPIQUE DE COUPES DE CCRCC**

Dès l'arrivée des pièces opératoires dans le service d'anatomopathologie, des échantillons de tumeur sont découpés manuellement à l'aide d'un scalpel stérile et deux zones macroscopiquement distinctes et représentatives sont échantillonnées et stockées dans du HBSS stérile à 4°C. Les deux échantillons sont immergés côte à côte dans une solution chaude (liquide) d'agarose à 2%, au sein d'un puits en plaque 6-puits, et immédiatement placés sur la glace pendant 5 minutes. Le bloc d'agarose solidifié contenant le tissu est retiré du puits et préparé pour être coupé à l'aide du Vibratome VT1200 (Leica Microsystems, Wetzlar, Allemagne) dans des conditions stériles. L'épaisseur des tranches de tissu est fixée à 400 µm. L'amplitude de vibration est réglée entre 2,95 mm. La vitesse de tranchage, comprise entre 0,04 et 0,1 mm/s, est optimisée en fonction de la densité et du type de tissu ; une vitesse de tranchage plus lente est utilisée sur des tissus plus mous et vice versa. Les tranches sont déposées sur des inserts (Millicell Cell Culture Insert, Réf : PICMORG50) pendant 48h. Plusieurs tranches de chaque échantillon ont été placées dans chaque insert afin d'occuper le maximum de surface sur l'insert et d'optimiser la quantité de tissu traité. Les inserts organotypiques sont des membranes en téflon avec des pores de 0,4µm qui permettent des échanges régulés de gaz et de nutriments entre le tissu et le milieu ainsi que la préservation de la structure tissulaire 3D en culture. La culture des tissus est réalisée à 37°C dans un incubateur humidifié à 5% de CO<sub>2</sub>, en utilisant 3ml de milieu DMEM complété par 20% de FBS inactivé (GIBCO), 100 U/ml de pénicilline (Invitrogen) et placé dans un agitateur à faible vitesse pour faciliter les échanges de gaz et de fluides avec le milieu.

### 3.3. TRAITEMENTS

Les coupes sont incubées avec les molécules pendant 48h. Les molécules sélectionnées sont présentées dans le Tableau 8. Le sunitinib et le pazopanib sont des inhibiteurs multi-cibles de récepteurs à tyrosine kinase ; le temsirolimus, le KU-60019 et le silmitasertib sont des inhibiteurs de tyrosine kinase ciblant respectivement mTOR, ATM et CK2, qui sont impliqués dans diverses voies de signalisation à l'œuvre dans le ccRCC (Figure 9) (Roelants *et al.*, 2018b). Enfin, le pictilisib et le saracatinib sont des molécules inhibitrices multi-cibles, impactant respectivement des protéines de la famille PI3K et la protéine kinase SRC, dont les implications dans la biologie du ccRCC ont été démontrées également (Giacosa *et al.*, 2021). Conformément à cette étude, les molécules ont une concentration de 10 $\mu$ M dans le milieu de culture (Tableau 8).

Stratégie	Molécule	Code	Cibles principales	Autres cibles	Voies de signalisation	Concentration dans le milieu
monothérapie	Sunitinib (SU11248)	S	PDGFRB	FLT3, IRE1 $\alpha$ , Kit	Voie PDGFR	10 $\mu$ M
monothérapie	Pazopanib	P	VEGFR1, VEGFR2, VEGFR3	c-Kit, PDGFR, FGFR	Voie VEGFR	10 $\mu$ M
monothérapie	Temsirolimus (CCI-779)	T	mTOR		Voie mTOR	10 $\mu$ M
combinaison	KU-60019	KU	ATM		Voie ATM/ATR	10 $\mu$ M
	Silmitasertib (CX-4945)	CX	CK2		Variées	5 $\mu$ M
combinaison	Pictilisib (GDC-0941)	GDC	p110 $\alpha$ , p110 $\beta$ , p110 $\delta$ , p110 $\gamma$	mTOR	Voie PI3K	10 $\mu$ M
	SARACATINIB (AZD0530)	SRC	Src, Lck, Fyn, Lyn	EGFR (L861Q), c-YES, EGFR (L858R)	Voie Src	10 $\mu$ M

Tableau 8. Caractéristiques des molécules utilisées dans le projet de thèse lié aux impacts des traitements

Les caractéristiques présentées sont en partie issues du profil des molécules sur le site sellectchem.com.

## 3.4. BIOLOGIE CELLULAIRE

### 3.4.1. DISSOCIATION TISSULAIRE

Dans cette thèse, plus de 60 prélèvements de tumeurs ont été utilisés et dissociés. Les deux projets expérimentaux de ce travail de thèse ont requis des adaptations du protocole de dissociation. Alors que la dissociation enzymatique et mécanique dans l'axe 1 (atlas de cellules uniques dans le ccRCC) suivait les instructions de Miltenyi Biotec (Bertrand *et al.*, 2020) ; la dissociation de coupes tumorales dans le projet d'étude de l'impact des traitements, a nécessité une dissociation plus douce et manuelle afin de ménager les coupes tumorales plus fragiles. Une lyse des globules rouges a été réalisée dans tous les protocoles. Deux centrifugations à 800 tour/minute ont été réalisées dans les dissociations du projet 1 afin d'éliminer les débris et les cellules mortes, tandis qu'une seule a été réalisée pour les dissociations de coupes tumorales, afin de ne pas perdre trop de cellules. Une déplétion des cellules positives pour le marqueur CD45 (cellules immunitaires) est réalisée à l'aide du kit Miltenyi (CD45 MicroBeads, humain) pour enrichir les échantillons en cellules malignes et stromales. Dans l'axe 2 (étude des traitements), les cellules mortes ont été éliminées avec le kit Dead cell removal de chez Miltenyi. La viabilité cellulaire a été estimée au microscope (systèmes Leica, Olympus IX70) en utilisant du bleu Trypan avant de procéder à la capture ou à la congélation.

### 3.4.2. CONGELATION ET DECONGELATION DE CELLULES

Des cellules de ccRCCs dissociés ont été congelées dans le projet d'atlas afin de d'observer l'impact de la congélation sur leur transcriptome. Les cellules de nombreuses tumeurs ont également été dissociées, stockées et congelées en vue de potentiels projets ultérieurs. Près de 300 000 cellules sont transférées dans 1ml de solution de congélation composée de 70% de RPMI 1640, 20% de sérum fœtal bovin (FBS), 10% de DMSO (Sigma) et immédiatement placées à -20°C pendant 1h avant que leur température ne soit progressivement abaissée (-1°C/min) jusqu'à -80°C. Les cellules sont ensuite transférées dans l'azote liquide pour un stockage à long terme. Pour la décongélation, les échantillons congelés sont progressivement décongelés à la main (à 37°C) et immédiatement dilués dans 14 ml de RPMI 1640 dans un tube Falcon de 15 ml.

### 3.4.3. MULTIPLEXAGE PAR CELL-HASHING

Dans le projet d'étude de l'impact des traitements, les cellules ont été marquées par Cell-hashing (voir la section **1.4.3.1.3.2. CITE-seq**) selon les recommandations du protocole de Biolegend (TotalSeq™-A Antibodies and Cell Hashing with 10x Single Cell 3' Reagent Kit v3



or v3.1 (Single Index) Protocol). Des anticorps dits « *hashtag* », dirigés contre les épitopes ubiquitaires B2M et CD298, ont été utilisés pour marquer tous les types cellulaires. Chaque lot d'anticorps est associé à un index de séquençage, qui permet d'identifier lors de l'analyse des données l'échantillon biologique d'origine de chaque cellule. Les cellules sont ensuite regroupées en proportions égales avant de procéder à leur capture.

#### 3.4.4. CYTOMETRIE

Pour réaliser les mesures de cytométrie en flux, les cellules ont été co-marquées avec des anticorps monoclonaux anti-CD45-FITC/CD34-PE humains et 7-AAD (BD Biosciences, San Jose, CA) pendant 15 minutes à -4°C dans l'obscurité. Les cellules ont ensuite été lavées deux fois dans du PBS (« *Phosphate-Buffered Saline* »). Les mesures sont réalisées avec un cytomètre LSR Fortessa X-20 (Becton Dickinson) via le logiciel d'acquisition BD FACSDiva (v8). Les données sont ensuite analysées avec le logiciel FlowLogic (v7.2.1).

#### 3.4.5. CAPTURE DE CELLULES UNIQUES

Les cellules dissociées ont été capturées avec le dispositif *Chromium* de chez 10X Genomics, selon les protocoles de préparation de bibliothèques de séquençage. Lors de ce processus, les ARNs de chaque cellule sont marqués avec un code-barres commun (« *10X barcode* ») et un identifiant moléculaire unique (UMI), qui permettent respectivement d'identifier la cellule d'origine des transcrits et de s'affranchir de leur biais d'amplification (voir la section **1.4.3.1.2. Le Chromium de 10X Genomics**). Dans l'axe 1, la capture a été réalisée selon les recommandations du fabricant de sorte à obtenir entre 2500 et 5000 cellules par échantillon. Ce choix de fourchette résulte d'un compromis entre la volonté d'obtenir suffisamment d'événements ; le nombre de doublets (co-capture de plusieurs cellules au lieu d'une) générés mécaniquement par une capture trop importante ; et le coût de séquençage (qui augmente avec le nombre de cellules, à raison de ~20,000 *reads* recommandés par cellule). Dans l'axe 1, la capture a été réalisée de sorte à obtenir près de 24,000 cellules pour chaque tumeur, soit environ 4000 cellules par condition de traitement (au nombre de six par expérimentation). Ce choix s'expliquait par des coûts de préparation de la bibliothèque très réduits grâce à l'emploi du cell-hashing et au recours à un seul kit de préparation.

## 3.5. BIOLOGIE MOLÉCULAIRE

### 3.5.1. MARQUAGE IMMUNO-HISTOCHIMIQUE

Pour les examens immuno-histochimiques, les échantillons de ccRCC ont été fixés pendant 24h dans du paraformaldéhyde-PBS à 4% et inclus en paraffine selon les protocoles standards. Un marquage à l'hématoxyline et à l'éosine (H&S) est réalisé sur les lames, et leur examen et photographies sont réalisés sur le logiciel CaseViewer v2.2.

### 3.5.2. LIBRAIRIES DE SEQUENÇAGE

Les librairies de séquençage ont été réalisées de façons différentes selon les projets de ce travail de thèse. Dans l'axe 1, les étapes ont été réalisées conformément aux instructions du fabricant (Chromium™ Single Cell 3' Library & Gel Bead Kit v2). Pour le projet visant à étudier les impacts de traitements, la préparation de la librairie incluait des étapes supplémentaires et des modifications liés à l'utilisation du cell hashing. Les étapes ont été réalisées conformément à un second protocole fourni par Biolegend (TotalSeq™-A Antibodies and Cell Hashing with 10x Single Cell 3' Reagent Kit v3 or v3.1 (Single Index) Protocol). La qualité des ADNc des librairies *single-cell* a été contrôlée à l'aide d'un bioanalyseur (Agilent Technologies). Les librairies ont été séquencées sur un instrument Illumina HiSeq 4000 pour le premier projet, et sur un instrument Illumina NovaSeq compatible avec la technologie 10X pour le second projet, avec dans les deux cas e recours à un séquençage paillé (2x100) (Illumina, San Diego, CA).

## 3.6. TRAITEMENT BIO-INFORMATIQUE

Les approches d'analyse bio-informatiques sont transversales entre les projets. En particulier, les étapes de prétraitement des données sont standards et généralement identiques d'un jeu de données à l'autre, à l'exception de mises à jour d'outils. Les données brutes de séquençages sont prises en charge avec le pipeline d'analyse CellRanger (10x Genomics Cell Ranger 3.1.0). La majorité des analyses a ensuite été réalisée dans un environnement R (Team, 2021), avec le langage de programmation du même nom. R est un langage dédié aux statistiques et au traitement des données (R core Team, 2021). Très utilisé en biologie, notamment en génomique, il possède une grande bibliothèque d'outils comme la suite Bioconductor, très utilisée en bio-informatique puisqu'elle fournit de nombreux outils dédiés à l'analyse de données des sciences de la vie (Huber et al. 2015). Dans les analyses de ce travail de thèse, des suites d'outils d'analyses intégrées tels quel Seurat (Butler *et al.*, 2018) (v3 & v4) ou Monocle3

(Trapnell and Cacchiarelli, 2014) ont été particulièrement utilisées pour le prétraitement et la manipulation des données de scRNA-seq.

### 3.6.1. PRETRAITEMENT, CONTROLES DE QUALITE, INTEGRATION ET NORMALISATION.

Les lectures de séquençage brutes (ou « *reads* ») sont démultiplexées, alignées sur le génome de référence humain GRCh38, puis quantifiées en comptant les UMIs uniques avec le pipeline Cell Ranger (v3.0.2, 10x Genomics). Les matrices de comptes brutes (« *raw count matrix* ») des échantillons individuels sont ensuite rassemblées en une matrice finale. L'algorithme d'identification des cellules de Cell Ranger, qui repose sur la méthode EmptyDrops (Lun *et al.*, 2019), permet d'écarter les cellules aberrantes et d'établir un premier ensemble de cellules valide. Par la suite, l'outil scater (v1.10.1) est utilisé dans l'environnement R pour éliminer les cellules aberrantes (ou « *outliers* »), en mesurant notamment les proportions de lectures correspondant aux gènes mitochondriaux et ribosomaux (McCarthy *et al.*, 2017). En parallèle, les événements correspondant en réalité à deux ou plus cellules (appelés doublets ou multiplets) ont été filtrés indépendamment dans chaque matrice individuelle à l'aide du paquet R DoubletFinder (v.2.0.2) (McGinnis, Murrow and Gartner, 2019). Les cellules ayant moins de 200 gènes détectés et les gènes détectés dans moins de 10 cellules ont été supprimées. Les phases du cycle cellulaire ont été attribuées aux cellules à l'aide du logiciel Scran R (Yip, Sham and Wang, 2018) (v1.10.2) et du logiciel Seurat v3.1. Les données ont été normalisées à l'aide de la fonction *sctransform* implémentée dans Seurat. Quand nécessaire, une correction des données pour les biais techniques (« *effet batch* ») a été réalisée en utilisant la procédure Seurat basée sur les fonctions *FindIntegrationAnchors* et *IntegrateData* avec les paramètres suivants : `anchor.features = 3000` et `dims = 30`.

### 3.6.2. REDUCTION DE LA DIMENSIONNALITE ET CLASSIFICATION

Dans les matrices d'expression, les 3000 gènes les plus variables ont été sélectionnés à l'aide de la fonction *FindVariableGenes* et utilisés pour effectuer une analyse en composantes principales (ACP) avec la fonction *RunPCA* implémentée dans Seurat. Les cellules ont ensuite été classifiées en groupes, ou « *clusters* », en utilisant les fonctions *FindNeighbors* et *FindClusters* sur les 30 composantes principales de l'ACP, avec des paramètres par défaut. Enfin, nous avons utilisé la méthode UMAP (« *uniform manifold approximation and projection* ») implémentée dans Seurat pour projeter les cellules individuelles dans un espace en deux dimensions.

### 3.6.3. ANNOTATION DES CLUSTERS ET IDENTIFICATION DES GENES DIFFERENTIELLEMENT EXPRIMÉS

Les *clusters* ont été annotés en type cellulaire à l'aide d'un ensemble de gènes marqueurs connus, complétés par des gènes qui étaient significativement différentiellement exprimés (« *differentially expressed genes* » ou DEGs) entre eux. Dans l'axe 1, les DEGs pour chaque *cluster* ont été identifiés avec la fonction *FindAllMarkers* implémentée dans Seurat, avec des paramètres par défaut. Dans une seconde partie de l'analyse, les DEGs entre les clusters de cellules malignes dans la tumeur RCC2 ont été identifiés en réalisant de multiples comparaisons des clusters deux à deux, avec la fonction *FindMarkers* implémentée dans Seurat. Dans l'axe 2, les gènes exprimés de manière différentielle ont été identifiés avec l'outil *presto* (<https://github.com/immunogenomics/presto>).

### 3.6.4. ANALYSES

#### 3.6.4.1. Analyse pseudo-bulk

Une matrice de comptage « *pseudobulk* », résultant de l'agrégation des valeurs d'expression de toutes les cellules du jeu de données, a été utilisée en première approximation pour calculer les corrélations (Pearson) entre échantillons. La matrice de comptage « *pseudobulk* » a également été utilisée pour effectuer une analyse en composantes principales basée sur l'outil FactoMineR (Lê, Josse and Husson, 2008).

#### 3.6.4.2. Analyse des variations du nombre de copies

L'outil R *inferCNV* a été utilisé pour inférer les variations du nombre de copies (« *Copy number variations* » ou CNV) dans les cellules malignes (*InferCNV* du projet Trinity CTAT : <https://github.com/broadinstitute/InferCNV>). L'inférence a été réalisée séparément sur chaque échantillon individuel, en utilisant les cellules épithéliales tubulaires du jeu de données comme référence. Le paramètre limite (« *cutoff* ») utilisé avait une valeur de 0,1, et les paramètres *denoise*, *cluster\_by\_groups* et *HMM* la valeur « TRUE ». L'outil renvoie une matrice avec un score CNV correspondant à l'intensité relative de l'expression génique pour une cytobande particulière, dans une cellule donnée.

Un score global de CNV (« *CNV Burden score* ») a été calculé pour chaque cellule et correspond à la moyenne pondérée des scores CNV absolus, en tenant compte de la taille de la cytobande, selon la formule suivante :

$$CNV \text{ Burden score} = \frac{\sum_{i=1}^n |C_i| \cdot S_i}{\sum_{i=1}^n S_i}$$

avec n étant le nombre de cytobandes, C le score CNV correspondant de la cytobande i, et S la taille de la cytobande i.

Une analyse différentielle du score CNV a été réalisée en comparant les scores CNV des cellules épithéliales tubulaires normales et des cellules malignes à l'aide d'un test de Wilcoxon. La p-value a ensuite été ajustée avec la méthode de Benjamin Hochberg et les scores CNV significatifs dans les cellules malignes ont été considérés pour la suite de l'analyse. Un arbre phylogénétique a été construit à partir des résultats de l'outil inferCNV en utilisant l'outil Uphyloplot2 (Kurtenbach *et al.*, 2021) avec le paramètre  $c = 1$ .

#### **3.6.4.3. Inférence de trajectoire de cellule unique et analyse de pseudo temps**

Les trajectoires dans les cellules malignes ont été inférées avec l'outil Monocle3 (Trapnell and Cacchiarelli, 2014) à partir des données scRNA-seq. Les cellules ont été partitionnées selon leur échantillon. Les pseudo-temps (« *pseudotime* ») ont été estimés à l'aide de la fonction *pseudotime* en indiquant comme cellules de départ celles présentant le moins d'aberrations chromosomiques, c'est-à-dire le score de CNV global le plus faible.

#### **3.6.4.4. Analyses fonctionnelles**

L'outil AMEN (Chalmel and Primig, 2008) a été utilisée pour identifier les termes significativement enrichis dans les ensembles de gènes, à partir d'un test hypergéométrique. Un terme spécifique est considéré comme enrichi dans un groupe de gènes co-exprimés si la *pvalue* ajustée (procédure de Benjamini-Hochberg) est inférieure ou égale à 0,05 et si le nombre de gènes portant une annotation donnée est supérieur ou égal à 3.

### **3.6.4.5. Traitement des données du TCGA-KIRC et analyses de survie**

#### **3.6.4.5.1. Données d'ARN bulk**

La matrice d'expression brute des comptages provenant du TCGA-KIRC (« *The Cancer Genome Atlas Kidney Clear Cell Carcinoma* ») et les données cliniques associées (stade de la tumeur, survie, dernier jour de suivi, etc.) ont été téléchargées à l'aide de l'outil R TGCABiolink et de la plateforme du Broad Institute TCGA-GDAC (<http://gdac.broadinstitute.org/>), respectivement. Les annotations pronostiques des gènes ont été extraites de la base de données TCGA-KIRC.

#### **3.6.4.5.2. Analyse de survie**

L'analyse de survie a été réalisée à l'aide de l'outil R Survival (Lin and Zelterman, 2002) (version 3.2-13). Pour l'analyse pronostique des gènes, les patients ont été classés dans les groupes d'expression faible ou élevée en fonction de la p-value la plus faible d'un test Log Rank. Comme défini par le HCA (« *Human Protein Atlas* »), une p-value < 0,001 reflète des différences significatives dans les résultats de survie entre les deux groupes, et définit le gène comme pronostique.

#### **3.6.4.5.3. Sélection d'ensembles de gènes favorables et défavorables**

Les gènes significativement favorables et défavorables pour la survie ont été définis en comparant la probabilité de survie à 2000 jours entre les deux groupes d'expression.

### **3.6.5. DEPOT DES DONNEES**

Les données de séquençage brutes sont disponibles sur le portail de l'EGA (European Genome-Phenome Archive) avec le numéro d'accès EGAS00001006534. Les données de scRNA-seq prétraités sont disponibles sur le portail ArrayExpress sous le numéro d'accès E-MTAB-11310. Toutes les données sont également disponibles via le portail Web UncoVer (The Uro-oncogenomics Viewer à l'adresse <https://uncover.genouest.org>).



# 4. RÉSULTATS





## 4.1. ATLAS DE L'HÉTÉROGÉNÉITE DANS LE CCRCC ET VALEUR PRONOSTIQUE

### 4.1.1. INTRODUCTION

L'hétérogénéité intra-tumorale (HIT) renvoie aux cellules malignes ainsi qu'aux cellules non malignes du microenvironnement (cellules immunitaires, stromales et endothéliales), qui toutes ont des rôles dans le comportement de la tumeur cancéreuse (voir la section 1.1.7). Avec l'avènement du scRNA-seq, l'HIT a été décrite dans les principaux cancers (cancer du sein, du poumon, du colon, de la prostate, de l'estomac, du foie et d'autres) (Suvà and Tirosh, 2019; G. Sun *et al.*, 2021), permettant de bâtir des atlas de cellules uniques. Ceux-ci portent la promesse de cartographier finement la diversité cellulaire afin d'explorer de nouvelles pistes de soins qui prendront par exemple en compte l'hétérogénéité spécifique de chaque tumeur pour améliorer le diagnostic et la prise en charge du patient. Par exemple, parmi les problématiques de diagnostic, actuellement, les données montrent que 10 à 25% des patients présentant de faibles risques d'un point de vue clinique finissent par rechuter en dépit de facteurs pronostiques favorables, et auraient nécessité en réalité une surveillance accrue ou une prise en charge thérapeutique (Sorbellini *et al.*, 2005; Rini, 2009; Nerich *et al.*, 2014). Dans le ccRCC, quelques atlas sont parus depuis la première étude d'ampleur en 2018 (Kim *et al.*, 2016; Young *et al.*, 2018; Bi *et al.*, 2021; Borcherding *et al.*, 2021; Braun *et al.*, 2021; Krishna *et al.*, 2021; Su *et al.*, 2021). Cependant, bien qu'ils décrivent globalement les mêmes types cellulaires, ils peuvent être incomplets (pour des raisons de sélection de cellules ou de plans expérimentaux défavorisant un type cellulaire) et aucune approche n'a prouvé qu'ils se recouvrent en termes de profils transcriptomiques. Il est d'ailleurs raisonnable de soupçonner qu'ils présentent des variations transcriptomiques significatives, y compris pour un type cellulaire donné. Premièrement, la technique de scRNA-seq est sensible à la variabilité inter-laboratoire (Tung *et al.*, 2017; Hicks *et al.*, 2018), ce qui induit dans les atlas des différences transcriptionnelles d'origine expérimentale et non biologique. Ensuite, ces atlas n'ont pas intégré l'ensemble des sous-types de cancers du rein, et l'hétérogénéité intra-tumorale elle-même est potentiellement plus vaste que ce que ces atlas ont montré jusqu'ici. Enfin, ils contiennent à ce jour une très grande majorité de cellules immunitaires, au détriment des autres cellules, à la fois dans le microenvironnement et dans le compartiment malin. Or, les cellules malignes, moteurs de l'HIT, divergent davantage d'une tumeur à l'autre et au sein d'une même tumeur que les cellules du microenvironnement (Gerlinger *et al.*, 2012; Swanton, 2012). De ce

point de vue, leur diversité inter et intra-tumorale, au contraire des cellules immunitaires, est particulièrement sous-représentée. Dans ce contexte, le premier projet dans lequel je me suis inséré au cours de ma thèse avait pour objectif de s'intéresser plus spécifiquement aux cellules malignes dans le ccRCC afin de bâtir un atlas dédié, et d'étudier leur potentiel diagnostique. L'hypothèse de travail était que la caractérisation de l'HIT, en particulier certains clones malins, pouvait être utilisée pour affiner la stratification des patients. En particulier, nous voulions examiner si cette information permettait d'améliorer la stratification des tumeurs de bas grade qui finissent par rechuter, chez les patients « *low-risk progressors* ».

**4.1.2. MANUSCRIT SOUMIS**



# **Single-cell deconvolution of a specific malignant cell population as a poor prognostic biomarker in low-risk clear cell renal cell carcinoma patients**

Saout J.<sup>1</sup>, Lecuyer G.<sup>1</sup>, Leonard S.<sup>1,2,3</sup>, Evrard B.<sup>1</sup>, Kammerer-Jacquet SF.<sup>1,4</sup>, Noel L.<sup>1</sup>, Khene ZE<sup>4</sup>, Mathieu R.<sup>1,5</sup>, Brunot A.<sup>6</sup>, Rolland AD.<sup>1</sup>, Bensalah K.<sup>5</sup>, Rioux-Leclercq N.<sup>1,4,⊠</sup>, Lardenois A.<sup>1,⊠</sup>, Chalmel F.<sup>1,\*</sup>

<sup>1</sup> Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) - UMR\_S 1085, F-35000 Rennes, France

<sup>2</sup> LabEx IGO “Immunotherapy, Graft, Oncology”, F-35043 Nantes, France

<sup>3</sup> Univ Rennes, INSERM, EFS, UMR S1236, F-35000, Rennes, France

<sup>4</sup> Pathology Department, University Hospital of Rennes, Rennes, France

<sup>5</sup> Urology Department, University Hospital of Rennes, Rennes, France

<sup>6</sup> Department of Oncology, University Hospital of Rennes, Rennes, France

\* To whom correspondence should be addressed

⊠ These authors contributed to the manuscript equally

**Correspondence:** frederic.chalmel@inserm.fr

## **Abstract**

**BACKGROUND:** Intra-tumor heterogeneity (ITH) is a key feature in clear-cell renal cell carcinomas (ccRCCs) which impacts outcomes such as aggressiveness, response to treatments or recurrence. In particular, it may explain tumor relapse after surgery in clinically low-risk patients who did not benefit from adjuvant therapy. Recently, single-cell RNA-seq (scRNA-seq) has emerged as a powerful tool to unravel ITH and might enable better assessment of clinical outcomes in ccRCC.

**OBJECTIVE:** To explore ITH in ccRCC with a focus on malignant cells (MCs), and assess its relevance to improve prognosis for low-risk patients.

**DESIGN, SETTING AND PARTICIPANTS:** We performed scRNA-seq on tumor samples from five untreated ccRCC patients. Data was complemented with a published dataset composed of pairs of matched normal and ccRCC samples.

**INTERVENTION:** Radical or partial nephrectomy on untreated ccRCC patients.

**OUTCOME MEASUREMENTS AND STATISTICAL ANALYSIS:** Viability and cell type proportions were determined by flow cytometry. Following scRNA-seq, functional analysis was performed and tumor progression trajectories were inferred. A deconvolution approach was applied on an external cohort and Kaplan-Meier survival curves were estimated with respect to malignant clusters prevalence.

### **RESULTS AND LIMITATIONS:**

We analyzed 54,812 cells and identified 35 cell subpopulations. The ITH analysis revealed that each tumor contained various degrees of clonal diversity. The transcriptomic signatures of MCs in one particularly heterogeneous sample were used to design a deconvolution-based strategy which allowed to risk-stratify 310 low-risk ccRCC patients.

**CONCLUSIONS:** We described ITH in ccRCCs and used this information to establish significant cell population-based prognostic signatures and better discriminate ccRCC patients. This approach has the potential to improve the stratification of clinically low-risk patients and their therapeutic management.

**PATIENT SUMMARY:** We sequenced the RNA content of individual cell subpopulations composing ccRCCs and identified specific malignant cells whose genetic information can be used to predict tumor progression.



## Introduction

Clear cell renal cell carcinoma (ccRCC) is the most frequent and lethal subtype of kidney cancer which represents 2.2% of all cancers and 180,000 deaths annually, worldwide [1]. It has a poor prognosis, as 45% of diagnosed patients are or become metastatic [2], and 30% of those treated with surgery eventually relapse, including 10-25% of clinically low-risk patients with progressing disease (called low-risk progressors) [3–5]. Recent targeted therapies have provided significant benefits for patients in terms of progression-free and overall survival [6]. However, their efficacy is limited with 30% of responsive patients, contrasting responses within tumors, and long-term relapses in most cases [7–9]. These pitfalls are not only due to inter- but also, importantly, to intra-tumor heterogeneity (ITH) which underlies the emergence of novel malignant clones, allowing the tumor to proliferate, escape treatments or relapse after surgery [10–13]. Single-cell RNA sequencing (scRNA-seq) has become the method of choice to decipher ITH [14] by simultaneously profiling the individual transcriptome of thousands of cells from a dissociated tissue. It has contributed to demonstrate that ITH is critical for tumor maintenance and progression [14]. In ccRCCs, a pioneering single-cell study provided the first evidence of ITH and revealed differences in targetable signaling pathways between a primary tumor and its metastasis [15]. Other studies have used scRNA-seq to explore renal carcinomas [16–28], providing new insights into tumor subtypes [26], origin and differentiation of RCC cells [25,28], or characterization of the immune landscape [16,22,27]. Most of the datasets used by these studies comprise a large proportion of immune cells (ICs) and only a small fraction of malignant cells (MCs). Hence, there is still a need for single-cell atlases focusing on MC heterogeneity which might provide new insights into the ability of ccRCC to proliferate, escape treatments, metastasize, or relapse. Here we designed and validated a scRNA-seq protocol to characterize the MCs in ccRCC tumors from five patients, which was complemented with a scRNA-seq dataset composed of three pairs of ccRCC and adjacent normal kidney samples [17]. The subsequent analysis revealed contrasted degrees of heterogeneity within tumors. One particular tumor displayed extensive clonal diversity, recapitulating aggressive features linked to epithelial-mesenchymal transition, cell migration, angiogenesis and cell adhesion. By applying an *in silico* deconvolution approach to a bulk RNA-seq dataset of ccRCCs, we

found that the predicted proportion of a specific MC cluster was a poor prognosis marker in low-risk progressors. Our results provide a valuable resource for the community that has been made available via the Uro-oncogenomics Viewer database (<https://uncover.genouest.org>) (Lecuyer, Saout *et al.*, in preparation).

## **Patients and methods**

### **Tumor specimens**

Fresh tumor samples were collected from a total of five untreated patients who underwent either radical or partial nephrectomy for ccRCC at the urology department of the Rennes university-hospital center. Patients met the requirements of the local institutional ethics committee (see Supplementary methods). Detailed clinical information about each patient is provided in Table S1.

### **Single-cell preparation, library preparation, sequencing and analysis**

For each tumor, ~2-9 punches were pooled and dissociated, and a partial depletion of ICs was performed to enrich the samples in MCs (Fig. 1A; Supplementary methods). Construction of single-cell libraries (Chromium™ Single Cell 3' Library & Gel Bead Kit v2) was carried out according to the manufacturer's instructions (10x Genomics). Single-cell libraries were quality-controlled using the Agilent 2100 Bioanalyzer system and sequenced on an Illumina HiSeq 4000 instrument using a 2x100 paired-end sequencing protocol (Illumina, San Diego, CA). Data processing was performed in R as extensively detailed in Supplementary methods.

## Results

### Validation of the experimental procedure

We studied five samples (designated RCC1-5) each from separate untreated ccRCC patients with various clinical features, as well as one sample from a vena cava tumor thrombus from the RCC5 patient (RCC5\_t) and a sample of RCC4 dissociated cells that were frozen (RCC4\_f) (Fig. S1; Fig. S2; Table S1; Supplementary results [SR], Section 1). As our primary objective was to build a single-cell atlas of predominantly MCs, we designed and evaluated, by flow cytometry, a dedicated protocol that showed good dissociation, viability and cell type representativeness (Fig. 1A; Fig. S3; SR, SR-Section 2). Importantly, it allowed us to enrich the MC content of the samples by dramatically reducing the proportion of immune cells, and proved to only slightly affect viability in frozen dissociated cells (Fig. S3; Fig. S4).

### A single-cell atlas of ccRCCs composed of ~55,000 cells

After confirming our ability to efficiently dissociate the tumor and enrich MCs, we sequenced the transcriptome of individual cells of seven samples (RCC1-5, RCC4\_f and RCC5\_t). Our dataset was complemented with three pairs of matched healthy renal and ccRCC tumor tissues (termed RCC[6-8]\_n and RCC[6-8], respectively) [17] (Table S1; Table S2). Following data processing, a pseudobulk analysis revealed differences between both datasets (Fig. S5; SR, Section 3), prompting us to perform a batch effect correction. The final scRNA-seq dataset was composed of 54,812 cells (Fig. 1B; SR, Section 3) which were projected on 2D UMAP spaces (Fig. 1C-D).

### Single-cell RNA-seq uncovers the cellular composition in ccRCC

Cells were next partitioned into 35 cell clusters (termed c1-c35) (Fig. 2A) which were associated with eight broad cell types based on specific marker genes such as CA9 and NNMT for MCs [29,30] (Table S3; Fig. 2B-D; SR, Section 3). The balance between ICs and MCs across samples clearly distinguished

both datasets (Fig. 2E), stressing the relevance of our approach to enrich MCs. Remaining ICs in RCC4 and RCC4\_f were CD45- B cells, which explains their insensitivity to IC depletion (data not shown). To further characterize MCs, we inferred their copy number variations (CNV) which confirmed the expected 3p loss in all samples and revealed sample-specific CNVs (Fig. S6; Fig. S7; SR, Section 3). Overall, these results confirm that ccRCC exhibits both intra- and inter-tumor heterogeneity (SR, Section 3).

### **Malignant cells display clonal diversity at chromosomal and gene expression levels**

To deepen the analysis of MC diversity, we next focused the analysis on the 15,471 MCs and the 3,737 tubular epithelial cells (or TECs; the normal cells from which MCs originate) which were subsequently partitioned into 24 clusters (termed c1-c24) (Fig. 3A; Table S4; SR, Section 4). With the exception of TECs (c1, c12 and c15), the vast majority of cell clusters did not mix across samples suggesting that each tumor had its own evolutionary history (Fig. 3A-B), as highlighted by the inter-tumor heterogeneity in chromosomal aberrations (Fig. S6; SR, section 3). Tumors also harbored various degrees of ITH (Fig. 3B-C). In particular in RCC2, MCs were distributed in five well delineated clusters; c3 (shared with RCC1, RCC3 and RCC4), c10, c23, c9 and c8, whereas other samples had a more homogeneous composition. Interestingly, the CNV analysis also revealed that the RCC2 sample carried higher intra-tumor CNV heterogeneity than the other tumors, in particular regarding 3q, 6p, 9q, 14q and 18p, that were lost in c8 but had contrasted status in other RCC2 malignant clusters (Fig. S7). With the exception of c16 and c24, the cell cycle analysis revealed that all clusters were composed of a vast majority of cells in G1 and 9-36% in G2M (Fig. 3D), in accordance with typical observation in solid tumors [31]. Along with CA9 and NNMT, we uncovered specific markers of MCs such as EGLN3, CP, MT3, NPTX2 and TRIB3 (Fig. 3E-F). Expression of cell cycle markers (CKS2, CKAP2L and TOP2A) confirmed the cycling status of c16. As expected, we also found high expression of TEC markers such as ALDOB, MIOX, QDPR and GATM in c1, c12 and c15. We also noticed the expression of PTPRC and CD34 in two small cell clusters (c24 and c22, respectively) corresponding to residual contaminants by immune cells and endothelial cells, respectively.

### **The RCC2 tumor clonal lineage recapitulates important steps of carcinogenesis**

Given the variable degree of ITH observed in each tumor, we next tried to reconstruct tumor lineages within each sample (Fig. S8). RCC2 displayed the longest cellular trajectory going from c3 through c10, c23 and c9, up to c8 (Fig. S8). The chromosomal imbalance analysis revealed that the CNV burden increased along the pseudotime, indicating progressive and cumulative chromosomal alterations, which provides additional elements in favor of a c3 to c8 lineage progression (Fig. 4B). Fifty-two regions were found to be significantly gained or lost along the RCC2 pseudotime, which was further supported by a phylogenetic analysis (SR, Section 5; Fig. S9). We also found 1,448 differentially expressed genes (DEGs) across RCC2 MC clusters which were subsequently partitioned into seven expression patterns (termed P1-7) (Fig. 4C; Table S5). The functional analysis demonstrated the accumulation of aggressive features along the RCC2 lineage, with a notable peak in c8, related to epithelial-mesenchymal transition (EMT, a crucial initial stage of the metastatic process), angiogenesis, cell migration and cell adhesion (Table S6-S7; SR, Section 6). Overall these results suggest that MCs from RCC2 are linked together through a common tumor clonal lineage from indolent (c3) to aggressive (c8) MC clusters showing the sequential acquisition of typical molecular features corresponding to important steps of tumor progression.

### **Deconvolution of specific malignant clusters is associated with tumor stages and survival in low-risk patients**

We next hypothesized that the RCC2 lineage could represent an *in vivo* model of ccRCC carcinogenesis which could be used as a reference. To verify this, we deconvolved the proportions of somatic cell types and RCC2 MC clusters on a massive bulk RNA-seq dataset (525 samples) of ccRCCs (Fig. 5A, Table S8). Substantial proportions of clusters c8 and c9 as well as non-immune microenvironment cells (endothelial cells, pericytes, myofibroblasts) were significantly detected, covering all four T stages. Despite their cellular proximity with MCs, TECs were not detected in ccRCC tumors, as expected.

Concerning MCs, the main deconvolved populations were c9 and c8 with opposite trends (Fig. 5A, Table S8), by respectively decreasing and increasing with tumor stages. We then evaluated the ability of deconvoluted cell types to predict survival in either all or in T1-T2 low-risk ccRCC patients of the TCGA cohort (SR, section 7, Fig. 5B, Table S9). When considering all patients, we found a significant association with poor prognosis of c8, and, to a lesser extent, of myofibroblasts and B cells. Conversely, c9, c10 and endothelial cells were three excellent predictors of good prognosis. When looking more specifically at low-risk patients, strikingly, only c8, c10 and c9, and, to a lesser extent, endothelial cells were still associated with survival.

## Discussion

On the basis of conventional diagnosis, a significant proportion of patients considered to be at low-risk are still wrongly excluded from adjuvant treatment. A major issue in the management of ccRCCs hence consists in improving patient stratification through early detection of those who will eventually relapse and could have benefited from specific surveillance and potential adjuvant treatments. Biomarkers can be used to stratify these patients, but they remain insufficient. Moreover, the differences between gene expression signatures and corresponding protein detection constitute an important gap for clinical diagnosis [32,33], prompting the use of alternative and complementary approaches (such as transcriptomics) in future diagnostic tools. In this context, it remains important to improve our understanding of tumor progression in ccRCCs in order to enhance patient care and consider personalized therapeutic strategies. To address these issues, we used scRNA-seq to decipher ITH in ccRCCs, with a specific focus on the malignant compartment, hypothesizing that this information could notably contribute to improve the detection of low-risk progressors.

Since published scRNA-seq datasets in ccRCCs are often focused on the characterization of the immune compartment [16,18,22,27], we established an experimental procedure to enrich the fraction of MCs by reducing the fraction of ICs. We also examined several punches per tumor to better capture ITH as suggested by Gerlinger and colleagues [13]. Samples from five ccRCC patients enabled us to assemble an atlas composed of ~25,826 cells (including ~48.3% of MCs) which we complemented with 28,986 cells (including ~10.3% of MCs) published by Young *et al.* [17]. Despite common molecular events such as the truncal 3p loss, the analysis of MCs revealed an important heterogeneity across samples. Various degrees of ITH in the distinct samples were observed, from low (RCC1, 6 and 7) to very high (RCC2 and 5 to a lower extent). We also noticed that both the inter- and intra-tumor heterogeneities in the samples of Young *et al.* were lower than in our dataset, possibly due to the difference in sequencing depths, the number of punches per tumor and/or the higher proportion of MCs in our samples. Those differences appear as critical features to properly capture ITH. Together these results suggest that ITH is not necessarily the rule, as pointed out in hepatocellular carcinoma [34]. Thus, future cohorts must



include more tumors along with their paired metastases in order to draw a complete picture of ccRCC heterogeneities.

Among the different patients, we could identify a unique tumor sample (RCC2), displaying a yet never observed ITH consisting of five distinct MC states. Those were connected through a common lineage characterized by the progressive acquisition of highly aggressive features related to EMT, angiogenesis, cell migration and cell adhesion, as well as specific chromosomal abnormalities. RCC2 was a localized and apparently low-grade tumor (pT2b) without clinical/histological markers of aggressiveness. Yet, the RCC2 patient is currently under surveillance after having developed a suspicious pancreatic nodule, a preferred site for ccRCC metastasis. We believe that the RCC2 MC lineage can therefore represent a valuable resource for the community as an *in vivo* and representative model of tumor progression.

A landmark study used cell type deconvolution from TCGA data to successfully determine three groups of ccRCC patients according to lymphocyte T cell states and demonstrated their significant association with prognosis [20]. Here we used a similar approach based on MC subpopulations composing the RCC2 tumor. Our results proved useful to predict survival and confirmed previous findings encouraging the consideration of ITH to identify predictive biomarkers [35]. The ability of this method to classify ccRCC patients from bulk RNA-seq is remarkable, especially the capacity of the predicted proportion of c8 and c9 MC clusters to serve as poor and good prognostic factors, respectively. In c9, despite the expression of aggressive features, the marked organization of adaptive immunity seems to confer an overall good clinical outcome. Importantly, some deconvoluted cell types constituting the tumor-microenvironment are already known to be associated with survival [36–38]. Here, not only do we confirm these results but we go further by obtaining refined and resolute malignant signatures with superior predictive power, especially in low-risk patients. Interestingly, the proportion of T1-T2 patients (~10%) with a relatively high proportion of c8 is similar to the proportions of clinically low-risk progressors (10-25%), suggesting a potential overlap between the predicted and actual populations with poor outcomes. Our results suggest that low-grade tumors exceeding ~25% of c8 cells warrant reinforced surveillance. Hence, we present here the first proof of concept for using MC expression signatures to improve clinical management of ccRCC patients. Importantly a tumor that better recapitulates the stages of

carcinogenesis could also complement and/or update our current deconvolution-based strategy. In the future it will be interesting to develop nomograms including relevant MC clusters and compare their accuracy with other conventional nomograms such as one recently developed in ccRCCs, based on scRNA-seq data [39]. Finally, it will be necessary to explore, in dedicated cohorts, to what extent these deconvolution-based models can predict other features such as relapse, metastasis and treatment response. Therefore, we strongly encourage the use of transcriptomics in future clinical routine, especially since low-cost technologies such as BRB-seq [40] are now available and easily deployable, in order to enable the use of such predictive strategies and develop a personalized medicine. In this study we have assembled a valuable atlas of ccRCC MCs; we are committed to expanding this dataset and to making it available online via the UncoVer resource. Thanks to the single-cell profiling of a rare tumor showing a high ITH, we have established a reliable tumor progression model for low-risk patients based on the deconvolution of specific MC subpopulations which could contribute to improving patient care.

## **Author's contributions**

FC, designed the study. JS, BE, SFKJ, RM, ABt, ADRd, KB and NRL acquired the data. JS, GL, SL, ADR, AL and FC analyzed and interpreted the data. JS and FC drafted the manuscript. GL, SL, BE, FC, AL, ADR and NRL participated in the critical revision of the manuscript for important intellectual content. JS, GL, SL, AL and FC performed the statistical analyses. FC, AB and NRL obtained fundings. LN provided administrative and technical support. FC supervised the study.

## **Funding**

This study was funded by La ligue contre le cancer committees 22, 35 and 81 and by the clinical and translational research committee (CORECT) of the Rennes University Hospital Center. S.L. was supported by the chair “Cancer & Innovation” of Rennes 1 Foundation (<https://fondation.univ-rennes1.fr/>) and a specific grant from the LabEx IGO program (n° ANR-11-LABX-0016) funded by the «Investment into the Future» French Government program, managed by the National Research Agency (ANR).

## **Conflict of interest**

The authors have no conflict of interest to declare.

## **Acknowledgements**

We thank all the staff of the Urology Department of the Rennes University Hospital for their expert assistance, and the participating patients. We thank the GenOuest bioinformatics facility for hosting the software as well as all members of the Research Institute for Environmental and Occupational Health (IRSET) for stimulating discussions. We thank the CytomeTri platform from Biosit (SFR UMS CNRS 3480 - INSERM 018), especially Laurent Deleurme and Marion Mandon for their precious help. We thank the H2P2 platform (Biosit) for their services. We also thank the Institut national de la santé et de la recherche médicale (Inserm), the Université de Rennes 1, and the French School of Public Health (EHESP) for supporting this work. For their valuable help, our special thanks also go to Catherine

Lavau, William Taylor, Stéphane Boutet (10X Genomics), Marjorie Gournay, Mikael Roussel and Frédéric Dugay.

## References

- [1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians* 2021;71:209–49. <https://doi.org/10.3322/caac.21660>.
- [2] Hsieh JJ, Purdue MP, Signoretti S, Swanton C, Albiges L, Schmidinger M, et al. Renal cell carcinoma. *Nat Rev Dis Primers* 2017;3:17009.
- [3] Nerich V, Hugues M, Nai T, Stein U, Hon TNT, Montcuquet P, et al. Clinical impact of targeted therapies in patients with metastatic clear-cell renal cell carcinoma. *OncoTargets and Therapy* 2014;365. <https://doi.org/10.2147/ott.s56370>.
- [4] Rini BI. Metastatic renal cell carcinoma: many treatment options, one patient. *J Clin Oncol* 2009;27:3225–34.
- [5] Sorbellini M, Kattan MW, Snyder ME, Reuter V, Motzer R, Goetzl M, et al. A postoperative prognostic nomogram predicting recurrence for patients with conventional clear cell renal cell carcinoma. *J Urol* 2005;173:48–51.
- [6] Ross K, Jones RJ. Immune checkpoint inhibitors in renal cell carcinoma. *Clin Sci* 2017;131:2627–42.
- [7] Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Rixe O, et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med* 2007;356:115–24.
- [8] Molina AM, Lin X, Korytowsky B, Matczak E, Lechuga MJ, Wiltshire R, et al. Sunitinib objective response in metastatic renal cell carcinoma: analysis of 1059 patients treated on clinical trials. *Eur J Cancer* 2014;50:351–8.
- [9] Crusz SM, Tang YZ, Sarker S-J, Prevoo W, Kiyani I, Beltran L, et al. Heterogeneous response and progression patterns reveal phenotypic heterogeneity of tyrosine kinase inhibitor response in metastatic renal cell carcinoma. *BMC Med* 2016;14:185.
- [10] López JI, Angulo JC. Pathological Bases and Clinical Impact of Intratumor Heterogeneity in Clear Cell Renal Cell Carcinoma. *Curr Urol Rep* 2018;19:3.
- [11] McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell* 2015;27:15–26.
- [12] Raatz M, Shah S, Chitadze G, Brüggemann M, Traulsen A. The impact of phenotypic heterogeneity of tumour cells on treatment and relapse dynamics. *PLoS Comput Biol* 2021;17:e1008702.
- [13] Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012;366:883–92.
- [14] Suvà ML, Tirosh I. Single-Cell RNA Sequencing in Cancer: Lessons Learned and Emerging Challenges. *Mol Cell* 2019;75:7–12.
- [15] Kim K-T, Lee HW, Lee H-O, Song HJ, Jeong DE, Shin S, et al. Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma. *Genome Biol* 2016;17:80.
- [16] Bi K, He MX, Bakouny Z, Kanodia A, Napolitano S, Wu J, et al. Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma. *Cancer Cell* 2021;39:649–61.e5.
- [17] Young MD, Mitchell TJ, Vieira Braga FA, Tran MGB, Stewart BJ, Ferdinand JR, et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science*

- 2018;361:594–9.
- [18] Borcherding N, Vishwakarma A, Voigt AP, Bellizzi A, Kaplan J, Nepple K, et al. Mapping the immune environment in clear cell renal carcinoma by single-cell genomics. *Commun Biol* 2021;4:122.
- [19] Fendler A, Bauer D, Busch J, Jung K, Wulf-Goldenberg A, Kunz S, et al. Inhibiting WNT and NOTCH in renal cancer stem cells and the implications for human patients. *Nat Commun* 2020;11:929.
- [20] Hu J, Chen Z, Bao L, Zhou L, Hou Y, Liu L, et al. Single-Cell Transcriptome Analysis Reveals Intratumoral Heterogeneity in ccRCC, which Results in Different Clinical Outcomes. *Mol Ther* 2020;28:1658–72.
- [21] Argüello RJ, Combes AJ, Char R, Gigan J-P, Baaziz AI, Bousiquot E, et al. SCENITH: A Flow Cytometry-Based Method to Functionally Profile Energy Metabolism with Single-Cell Resolution. *Cell Metab* 2020;32:1063–75.e7.
- [22] Braun DA, Street K, Burke KP, Cookmeyer DL, Denize T, Pedersen CB, et al. Progressive immune dysfunction with advancing disease stage in renal cell carcinoma. *Cancer Cell* 2021;39:632–48.e8.
- [23] Liu K, Gao R, Wu H, Wang Z, Han G. Single-cell analysis reveals metastatic cell heterogeneity in clear cell renal cell carcinoma. *J Cell Mol Med* 2021;25:4260–74.
- [24] Yu Z, Lu W, Su C, Lv Y, Ye Y, Guo B, et al. Single-Cell RNA-seq Identification of the Cellular Molecular Characteristics of Sporadic Bilateral Clear Cell Renal Cell Carcinoma. *Front Oncol* 2021;11:659251.
- [25] Zhang Y, Narayanan SP, Mannan R, Raskind G, Wang X, Vats P, et al. Single-cell analyses of renal cell cancers reveal insights into tumor microenvironment, cell of origin, and therapy response. *Proc Natl Acad Sci U S A* 2021;118. <https://doi.org/10.1073/pnas.2103240118>.
- [26] Su C, Lv Y, Lu W, Yu Z, Ye Y, Guo B, et al. Single-Cell RNA Sequencing in Multiple Pathologic Types of Renal Cell Carcinoma Revealed Novel Potential Tumor-Specific Markers. *Front Oncol* 2021;11:719564.
- [27] Krishna C, DiNatale RG, Kuo F, Srivastava RM, Vuong L, Chowell D, et al. Single-cell sequencing links multiregional immune landscapes and tissue-resident T cells in ccRCC to tumor topology and therapy efficacy. *Cancer Cell* 2021;39:662–77.e6.
- [28] Young MD, Mitchell TJ, Custers L, Margaritis T, Morales-Rodriguez F, Kwakwa K, et al. Single cell derived mRNA signals across human kidney tumors. *Nat Commun* 2021;12:3896.
- [29] Baniak N, Flood TA, Buchanan M, Dal Cin P, Hirsch MS. Carbonic anhydrase IX (CA9) expression in multiple renal epithelial tumour subtypes. *Histopathology* 2020;77:659–66.
- [30] Tang S-W, Yang T-C, Lin W-C, Chang W-H, Wang C-C, Lai M-K, et al. Nicotinamide N-methyltransferase induces cellular invasion through activating matrix metalloproteinase-2 expression in clear cell renal cell carcinoma cells. *Carcinogenesis* 2011;32:138–45.
- [31] Terz JJ, Lawrence W Jr, Cox B. Analysis of the cycling and noncycling cell population of human solid tumors. *Cancer* 1977;40:1462–70.
- [32] Buccitelli C, Selbach M. mRNAs, proteins and the emerging principles of gene expression control. *Nat Rev Genet* 2020;21:630–44.
- [33] Gygi SP, Rochon Y, Robert Franza B, Aebersold R. Correlation between Protein and mRNA Abundance in Yeast. *Molecular and Cellular Biology* 1999;19:1720–30. <https://doi.org/10.1128/mcb.19.3.1720>.
- [34] Xue R, Li R, Guo H, Guo L, Su Z, Ni X, et al. Variable Intra-Tumor Genomic Heterogeneity of Multiple Lesions in Patients With Hepatocellular Carcinoma. *Gastroenterology* 2016;150:998–1008.
- [35] Gulati S, Martinez P, Joshi T, Birkbak NJ, Santos CR, Rowan AJ, et al. Systematic evaluation of the prognostic impact and intratumour heterogeneity of clear cell renal cell carcinoma biomarkers. *Eur Urol* 2014;66:936–48.
- [36] Edeline J, Mottier S, Vigneau C, Jouan F, Perrin C, Zerrouki S, et al. Description of 2 angiogenic phenotypes in clear cell renal cell carcinoma. *Hum Pathol* 2012;43:1982–90.
- [37] Ambrosetti D, Coutts M, Paoli C, Durand M, Borchiellini D, Montemagno C, et al. Cancer-associated fibroblasts in renal cell carcinoma: implication in prognosis and resistance to anti-angiogenic therapy. *BJU Int* 2022;129:80–92.

- [38] Li S, Huang C, Hu G, Ma J, Chen Y, Zhang J, et al. Correction: Tumor-educated B cells promote renal cancer metastasis via inducing the IL-1 $\beta$ /HIF-2 $\alpha$ /Notch1 signals. *Cell Death Dis* 2022;13:415.
- [39] Xia Z-N, Wu J-G, Yao W-H, Meng Y-Y, Jian W-G, Wang T-D, et al. Identification of a differentiation-related prognostic nomogram based on single-cell RNA sequencing in clear cell renal cell carcinoma. *Sci Rep* 2022;12:10973.
- [40] Alpern D, Gardeux V, Russeil J, Mangeat B, Meireles-Filho ACA, Breysse R, et al. BRB-seq: ultra-affordable high-throughput transcriptomics enabled by bulk RNA barcoding and sequencing. *Genome Biol* 2019;20:71.

## Figure legends

**Figure 1. Single-cell analysis of ccRCC.** (A) Flowchart depicting the experimental design for each ccRCC sample. Single-cell suspensions were assessed by flow cytometry at two steps in the protocol : after tumor dissociation (step 1), and, after low centrifugations and immune cell removal (step 2). (B) Bar plots representing the number of cells and the median number of detected genes for each sample. Each sample corresponds to a distinct color, with normal samples shown in shades of gray. *\_f* = frozen sample; *\_t* = thrombus sample; *\_n* = normal kidney. (C) UMAP plot of all cells from pooled samples colored according to panel B, in which nearby points represent transcriptionally similar cells. (D) Deconvolution of the global UMAP plot (panel C) per patient. In each UMAP plot, the cells are colored by sample according to panel B.

**Figure 2. Cellular diversity in ccRCC.** (A) UMAP plot representing the cell type distribution (main colors) and clustering (shades of color) of all cells. (B-C) Expression levels of the CA9 and NNMT malignant cell markers. (D) Dot plot representing the expression of known cell type markers in each cell cluster. MCs = Malignant cells ; TECs = Tubular epithelial cells ; CDEC = Collecting duct epithelial cells. (E) Stacked bar plot illustrating the distribution of the main cell types in each sample. RCC4\_*f* = frozen sample; *\_t* = thrombus sample; *\_n* = normal kidney sample.

**Figure 3. Characterisation of ccRCC malignant cells and normal tubular epithelial cells.** (A) UMAP plots and clustering of all malignant cells. (B) Deconvolution of the global UMAP plot (panel A) per patient. Cells are colored by sample as indicated previously (Figure 1B). MCs = Malignant cells ; TECs = Tubular epithelial cells; TCs = Thrombus cells. (C) Distribution of clusters per sample. *\_t* = thrombus sample; *\_n* = normal kidney. (D) Distribution of samples per cluster. (E) Distribution of cells with S, G1 or G2M status per cluster. (F) Dot plot representing the expression of differentially expressed genes (DEGs) in malignant cell clusters. (G) Expression of DEGs as UMAP plots. Colors correspond to the scaled expression of panel F.

**Figure 4. Intra-tumor heterogeneity of RCC2 malignant cells.** (A) UMAP plot showing pseudotime values as computed by Monocle3 in shades of red and the lineage trajectory in RCC2 malignant cells as a black line. Here, the pseudotime is an abstract value describing the distance of a cell from the beginning of the lineage (green dot) in terms of transcriptomic variations, hence partially depicting tumor progression. (B) Violin plots of proportions of RCC2 MC clusters (left) and of the CNV burden (right) along the pseudotime. (C) Heatmap representation of the scaled expression values of 1,448 genes that are differentially expressed between RCC2 malignant cell clusters. These were ordered according to the pseudotime and were further colored given their CNV burden. DEGs are organized in seven patterns according to their expression in RCC2 malignant clusters. Enriched GO terms associated with each gene expression pattern are presented along with their associated *p values*. Shades of gray in annotation boxes on the left of heatmap correspond to different levels of significance for enrichment. Selection of genes for EMT and outcome (Favorable and Unfavorable) annotations is detailed in supplementary methods.

**Figure 5. Deconvolution and survival analysis.** (A) Density plots showing the distribution of percentages of the most prevalent deconvoluted cell types and RCC2 malignant cells in ccRCC TCGA patients divided into T stages. Malignant clusters c8, c9 and c10, correspond to relative malignant clusters as defined in supplementary methods. MCs = Malignant cells; TECs = Tubular epithelial cells. (B) Kaplan-Meier survival curves of patients grouped based on the low or high prevalence of different cell types. The analysis was performed for either all ccRCC patients (upper curves) or on lower risk (T1-T2) patients (lower curves). Cell types with significant association (log rank *p value* inferior to 0.01, in bold) in all patients are presented.

## Supplementary figure legends

**Fig S1. Histology of the six primary ccRCC specimens used in the study.** Sections were obtained from formalin-fixed, paraffin-embedded tissue stained with hematoxylin and eosin. ISUP = International Society of Urological Pathology ; pT = Pathological tumor.

**Fig. S2. Processing of macroscopic tumor samples.** (A) Macroscopic view of tumor samples. (B) Up to 9 distinct macroscopic areas were sampled in each tumor specimen. (C) Each punch was separated into three parts : a quarter for snap freezing, a quarter for tissue fixation, and one half which was pooled with other half punches for tissue dissociation.

**Fig. S3. Assessment of quality at different dissociation steps as illustrated with the RCC4 sample.** (A) Microscopic pictures (x10) of single cells after dissociation. (B) Flow cytometry cytograms of single cell suspensions. From left to right are the side scatter versus forward scatter gating, the live cell gating (side scatter versus 7AAD staining), and the proportions of different cell types based on the expression of CD34 (endothelial marker) and CD45 (immune cell marker). Malignant and stromal cells are negative for CD34 and CD45, hence located in the bottom left section. (C-D) Cell count and cell viability in each sample. (E-F) Proportions and cell viability of each monitored cell population before and after immune cell depletion. (G) Proportions of immune cells measured by flow cytometry after tumor dissociation or estimated by immunohistochemistry (IHC, prior to dissociation). (H) Cell viability of fresh versus frozen single cells (n=6).

**Fig. S4. Illustration of the automatic counting method performed on IHC sections (LCA immune cell labeling) using ImageJ software.** (A) Raw picture. (B) Picture with detection (red) of labeled immune cells. (C) Picture with detection (red) of unlabeled non-immune cells.

**Fig. S5. Pseudo-Bulk analysis of scRNA-seq data.** (A) Correlation matrix comparing all samples.  $_f$  = frozen sample;  $_t$  = thrombus sample;  $_n$  = normal kidney. (B) PCA space projection of all samples.  $_f$  = frozen sample;  $_t$  = thrombus sample;  $_n$  = normal kidney.

**Fig. S6. Heatmap representation of chromosome gains or losses estimated by inferCNV in all tumors.** (A) CNV in malignant cells.  $f$  = frozen sample;  $_t$  = thrombus sample;  $_n$  = normal kidney. (B) CNV in normal cells.  $f$  = frozen sample;  $_t$  = thrombus sample;  $_n$  = normal kidney.

**Fig. S7. CNVs in malignant and tubular epithelial cells.** (A) UMAP plot and clustering of all malignant cells. Gray cells correspond to tubular epithelial cells whereas colored cells correspond to malignant cells. (B) Dot plot showing computed gains (red) or losses (blue) of chromosome arms in malignant cell clusters. (C) UMAP plots presenting the distribution of CNVs across malignant cells.

**Fig. S8. Lineage trajectories in malignant and tubular epithelial cells of each tumor, as computed with Monocle3.** Lineages begin at the green dot corresponding to the cells with the least chromosomal aberrations. In the resulting trajectories, the degrees of branching match the levels of heterogeneity.

**Fig. S9. RCC2 malignant cells hierarchy.** (A) Heatmap of CNV in RCC2 malignant cell clusters. (B) Phylogenetic tree of RCC2 malignant cells. The phylogenetic tree was constructed using the Uphyloplot2 R package and is based on imputed CNVs from inferCNV. Percentages in brackets indicate the overlap between malignant cells defined with Seurat and Uphyloplot2. (+) : chromosomal gain; (-) : chromosomal loss.



## Supplementary table legends

**Table S1. Clinical features of the patients' tumors.** A total of 8 tumors were included in this study, from current (Saout et al.) and additional (Young et al.) datasets. ccRCC = clear cell renal cell carcinoma ; ISUP = International Society of Urological Pathology ; pTNM = Tumor-Node-Metastasis.

**Table S2. Characteristics of single-cell RNA-seq samples.** Both experimental and sequencing data are presented for each sample. RCC4\_f = RCC4 frozen cells ; RCC5\_t = RCC5 tumor thrombus ; RCC[6-8]\_n = RCC[6-8] adjacent normal tissue ; IC = Immune Cells ; LCA = Leukocyte Common Antigen ; IHC = ImmunoHistoChemical. Step 1 and step 2 refer to the malignant cell enrichment protocol steps described in the Patients and Methods section. Step 1 is after tumor dissociation and step 2 is after low speed centrifugations and immune cells removal.

**Table S3. Top differentially expressed genes (DEGs) between the 34 clusters in the total dataset.** DEGs were identified using the FindAllMarkers command from R package Seurat with default parameters. Pct.1 = percentage.

**Table S4. Top differentially expressed genes (DEGs) in the 24 clusters of malignant and tubular epithelial cells.** DEGs were identified using the FindAllMarkers command from R package Seurat with default parameters.

**Table S5. Top differentially expressed genes (DEGs) in the 5 clusters of malignant cells from RCC2.** DEGs were identified by pairwise comparisons between clusters using the FindMarkers command from R package Seurat.

**Table S6. Top enriched gene ontology (GO) terms associated with gene expression patterns in RCC2 malignant cell clusters.** Gene expression patterns were found to be enriched in 884 biological processes terms from the gene ontology. Custom terms related to TCGA prognostic and EMT involvement categories were tested for enrichment (see supplementary methods).  $r$  is the number of genes related to the GO term in a given pattern whereas  $R$  is the total number of genes related to the GO term.

**Table S7. Details of annotations of RCC2 differentially expressed genes (DEGs).** Genes are presented as being associated with different gene ontology (GO) terms and annotations. Genes in the EMT column are either tumor suppressors ("Tsup"), oncogenes ("Onco"), both ("Dual"), related to EMT ("other") or not ("unrelated"). The number of pubmed references was determined with the search "renal cell carcinoma" and the gene name in PubMed. The TDL category designates genes expressing surface proteins (GESPs) with their level of Targetability development ("TDL" or "Target Development Level"). These genes encode proteins through which approved drugs act ("Tclin") (i.e., mode-of-action drug targets), or are either known to bind small molecules with high potency ("Tchem"), have well-studied biology ("Tbio"), or are understudied ("Tdark"), as described elsewhere (PMID: 31898878). Genes were identified as favorable or unfavorable according to the processing of the Cancer Genome Atlas Kidney Renal Cell Carcinoma (TCGA-KIRC) data described in Supplementary Material.

**Table S8. Results of deconvolution of cell types in TCGA-KIRC cohort bulk RNA-seq data, using the Music R package.** Cell type estimations are indicated in percentages. Among malignant cell clusters, "Others" correspond to few RCC2 malignant cells that were in clusters other than c3, c10, c23, c9 or c8. Relative malignant cell clusters are initial percentages of the considered malignant cells divided by the total number of RCC2 malignant cells.

**Table S9. Survival analysis results.** Both the Cox Proportional-Hazards Model and Log Rank test methods were used to evaluate the significance of associations of deconvoluted cell types with survival. An association was considered significant when the p value was inferior to 0.01.

### **Supplementary tables**

Tables' dimensions are too important to fit this manuscript. They are available at the following address

instead:

[https://docs.google.com/spreadsheets/d/1CV8TzDwAdTZGy3wXOhUIgaR4Z5TRilrIDV2opUlq\\_fk/e\\_dit#gid=351500582](https://docs.google.com/spreadsheets/d/1CV8TzDwAdTZGy3wXOhUIgaR4Z5TRilrIDV2opUlq_fk/e_dit#gid=351500582)

# Figures

Fig. 1

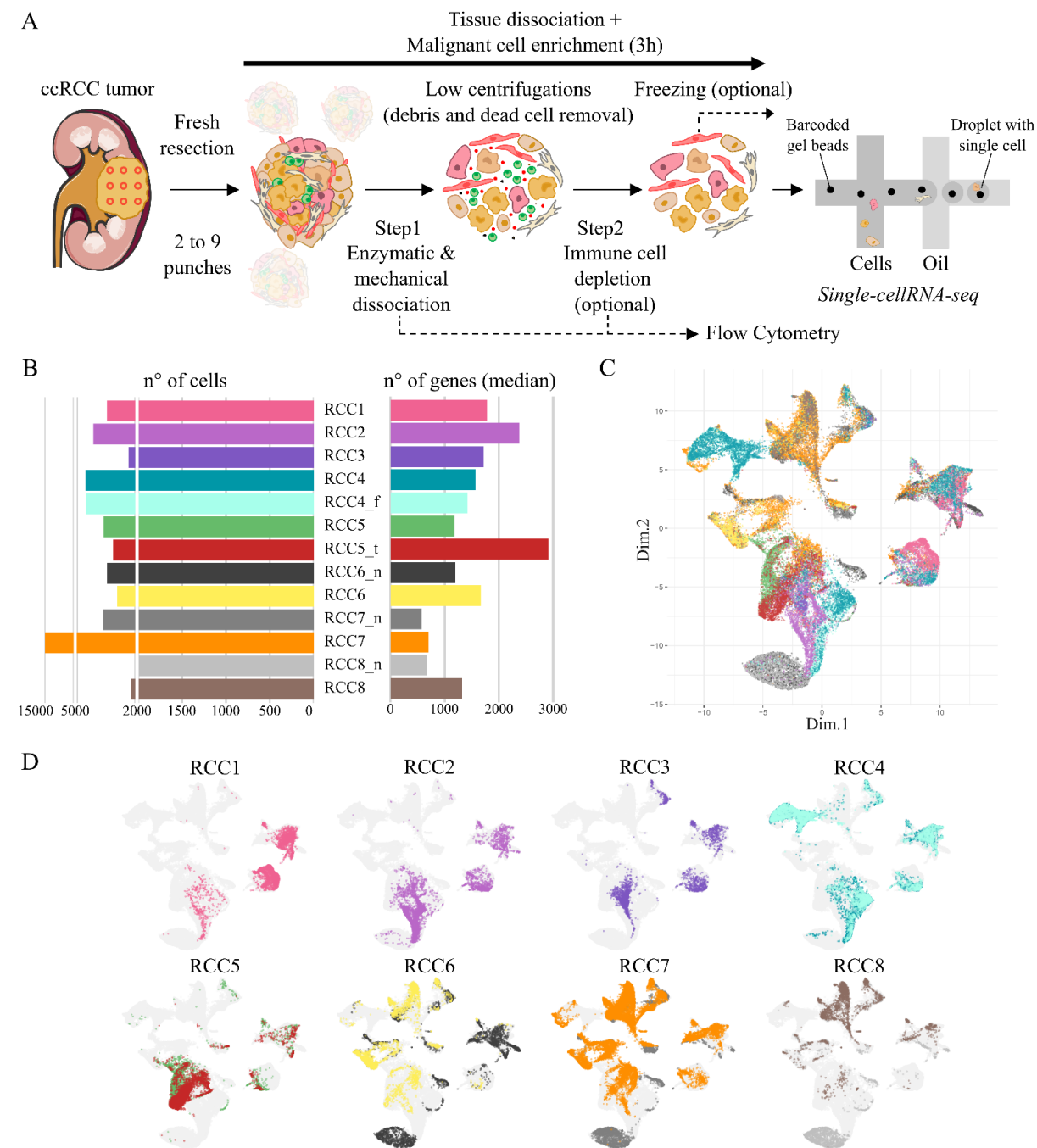


Fig. 2

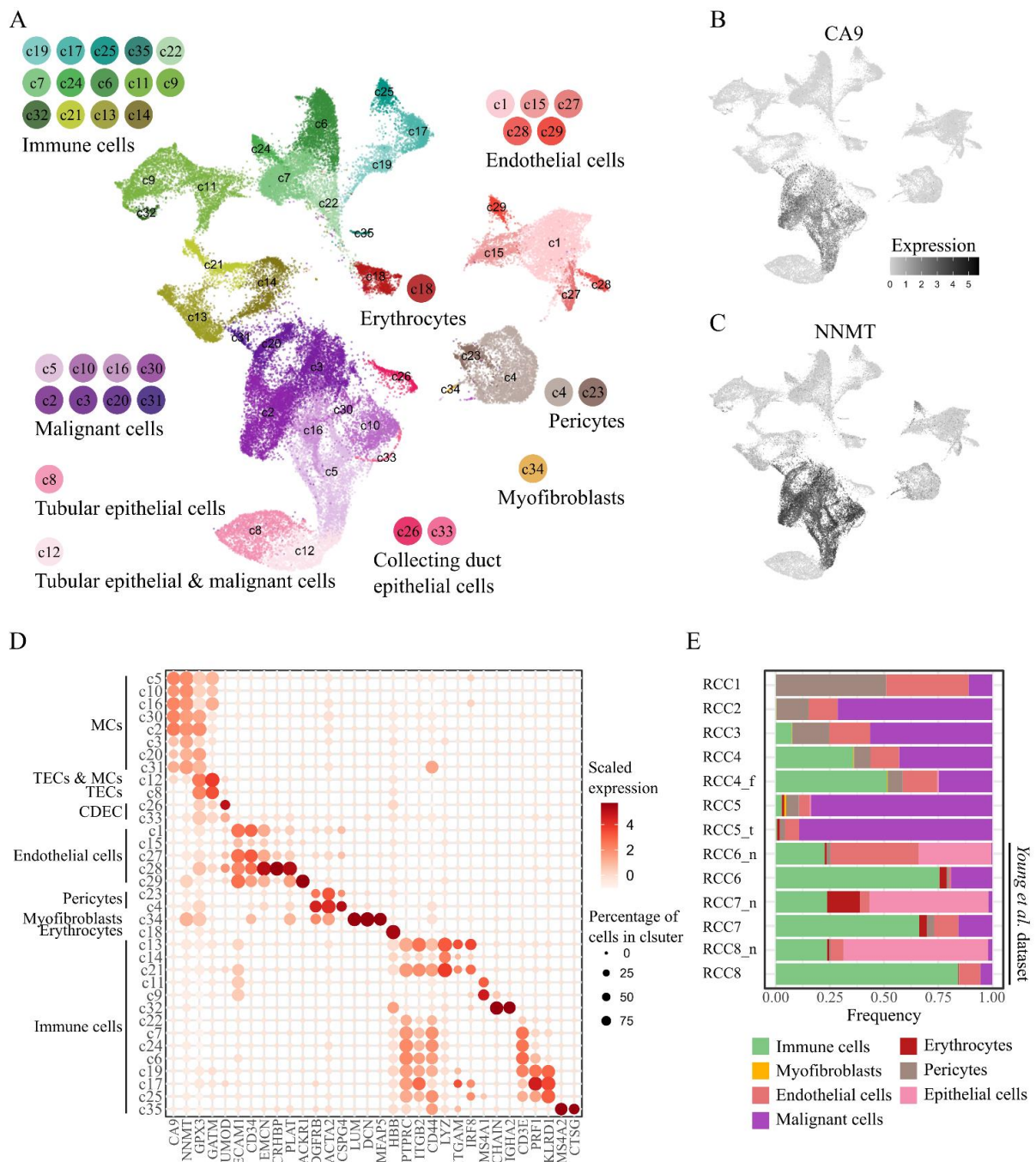


Fig. 3

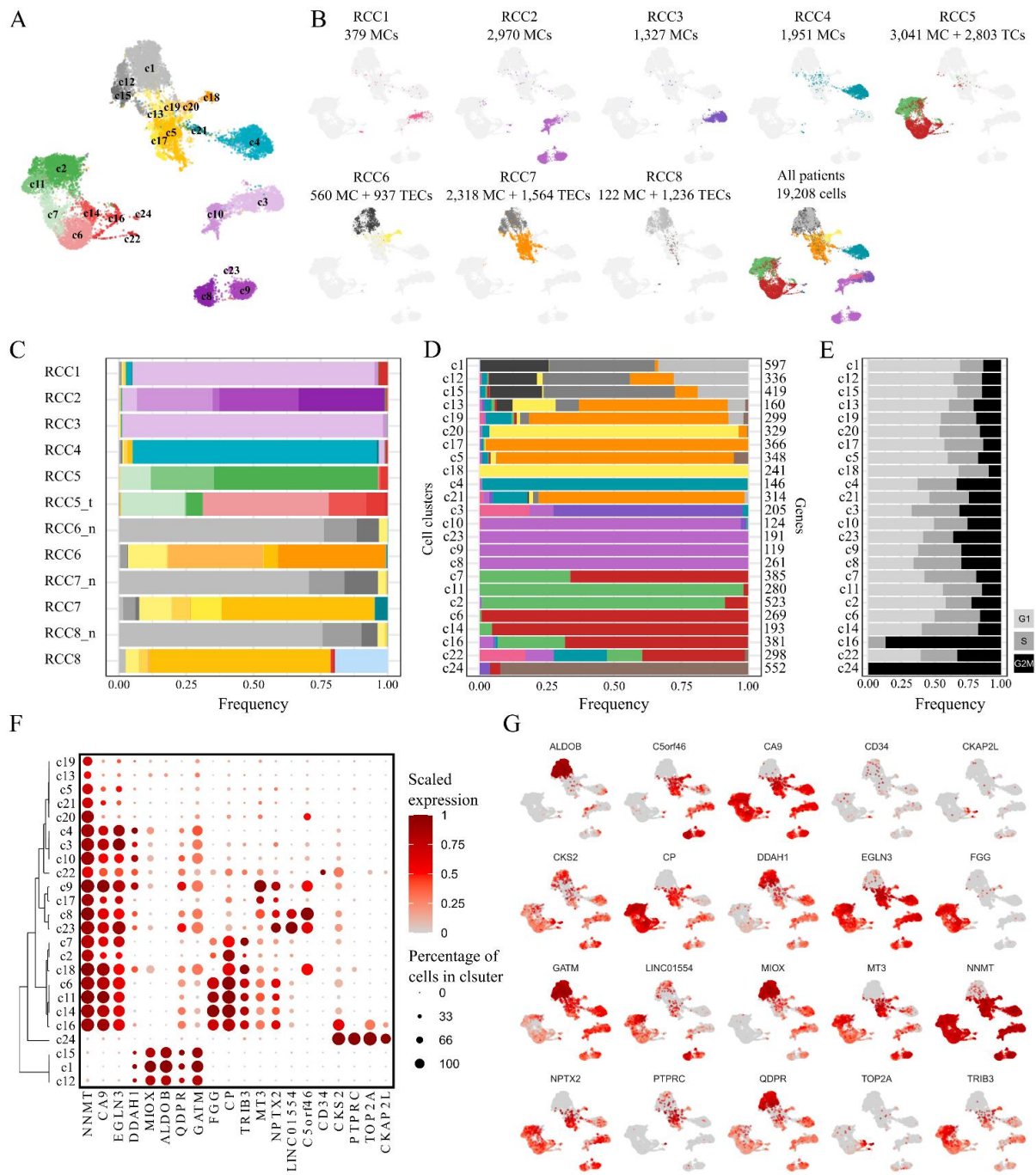




Fig. 4

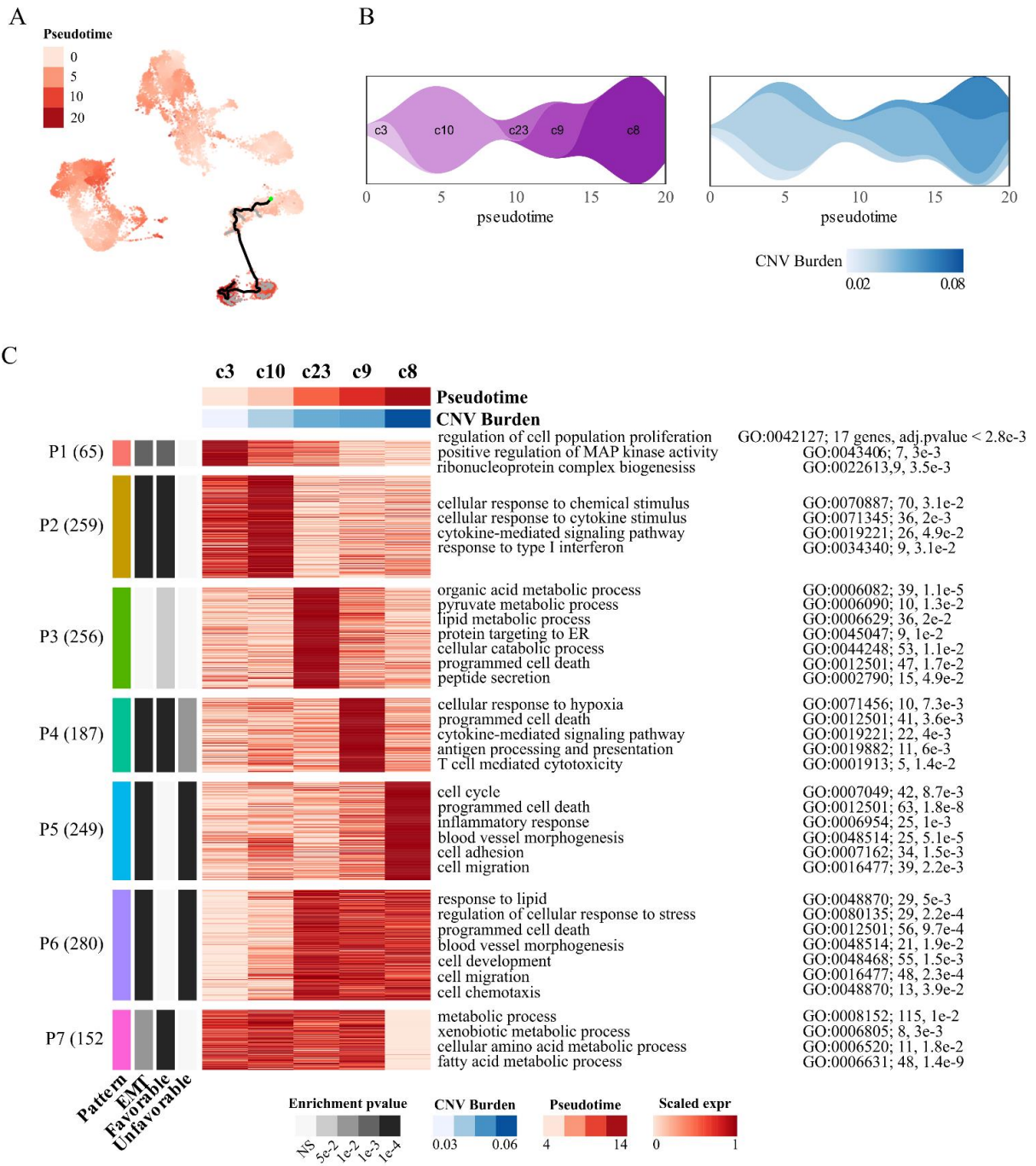
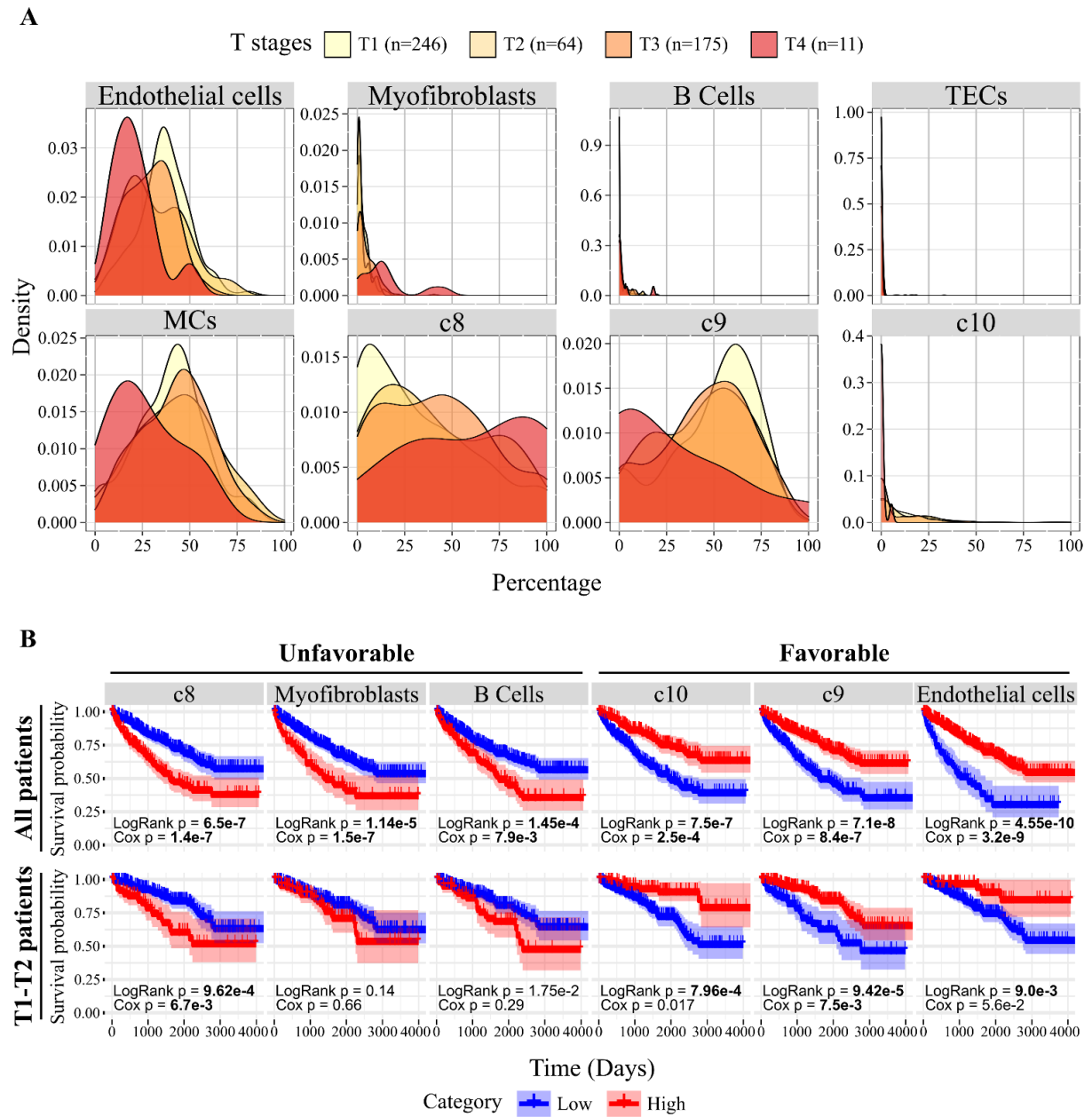


Fig. 5



## Supplementary figures

Fig. S1

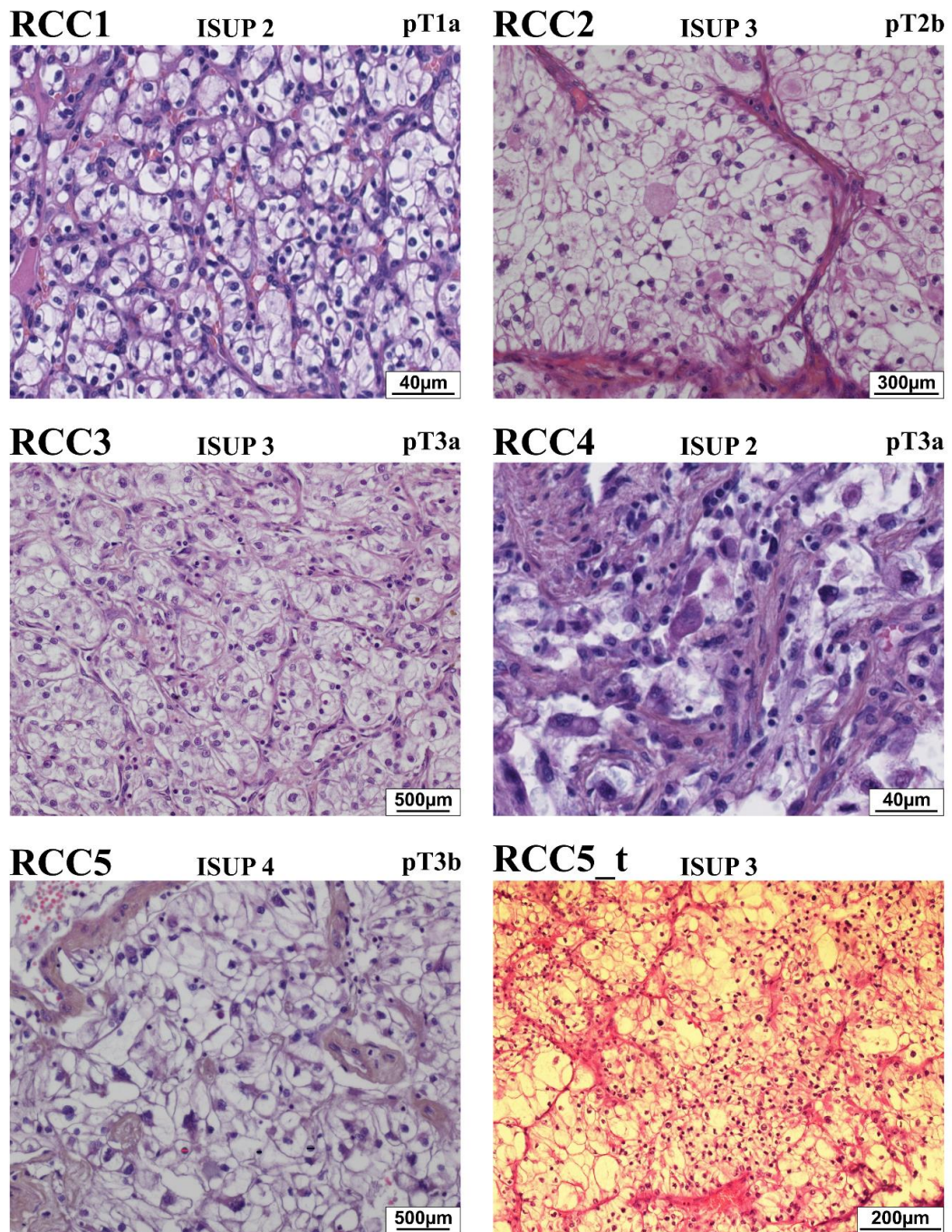
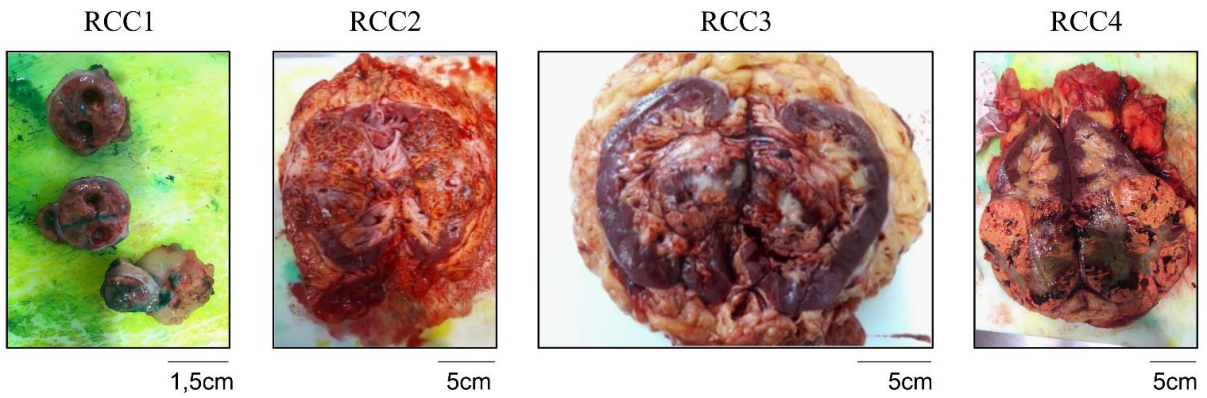


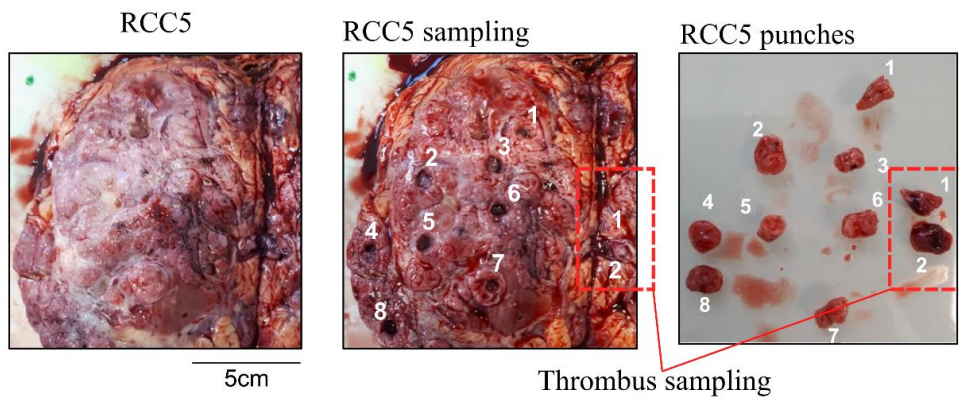


Fig. S2

A



B



C

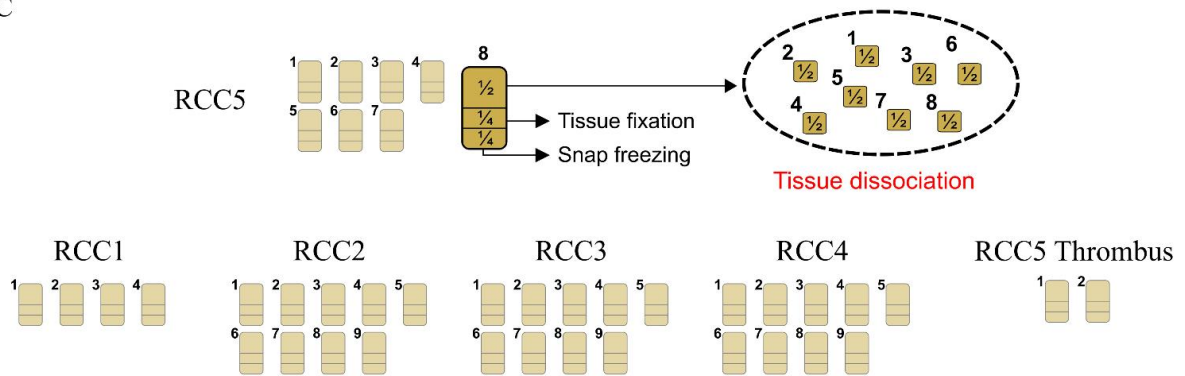


Fig. S3

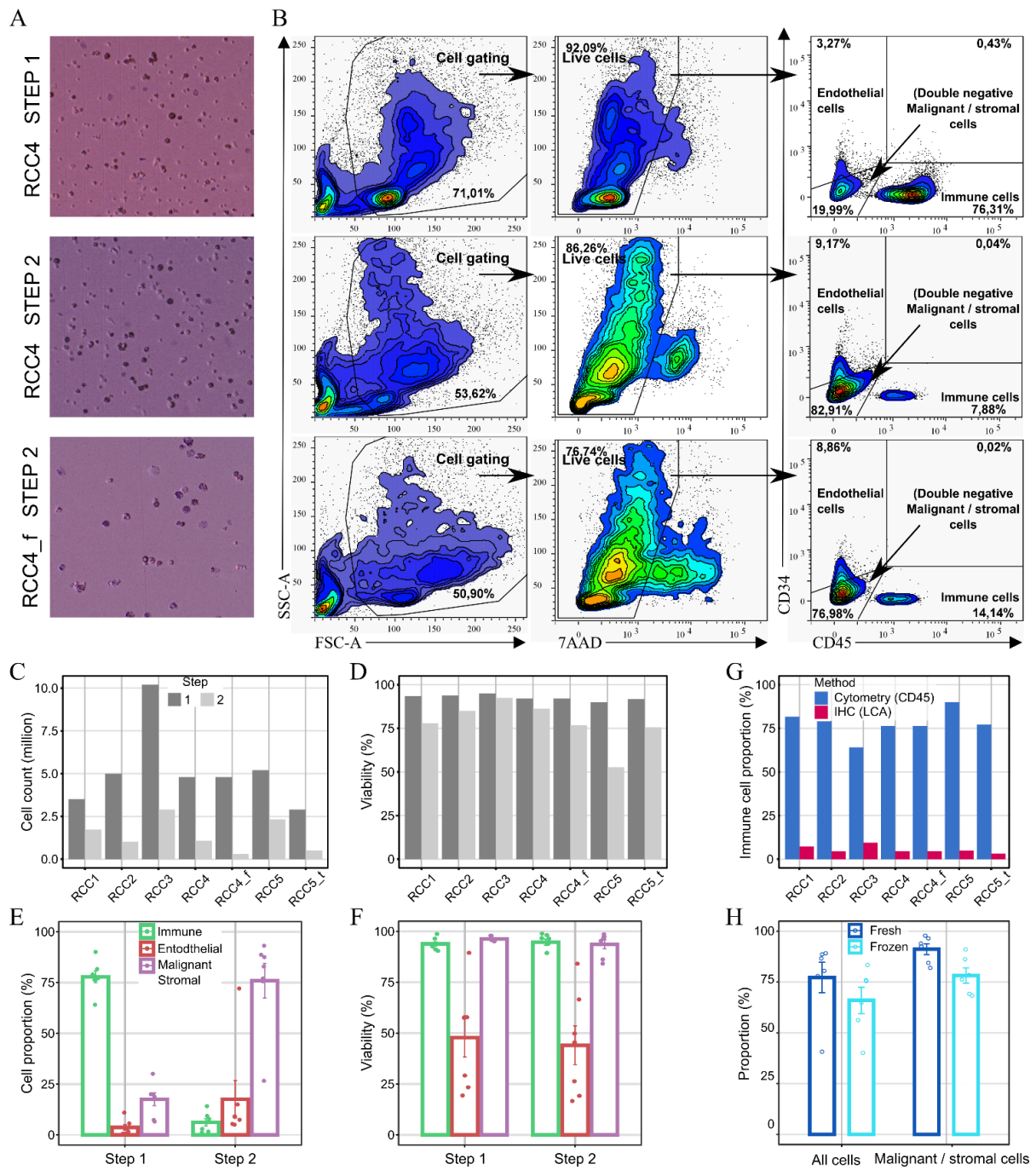




Fig. S4

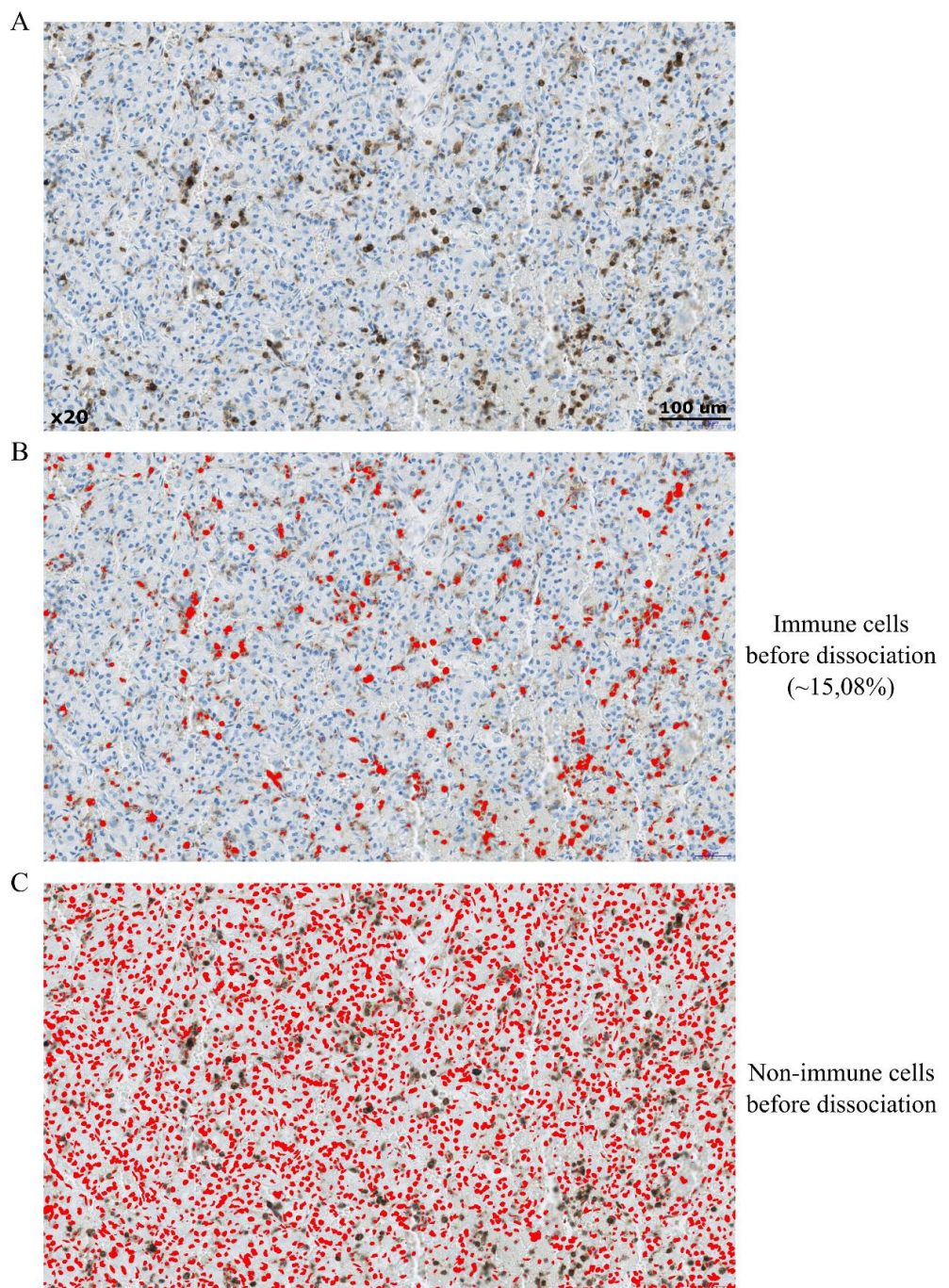
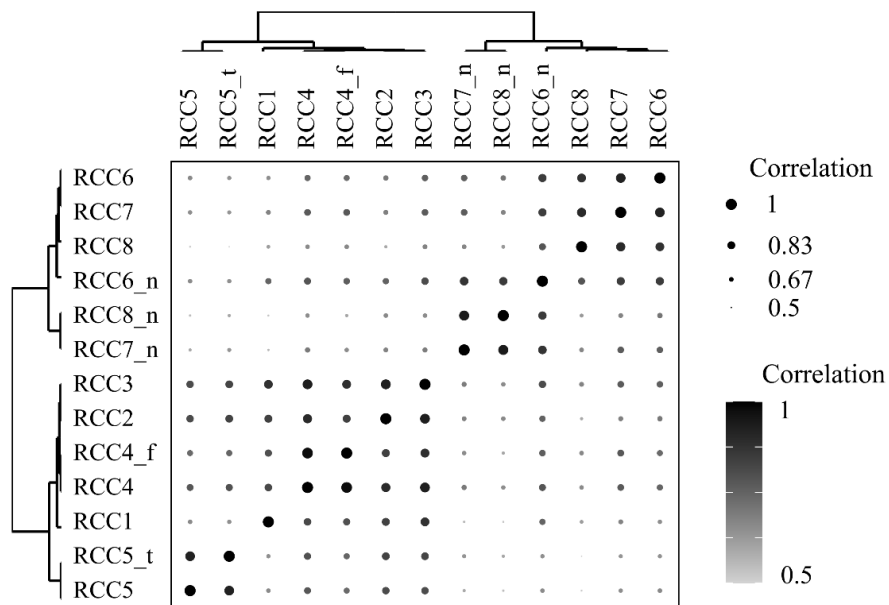


Fig. S5

A



B

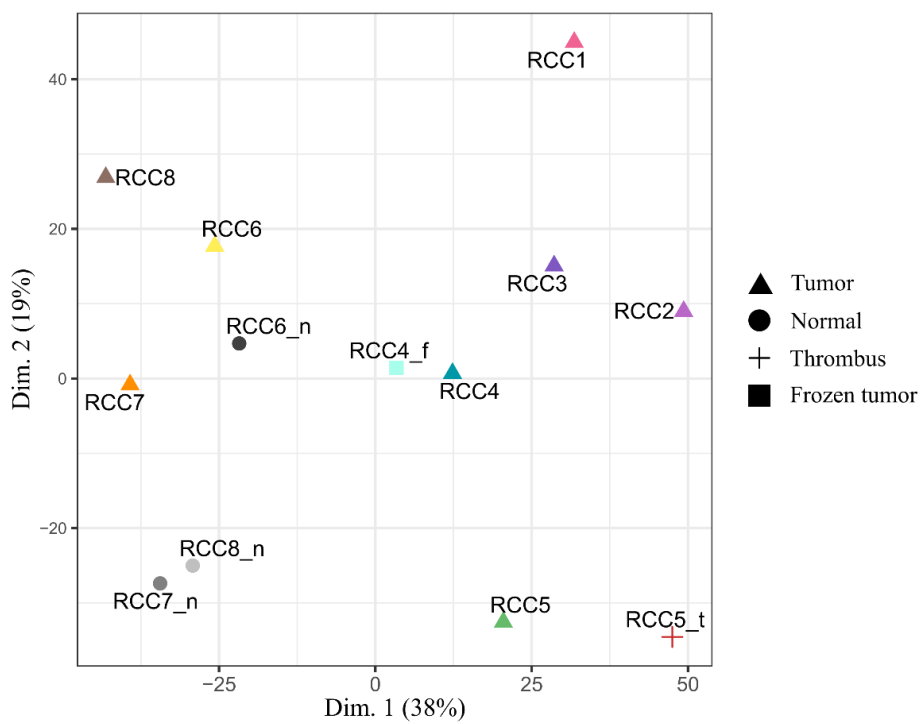
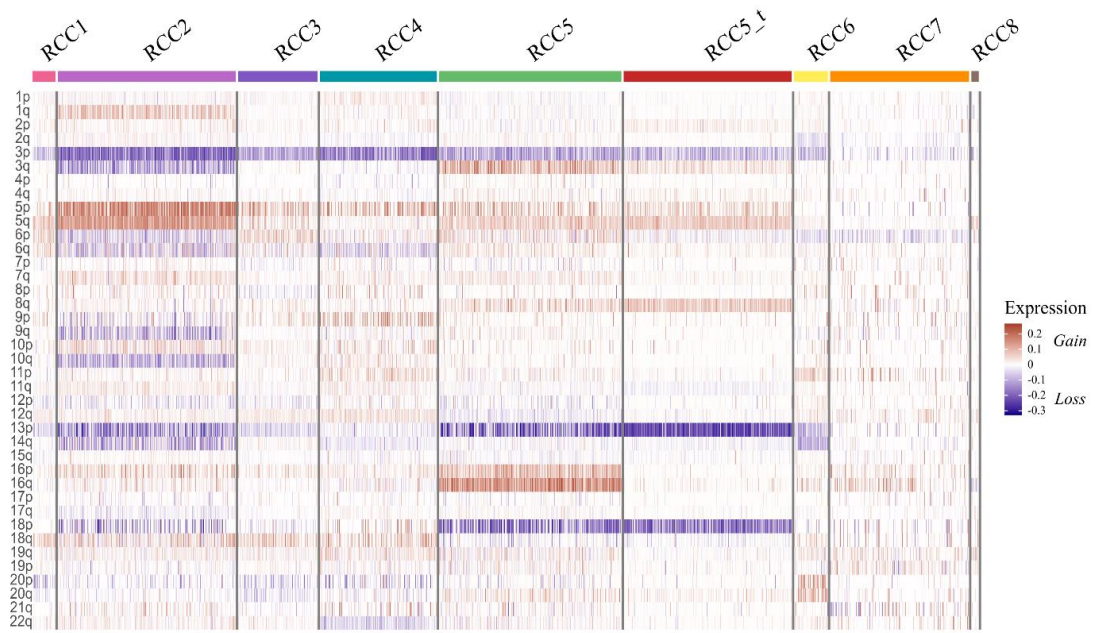


Fig. S6

A



B

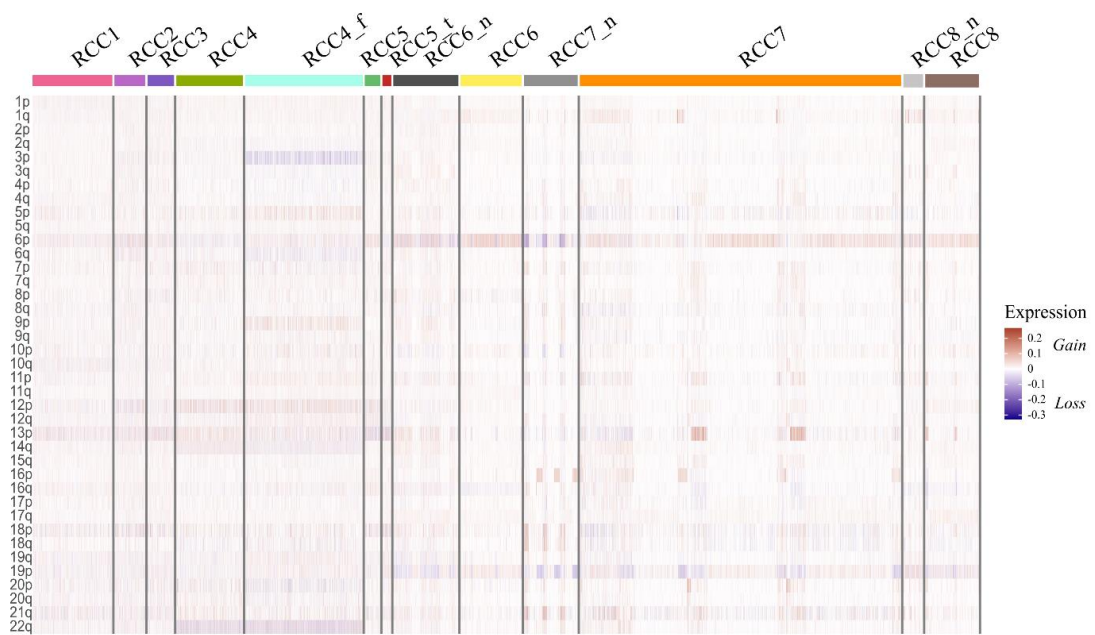


Fig. S7

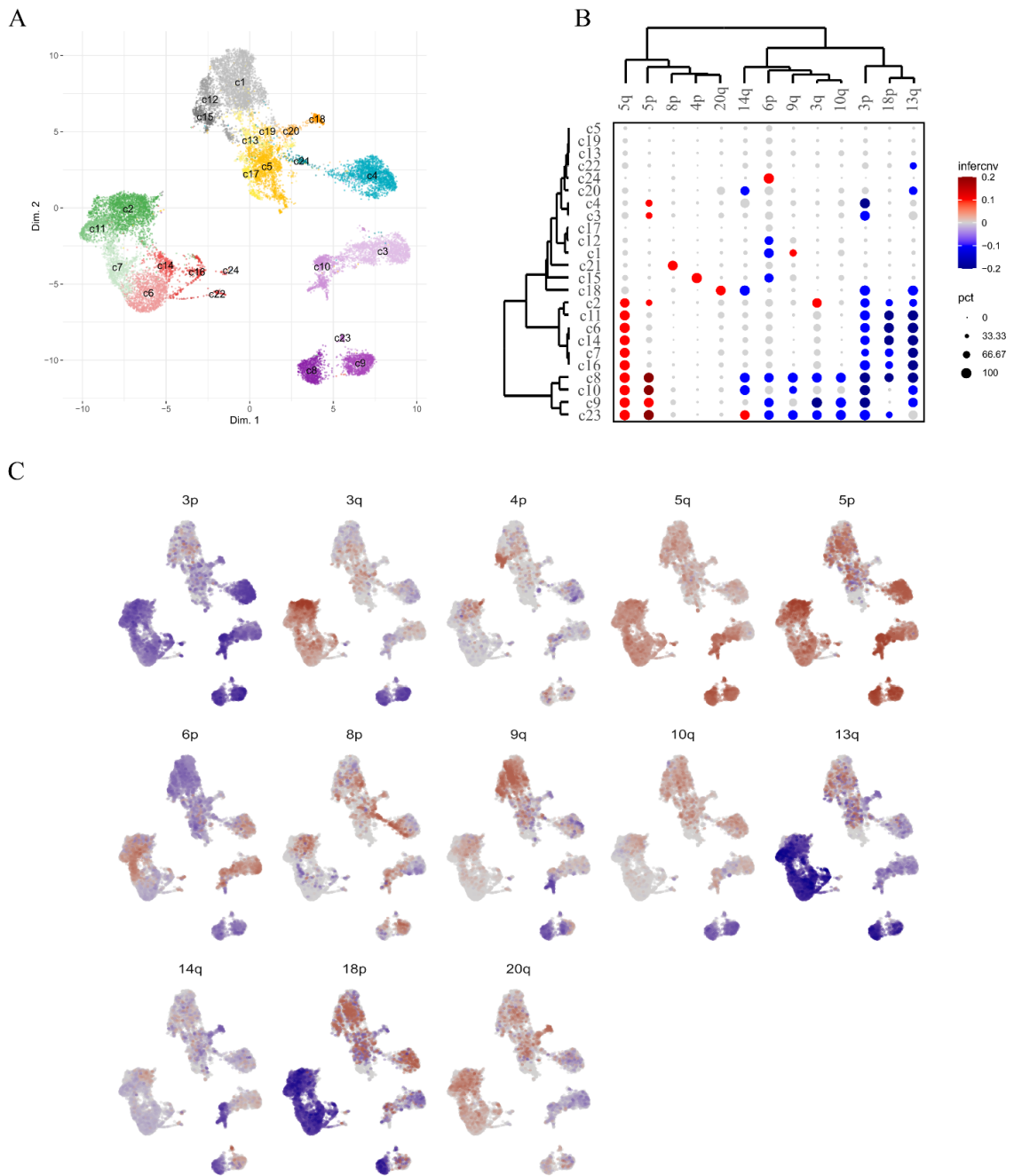




Fig. S8

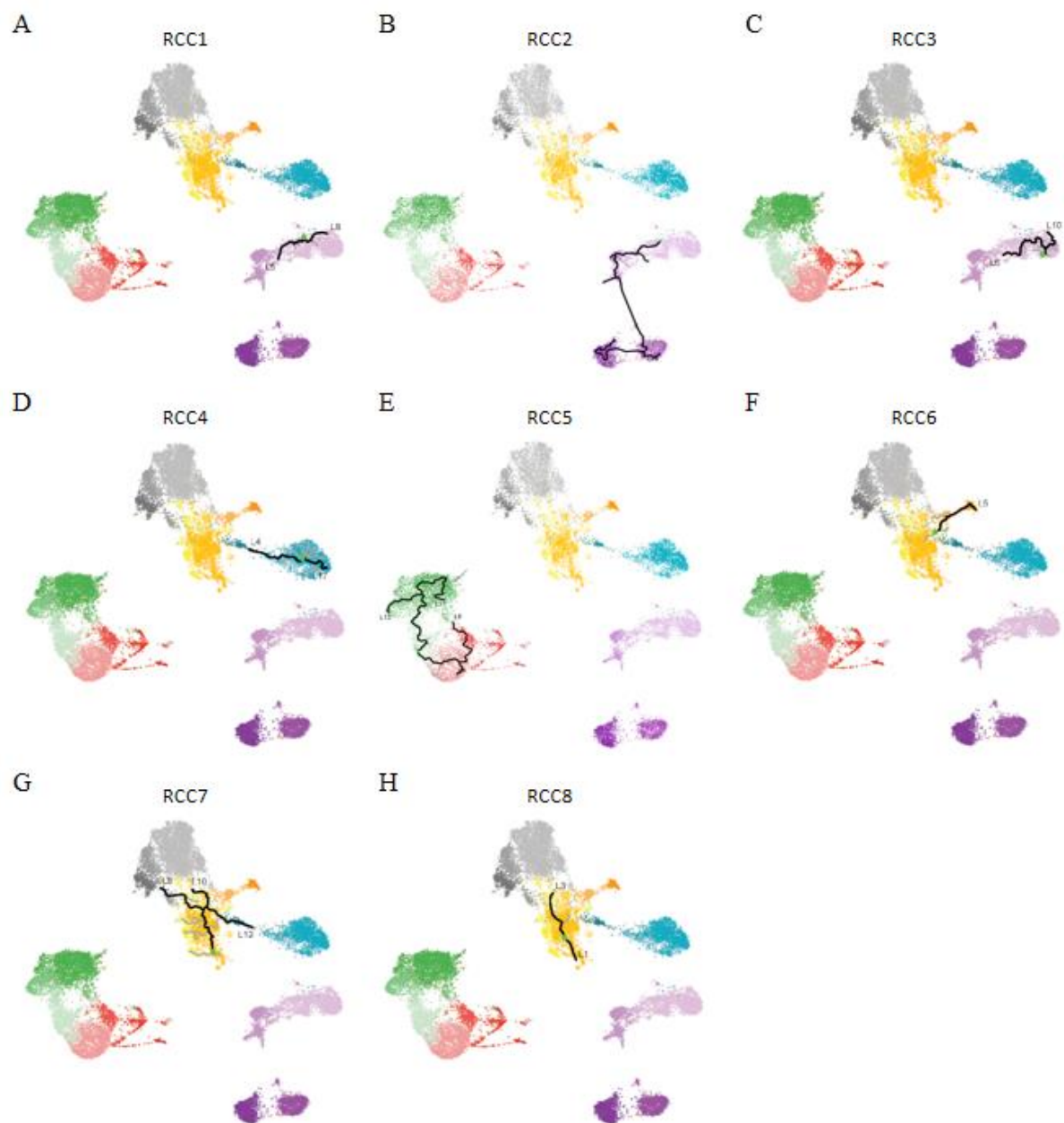
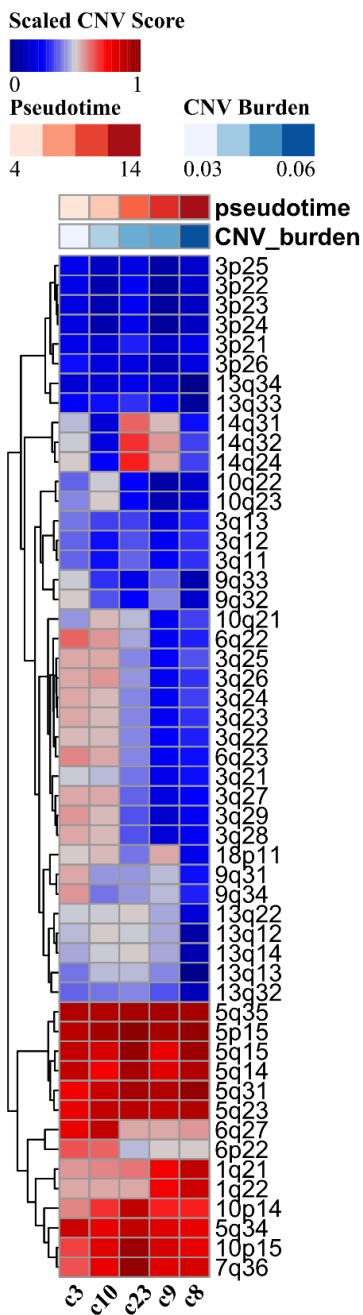
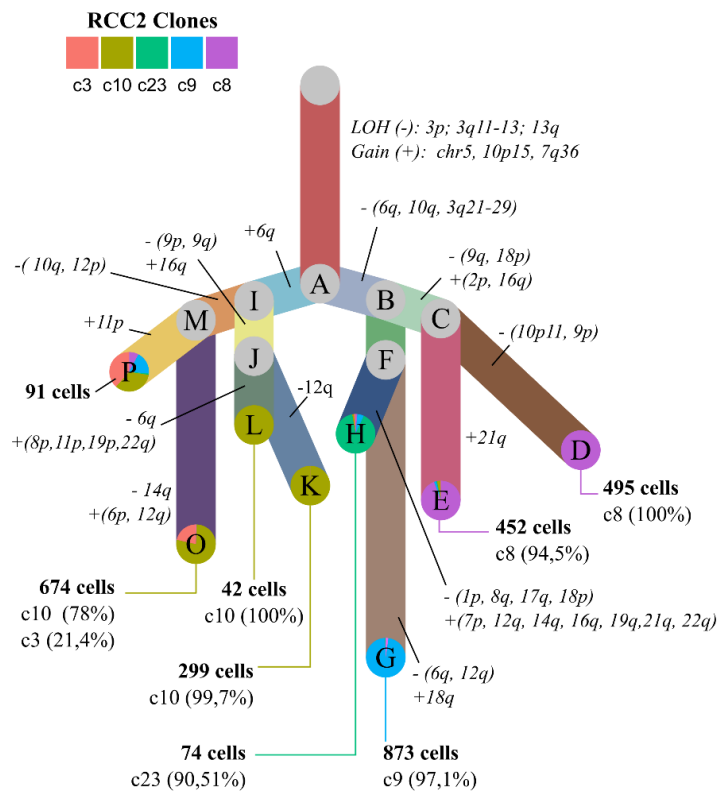


Fig. S9

A



B





# Supplementary Materials

## **Single-cell deconvolution of a specific malignant cell population as a poor prognostic biomarker in low-risk clear cell renal cell carcinoma patients**

Saout J., Lecuyer G., Leonard S., Evrard B., Kammerer-Jacquet SF., Noel L., Khene ZE., Mathieu R., Brunot A., Rolland AD.<sup>1</sup>, Bensalah K., Rioux-Leclercq N., Lardenois A., Chalmel F.

### **Supplementary methods**

#### **Ethics statement**

Human samples were obtained through the *Centre de Ressources Biologiques* (CRB) Santé of Rennes (ID in the national biobank infrastructure: BB-0033-00056). The research protocol was conducted according to French legal guidelines and fulfilled the requirements of the local institutional ethics committee. Informed consent was signed by each patient.

#### **Histological staining**

Additional samples of each ccRCC were fixed for 24h in 4% paraformaldehyde-PBS and paraffin embedded following standard protocols. Microscopic H&E features were acquired on the CaseViewer v2.2 software.

### **Tissue dissociation and tumor cell enrichment assessment**

For each tumor, punches were performed in locally distinct tumor areas without necrosis, fibrosis or edema. Punches were pooled and dissociated following a previously described protocol [1]. Briefly, samples were first subjected to enzymatic and mechanical disaggregation (tumor dissociation kit, Miltenyi Biotec) before undergoing red blood cell lysis and low-speed centrifugation steps to remove debris and dead cells while maintaining tumor cells (Fig. 1A, step1). Partial depletion of CD45 (an immune cell marker) positive cells was performed using the Miltenyi kit (CD45 MicroBeads, human) to enrich the samples in malignant/stromal cells and remove circulating and tumor-infiltrating immune cells (step2).

### **Microscopic observations**

Cell suspension quality was assessed using an Olympus IX70 microscope from Olympus-Lifesciences. Cells were stained with Trypan blue to estimate mortality rate and cells were counted with a hemocytometer.

### **Freezing and thawing of cells**

Approximately 300,000 cells were transferred in 1ml freezing solution consisting of 70% RPMI 1640, 20 % fetal bovine serum (FBS), 10 % DMSO (Sigma) and immediately placed at -20°C for 1h before their temperature was progressively lowered (-1°C/min) to -80°C for storage. Cells were then transferred to liquid nitrogen for long-term storage. Frozen samples were hand-thawed (at 37°C) and immediately diluted with 14 ml RPMI 1640 in a 15 ml conical tube. With the exception of RCC4\_f, scRNA-Seq was performed on freshly harvested samples without prior freezing.

## **Flow cytometry**

Cells were collected at two stages in the protocol: 1- after tissue disaggregation, RBC lysis, and low-speed centrifugation steps; and, 2- after partial CD45 depletion. Cells were preserved on ice for up to 2 hours before co-staining with anti-human CD45-FITC/CD34-PE monoclonal antibodies and 7-AAD (BD Biosciences, San Jose, CA) for 15 minutes at -4°C in the dark. Cell suspensions were washed twice in PBS. Samples were acquired with a LSR Fortessa X-20 (Becton Dickinson) cytometer using BD FACSDiva software v8. The data were analyzed using FlowLogic v7.2.1 software.

## **Estimation of tumor-infiltrating immune cells**

The proportion of infiltrating leukocytes was estimated in tumors from scanned images of H&E and LCA (Leukocyte Common Antigen which recognizes CD45; DAKO - ref: M0701) immunohistochemical staining, using the ImageJ suite. Five counting zones were randomly selected within the tumor tissue, and an average proportion of stained cells was calculated for each tumor. Unstained cells were considered as non-immune and identified as blue nuclei with the following color parameters : Hue ranging from 128 to 192; Saturation ranging from 32 to 255; and Brightness ranging from 128 to 255. Stained cells appearing as intense brown cells were scored as Immune cells and identified with the following color parameters : Hue ranging from 0 to 255; Saturation ranging from 0 to 255; and Brightness ranging from 0 to 128. For both counts, the image was changed to 8-bit. Watershed function was applied to distinguish cells within aggregates. Particles corresponding to either immune or non-immune cells were then automatically counted using the Analyze particle function, excluding particles under 35 and 50 pixels<sup>2</sup> for immune cells and non-immune cells, respectively.

## **Data analysis of a single-cell atlas composed of 13 samples**

*Pre-processing, quality controls, integration and normalization.* Demultiplexed raw sequencing reads were processed, mapped to the GRCh38 human reference genome and quantified based on unique

molecular identifiers (UMIs) with the Cell Ranger pipeline (v3.0.2, 10x Genomics). The resulting raw count matrices from the 7 individual samples were then merged. The cell-calling algorithm from Cell Ranger, which relies on the EmptyDrops method [2], was also used to establish a first set of potential valid cells. Subsequently, the scater R package. (v1.10.1) was used to remove outlier cells by using several cell features including the proportions of reads mapping to mitochondrial and ribosomal genes [3]. In parallel, doublets were filtered out independently in each individual matrix by using the DoubletFinder R package (v.2.0.2) [4]. Next, raw count matrix files from the scRNA-seq dataset published by Young and colleagues were downloaded from their online supplementary data [5]. Cells with fewer than 200 detected genes and genes detected in fewer than 10 cells were removed. Cells from both datasets were assigned a cell cycle phase by using both the Scran R package (v1.10.2) [6] and the Seurat v3.1.5 package [7]. Each dataset was individually normalized using the NormalizeData function implemented in Seurat. The two datasets were merged and batch-effect corrected using the Seurat procedure based on the FindIntegrationAnchors and IntegrateData functions with the following parameters: anchor.features=3000 and dims=30.

***Pseudobulk analysis.*** A “pseudobulk” count matrix resulting from the aggregation of counts of all selected cells of each sample was used to calculate pearson correlation coefficients and perform pairwise comparisons between samples. The “pseudobulk” count matrix was also used to perform a principal component analysis based on the FactoMineR package [8].

***Reduction of dimensionality and clustering.*** The top-3000 most varying genes were selected by using the FindVariableGenes function and used to perform a principal component analysis with the RunPCA function implemented in Seurat. Cells were then clustered by using the FindNeighbors and FindClusters functions on the top-30 principal components, with default parameters. Finally, we used the uniform manifold approximation and projection (UMAP) method implemented in Seurat to project single cells in a reduced 2D space.

***Cluster annotation and differential gene expression analysis.*** Cell clusters were annotated using a set of known marker genes complemented with genes that were significantly differentially expressed

between cell clusters. Differentially expressed genes were identified with the FindAllMarkers function implemented in Seurat, with default parameters.

### **Copy Number Variation analysis and phylogenetic tree construction**

The inferCNV R package (InferCNV of the Trinity CTAT Project : <https://github.com/broadinstitute/InferCNV>) was used to evaluate copy number variations (CNV) in malignant cells based on single-cell RNA-seq raw counts (InferCNV of the Trinity CTAT Project: <https://github.com/broadinstitute/InferCNV>). Inference was run separately on each individual sample, using tubular epithelial cells as reference. The cutoff parameter was set to 0.1, and parameters denoise, cluster\_by\_groups and HMM were set to TRUE. The tool returns a matrix with a CNV score corresponding to the relative intensity of gene expression for a particular cytoband in a given cell. We also calculated an overall CNV burden score for each cell corresponding to the weighted mean of the absolute CNV scores taking into account the cytoband size, using the following formula:

$$CNV \ Burden \ score = \frac{\sum_{i=1}^n |C_i| \cdot S_i}{\sum_{i=1}^n S_i}$$

with n being the number of cytobands, C the corresponding CNV score of cytoband i, and S the size of cytoband i.

A differential CNV score analysis was performed by comparing the CNV scores of normal tubular epithelial cells and malignant cells using a Wilcoxon test. The *p value* was then adjusted with the Benjamin Hochberg method and significant CNV scores in malignant cells were considered for further analysis. The phylogenetic tree was constructed from the outputs of inferCNV by using Uphyloplot2 [36] with parameter c = 1.

## **Single-cell RNA data analysis of malignant cells and tubular epithelial cells**

*Selection of malignant cells and tubular epithelial cells.* Cell populations belonging to clusters associated with tubular epithelial cells (c8, c12) and/or malignant cells (c2, c3, c5, c10, c16, c20, c30, c31) were selected for further analysis. The Seurat object was reduced to those selected cells by using the subset function implemented in Seurat.

*Reduction of dimensionality and clustering.* A multi-factorial analysis (MFA) was performed with the single-cell transcriptomic data (based on the top-3000 most varying genes) and the single-cell CNV matrix inferred by InferCNV (based on the cytobands with variance>0) (see **Copy Number Variation analysis and phylogenetic tree construction**) by using the FactoMineR package [8]. Cells were then clustered by using the FindNeighbors and FindClusters functions on the top-30 principal components of the MFA, with default parameters. Finally, we used the uniform manifold approximation and projection (UMAP) method implemented in Seurat to project single cells in a reduced 2D space based on those top-30 MFA components.

*Cluster annotation and differential gene expression analysis.* Cell clusters were annotated using a set of known marker genes complemented with genes that were significantly differentially expressed between cell clusters. Differentially expressed genes were identified by pairwise comparisons between all clusters using the FindMarkers function implemented in Seurat, with default parameters.

*Single-cell trajectory inference and pseudotime analysis.* Malignant cell trajectories were inferred with Monocle3 [9] on the scRNA-seq data. Each tumor sample was used as a distinct partition. Pseudotimes were estimated using the pseudotime function by indicating as starting cells those with the least chromosomal aberrations, *i.e.* the lowest CNV burden (see **Copy Number Variation analysis and phylogenetic tree construction**).

## **Processing of the Cancer Genome Atlas Kidney Renal Cell Carcinoma (TCGA-KIRC) data**

*Bulk RNA datasets of ccRCCs.* The raw count expression matrix from The Cancer Genome Atlas Kidney Renal Cell Carcinoma (TCGA-KIRC) and the associated clinical data (tumor stage, survival, last day of follow-up etc.) were downloaded using the TGCABiolink R package and the Broad Institute TCGA-GDAC (<http://gdac.broadinstitute.org/>), respectively. The gene prognostic annotations were derived from the TCGA-KIRC database.

*Survival analysis.* Survival analysis was performed using the Survival R package [10] (version 3.2-13). First, clinical data were censored according to the last day of follow-up and whether the patient was still alive or not. For the gene prognostic analysis, patients were categorized in either low or high expression groups according to the lowest Log Rank Test p-value. As defined by the Human Protein Atlas, *p value* < 0.001 reflects significant differences in survival outcomes between the two groups and defines the gene as prognostic.

*Selection of favorable and unfavorable sets of genes.* Significantly favorable and unfavorable genes for survival were defined by comparing the survival probability at 2000 days between the two expression groups.

## **Selection of gene sets of interest**

In addition to the sets of favorable and unfavorable (see previous section) genes, we also selected genes involved in the epithelial-mesenchymal transition (EMT), using a comprehensive collection of EMT-related genes via a web-accessible resource called dbEMT [11]. This database consists of a core dataset of 1,184 human genes involved in the EMT retrieved from 2,665 PubMed abstracts.

## **Functional analysis**

The AMEN suite of tools [12] was used to calculate the Fisher exact probability and the Gaussian hypergeometric test was used to identify significantly enriched terms from the Gene Ontology as well

as from the selected sets of genes (see previous section). A specific term was considered enriched in a group of coexpressed genes if the adjusted (Benjamini–Hochberg procedure) *p* value was  $\leq 0.05$  and the number of genes bearing a given annotation was  $\geq 3$ .

### **RNA deconvolution and survival model**

**Bulk RNA deconvolution.** Deconvolution was performed with the MuSiC R package v0.2.0 [13] on HTSEQ raw counts of TCGA bulk data, using single-cell RNA-seq data from either RCC2 malignant cell clusters or somatic cells across all patients as a reference. A proportion of putative RCC2 malignant clusters was then inferred in each TCGA cohort patient. Malignant cell clusters called “Relative” (such as “Relative c8”) correspond to the proportion of the considered malignant cells as compared to all RCC2 malignant cells.

**Survival analysis.** Survival analysis was performed using the Survival R package (version 3.2-13) as previously described (**The Cancer Genome Atlas Kidney Renal Cell Carcinoma (TCGA-KIRC) processing** section). Two statistical methods were implemented for the deconvolution-based survival analysis: the Cox Proportional-Hazards Model method, and the Log Rank Test in which patients were categorized as having either a low or high proportion of a given cell type according to the lowest Log Rank Test p-value - method similar to that implemented in The Human Protein Atlas [14]. P-value  $< 0.01$  reflects significant differences in survival outcomes between the two groups and defines the cell type as prognostic. Finally, survival analysis curves were produced using the Kaplan-Meier model.

### **Heatmap and spot plot**

Spot plots were generated using the FlexDotPlot package [15]. Heatmaps were generated with the pheatmap package v1.0.12 (<https://CRAN.R-project.org/package=pheatmap>).



**Data availability**

Raw data are available on the EGA (European Genome-Phenome Archive) portal under accession number EGAS00001006534. Processed single-cell RNA-seq and available on ArrayExpress portal under accession number E-MTAB-11310. All processed data are also available via the web-based portal UncoVer (The Uro-oncogenomics Viewer at <https://uncover.genouest.org>) (Lecuyer, Saout *et al.*, in preparation).

## **Supplementary results**

### **Section 1 : Tumors displayed various clinical features**

Tumors were classified from pT1a to pT3b and from ISUP grade 2 to 4, according to the TNM and 2017 WHO/ISUP classification systems, respectively (Table S1; Fig. S1). Four samples came from distinct primary ccRCC patients (RCC1 to RCC4) and two samples consisting of matched primary tumor and distant vena cava tumor thrombus came from the same patient (RCC5 and RCC5\_t, respectively). Macroscopic examinations revealed various appearances with heterogeneous patterns of vascularization, hemorrhagic and necrotic remodeling (Table S1; Fig. S2A). These features of heterogeneity led us to sample each tumor in up to 9 distinct macroscopic areas, depending on the size of the lesion (Fig. S2B-C).

### **Section 2 : Experimental approach to enrich malignant cells in cell suspensions from ccRCC tumors**

Tumor samples were processed following a preset protocol including tissue dissociation followed by low-speed centrifugation, referred to as step1; and depletion of CD45+ cells so as to remove circulating and infiltrating immune cells (IC) and enrich malignant cells (MCs), referred to as step-2 (Fig. 1A). In order to assess the impact of cryo-preservation, we froze part of the RCC4 cell suspension (RCC4\_f) and then processed it the same way as the fresh samples. As scRNA-seq is a sensitive technique requiring thorough cell dissociation with a high preservation of cell viability, we first evaluated cell viability and proportions of various cell subsets at the different steps of the protocol as illustrated with RCC4 in Fig. S3A-B. Microscopic examination at each step revealed well individualized and diverse cells, in terms of shape and size (Fig. S3A). To further confirm the distribution and integrity of different cell types, we used flow cytometry to analyze immune cells (CD45+), endothelial cells (CD34+), stromal and MCs (CD45-/CD3-) and viability (7AAD) (Fig. S3B). Cell counts decreased from step 1 to step 2 in each sample, sometimes dramatically (Fig. S3C), while global cell viability slightly decreased (Fig. S3D). However, the proportion of viable cells remained the same in each of the monitored cell types (Fig.

S3F), suggesting the decrease in cell viability could be due to the removal of ICs given their relatively high rate of viability. ICs are globally predominant in the samples (Fig. S3B-E) and highly viable (above 90%) in each sample. Malignant, stromal and endothelial cells were efficiently enriched by the immune depletion (Fig. S3B-E). With the exception of RCC5, the most clinically advanced tumor (pT3b), the cell viability rate exceeded the 70% minimum viability threshold recommended for single-cell RNA sequencing experiments (Fig. S3D). Among the different cell types, endothelial cells were the most impacted by the tissue dissociation, with less than 50% cell viability at step 1 and 2, whereas malignant and stromal cells remained highly viable. To assess the impact of cryopreservation, we froze dissociated single cells from 6 ccRCCs, including RCC4\_f already frozen and thawed, and observed a slight decrease in their overall viability, suggesting that cells may be harmed by freezing (Fig. S3H). However, malignant and stromal cells still presented a viability rate above 75%. To better document the nature of the depleted ICs, we compared their proportion by flow cytometry at step 1 and immunohistochemistry prior to dissociation (Fig. S3D and Fig. S4). In comparison, the estimated proportion of ICs by flow cytometry was ~8-23 times higher than by IHC (ranging from 3–10%) (Table S2). This difference could result from the resistance and individualized state of ICs, facilitating their selection during dissociation, or contrasted sensitivities of the two techniques. However, given the large difference, these results suggest that a substantial proportion of ICs detected by flow cytometry may correspond to circulating ICs coming from blood circulation in the tumor sample.

### **Section 3 : Single-cell RNA-seq uncovers the cellular composition of ccRCC tumors**

The final set of samples was composed of seven samples from our lab (RCC1-5, RCC4\_f and RCC5\_t) complemented with three ccRCC samples (RCC6-8) and three healthy samples (RCC[6-8]\_n) published by Young and colleagues [5]. After quality controls, the entire dataset contained the expression levels of 37,832 genes in 54,812 cells (Fig. 1B). The number of detected cells per sample ranged from 2,354 and 4,549 in our dataset (sd = 796), and from 2,001 to 14,769 in the additional dataset (sd = 4,912) (Table S2). The median number of detected genes per sample ranged from 572 to 2,917, with more

genes in our dataset as compared to that of Young et al. (median of 1,719 versus 868), probably as a result of deeper sequencing.

In order to get an overview of the transcriptomic proximity between samples, we first aggregated their cells' transcriptomes and performed pairwise comparisons between the resulting pseudobulk samples (Fig. S5A). Samples sharing the same experimental and biological background were strongly correlated and distributed in three groups, consisting of samples from our dataset (RCC1 to RCC5), and tumor or normal samples from additional dataset (RCC6 to RCC8). From this "pseudobulk" expression matrix, we projected samples on the first two components of the PCA (Fig. S5B). Accordingly, the first component distinguishes samples depending on which dataset they belong to, suggesting differences in their experimental processing. Interestingly, strong correlations are observed between biologically related samples such as RCC4/RCC4\_f (0.99) and RCC5/RCC5\_t (0.96). Given the differences between tumor samples, the dataset was batch-effect corrected using *IntegrateData* function in Seurat.

The transcriptomic signal was then analyzed at the single cell resolution by performing dimensionality reduction on the single-cell count matrix and projecting single cells on a UMAP (Uniform Manifold Approximation and Projection) 2D space (Fig. 1C). The resulting single-cell transcriptomic atlas gives a general idea of the overall cell diversity of the ccRCC samples. We annotated single cells according to their sample of origin, which revealed various levels of intra-tumor heterogeneities, as each sample contained three or more broad cell populations. In accordance with the pseudobulk analysis, RCC5 primary and thrombus derived cells, although distinguishable, appeared to be similarly distributed. More remarkably, fresh and frozen cells from RCC4 perfectly overlapped (Fig. 1D), demonstrating that the freezing step in our experimental protocol has little to no impact on single-cell transcriptomes. This indicates the feasibility of performing scRNA-Seq on frozen ccRCC dissociated cells, enabling the simultaneous analysis of samples that would normally be acquired at different time points, such as matched primary tumor and metastasis samples. Alternatively, single-nuclei RNA-seq techniques may be optimal to analyze frozen samples with only a slight decrease in sensitivity [16].

To identify cell types present in ccRCC, we performed clustering on the top-30 components of the PCA matrix, and identified 35 cell clusters (denominated c1-c35) (Fig. 2A) which were annotated based on

the expression of well known cellular markers (Fig. 2B-D). MCs (associated with eight cell clusters) were identified based on the expression of CA9 (Fig. 2B) and NNMT (Fig. 2C), two well-known canonical markers [17,18]. Fourteen clusters were classified as ICs as they expressed canonical markers such as PTPRC, MS4A1 or IGHA2. The remaining clusters were classified as endothelial cells based on CD34 and PECAM1 (c1, c15, c27, c28, c29); pericytes based on ACTA2 and PDGFRB (c4, c23); myofibroblasts based on LUM, DCN and residual expression of pericytes markers (c34); collecting duct epithelial cells based on UMOD (c26, c33); tubular epithelial cells based on GATM (c8); erythrocytes based on HBB (c18); and finally cluster c12 exhibiting both expression of tubular epithelial cell markers (GATM, GPX3) and MC markers (CA9 and NNMT) that corresponds to an intermediate status between the two cell types.

To confirm the identity of MCs and further characterize them, we used inferCNV (inferCNV of the Trinity CTAT Project. <https://github.com/broadinstitute/inferCNV>) and checked the 3p chromosomal arm status across normal and MCs, as its deletion is a key initial event resulting in VHL loss of function in more than 90% of ccRCCs [19,20]. This approach confirmed the 3p deletion and revealed its occurrence in almost all of the cells identified as malignant (Fig. S6), although the sequencing depth and coverage were too weak to categorically confirm this observation in the RCC7 MCs. Interestingly, we also found some of the most common CNVs such as 5q and 5p gains (found in 60% and 26% of ccRCC, respectively) or 13q loss (24% of ccRCC) [20] (Fig. S7). Hence, the inference appeared to accurately reflect the CNV landscape of tumor cells and constitutes an additional marker of heterogeneity in the MCs. Interestingly, we also uncovered poorly described chromosomal aberrations with an heterogeneous distribution across the tumors, such as a very specific gain in chromosome 16 for RCC5 primary MCs or a significant loss in 18p for RCC2 and RCC5.

#### **Section 4: Generation of a single-cell atlas of malignant cells in ccRCCs**

We assembled a single-cell atlas of MCs by selecting all the cells belonging to clusters associated with MCs (c2, c3, c5, c10, c16, c20, c30, c31) and normal tubular epithelial cells (TECs) (c8, c12) and

reanalyzed the resulting cells (15,471 MCs and 3,737 TECs). The dimensionality reduction was performed by using a multi-factorial analysis (MFA) based on the scRNA-seq data and the inferred CNV data. Finally cells were projected on a 2D space (UMAP) and partitioned into 24 cell clusters.

### **Section 5 : CNV heterogeneity in RCC2 malignant cells**

Losses in 3p21-26 and 13q33-34, as well as gains of 1q21-22, 7q36, 10p14-15 and 5q were observed in all RCC2 MC clusters, designating them as early events in its initiation (Fig. S9A). Although the rate of gain was approximately the same, the losses increased with pseudotime, culminating in c8, in accordance with the global CNV burden. Based on this CNV accumulation we used UphyloPlot2 [21] to construct a phylogenetic tree (Fig. S9B), which revealed a branched evolution from a common ancestor (A) to distinct malignant clones (P, O, L, K, H, G, E and D). Interestingly, we observed that CNV-based clones overlap with transcriptomic clusters. In accordance with previous observations, the loss of 3p and the gain of 5q appeared to be common to all clones whereas some specific events were restricted, like the numerous specific aberrations in branch H, corresponding to c23. The phylogenetic tree, although probably incomplete, gives a broad idea of the diversity and the relationships between the MCs within RCC2, further supporting the idea of diverse behaviors.

The 3p arm, lost in all MCs, carries several genes involved in ccRCC carcinogenesis, such as SETD2, BAP1 and PBRM1, whose losses and their consequences are widely observed and well documented in ccRCC [19]. The 3q arm was also lost to various extents in certain MC clusters; larger proportions were affected in c23, c8 and c9 than in c3 and c10. In particular, band 3q26, lost at the end of the tumor lineage, carries the PIK3CA oncogene, which was found to be mutated in different ccRCC cohorts [22]. Interestingly, PIK3CA is part of the same biological pathway as PTEN, a common tumor suppressor on band 10q23, which happens to also be lost specifically in c23, c9 and c8. Additionally, part of c8, harbors a unique 9p loss (Fig. S9B), which carries CDKN2A (9p21), an important suppressor gene. Interestingly, the loss of 9p21 is observed in larger proportions of metastases than primary tumors which makes it a

likely cause of metastasis in ccRCC [22]. Altogether, RCC2 MC clusters carry the most chromosomal alterations, in particular in c8, which could infer to it particularly aggressive properties.

## **Section 6: Dynamics of the transcriptional landscape across the RCC2 tumor lineage**

In total we identified 1,448 genes differentially expressed along the RCC2 tumor lineage which were subsequently distributed into 7 expression patterns (P1-7) (Fig. 4C, Table S5). Patterns P1, P3, P4 and P5 include genes that are preferentially expressed in c3, c23, c9 and c8, respectively, whereas P2, P6 and P7 contain genes whose expression profiles are shared by several adjacent MC clusters: P2 contains genes that are highly expressed in c3 and to a greater extent in c10; P6 includes genes expressed in c23 onwards; and, P7 regroups genes that are specifically down-regulated in c8. The functional analysis revealed a progression towards an increasingly unfavorable score (corresponding to the significant over-representation of unfavorable genes within a given expression pattern) from P1 to P6, which reflects a gain in aggressive features along the tumor clonal lineage. Conversely the favorable score was globally higher in the first expression patterns (P1-4) and in P7 than in P5-6. With the exception of P3, all expression patterns displayed a significant enrichment of genes involved in the EMT. Gene ontology (GO) analysis revealed 884 GO terms that were significantly enriched in at least one expression pattern (Table S6). P1 was significantly associated with genes involved in positive regulation of MAP kinase activity (7 genes,  $p$  values  $< 3e-3$ ), regulation of cell proliferation (17,  $2.8e-3$ ) and ribonucleoprotein complex biogenesis (9,  $3.5e-3$ ); P2 was found to be enriched in genes associated with cellular response to cytokine stimuli (36,  $2e-3$ ) and response to type I interferon (9,  $3.1e-2$ ), indicating innate immune activity in c3 and c10; genes belonging to P3 were over-represented in organic acid (39,  $1.1e-5$ ), pyruvate (10,  $1.3e-2$ ) and lipid metabolic processes (36,  $2e-2$ ) as well as protein targeting to the endoplasmic reticulum (9,  $1e-2$ ), peptide secretion (15,  $4.9e-2$ ) and programmed cell death (47,  $1.7e-2$ ); P4 genes were significantly associated with cytokine-mediated signaling pathways (22,  $4e-3$ ), antigen processing and presentation (11,  $6e-3$ ), and more precisely T cell mediated cytotoxicity (5,  $1.4e-2$ ) suggesting the organization of an adaptive immune response in c9. Interestingly, P4 was also

characterized by genes involved in programmed cell death (41, 3.6e-3), and is further linked to cellular responses to hypoxia (10, 7.3e-3), the initiation stage of angiogenesis [23]. As expected, P5 was found to be enriched in terms linked to aggressive tumor states, including inflammatory responses (25, 1e-3), blood vessel morphogenesis (25, 5.1e-5), cell cycle (42, 8.7e-3), cell migration (39, 2.2e-3), programmed cell death (63, 1.8e-8) and cell adhesion (34, 1.5e-3); P6 was over-represented in genes involved in programmed cell death (56, 9.7e-4), blood vessel morphogenesis (21, 1.9e-2) cell migration (48, 2.3e-4), cell development (55, 1.5e-3) and cell chemotaxis (13, 3.9e-2) illustrating a more aggressive state from c23 cluster onwards. Finally, P7 was characterized by a significant association with genes involved in global metabolic processes (115, 1e-2), and more specifically xenobiotic (8, 3e-3), cellular amino acid (11, 1.8e-2) and fatty acid (48, 1.4e-9) metabolic processes, suggesting that c8 has an altered metabolic state and may respond differently to therapy.

Gene level analysis corroborates previous observations regarding the aggressiveness of patterns P5 and P6 (Table S7). Amongst the 1,448 DEGs, we found well-known oncogenes already described in renal cell carcinoma such as the transcription factors JUN/JUNB (P5) and AR (P6), the epithelial growth factor receptor (EGFR, P5) or CXCR4 (P6). A total of 181 genes were linked to EMT, including 81 in the P5 and P6 patterns. There were 464 genes associated with at least one GO term among angiogenesis, cell migration, cell adhesion or programmed cell death, including 202 for the P5 and P6 patterns. Twenty-three of them were associated with all terms, including 13 genes in the P5 and P6 patterns, such as CAV1 [24] and CCL2 [25] that have been reported in ccRCCs. Among genes that haven't been previously described in RCCs, 98 were unfavorable in the TCGA cohort. For instance, RRAS and CALR were associated with the regulation of cell migration and adhesion as well as with angiogenesis (RRAS) and programmed cell death (CARL), warranting deeper characterization of their impact in ccRCC. A total of 16 genes coded proteins that are targeted by FDA-approved drugs, including three gene products already targeted in kidney cancer (AXL, NTRK2 and VEGFA). AXL, in particular, showed peak expression in c23, c9 and c8, suggesting its inhibition with cabozantinib could counter the effects of the most aggressive MC clusters in RCC2. Similarly, 19 gene products, although not targeted by drugs currently used to treat ccRCCs, are reported to facilitate the action of FDA-approved drugs ("Tclin"),



warranting further characterization of their status in ccRCCs. Among other notable genes, P6 contains MMP7, the metalloproteinase involved in extracellular matrix (ECM) degradation that promotes the invasion of MCs into the ECM [26], and PI3KR1 (Phosphatidylinositol 3-kinase regulatory subunit alpha), which is part of the highly activated PI3K/AKT pathway promoting cell proliferation in renal cell carcinomas [27]. P5 contains TGFBI (transforming growth factor beta induced), a cell adhesion inhibitor in the ECM [28] and MDK which is involved in cancer progression and EMT, notably modulating cell adhesion and migration [29]. Altogether, the functional analysis (Table S6) and the RCC2 DEGs annotations (Table S7) demonstrate the heterogeneous and aggressive states of the RCC2 tumor from the molecular standpoint. We also illustrate here a global method to identify relevant targets, warranting single-cell transcriptomics profiling efforts to support the design of novel targeted therapies.

## **Section 7: Deconvolution of RCC2 malignant cell subpopulations are associated with patient survival**

Based on log rank tests run on all the patients, we established groups having low versus high proportions of the deconvoluted cell types by selecting the percentile displaying the most significant outcome difference (Table S9), for instance in c8 (percentile: 43.69; mean percentage in low group: 17.13%, CI[15.30-19.04] versus 70.10% [68.12-72.01] in high group), c9 (48.2; 23.70% [21.8,25.61] versus 64.73% [63.42,66.04]), or endothelial cells (19.80; 12.72 [12.28,13.16 versus 37.83 [35.65,40.01]). Outcome was significantly unfavorable ( $p$  value < 0.01) among all the patients for those having high proportions of c8 among MCs (called “Relative c8”, see supplementary materials;  $p$  < 6.5e-7), and to a lesser extent myofibroblasts (<1.14e-5) and B cells (<1.5e-5) (Fig. 5B). Conversely, high proportions of relative c9 (<7.1e-8), endothelial cells (<4.55e-10) as well as relative c10 (<7.5e-7), despite its low levels, were significantly favorable factors. These associations with survival were also significant according to the Cox regression test. In high-risk patients (T3-T4), all previous cell types but B cells had a significant association with survival. However, these patients most probably received therapies impacting their survival, which confounds the analysis. Very interestingly though, when considering

low-risk patients (T1-T2), endothelial ( $p < 9.0e-3$ ), but most significantly relative c8 ( $< 9.62e-4$ ) and c9 ( $< 9.42e-5$ ) remained predictive of survival, further discriminating them. For instance, patients displaying a high relative c8 had an approximately 25% lower 5-year survival probability than those with a low relative c8. Taking into account the confidence intervals, patients displaying more than 23.5% of relative c8 might be considered at higher risk than expected amongst T1-T2 patients. Overall the relative proportions of c8 and c9 among MCs constituted the two best poor and good prognostic factors, respectively, for survival in low-risk patients.

## Supplementary references

- [1] Bertrand E, Judikael S, Aurelie L, Solène-Florence K-J, Nathalie R-L, Frederic C. Human Kidney Tumor Dissociation for single-cell genomics. *Protocols.io* 2020. <https://www.protocols.io/view/human-kidney-tumor-dissociation-for-single-cell-g-2n2gdge> (accessed May 15, 2022).
- [2] Lun ATL, Riesenfeld S, Andrews T, Dao TP, Gomes T, participants in the 1st Human Cell Atlas Jamboree, et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol* 2019;20:63.
- [3] McCarthy DJ, Campbell KR, Lun ATL, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* 2017;btw777. <https://doi.org/10.1093/bioinformatics/btw777>.
- [4] McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* 2019;8:329–37.e4.
- [5] Young MD, Mitchell TJ, Vieira Braga FA, Tran MGB, Stewart BJ, Ferdinand JR, et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* 2018;361:594–9.
- [6] Lun ATL, McCarthy DJ, Marioni JC. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res* 2016;5:2122.
- [7] Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;36:411–20.
- [8] Kostov B, Bécue-Bertaut M, Husson F. Multiple Factor Analysis for Contingency Tables in the FactoMineR Package. *The R Journal* 2013;5:29. <https://doi.org/10.32614/rj-2013-003>.
- [9] Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* 2019;566:496–502.
- [10] Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. *Statistics for Biology and Health* 2000. <https://doi.org/10.1007/978-1-4757-3294-8>.
- [11] Zhao M, Liu Y, Zheng C, Qu H. dbEMT 2.0: An updated database for epithelial-mesenchymal transition genes with experimentally verified information and precalculated regulation information for cancer metastasis. *Journal of Genetics and Genomics* 2019;46:595–7. <https://doi.org/10.1016/j.jgg.2019.11.010>.

- [12] Chalmel F, Primig M. The Annotation, Mapping, Expression and Network (AMEN) suite of tools for molecular systems biology. *BMC Bioinformatics* 2008;9:86.
- [13] Wang X, Park J, Susztak K, Zhang NR, Li M. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nat Commun* 2019;10:380.
- [14] Thul PJ, Lindskog C. The human protein atlas: A spatial map of the human proteome. *Protein Science* 2018;27:233–44. <https://doi.org/10.1002/pro.3307>.
- [15] Leonard S, Lardenois A, Tarte K, Rolland A, Chalmel F. FlexDotPlot: a universal and modular dot plot visualization tool for complex multifaceted data n.d. <https://doi.org/10.1101/2020.04.03.023655>.
- [16] Bakken TE, Hodge RD, Miller JA, Yao Z, Nguyen TN, Aebermann B, et al. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS One* 2018;13:e0209648.
- [17] Baniak N, Flood TA, Buchanan M, Dal Cin P, Hirsch MS. Carbonic anhydrase IX (CA9) expression in multiple renal epithelial tumour subtypes. *Histopathology* 2020;77:659–66.
- [18] Tang S-W, Yang T-C, Lin W-C, Chang W-H, Wang C-C, Lai M-K, et al. Nicotinamide N-methyltransferase induces cellular invasion through activating matrix metalloproteinase-2 expression in clear cell renal cell carcinoma cells. *Carcinogenesis* 2011;32:138–45.
- [19] Hakimi AA, Pham CG, Hsieh JJ. A clear picture of renal cell carcinoma. *Nat Genet* 2013;45:849–50.
- [20] Moore LE, Jaeger E, Nickerson ML, Brennan P, De Vries S, Roy R, et al. Genomic copy number alterations in clear cell renal carcinoma: associations with case characteristics and mechanisms of VHL gene inactivation. *Oncogenesis* 2012;1:e14.
- [21] Kurtenbach S, Cruz AM, Rodriguez DA, Durante MA, Harbour JW. UPhyloplot2: visualizing phylogenetic trees from single-cell RNA-seq data. *BMC Genomics* 2021;22:419.
- [22] Bui TO, Dao VT, Nguyen VT, Feugeas J-P, Pamoukdjian F, Bousquet G. Genomics of Clear-cell Renal Cell Carcinoma: A Systematic Review and Meta-analysis. *Eur Urol* 2022;81:349–61.
- [23] Krock BL, Skuli N, Simon MC. Hypoxia-induced angiogenesis: good and evil. *Genes Cancer* 2011;2:1117–33.
- [24] Ruan H, Li X, Yang H, Song Z, Tong J, Cao Q, et al. Corrigendum to “Enhanced expression of caveolin-1 possesses diagnostic and prognostic value and promotes cell migration, invasion and sunitinib resistance in the clear cell renal cell carcinoma” [Exp. Cell Res. (2017) 269-278]. *Exp Cell Res* 2018;362:244.
- [25] Arakaki R, Yamasaki T, Kanno T, Shibasaki N, Sakamoto H, Utsunomiya N, et al. CCL2 as a potential therapeutic target for clear cell renal cell carcinoma. *Cancer Med* 2016;5:2920–33.
- [26] Cabral-Pacheco GA, Garza-Veloz I, la Rosa CC-D, Ramirez-Acuña JM, Perez-Romero BA, Guerrero-Rodriguez JF, et al. The Roles of Matrix Metalloproteinases and Their Inhibitors in Human Diseases. *International Journal of Molecular Sciences* 2020;21:9739. <https://doi.org/10.3390/ijms21249739>.
- [27] Guo H, German P, Bai S, Barnes S, Guo W, Qi X, et al. The PI3K/AKT Pathway and Renal Cell Carcinoma. *J Genet Genomics* 2015;42:343–53.
- [28] Corona A, Blobel GC. The role of the extracellular matrix protein TGFBI in cancer. *Cell Signal* 2021;84:110028.
- [29] Filippou PS, Karagiannis GS, Constantinidou A. Midkine (MDK) growth factor: a key player in cancer progression and a promising therapeutic target. *Oncogene* 2020;39:2040–54. <https://doi.org/10.1038/s41388-019-1124-8>.

### 4.1.3. DISCUSSION ET PERSPECTIVES

Dans ce projet, j'ai établi un atlas de cellules uniques dans le ccRCC contenant plus de 54,000 cellules, et décrivant l'hétérogénéité tumorale, notamment dans le compartiment malin. Cette étude nous a mené au constat d'une hétérogénéité significative, à la fois inter et intra-tumorale. L'atlas résultant comprend 35 clusters de cellules, soit une diversité importante à la fois dans le microenvironnement tumoral et dans les cellules malignes, qui comprennent en réalité 24 sous-clusters distincts, identifiés après leur extraction et ré-analyse. Ensuite, la comparaison des répartitions de types cellulaires entre notre étude et l'étude intégrée illustre bien le fait qu'une majorité de cellules immunitaires soit retrouvée dans les études de scRNA-seq quand elles ne les éliminent pas expérimentalement. Selon notre expérience, cela est dû à la grande résistance des cellules immunitaires, notamment les lymphocytes, et au fait qu'elles soient déjà intrinsèquement individuelles, ce qui facilite leur sélection lors de la dissociation tissulaire. Enfin, j'ai observé que le jeu de données additionnel était de moins bonne qualité, avec une répartition très hétérogène de la profondeur de séquençage et du nombre de cellules par individu. De façon remarquable, la projection sur la représentation UMAP des cellules malignes provenant de cette étude révèle une relative homogénéité. Cela peut être une réalité biologique mais peut également résulter d'une faible profondeur de séquençage, qui a tendance à priver les cellules de leur spécificité transcriptionnelle dans l'analyse. Cela pourrait aussi résulter du moindre nombre d'échantillons prélevés par tumeur dans cette étude, alors que notre protocole comprend les prélèvements de deux à neuf zones macroscopiquement distinctes sur la tumeur. Peut-être que cette particularité nous a permis de récolter une plus grande diversité de cellules, en général.

Nous n'avons pas observé de limites à l'HIT, ce qui suppose que d'autres prélèvements sont nécessaires pour dresser une image complète, incitant à mettre en place des cohortes plus importantes. Nous avons récemment fait une méta-analyse de plusieurs études de scRNA-seq dans les cancers urologiques (**Axe 3** de cette thèse) qui pourra peut-être permettre de répondre à cette problématique. Il est aussi possible de mettre en place de telles cohortes expérimentalement, le principal enjeu étant financier. Les techniques de multiplexage des cellules uniques telles que le Cell Hashing (Stoeckius *et al.*, 2018) ou la solution CellPlex développée par 10X Genomics (Mylka *et al.*, 2022) permettent d'économiser les coûts de réactifs nécessaires à la préparation des bibliothèques de séquençage, et de reporter les budgets sur l'augmentation des cohortes et le coût du séquençage. Pour aller plus loin dans la réduction des coûts, il est aussi possible d'avoir recours au sciRNA-seq (« *Single cell combinatorial indexing RNA sequencing* »). Le protocole est plus accessible, car il s'affranchit de la microfluidique et

de la capture en gouttelette, et ne nécessite que du matériel conventionnel de biologie moléculaire. Avec des dilutions et indexages successifs des cellules en plaques 96-puits, le sciRNA-seq permet de générer des données sur des millions de cellules à un coût relativement faible puisqu'il s'affranchit aussi de l'achat de kits commerciaux très coûteux. Le seul budget significatif est celui des index nécessaires à marquer les cellules individuelles. Cependant, puisque ces index permettent également d'identifier les échantillons d'origine des cellules, nous avons commencé à implémenter la technique en collaboration avec une autre équipe sur le site rennais dans l'objectif de mettre en commun l'achat de ces index et de multiplexer des échantillons variés. De cette façon, il sera possible de réaliser en routine du scRNA-seq très haut-débit et moins coûteux. Cela permettra notamment de séquencer suffisamment de tumeurs et d'obtenir une image représentative de la diversité globale dans le ccRCC à l'échelle de la cellule unique. Il faut noter toutefois que nous manquons de recul sur la technique de sciRNA-seq et devons encore valider l'approche.

Dans cette étude, nous avons eu la chance d'obtenir une tumeur particulièrement hétérogène et présentant une progression tumorale intéressante, avec les gains de caractéristiques particulièrement agressives, telles que l'EMT, la migration et la prolifération cellulaire ou encore la régulation négative de l'apoptose. À ma connaissance, aucune étude de scRNA-seq n'a révélé un tel degré d'HIT dans un ccRCC. Pour aller plus loin, nous avons utilisé les profils transcriptomiques de ces clones tumoraux pour prédire leur représentativité dans des données massives de RNA-seq disponibles au sein de l'atlas du génome du cancer (TCGA), ce qui a permis d'observer des associations significatives entre la représentativité de certains clones et la survie des patients. Cette approche a déjà été employée avec la prédiction de sous-populations de lymphocytes T (Hu *et al.*, 2020), en associant les profils de tumeur possédant plus de lymphocytes T CD8 activés avec un mauvais pronostic. Notre approche s'en démarque par le type cellulaire inféré et associé à la survie ; les cellules malignes. Cette particularité résulte de notre stratégie d'enrichissement et d'étude spécifique de ces cellules. Cela appelle cependant à de plus amples analyses. À l'avenir, il serait intéressant d'intégrer de nouvelles signatures clonales, complétant ou remplaçant celle que nous avons aujourd'hui. De même, il sera nécessaire d'établir un modèle plus précis, intégrant d'autres facteurs, par exemple cliniques, et de le comparer aux autres modèles existants tels que les signatures géniques classiques. Ces résultats montrent que les facteurs cliniques sont limités et que d'autres marqueurs pourraient venir en complément afin d'améliorer le pronostic et la stratification des patients. Ces résultats encouragent également à considérer l'utilisation de la transcriptomique en clinique pour compléter les facteurs pronostiques anatomopathologiques

(stade d'avancement, grade nucléolaire, facteurs de comorbidité). Alors que la révolution des NGS depuis les années 2000 a été très peu suivies par des applications concrètes en clinique, je pense que cela sera désormais davantage possible grâce à la technologie de BRB-seq (Alpern *et al.*, 2019) (voir la section 1.4.2.3.3. BRB-seq). C'est technologie de séquençage ARN *bulk* est globalement peu chère, grâce à l'utilisation d'UMI, et serait adaptée à l'application de notre approche. Selon moi, il serait possible, au sein des hôpitaux ou sur des plateformes dédiées, de préparer en routine des bibliothèques de BRB-seq de plusieurs dizaines de patients (à raison de 30€ par patient) afin d'obtenir leur profil transcriptomique *bulk* et, par exemple, de déterminer par déconvolution s'ils présentent des profils de cellules tumorales problématiques.



## 4.2. ÉTUDE DES IMPACTS D'INHIBITEURS DE KINASES SUR L'HÉTÉROGÉNÉITE INTRA-TUMORALE DU CCRCC PAR UNE APPROCHE COMBINANT UN MODÈLE DE CULTURE 3D ET SINGLE-CELL RNA-SEQ

### 4.2.1. INTRODUCTION

L'hétérogénéité intra-tumorale est impliquée dans l'acquisition de résistance aux thérapies par les cellules malignes dans les tumeurs sous traitement. L'étude des mécanismes de résistance acquise est donc un enjeu important pour délimiter des approches thérapeutiques plus adaptées. Alors que l'analyse scRNA-seq permet de déchiffrer l'hétérogénéité de tumeurs complexes (Suvà and Tirosh, 2019), son potentiel pour révéler des sensibilités thérapeutiques spécifiques de types cellulaires est très prometteur mais dépend grandement des modèles d'étude de ces thérapies. Les modèles traditionnels tels que les lignées et cultures cellulaires primaires en deux dimensions, davantage adaptées à certaines études mécanistiques, échouent à reproduire l'architecture et le métabolisme tumoraux (voir la section **1.3. Des Modèles d'étude des ccRCCs**) (Meijer, Jager and Gent, 2017). Des modèles plus avancés de culture en trois dimensions (sphéroïdes / organoïdes) ou de xénogreffes sur souris immunodéprimées présentent eux aussi des inconvénients (Meijer, Jager and Gent, 2017). Dans ce contexte, la culture organotypique de coupes est un modèle attractif, car il est facile et rapide à mettre en œuvre, récapitule l'organisation 3D du tissu tumoral ainsi que le microenvironnement, et coûte peu cher (Vaira *et al.*, 2010; Powley *et al.*, 2020). De tels modèles ont déjà été développés et éprouvés dans le ccRCC (Roelants *et al.*, 2018a, 2020; Giacosa *et al.*, 2021). Par ailleurs, comme évoqué dans les parties précédentes, le scRNA-seq a déjà permis de décomposer le ccRCC et de constater sa nature hétérogène. En particulier, dans le premier axe de cette thèse, nous avons enrichi les cellules malignes qui ont révélé des degrés d'hétérogénéité inter et intra-tumorale significatifs. Dans ce contexte, nous avons établi une collaboration avec l'équipe d'Odile Filhol-Cochet (CEA, Grenoble) qui a développé une approche innovante de culture organotypique de coupes de ccRCC pour tester les impacts d'agents inhibiteurs de kinases seuls et en combinaison. Ceux-ci comprennent trois molécules possédant une autorisation de mise sur le marché pour traiter le ccRCC métastatique : le Sunitinib (inhibiteur multi-cibles anti-PDGFR et VEGFR), le Pazopanib (inhibiteur multi-cibles anti-VEGFR) et le Temserolimus (anti-mTOR).



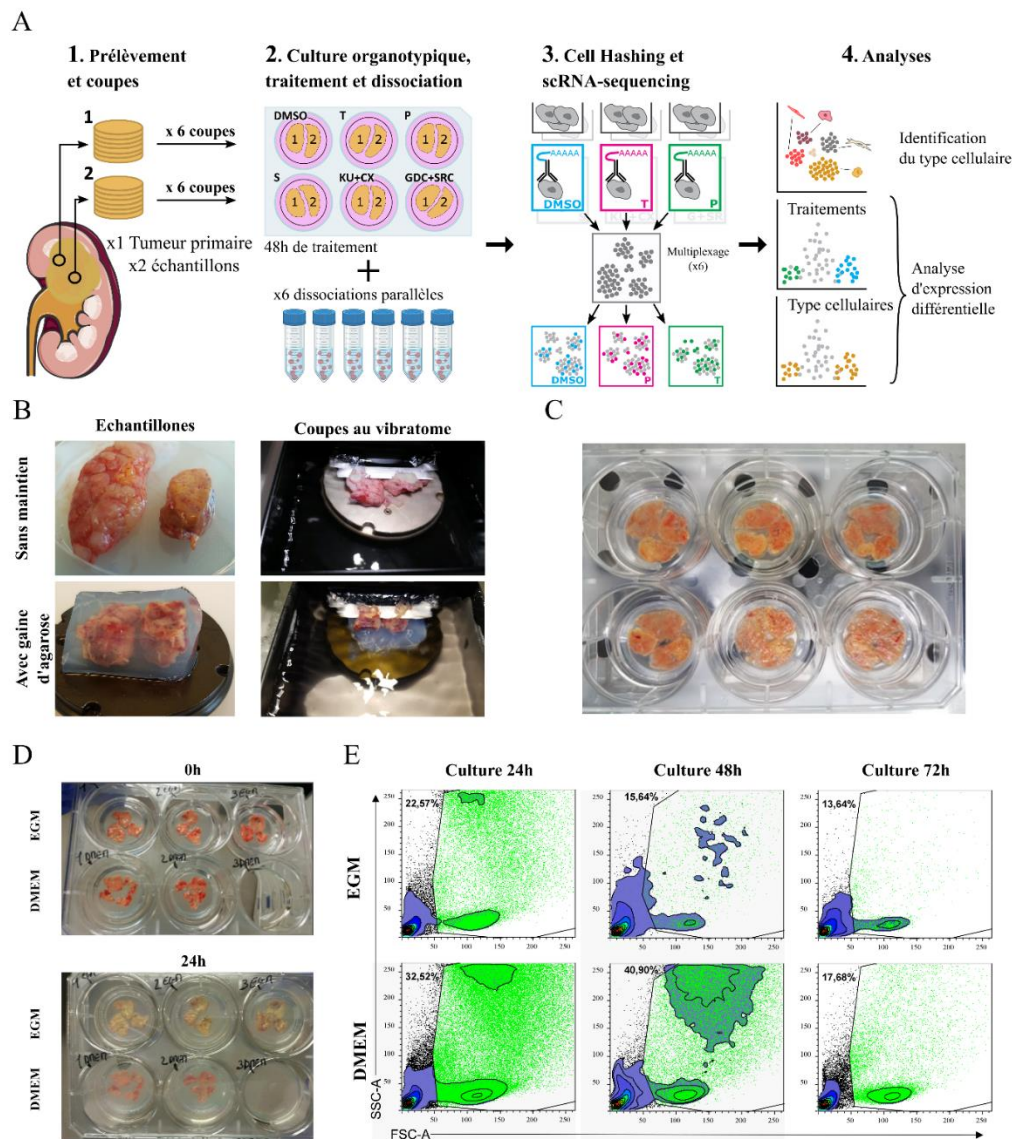


Figure 46. Mises au point du protocole de culture organotypique de coupes de ccRCC

(A) Schéma du protocole combinant la culture organotypique 3D, le multiplexage de cellules et le scRNA-seq. Deux échantillons sont collectés à partir d'un ccRCC fraîchement opéré. Les deux échantillons sont coupés en tranches de 400 $\mu$ m qui sont ensuite placées sur des inserts dans des plaques à 6 puits et incubées avec du milieu de culture et des inhibiteurs de kinases dilués pendant 48h. Le milieu est renouvelé à 24h. Six dissociations parallèles sont effectuées et les cellules uniques résultantes de chaque condition de traitement sont multiplexées avec des anticorps de *cell hashing* avant la préparation de la librairie et le séquençage des ADNc. Enfin, une série d'analyses est effectuée afin d'évaluer l'impact des médicaments sur le ccRCC à la résolution de la cellule unique. T = Tesmerolimus ; P = Pazopanib ; S = Sunitinib ; KU = KU-60019 ; CX = CX-4945 ; GDC = GDC-0941 (Pictilisib) ; SRC = Saracatinib (AZD0530). (B) Photographies illustrant les échantillons tumoraux coupés au vibratome avec ou sans agarose. (C) Photographie illustrant la culture organotypique. Des séries régulières de tranches provenant de trois prélèvements sont disposées dans six inserts de culture. (D) Comparaison des milieux de culture à base de DMEM (80% de DMEM, 20% de SVF, 100U/ml de Penicilline/Streptomycine) ou de EGM (80% EBM supplémenté, 10% SVF, 10% RPMI, 1% Glycérol, 100U/ml de Penicilline/Streptomycine). (E) Cytogrammes présentant les cellules (SSC-A - FSC-A) collectées à 24h, 48h, et 72h, après dissociations des tranches cultivées avec le milieu à base de DMEM ou à base de EGM.

Les combinaisons de thérapies mélangent le KU-60019 (inhibiteur de la protéine ATM) avec le silmitasertib (ou CX-4945, inhibiteur de la protéine CK2), et le pictilisib (inhibiteurs des isoformes PI3K de classe I) avec le saracatinib (anti-SRC). Ces molécules ciblent des voies de signalisation connues pour être suractivées ou marqueurs de mauvais pronostic dans le ccRCC (Roelants *et al.*, 2018a; Giacosa *et al.*, 2021). Alors que le modèle a été validé et a permis de rendre compte des effets synergiques de combinaisons de traitements par analyses histologiques et RNA-seq, nous avons souhaité aller plus loin et établir une stratégie innovante combinant la culture organotypique, le multiplexage *cell hashing*, et scRNA-seq, afin d'étudier les impacts de ces molécules sur l'HIT du ccRCC. L'hypothèse était que l'HIT détermine la nature des réponses aux thérapies ciblées, c'est-à-dire que ces réponses diffèrent selon les (sous-)types cellulaires présents au sein de la tumeur. Nous faisons également l'hypothèse que ces différences sont en partie détectables de façon précoce, par une approche de transcriptomique, dès 48h de traitement. L'objectif premier est de constituer une preuve de faisabilité de notre approche, du point de vue expérimental et analytique. Dans ce contexte, j'ai souhaité notamment prendre en compte l'hétérogénéité tissulaire dès l'étape expérimentale, optimiser la quantité de tissu cultivé, et valider par cytométrie mes résultats de dissociation. La cytométrie est également une approche haut-débit à l'échelle de la cellule unique, qu'il est particulièrement utile d'adosser aux expériences de scRNA-seq selon moi. Dans ce cadre, les objectifs suivants consistent à déterminer et caractériser les cellules sensibles ou résistantes au traitement étudié, ainsi que caractériser les réponses moléculaires aux thérapies, en rapport avec les voies de signalisation qu'elles ciblent. Ce projet est encore en cours d'analyses et les résultats présentés dans cette section sont préliminaires.

#### **4.2.2. RESULTATS SUPPLEMENTAIRES - MISES AU POINT EXPERIMENTALES**

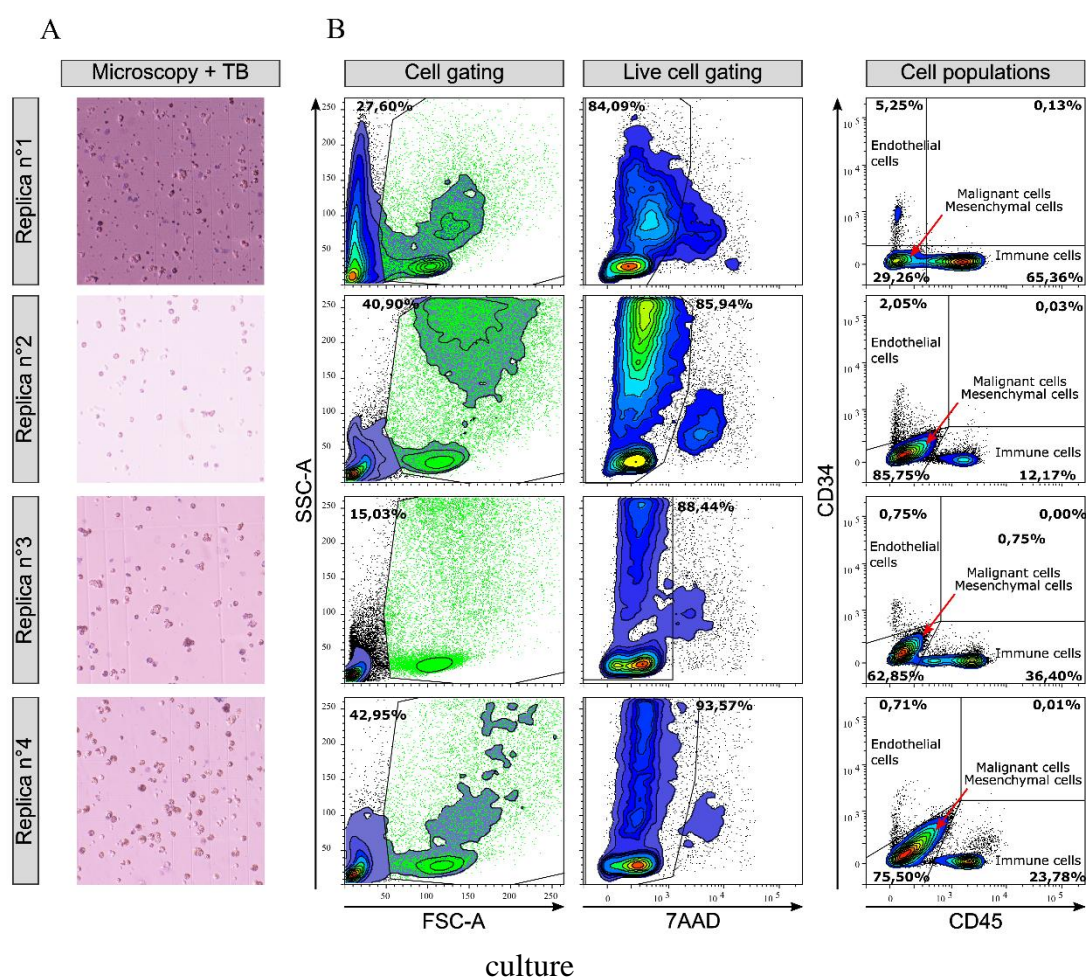
Dans ce projet, j'ai développé la culture organotypique de tranches de ccRCC dans le laboratoire. En m'appuyant sur les travaux de collaborateurs (Roelants *et al.*, 2020), j'ai conçu un protocole adapté, compatible avec le séquençage ARN sur cellules uniques (Figure 46A). Cette section présente les mises au point nécessaires et leurs résultats clefs.

##### **4.2.2.1. Conception et contraintes expérimentales**

La première étape dans ce projet a consisté à concevoir le plan expérimental. Premièrement, afin de capturer l'HIT, j'ai décidé de réaliser deux à trois prélèvements simultanés par tumeur (Figure 46B). Deuxièmement, la dissociation tissulaire induisant une

perte de cellules plus ou moins importante (selon la qualité du prélèvement et les degrés d'inflammation ou de remaniement œdémateux/hémorragiques), j'ai choisi de maximiser la quantité de tissu cultivé. Par ailleurs, je tenais à avoir les conditions de culture les plus homogènes possibles et donc à traiter ensemble, sur un même insert, les coupes issues des prélèvements d'une même tumeur. J'ai donc sélectionné des inserts de culture pour plaque 6-puits (30 mm de diamètre). L'étape suivante consiste à réaliser des coupes homogènes de ces prélèvements à l'aide d'un vibratome (Vibratome VT1200S - Leica Biosystems). Les protocoles en ligne et nos collaborateurs recommandant de ne pas dépasser 500  $\mu\text{m}$  d'épaisseur, j'ai choisi de réaliser des coupes de 400  $\mu\text{m}$ . Cette épaisseur reste importante par rapport aux standards de la littérature, mais répond à mon objectif de maximisation de la quantité de tissu.

Figure 47. Évaluation de la qualité de dissociation de coupes de ccRCC après 48h de



(A) Images de microscopie (x10) de cellules uniques après dissociation des tranches contenues dans un insert de culture. (B) Cytogrammes des cellules dissociées. De gauche à droite sont présentés la sélection des cellules (SSC-A - FSC-A), la sélection des cellules vivantes (SSC-A - 7AAD), et les mesures de proportion des types cellulaires avec les marqueurs CD34 (cellules endothéliales) et CD45 (cellules immunitaires). Les cellules malignes et stromales sont négatives pour le CD34 et le CD45, et sont donc situées dans le cadre inférieur gauche des cytogrammes.

#### 4.2.2.2. Résultats de mises au point

Au cours de ma thèse, il est important de préciser que j'ai pris en charge plus de 60 tumeurs, dont plus de la moitié était dédiée à la mise au point de ce protocole. J'ai d'abord testé de nombreux paramètres de coupes au vibratome. Les premiers essais portaient sur des échantillons directement collés sur le support métallique prévu dans le dispositif (Figure 46B). Bien que cela ait permis de réaliser les premières coupes et de faire quelques essais de dissociation, le processus ne fonctionnait pas toujours, en particulier quand le tissu était de mauvaise qualité ou trop remanié ; l'échantillon se décollait et n'était pas réutilisable. J'ai donc opté pour une seconde approche, consistant à inclure les échantillons dans un bloc refroidi d'agarose afin de les maintenir pendant la découpe (Figure 46B). Le principe est de chauffer une solution d'agarose à 2 %, de déposer les prélèvements dans un puits en plaque 6-puits, de couler l'agarose liquide sur les prélèvements, et de laisser refroidir le système sur glace jusqu'à solidification de l'agarose. L'ensemble constitué du bloc d'agarose et des prélèvements est ensuite collé sur le support magnétique puis coupé au vibratome. Cela a permis d'obtenir de bien meilleurs résultats et de s'affranchir de la limite de qualité des tissus. Les coupes étaient également plus reproductibles. Ces adaptations m'ont permis de réaliser en routine des séries allant jusqu'à 20 coupes, de façon simultanée sur plusieurs prélèvements. Le système de culture ainsi établi permet de tester jusqu'à six conditions de traitement sur des coupes tumorales reproductibles, représentant environ 200mm<sup>3</sup> de tissu (Figure 46C).

J'ai ensuite testé deux milieux de culture du ccRCC ; l'un à base de DMEM (« Dulbecco's Modified Eagle Medium ») (Roelants *et al.*, 2020), l'autre à base d'EGM (« Endothelial cell growth factor ») (communication interne) (Figure 46E-D). Dans les conditions établies, l'emploi du milieu DMEM résultait en une meilleure qualité du tissu (couleur et maintien) et un plus grand nombre de cellules après dissociation jusqu'à 72h. J'ai également essayé sur quelques expérimentations de prolonger la culture jusqu'à 96h mais de façon générale, la qualité des dissociations diminue significativement à 72h, et drastiquement à 96h, de sorte que la plupart des dissociations donnent trop peu de cellules pour des mesures de cytométrie. J'ai cependant validé par cytométrie en flux la qualité des dissociations à 48h, qui est la durée de culture utilisée par nos collaborateurs (Figure 47). L'examen microscopique au bleu de trypan montre que les cellules sont bien en bon état général, bien dissociées et diverses en termes de taille et de forme. La mesure du 7AAD, un marqueur de viabilité, en cytométrie, révèle un taux de viabilité moyen de 88,11 % (n=4, écart-type = 4,13), soit supérieur au seuil recommandé pour les expériences de scNRA-seq avec la technologie 10X (~70 %). La double coloration CD34 et CD45 confirme par ailleurs que les suspensions cellulaires sont diverses,

principalement composées de cellules stromales et malignes (CD34<sup>-</sup>/CD45<sup>-</sup>), immunitaires (CD45<sup>+</sup>), et d'une petite proportion de cellules endothéliales (CD34<sup>+</sup>).

Sur la base du protocole publié par nos collaborateurs (Roelants *et al.*, 2020), j'ai arrêté le temps de culture à 48h et validé par cytométrie la viabilité. J'ai ensuite simulé le protocole avec l'application de 6 conditions de traitement (Fig. 2B), soit 6 groupes de coupes traitées à dissocier en parallèle. À la suite de l'ensemble des étapes de mises au point décrites, nous obtenons des dissociations de qualité dans chaque condition de traitement, avec des cellules individuelles et variées (Fig. 2D), une excellente viabilité et un nombre moyen de cellules allant de 50 000 à 1 million en raison de la variabilité importante d'une tumeur à l'autre.

Par la suite, nous avons sélectionné quatre ccRCC primitifs (RCC1 à RCC4) dont trois présentaient un haut grade ( $>pT3a$  ;  $ISUP \geq 3$ ) et un risque raisonnable de rechute ou de développement métastatique. Au total, six conditions de traitement ont été appliquées, soit un contrôle (DMSO) et cinq stratégies thérapeutiques (Sunitinib, Pazopanib, Temsirolimus, combinaison CX4945 + KU60019 et combinaison GDC + saracatinib). Après dissociation des coupes, nous observons une viabilité moyenne de 88,11 % ( $n=4$ , écart-type = 4,13).

**4.2.3. MANUSCRIT EN PREPARATION**



# **Single-cell RNA-seq of 3D organotypic culture of ccRCCs treated with several kinase inhibitors revealed specific transcriptional responses in malignant cells**

Saout J.<sup>1,\*</sup>, Lecuyer G.<sup>1,\*</sup>, Evrard B.<sup>1</sup>, Leonard S.<sup>1</sup>, Khene ZE.<sup>3</sup>, Pillet C.<sup>4</sup>, Mathieu R.<sup>1,3</sup>, Bensalah K.<sup>3</sup>, Lardenois A.<sup>1</sup>, Rioux-Leclercq N.<sup>1,2,\*</sup>, Filhol-Cochet O.<sup>4</sup>, Chalmel F.<sup>#,1</sup>

<sup>1</sup> Univ Rennes, Inserm, EHESP, Irset (Institut de Recherche en Santé, Environnement et Travail), UMR\_S 1085, F-35000 Rennes, France.

<sup>2</sup> Department of Pathology, University Hospital of Rennes, Rennes, France.

<sup>3</sup> Department of Urology, University Hospital of Rennes, Rennes, France.

<sup>4</sup> Université Grenoble Alpes, Inserm, CEA, IRIG-Biology of Cancer and Infection, UMR\_S 1036, F-38000 Grenoble, France.

\* These authors contributed equally to this work

# To whom correspondence should be addressed

**Correspondance:** frederic.chalmel@inserm.fr



## **Abstract**

Clear cell renal cell carcinoma (ccRCC) is the most common and lethal subtype of kidney cancer. Despite some initial benefits of tyrosine kinase inhibitor (TKI) targeted therapies, tumors acquire resistance, eventually leading the cancer to progress. These pitfalls are partly due to intra-tumor heterogeneity (ITH), which can be deciphered with single-cell RNA sequencing (scRNA-seq). While those technologies hold many promises, their potential in investigating the impacts of targeted therapies is conditioned by the relevance of tumor models. Here we implemented a proof-of-concept study combining 3D organotypic culture of ccRCC slices and single-cell transcriptome profiling to study the impact of several targeted therapies, including TKIs. The analysis revealed important cell-specific and treatment-specific transcriptional shifts in most cell subpopulations, including malignant cells, associated with distinct signaling pathways that could be reflective of/contribute to therapy response or resistance. Together, our study provides a valuable methodological resource that could be broadly useful to further explore the response to targeted therapies in ccRCCs at the single-cell level.

## **Introduction**

Clear cell renal cell carcinoma (ccRCC) is the most common and lethal subtype of kidney cancer, accounting for 2.2% of all cancers and nearly 180,000 deaths per year, worldwide [1,2]. It has a poor prognosis, particularly for 45% of patients who are or will become metastatic, when diagnosed [2,3]. These past decades, the development of targeted therapies including anti-angiogenic tyrosine-kinase inhibitors (TKIs), and more recently immune checkpoint inhibitors, has provided significant benefits for those patients in terms of progression-free

survival and overall survival [4]. However, these therapies only have partial efficiency, with 30% of responsive patients, contrasting responses across tumor regions within the same patient, and a long-term tumor escape in most cases [5–7]. More generally, an estimated 90% of currently available targeted therapies have a positive effect in only around 40% of patients [8]. These pitfalls are not only due to inter- but also importantly to intra-tumor heterogeneity (ITH), which has been demonstrated to be a typical feature of ccRCCs [9,10] (Saout *et al.*, *submitted*). A major challenge in oncology research remains to study the relationship between ITH and therapy response or resistance. While scRNA-seq can be used to determine the cellular composition of complex tumors, one of the most critical points is to select the proper experimental model relevant for the therapeutic molecules to be tested. As far as targeted therapies are concerned, 3D organotypic culture of tumor slices is an attractive and cost-effective option to be considered as it preserves the tumor architecture and spatial interactions within the microenvironment [11]. Robust protocols already exist for many cancers, including ccRCCs [12]. In particular, two studies have demonstrated the ability of organotypic culture of ccRCC slices to reveal the synergistic effects of kinase inhibitors used in combination [13,14]. To the best of our knowledge such culture models of ccRCCs have never been combined with single-cell transcriptome profiling to study the impact of therapies at a single-cell resolution.

In this proof-of-concept study we exposed primary organotypic slice cultures from three ccRCC tumors to three gold-standard targeted therapies (sunitinib, pazopanib, and temsirolimus) and two promising combination therapies (KU-60019 + CX-4945, and GDC-0941 + Saracatinib) [13,14] followed by scRNA-seq in order to decipher the transcriptional response of each cell population to those molecules. We demonstrated the feasibility of our approach and characterized ccRCC cultured slices at the single-cell resolution for the first time. We also showed that malignant cells display drug-specific transcriptomic responses.

## **Patients and Methods**

Materials and methods are detailed in the supplementary materials.

### **Tumor specimens**

Fresh tumor samples were collected from a total of four untreated patients who underwent radical nephrectomy for ccRCC at the urology department of the Rennes University-Hospital center. Patients met the requirements of the local institutional ethics committee (see Supplementary methods). Detailed clinical information from each patient is provided in Table S1.

### **Culture, single-cell capture, library preparation, sequencing and analysis**

Upon reception, tumors were sampled in two distinct macroscopic areas. Samples were sliced using a Vibratome VT1200 (Leica Microsystems, Wetzlar, Germany), deposited on inserts (Millicell Cell Culture Insert, Réf: PICMORG50) and cultivated with drugs for 48h (see Supplementary methods). Cultured slices were subjected to an enzymatic and mechanical dissociation. Cell suspension underwent a partial depletion of CD45 positive cells using the Miltenyi kit (CD45 MicroBeads, human) to enrich the samples in malignant/stromal cells and remove circulating and tumor-infiltrating immune cells. Single cells were then multiplexed by Cell Hashing and captured using the 10X Chromium technology (see Supplementary methods). The construction of single-cell libraries were carried out according to 10X Genomics and Biologend's instructions (TotalSeq<sup>TM</sup>-A Antibodies and Cell Hashing with 10x Single Cell 3' Reagent Kit v3 or v3.1 [Single Index] Protocol). Single-cell libraries were quality-controlled using a bioanalyzer (Agilent Technologies) and sequenced on an Illumina NovaSeq 6000

instrument using a 2x100 paired-end sequencing protocol (Illumina, San Diego, CA). Data were processed and analyzed using a standard workflow, including Cell Ranger (v3.0.2, 10x Genomics) and Seurat (v4) [15] (see Supplementary methods).

## Results

### Experimental design

To evaluate the impact of targeted therapies, we have put in place a strategy combining an organotypic culture model that recapitulates well the tumor 3D architecture, microenvironment and physiology, and scRNA-sequencing by multiplexing samples using the cell-hashing method (Fig. 1A; Table S2) (Supplementary results). Although the culture system implemented here has already been described, it has never been validated by flow cytometry (Fig. 1B). We demonstrated that ccRCC slices cultured for 48h followed by tissue dissociation displayed good overall single-cell diversity and viability, including of malignant cells (Supplementary results, Table S3). To further enrich the malignant cells and specifically study their responses to kinase inhibitors, we performed a partial depletion of immune cells, as recently suggested/encouraged by previous studies [16] (Saout *et al.*, *submitted*).

Once validated, the culture model was applied on tumor slices from four ccRCC patients (named RCC1-4) exposed to six distinct conditions (Fig. S1A-C) including: *i*) DMSO as a control; *ii-iv*) three targeted therapies (sunitinib, temsirolimus, pazopanib; and, *v-vi*) two combination therapies combining KU-60019 with silmitasertib (KU+CX) and pictilisib with saracatinib (GDC+SRC) (Fig. 1A, Table S2). Clinical characteristics of patients and drug characteristics are provided in Table S1, Table S2 and Supplementary results, respectively.

Following culture, tumor slices were then dissociated and samples were multiplexed by using the Cell-Hashing method. Single-cell libraries were prepared and sequenced (Supplementary results). Quality control of scRNA-seq data revealed low quality samples for all those originating from one specific patient (RCC4) (Fig. S2). RCC4 samples were consequently removed from the analysis resulting in a final dataset composed of 18 scRNA-seq samples (three tumors exposed to six experimental conditions) (Supplementary results).

### **Organotypic culture of ccRCC slices followed by scRNA-seq recapitulate cell diversity**

The total dataset was composed of 34,512 cells that are evenly distributed across the 18 samples, with numbers ranging from 1,332 to 2,649 (Fig. 2A). Cells were next partitioned into 21 clusters (Fig. 2B) which were further associated with five broad cell types including pericytes, fibroblasts, malignant, endothelial and immune cells, based on the expression of well-known marker genes (Fig. 2C-D). Malignant cells and pericytes were by far the most represented cell types in the dataset, with 16,641 and 16,357 cells, respectively, distributed in 18 out of 21 clusters. In comparison, endothelial cells (cluster c12; 1105 cells) and fibroblasts (c20; 354 cells) were substantially detected but significantly less represented. There were only 55 immune cells in the smallest cluster (c21), consistent with their partial depletion in the samples. The distribution of cells per tumor revealed a relative homogeneity in endothelial cells (c20) (Fig. 2E) but, in contrast, an inter-tumor heterogeneity concerning pericytes (c1, c4, c6, c7, c11, c13, c15, c18, c19) and malignant cells (c2, c3, c5, c8, c9, c10, c14, c16, c17).

## **Single-cell RNA-seq revealed contrasted cell population responses to kinase inhibitors**

Looking at the distribution of the 34,512 cells, sunitinib and temsirolimus seemed to have little impact as they overlap with DMSO (Fig. 2E, Fig. S3, Fig. S4). With the exception of a transcriptomic shift in pericytes in RCC1, Pazopanib had also little effect on cells' transcriptome. Strikingly, KU+CX and, to a lesser extent, GDC+SRC induced major transcriptional shifts in malignant cells in RCC2 and RCC3 (Fig. S3, Fig. S4). To evaluate the impacts of kinase inhibitors, we performed a differential expression analysis between DMSO and the five other experimental conditions in each broad cell population for each patient. This analysis identified a total of 4,047 unique differentially expressed genes (DEGs) among which 1,226 DEGs were up or down regulated in at least two patients (Fig. 3A). Strikingly, KU+CX and, to a lesser extent, GDC+SRC, had significantly more up and down regulated genes than the other drugs, all cell types considered. The largest number of DEGs was found in endothelial cells, which contrasted with previous examination of cell distribution that suggested a relative homogeneity across experimental conditions. To refine the analysis, we next selected the 1,226 DEGs common to at least two patients (Fig. S5). KU+CX combination induced the most significant up and down regulations in both pericytes, endothelial and malignant cells. Interestingly, a substantial proportion of these DEGs seemed to be shared across the three cell types whereas smaller groups were specific to each of them. Overall these results suggest a significant impact of combination treatments on ccRCC slices, with the most manifest effects on malignant cells.

## Functional analysis revealed differential drug responses in malignant cells

To investigate these transcriptional effects into further detail, we next selected the 403 DEGs in malignant cells of at least two patients and observed a marked differential expression in combination treatments (Fig. 3B). Among these, we then selected DEGs that were specific to each experimental condition and intersected them (Fig. 3C-D, Table S4). Most of these DEGs were specifically up or down regulated in one treatment condition (Fig. 3D). In total 37 DEGs were up-regulated in both KU+CX and GDC+SRC conditions, whereas only four were commonly down-regulated. Interestingly, temsirolimus had 22 and 19 shared up-regulated DEGs with KU+CX and GDC+SRC, respectively, suggesting potential common targeted pathways. We next performed a functional analysis on these therapy-specific gene groups (Fig. 3E). Some common features of carcinogenesis were found to be enriched in KU+CX, GDC+SRC and Pazopanib conditions, such as the regulation of cell population proliferation (adjusted  $p$ -value $<0.05$  with 41, 34, and 13 related genes, respectively), programmed cell death ( $p<1.10e-2$ ; 75, 53 and 17), cell communication ( $p<1.10e-2$ ; 116, 91 and 31) and cell motility ( $p<1.10e-3$ ; 48, 42 and 17). GDC+SRC and Pazopanib also significantly impacted cell cycle ( $p<5.10e-2$  ; 37 and 14 DEGs). DEGs associated with Temsirolimus exposure were only enriched in two pathways including NFG-stimulated transcription ( $p<1.10e-2$ , 3 genes) and Orexin receptor pathway ( $p<1.10e-2$ , 4 genes). Interestingly, Sunitinib had a significant impact on the VEGFA-VEGFR2 signaling pathway ( $p<5.10e-2$ ; 14), one of its main targets, although the pathway was even more impacted by KU+CX ( $p<1.10e-6$ ; 29) and GDC+SRC ( $p<1.10e-6$ ; 27). Several canonical pathways known for their implication in cancer evolution were also differentially enriched across experimental conditions, such as JAK/STAT pathway for GDC+SRC ( $p<5.10e-2$ ; 7), TGF-beta signaling pathway for G+S and Pazopanib ( $p<5.10e-2$ ; 5 and 5), the MAPK signaling pathway for Pazopanib ( $p<5.10e-2$ ; 4) for Pazopanib. Among other notable results, K+C, G+S and P were also significantly enriched in genes involved in

the response to chemical ( $p < 1.10e-3$ ; 117, 98 and 27) and more specifically cytokine ( $p < 2.10e-5$ ; 39, 44 and 16). Overall, these results demonstrated differential responses to kinase inhibitors in ccRCC malignant cells, with a more significant impact of combination therapies and Pazopanib.

## Discussion

Altogether, this proof-of-concept study described a promising experimental and analytical approach that allowed us to decipher cell type-specific drug responses in ccRCC tissue from individual patients. The dissociation of cultivated ccRCC slices in our model preserved the cell diversity and viability, and allowed us for multiplexed single-cell capture. The cell diversity was confirmed at the single-cell resolution both with flow cytometry and scRNA-seq, and was consistent with previous studies describing unaltered ccRCC intra-tumor heterogeneity [17,18; Saout *et al.*, *submitted*]. Our protocol proved efficient, although we were not able to go further with the RCC4 sample. We examined the cells on microscopic pictures right before the single-cell capture and noticed their unique dark appearance as compared to other RCCs (Fig. S1). Further investigation would be required to understand the state of this tumor and its cells under microscope. Moreover, projection of cells color-coded according to their treatment status on UMAPs revealed drug-specific transcriptional shifts in malignant cells. Hence, our approach provided a useful setting to evaluate further drug response and resistance, and to develop applications beyond previous traditional analyses [13]. A specific focus on malignant cells allowed us to highlight GDC+SRC and KU+CX pronounced impacts. This may be due to their synergistic actions that target a significant number of pathways, and probably induce a wider range of molecular responses. In previous studies, these targeted therapies were shown to be more efficient than conventional treatment to stop the proliferation of malignant cells [13,14].



Future analysis includes going into details of DEGs in malignant sub-populations. The functional analysis revealed significantly enriched pathways which would require further investigation. Interestingly, DEGs in KU+CX condition were enriched in genes involved in response to oxidative stress ( $p < 5.10e-2$ , 16), which corroborates the results of another study suggesting the induction of an overproduction of reactive oxygen species in ccRCC in response to the combination therapy [13]. To conclude, this study is a valuable resource for the community and a basis for investigating and understanding the relationships between intra-tumor heterogeneity and small molecule inhibitors.

## **Author contributions**

FC supervised the research and designed the study. JS, GL and FC wrote the manuscript. All authors contributed to the manuscript. GL, PR and FC processed and analyzed the data. SL contributed to the analysis. JS, NRL, AL and FC assisted in the manual curation. TAD, LN, PR, MB, OC and OS contributed to the development and deployment of the UncoVer website. All authors approved the final version of the manuscript.

## **Acknowledgements**

We thank all the staff of the Urology Department of the Rennes University Hospital for their expert assistance, and the participating patients. We thank the GenOuest bioinformatics facility for hosting the software as well as all members of the Research Institute for Environmental and Occupational Health (IRSET) for stimulating discussions. We thank the CytomeTri platform from Biosit (SFR UMS CNRS 3480 - INSERM 018), especially Laurent Deleurme for his precious help. Our special thanks go to the members of the H2P2 platform (Biosit) for their

precious help in the development of the culture models. We also thank the Institut national de la santé et de la recherche médicale (Inserm), the Université de Rennes 1, and the French School of Public Health (EHESP) for supporting this work. For their valuable help, our special thanks also go to Stéphane Boutet (10X Genomics), Marjorie Gournay, Solène-Florence Kammerer-Jacquet and Frédéric Dugay.

## **Funding**

This study was supported by La ligue contre le cancer committees 22, 35 and 81; the clinical and translational research committee (CORECT) of the Rennes University Hospital Center; the French National Institute of Health and Medical Research (Inserm); the University of Rennes 1; and the French School of Public Health (EHESP). S.L. was supported by the chair “Cancer & Innovation” of Rennes 1 Foundation (<https://fondation.univ-rennes1.fr/>) and a specific grant from the LabEx IGO program (n° ANR-11-LABX-0016) funded by the «Investment into the Future» French Government program, managed by the National Research Agency (ANR).

Conflict of Interest: none declared.

## Figure legends

**Figure 1. Experimental design.** (A) Flowchart of the protocol combining 3D organotypic culture, cell hashing multiplexing and scRNA-seq. First two samples are collected from a fresh ccRCC nephrectomy resection. Both samples are cut in 400 $\mu$ m slices which are then placed on inserts in 6-well plates and incubated with culture medium and diluted drugs for 48h. Six parallel dissociations are performed and resulting single cells from each treatment condition are multiplexed with hashtag antibodies before the library preparation and scRNA-sequencing. Finally, a set of analyzes is performed in order to evaluate the impacts of drugs on ccRCC at the single-cell resolution. T = Temsirolimus; P = Pazopanib; S = Sunitinib; KU = KU-60019; CX = CX-4945; GDC = GDC-0941 (Pictilisib); SRC = Saracatinib (AZD0530). (B) Assessment of the quality of slice dissociation after 48h of culture, without drug perturbations. On the left are presented microscopic pictures (x10) of single cells after slice dissociation. Right- Flow cytometry cytograms of single cell suspensions. From left to right are the cell gating (SSC-A (side scatter) versus FSC-A (forward scatter)), the live cell gating (SSC-A versus 7AAD staining), and the cell type proportions measurement with CD34 (endothelial cell) versus CD45 (immune cells) staining. Malignant and stromal cells are negative for CD34 and CD45, hence located in the bottom left frame of ‘cell populations’ cytograms.

**Figure 2. Single-cell analysis of ccRCC treated with TKIs.** (A) Number of cells and median number of detected genes by sample. (B) UMAP plot representing the cell type distribution (main colors) and clustering (shades of color) of all cells. (C) Dot plot representing the expression of known cell type markers in each cell cluster. (D) UMAP plots showing the expression levels of the CA9 (malignant cells), PLVAP (endothelial cells), PTPRC (immune cells), LUM (fibroblasts), and RGS5 (pericytes) cell markers. (E) Distribution of cells according to treatment conditions (left) and tumor sample (right). (F) Crossed proportions of cells distributed in clusters and treatment conditions.

**Figure 3. Differential gene expression analysis.** (A) Histograms showing the number of DEGs that were either Up (red) or Down expressed (blue) according to treatment condition and cell type. DEGs are either specific to one patient (dark red/blue), or common to two (intermediate red/blue) or three patients (light red/blue) that are specific to one patient. (B) Heatmap representation of 403 DEGs between 18 samples in malignant cells. Samples consist in the crossing of treatment conditions and patients. in malignant cells. (C) Heatmap representation of 549 DEGs between five treatment conditions in malignant cells. (D) Intersect Plot of DEGs between treatment conditions in malignant cells. The plot shows the numbers DEGs that were either specific or common to indicated categories. For example, there are 59 specifically up-regulated genes in GDC+SRC (“G+S UP”) treatment conditions, and 18 common DEGs between G+S UP and K+C UP. (E) Spotplot showing the gene ontology terms enrichment in each treatment. DEGs upon which enrichment analysis was performed were either up regulated in at least two patients, or down regulated in at least to patients. Selected

terms are from GO biological\_process or Wikipathways (indicated “WP”). Spot size reflects the proportion of DEGs (r) related to a GO term on total genes related to that term (R).

## Supplementary figure legends

**Figure S1. Tumor sample processing.** (A) Macroscopic view of radical nephrectomy specimens. The tumors show various edematous and hemorrhagic changes. (B) Pictures of organotypic ccRCC slice culture systems at 0h, 24h and 48h. (C) Microscopic pictures (x10 magnification) of pooled single cells after their labeling by hashtag antibodies. For each tumor sample (RCC1 to RCC4), six single cell suspensions, corresponding to the six treatment conditions, were labeled with hashtag antibodies and pooled.

**Figure S2. Quality control of single-cell RNA-seq data from 4 tumors.** The four columns on the right show quality metrics for each tumor and each treatment condition, i.e. for 24 total samples. The four columns on the left show the proportions of cells eliminated (black section of histograms) according to complementary bioinformatics tools.

**Figure S3. UMAPs of intersecting cells between patients and treatments.** On each UMAP, colored cells intersect between a tumor sample and a treatment condition. Cells colored in light gray in the background belong to other samples. Colors correspond to treatment conditions.

**Figure S4. UMAPs of intersecting cells between patients and cell types, colored by treatment.** On each UMAP, colored cells intersect between a tumor sample and a cell type annotation. Cells colored in light gray in the background belong to other samples. Colors correspond to treatment conditions. As in Fig. S3, Dark gray corresponds to DMSO, green to Sunitinib, purple to Pazopanib, orange to Temsirolimus, red to GDC+SRC, and blue to KU+CX.

**Figure S5. Differential expression analysis.** Heatmap representation of the scaled expression values of 1266 genes that are differentially expressed between samples, crossing treatment, patient ID, and cell types. DEGs shown are common to at least two patients.

## Supplementary figure legends

**Table S1. Clinical characteristics of patients participating in the study.** A total of 4 tumors were included in this study. ccRCC = clear cell renal cell carcinoma ; ISUP = International Society of Urological Pathology ; pTNM = Tumor-Node-Metastasis.

**Table S2. Characteristics of drugs used in the study.** Drugs's main characteristics presented in the table are taken from their profile page on selleckchem.com. The names of targets refer to HGNC nomenclature instead of protein names.

**Table S3. Cell viability and cell proportions measured by flow cytometry.** To validate our dissociation protocol on organotypic culture of ccRCC slices, we used flow cytometry on 4 tumor samples.

# Figures

Fig. 1

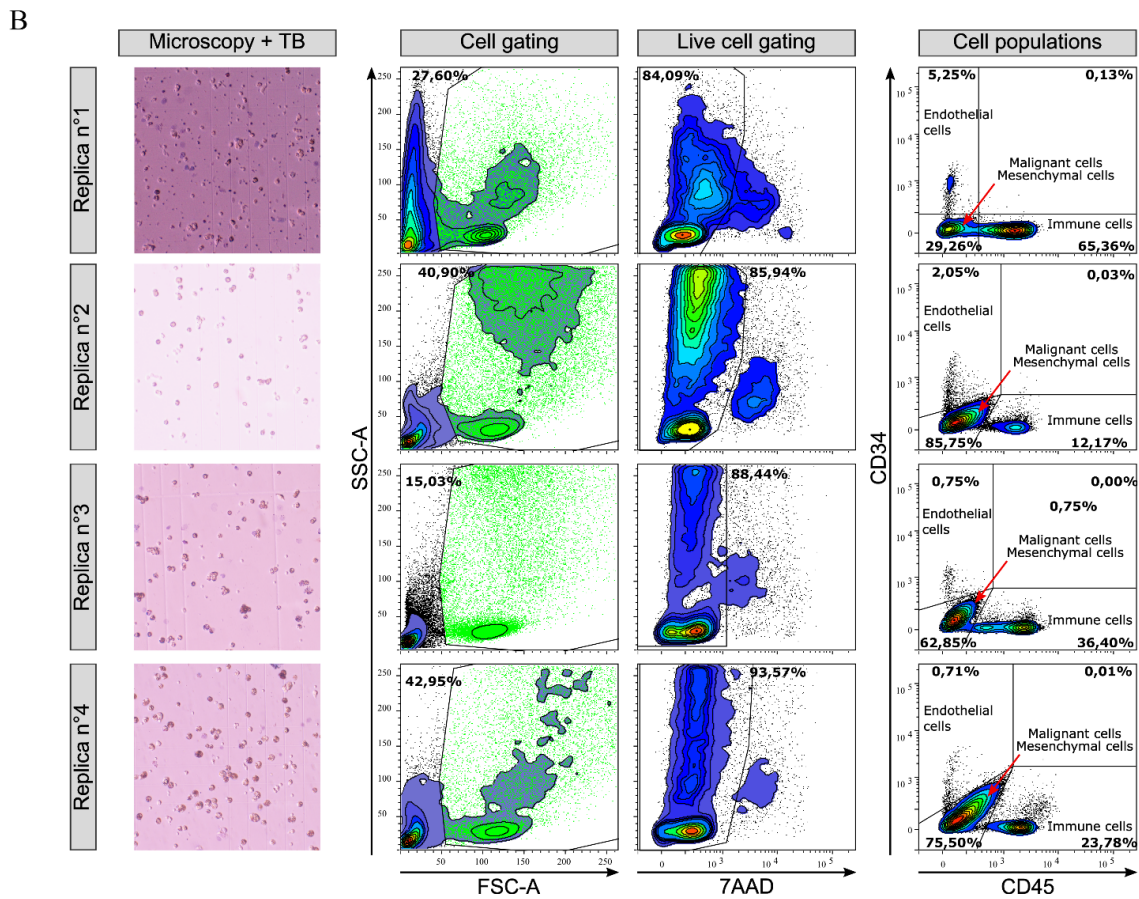
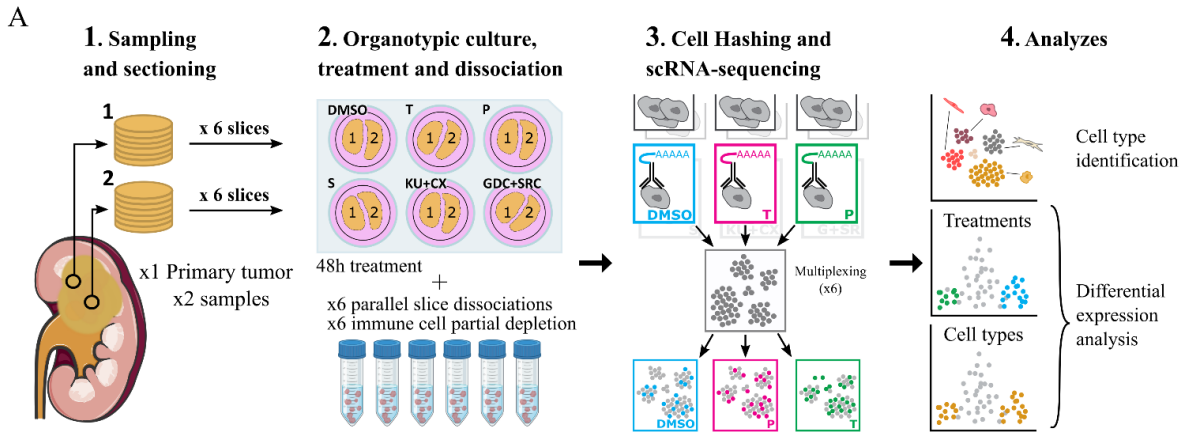


Fig. 2

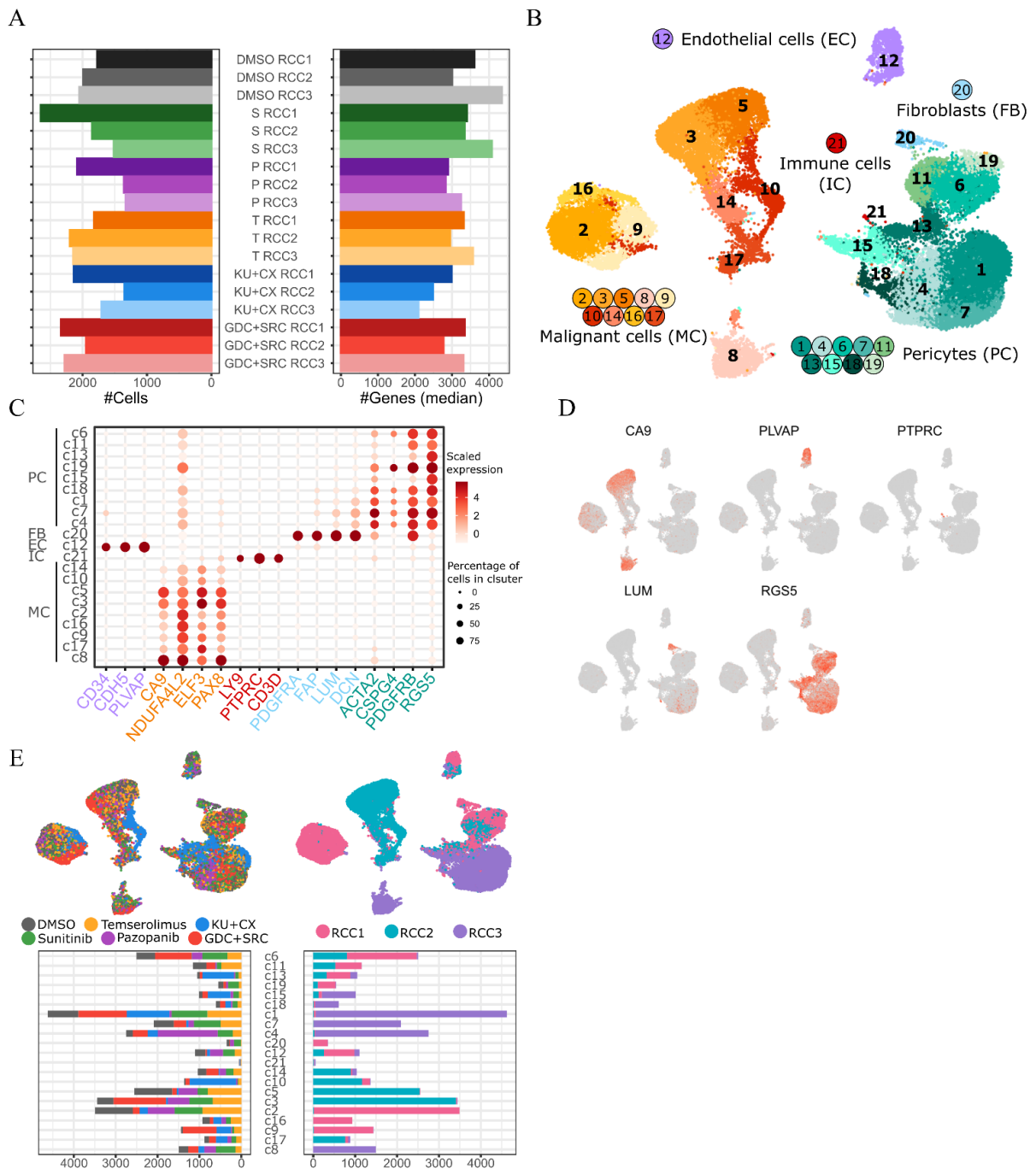
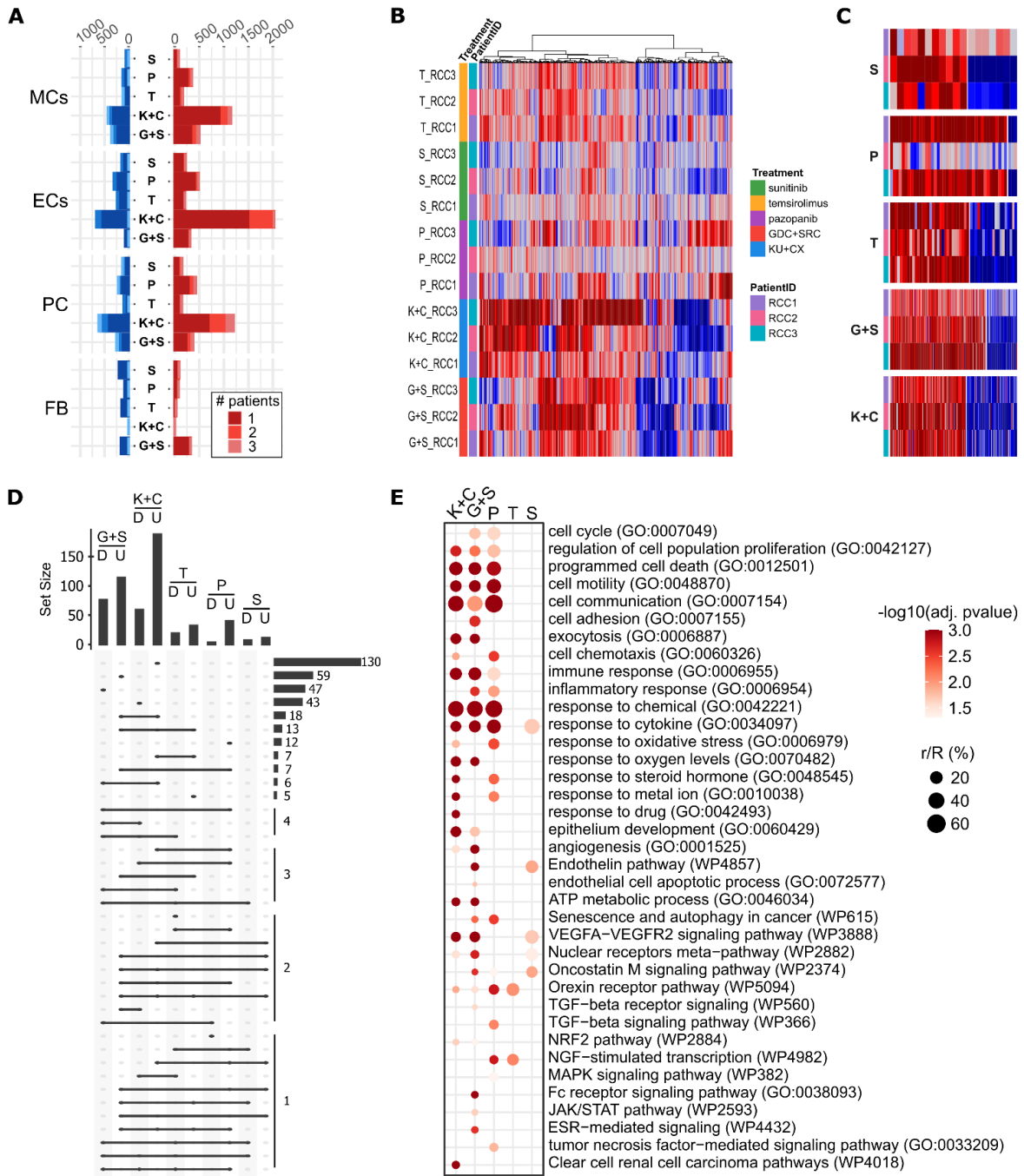


Fig. 3





# Supplementary figures

Fig. S1

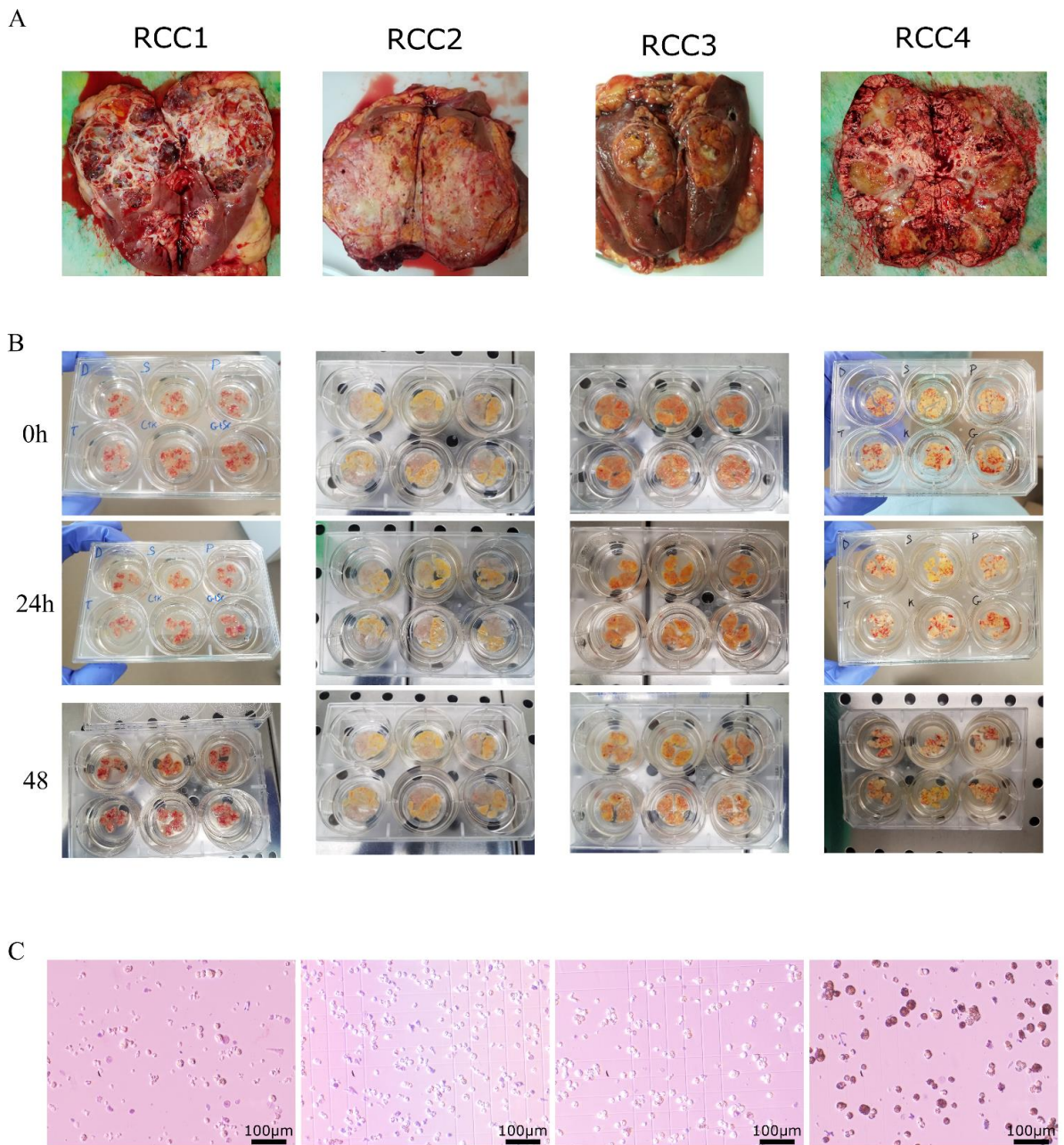


Fig. S2

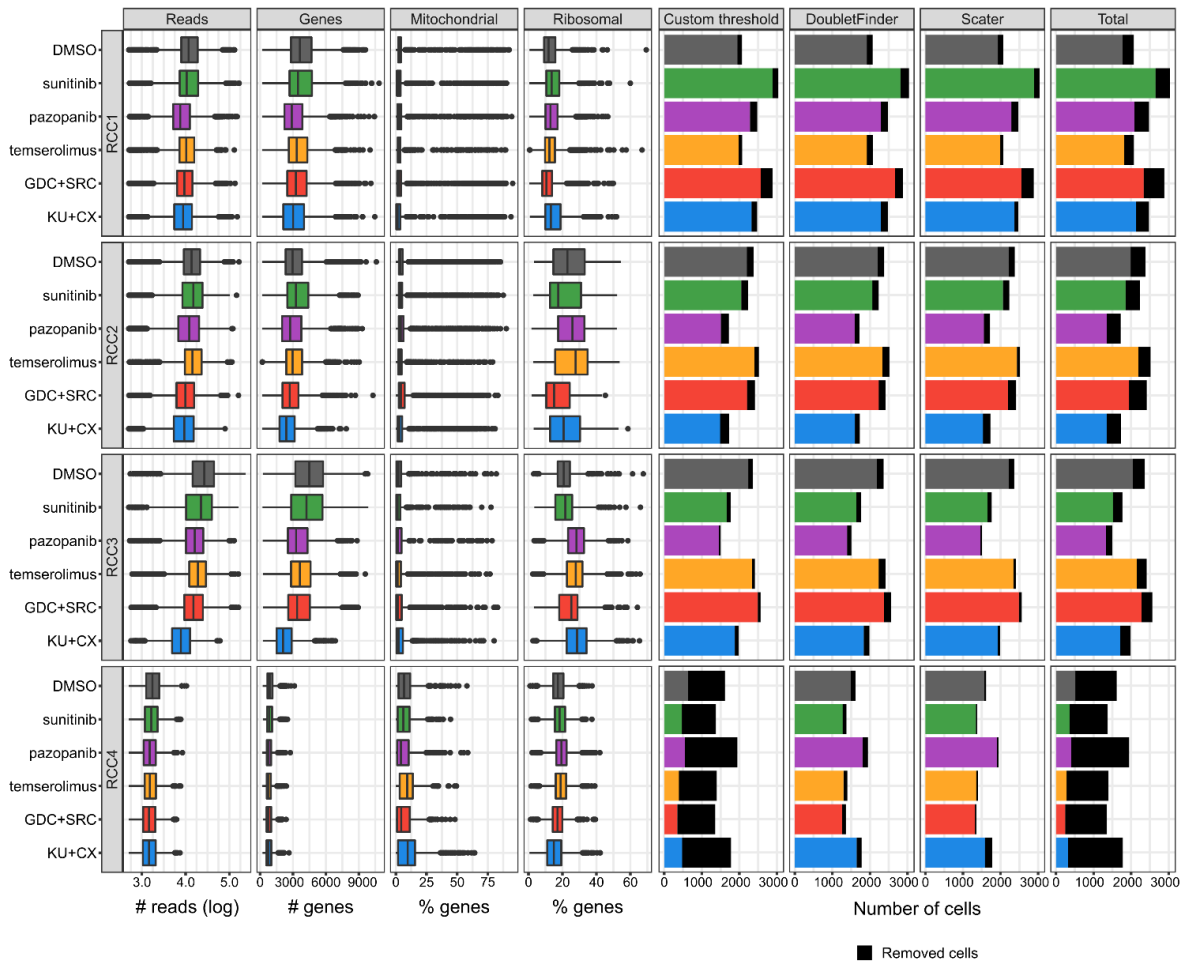


Fig. S3

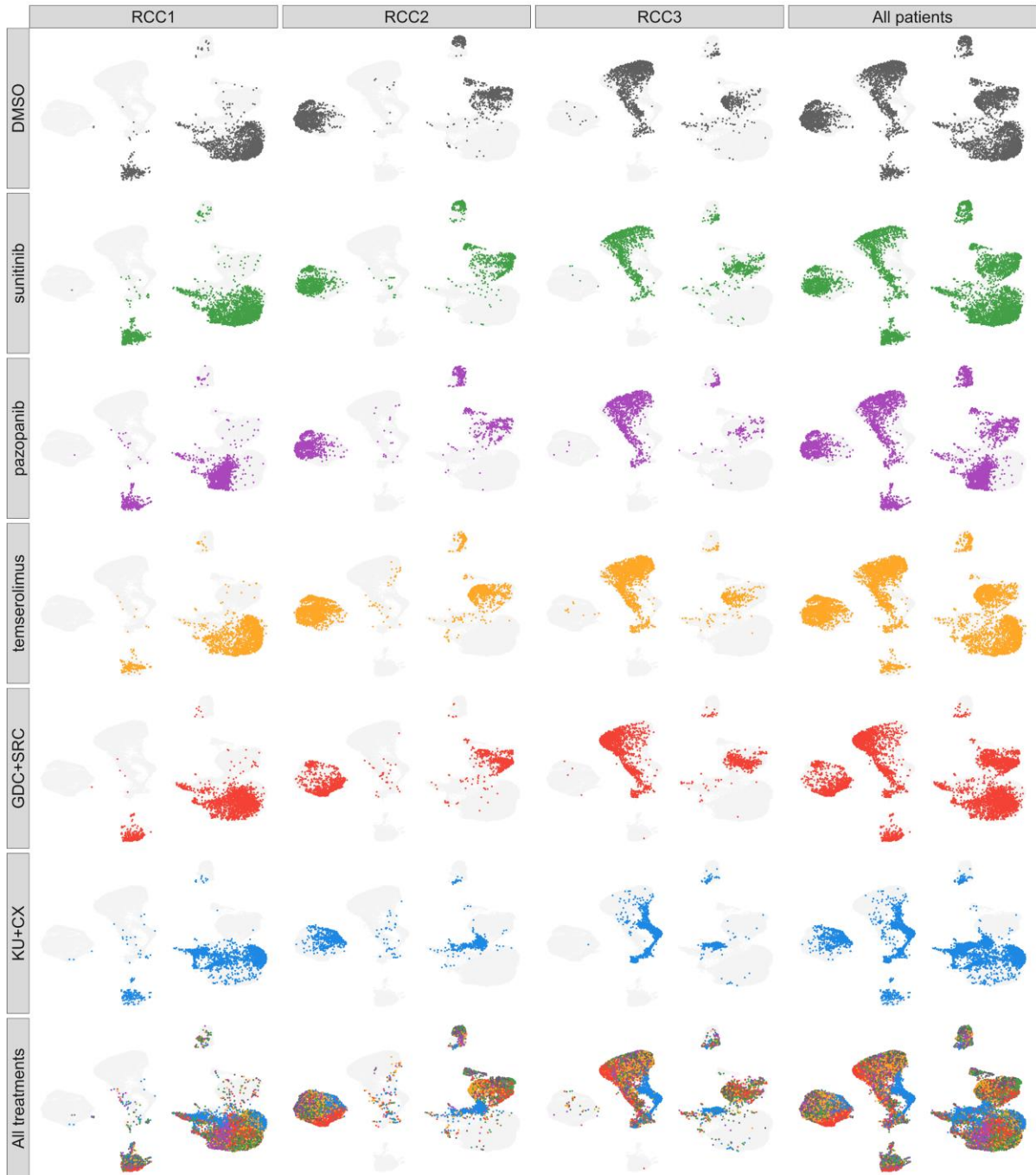


Fig. S4

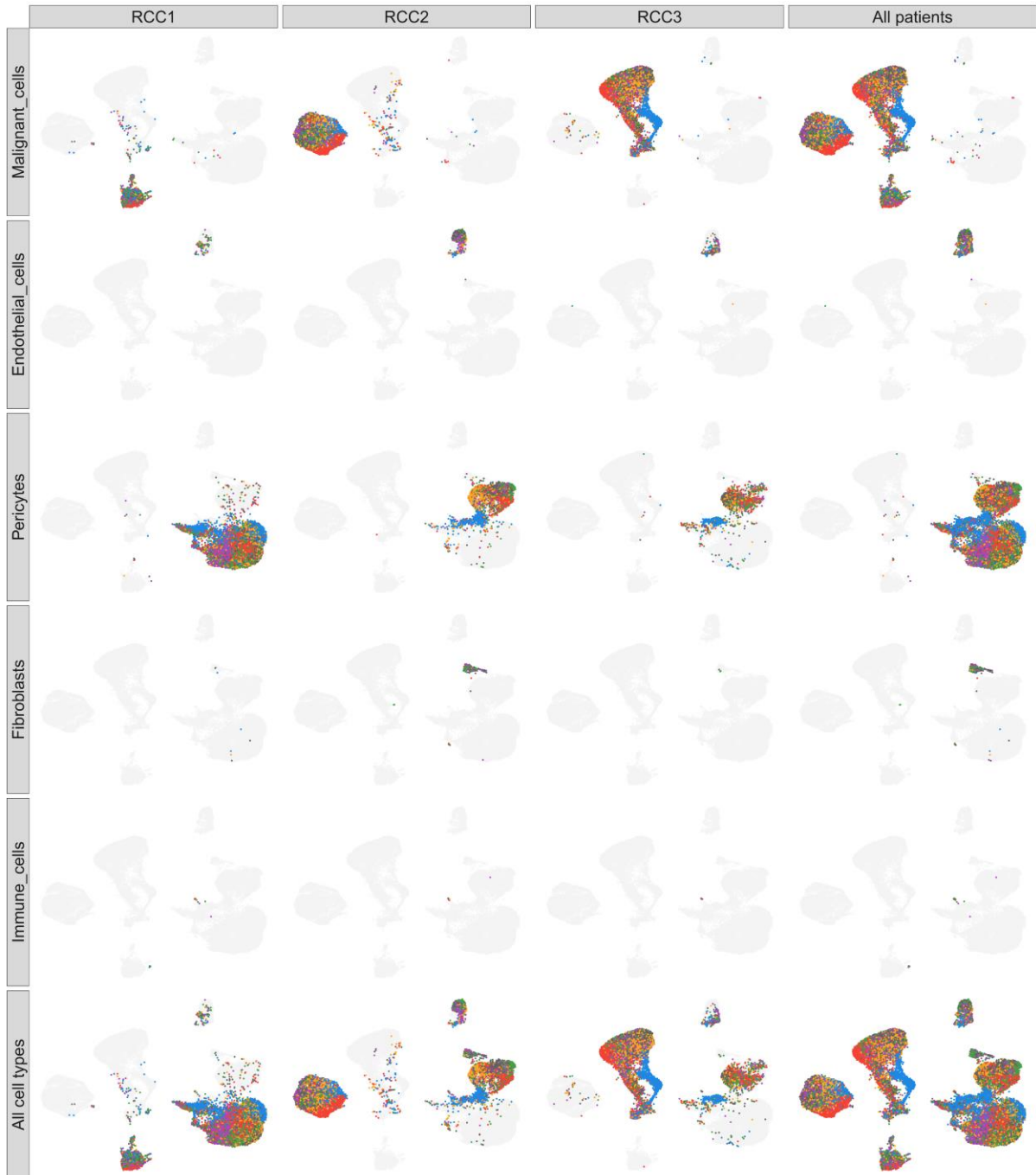
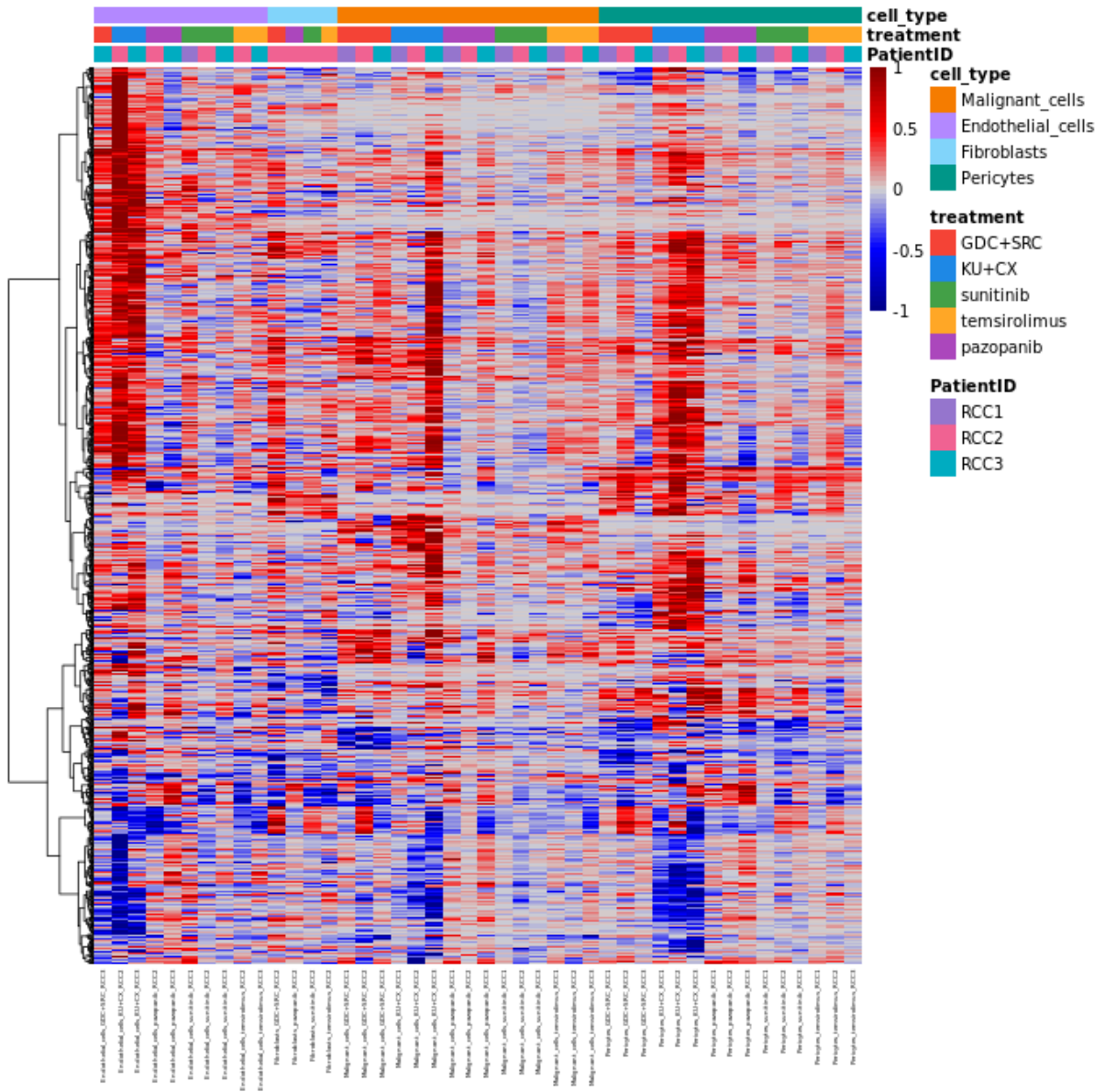


Fig. S5





## Supplementary tables

<b>Histological subtypes</b>	ccRCC	ccRCC	ccRCC	ccRCC
<b>Sex</b>	M	M	M	M
<b>Age</b>	56	51	77	66
<b>Size(cm)</b>	11	12	5	12
<b>ISUP</b>	3	4	1	4
<b>pTNM</b>	pT3a	pT3a	pT1b	pT3a
<b>Metastases</b>	no	no	no	no
<b>Sarcomatoid component</b>	no	yes	no	no
<b>Microscopic vascular embolism</b>	no	no	no	no
<b>Necrosis</b>	no	yes	no	no
<b>Hemorrhagic changes</b>	yes	yes	no	yes
<b>Embolism in renal vein</b>	no	no	no	no
<b>Hilar fat infiltration</b>	yes	yes	no	yes
<b>Perirenal fat infiltration</b>	no	no	no	yes
<b>Ganglion in Hil</b>	no	yes	yes	no

**Table S1. Clinical characteristics of patients participating in the study.** A total of 4 tumors were included in this study. ccRCC = clear cell renal cell carcinoma ; ISUP = International Society of Urological Pathology ; pTNM = Tumor-Node-Metastasis.

<b>Molecule</b>	<b>Short</b>	<b>Main targets</b>	<b>Other targets</b>	<b>Signaling pathways</b>
Sunitinib (SU11248)	S	PDGFRB	FLT3, ERN1, KIT	PDGFR signaling pathway
Pazopanib	P	FLT1, KDR, FLT4	KIT,PDGFR,FGFR3	VEGFR signiling pathway
Temsirolimus (CCI-779)	T	MTOR		mTOR signaling pathway
KU-60019	KU	ATM		ATM/ATR signaling pathway
Silmitasertib (CX-4945)	CX	CSNK2A1		
Pictilisib (GDC-0941)	GDC	PIK3CA, PIK3CB, PIK3CD, PIK3CG	MTOR	PI3K signaling pathway
SARACATINIB (AZD0530)	SRC	SRC, LCK, FYN, LYN, FGR	EGFR , YES1	Src signaling pathway

**Table S2. Characteristics of drugs used in the study.** Drugs's main characteristics presented in the table are taken from their profile page on selleckchem.com. The names of targets refer to HGNC nomenclature instead of protein names.

<b>Experiment</b>	<b>Cell viability 48h (%)</b>	<b>Immune cell proportion 48h (%)</b>	<b>endoth. cell proportion 48h (%)</b>
Replica n°1	84.09	65.36	5.25
Replica n°2	85.94	12.17	2.05
Replica n°3	88.84	36.4	0.75
Replica n°4	93.57	23.78	0.71
<b>Mean</b>	88.11	34.4275	2.19

**Table S3. Cell viability and cell proportions measured by flow cytometry.** To validate our dissociation protocol on organotypic culture of ccRCC slices, we used flow cytometry on 4 tumor samples.



# Supplementary materials

## Supplementary methods

### **Ethics statement and human tumor specimens**

Human samples were obtained from the processing of biological samples through the Centre de Ressources Biologiques (CRB) Santé of Rennes BB-0033-00056. The research protocol fulfilled the requirements of the local institutional ethics committee. Informed consent was signed by each patient. Fresh nephrectomy samples were collected from untreated patients who underwent radical nephrectomy for a ccRCC in the urology department of the university-hospital center of Rennes. Clinical characteristics of patients are provided in table S1.

### **Preparation of tumor slices**

Resections were manually minced using a sterile scalpel and two macroscopically distinct and representative areas were sampled and soaked in ice-cold sterile HBSS. The two samples were immersed side by side in a warm 2%-agarose solution within a 6-well plate well and immediately placed on ice for five minutes. The resulting solidified agarose block containing the tissue was removed from the well and prepared for further cuts using the Vibratome VT1200 (Leica Microsystems, Wetzlar, Germany) under sterile conditions. Tissue slice thickness was set to 400  $\mu\text{m}$ . Vibration amplitude was set between 2.95 mm. Slicing speed, ranging from 0.04 to 0.1mm/s, was optimized according to tissue density and type; slower slicing speed was used on softer tissues and vice versa.

### **Organotypic tissue culture**

Slices were deposited on inserts (Millicell Cell Culture Insert, Réf: PICMORG50) for 48h. in order to allow regulated exchanges of gas and nutrients between tissue and medium. Several slices of each sample were placed in each insert in order to occupy the maximum surface on the insert and optimize the quantity of treated tissue. Organotypic inserts are Teflon membranes with 0.4µm pores that allow preservation of 3D tissue structure in culture. Tissue culture was performed at 37°C in a 5% CO<sub>2</sub> humidified incubator using 1ml of DMEM media supplemented with 20% inactivated FBS (GIBCO), 100 U/ml penicillin (Invitrogen) and place in a rotor agitator to allow gas and fluids exchanges with the medium.

### **Treatments used**

Drugs used in this study were provided by SelleckChem (<https://www.selleckchem.com/>) and used at the concentration of 10µM.

### **Tissue dissociation and cell hashing labeling**

After 48h of treatment, tumor tissue slices of each treatment condition were individually subjected to gentle enzymatic and manual disaggregation (tumor dissociation kit, Miltenyi Biotec) using a pipette, before undergoing red blood cell lysis and a low-speed centrifugation at 800 rpm to remove debris (Fig. 1A, step1). Partial depletion of CD45 (an immune cell marker) positive cells was performed using the Miltenyi kit (CD45 MicroBeads, human) to enrich the samples in malignant/stromal cells and remove circulating and tumor-infiltrating immune cells (step2). Dead cell removal (Miltenyi) was performed on each sample individually before pooling them. Pooled cells were then labeled with hashtag-antibodies, following

manufacturer's instructions [19]. These are directed against ubiquitously expressed B2M and CD298 epitopes and allow to identify each treatment condition for a given cell. Cells were then pooled in equal proportions and viability was estimated under microscope (Leica systems, Olympus IX70) using Trypan Blue before proceeding to single-cell capture.

### **Single-cell capture, library preparation and sequencing**

The construction of single-cell libraries were carried out according to 10X Genomics and Biolegend's instructions (TotalSeq™-A Antibodies and Cell Hashing with 10x Single Cell 3' Reagent Kit v3 or v3.1 [Single Index] Protocol). Single-cell libraries were quality-controlled using a bioanalyzer (Agilent Technologies) and sequenced on an Illumina NovaSeq 6000 instrument using a 2x100 paired-end sequencing protocol (Illumina, San Diego, CA).

### **Data analysis of single-cell data**

*Data preprocessing.* Demultiplexed raw sequencing reads were mapped to the GRCh38 human reference genome and quantified based on unique molecular identifiers (UMIs) with the Cell Ranger pipeline (v3.0.2, 10x Genomics). The resulting raw count matrices from the three individual samples were then aggregated (Cell Ranger). Cells corresponding to different treatment conditions were demultiplexed using the HTODemux function implemented in Seurat v4.0. Individual matrices were then aggregated in a global matrix containing the sample source information.

*Data quality control and normalization.* Following the generation of the seurat object from the 24 samples (four tumors and six treatment conditions), cells expressing less than 500 genes

and having <2,000 UMIs or >100,000 UMIs were removed, as well as genes expressed in <10 cells. Subsequently, a variety of QC (quality control) metrics with selected variables corresponding to the numbers of UMIs and detected genes per cell, the proportions of UMIs mapping mitochondrial and ribosomal genes was performed to identify and remove low quality and dead cells using the scater R package. (v1.10.1 ). The dataset was normalized using the NormalizeData and the SCTransform functions implemented in Seurat.

***Reduction of dimensionality and clustering.*** The top-3000 most varying genes were used to perform a principal component analysis with the RunPCA function implemented in Seurat. Cells were then clustered by using the FindNeighbors and FindClusters functions on the top-50 principal components, with default parameters. Finally, we used the uniform manifold approximation and projection (UMAP) method implemented in Seurat to project single cells in a low-dimensional space.

***Cluster annotation and differential gene expression analysis.*** Cell clusters were associated to cell type by using a set of well-known cell marker genes (Fig. 2C) and manual curation by experts. Differentially expressed genes were identified with the wilcoxauc function implemented in presto (v.1.0.0) [20].

***Functional analysis.*** The AMEN suite of tools (ref) was used to calculate the Fisher exact probability and the Gaussian hypergeometric test was used to identify significantly enriched terms from the Gene Ontology. A specific term was considered enriched in a group of coexpressed genes if the adjusted (Benjamini–Hochberg procedure) *p value* was  $\leq 0.05$ .

## **Heatmap and spot plot**

Heatmaps were generated with the pheatmap package v1.0.12 (<https://CRAN.R-project.org/package=pheatmap>). Spot plots were generated using the ggplot package [16].

## **Data availability**

Our raw sequencing data were submitted to the XX under the accession number XXX. Processed single-cell RNA-seq data are available on XXX. All processed data are also available through UncoVer : The Uro Oncology Viewer. (UncoVer: <https://uncover.genouest.org>) database (Lecuyer, Saout et al., in preparation).

## **Supplementary results**

### **Cultured slice dissociation display good overall single-cell viability**

To evaluate the impacts of inhibitor molecules, we set a strategy combining an organotypic culture model that recapitulates well the tumor 3D architecture, microenvironnement and physiology, cell-hashing multiplexing and scRNA-sequencing (Fig. 1A; Table S2). Regarding organotypic culture, fresh nephrectomy samples were collected from untreated patients who underwent radical nephrectomy. Each tumor was sampled in two macroscopically distinct areas in order to capture the ITH (Fig. 1). Tumor slices were prepared and cultured 48h. Although the culture of ccRCC slices has already been validated, their dissociation has not been evaluated yet. Hence, we used flow cytometry and trypan blue staining to evaluate the dissociated cell viability and diversity after 48h of culture (Fig. 1B). Microscopic examination revealed well individualized and diverse cells, in terms of shape and size (n=4). Noteworthy, there is only few to no debris, suggesting an overall good shape of tumor cells before and during

dissociation. The live cell gating using 7AAD staining in flow cytometry revealed an average viability rate of 88,11%. Double CD34 and CD45 staining revealed that cell suspensions are mainly composed of stromal, malignant and immune cells, with a little endothelial cell representativeness (Table S3). Overall, we are able to efficiently dissociate ccRCC tumor slices after culture.

### **Five kinase inhibitors**

The molecules used in this work are kinase inhibitors targeting various pathways (Table SX). They include three gold-standard therapies, the multi-target Sunitinib (VEGFR and PDGFR inhibitor), Pazopanib (VEGFR inhibitor) and Temsirolimus (mTOR inhibitor), as well as two combination therapies [KU-60019 (ATM inhibitor) combined with CX-4945 (CK2 inhibitor); GDC-0941 (class I PI3K isoform inhibitor) combined with Saracatinib (SRC inhibitor).

### **Quality controls revealed a low quality sample**

To check the quality of our data, we measured the number of UMIs and genes detected per cell (Fig. S1). RCC1 to RCC3 samples displayed a number of UMIs ranging from 7,515 and 26,687 and a median of detected genes per cell ranging from 2,093 to 4,485. On the other hand, RCC4 showed 1,457 to 1,742 features and a median of 753 to 879 detected genes, which are low values typical of excessive mortality before capture. This was also confirmed in RCC4 cells by the high percentage of reads mapping mitochondrial genes, and the low percentage of reads mapping ribosomal genes. Consistently, when checking for outlier cells with DoubletFinder and Scater tools, we found a significant proportion of outliers in all RCC4 samples. Altogether, we decided to remove RCC4 samples from the analysis.

## References

- [1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- [2] Hsieh JJ, Purdue MP, Signoretti S, Swanton C, Albiges L, Schmidinger M, et al. Renal cell carcinoma. *Nat Rev Dis Primers* 2017;3:17009.
- [3] Makhov P, Joshi S, Ghatalia P, Kutikov A, Uzzo RG, Kolenko VM. Resistance to Systemic Therapies in Clear Cell Renal Cell Carcinoma: Mechanisms and Management Strategies. *Mol Cancer Ther* 2018;17:1355–64.
- [4] Ross K, Jones RJ. Immune checkpoint inhibitors in renal cell carcinoma. *Clin Sci* 2017;131:2627–42.
- [5] Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Rixe O, et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med* 2007;356:115–24.
- [6] Molina AM, Lin X, Korytowsky B, Matczak E, Lechuga MJ, Wiltshire R, et al. Sunitinib objective response in metastatic renal cell carcinoma: analysis of 1059 patients treated on clinical trials. *Eur J Cancer* 2014;50:351–8.
- [7] Crusz SM, Tang YZ, Sarker S-J, Prevoo W, Kiyani I, Beltran L, et al. Heterogeneous response and progression patterns reveal phenotypic heterogeneity of tyrosine kinase inhibitor response in metastatic renal cell carcinoma. *BMC Med* 2016;14:185.
- [8] What happened to personalized medicine? *Nature Biotechnology* 2012;30:1–1. <https://doi.org/10.1038/nbt.2096>.
- [9] Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012;366:883–92.
- [10] Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* 2013;501:338–45.
- [11] Meijer TG, Naipal KA, Jager A, van Gent DC. tumor culture systems for functional drug testing and therapy response prediction. *Future Sci OA* 2017;3:FSO190.
- [12] Roelants C, Pillet C, Franquet Q, Sarrazin C, Peillon N, Giacosa S, et al. Ex-Vivo Treatment of Tumor Tissue Slices as a Predictive Preclinical Method to Evaluate Targeted Therapies for Patients with Renal Carcinoma. *Cancers* 2020;12. <https://doi.org/10.3390/cancers12010232>.
- [13] Giacosa S, Pillet C, Séraudie I, Guyon L, Wallez Y, Roelants C, et al. Cooperative Blockade of CK2 and ATM Kinases Drives Apoptosis in VHL-Deficient Renal Carcinoma Cells through ROS Overproduction. *Cancers* 2021;13. <https://doi.org/10.3390/cancers13030576>.
- [14] Caroline R, Sofia G, Catherine P, Remi B, Pierre C, Olivier B, et al. Correction: Combined inhibition of PI3K and Src kinases demonstrates synergistic therapeutic efficacy in clear-cell renal carcinoma. *Oncotarget* 2019;10:2236.
- [15] Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;36:411–20.
- [16] Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 2014;344:1396–401.
- [17] Bi K, He MX, Bakouny Z, Kanodia A, Napolitano S, Wu J, et al. Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma. *Cancer Cell* 2021;39:649–61.e5.
- [18] Young MD, Mitchell TJ, Vieira Braga FA, Tran MGB, Stewart BJ, Ferdinand JR, et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* 2018;361:594–9.
- [19] Stoeckius M, Zheng S, Houck-Loomis B, Hao S, Yeung BZ, Mauck WM 3rd, et al. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol* 2018;19:224.
- [20] Korsunsky I, Nathan A, Millard N, Raychaudhuri S. Presto scales Wilcoxon and auROC analyses to millions of observations. *bioRxiv* 2019. <https://doi.org/10.1101/653253>.

#### 4.2.4. DISCUSSION ET PERSPECTIVES

Dans ce deuxième axe de recherche, j'avais pour objectif d'étudier les impacts de thérapies ciblées sur l'HIT dans le ccRCC, à l'échelle de la cellule unique. Pour cela, j'ai adapté un protocole de culture organotypique pour le rendre compatible avec mon approche. Les principaux problèmes rencontrés concernent l'exploitation des échantillons. En effet, la qualité des tumeurs et la quantité prélevée sont très variables. Pour commencer, beaucoup de tumeurs n'étaient pas exploitables car trop petites, ou présentaient trop de remaniements œdémateux ou hémorragiques. Ensuite, certains échantillons ne se maintenaient pas pendant la découpe au vibratome, avant que je n'améliore la technique en utilisant une gaine d'agarose. D'autres échantillons se révélaient de mauvaise qualité après 24h de culture et ne permettaient pas de prolonger la culture à 48h. Quand il était possible de dissocier des coupes cultivées, la qualité de la dissociation ne permettait pas toujours de récolter suffisamment de cellules et de réaliser de la cytométrie en flux. Après ces mises au point, j'ai ensuite appliqué la stratégie d'analyse par single-cell RNA-seq sur quatre tumeurs. Pour l'une d'entre elle, le RCC4, la préparation de la librairie de séquençage a mené à de très faibles quantités d'ADNc. En amplifiant davantage cet ADNc, nous sommes parvenus à une quantité suffisante et avons souhaité séquencer l'échantillon, ce qui a révélé de très faible quantité de lectures de séquençage et de gènes détectés. Pour cette raison, nous avons écarté l'échantillon lors de l'analyse. Je ne connais pas la raison de cet échec, mais il m'a semblé intéressant de l'inclure dans les données présentées et de discuter ce résultat. Comme pour chaque expérience, j'ai validé la qualité de la suspension cellulaire avant capture, sur la base de mes observations au microscope et de l'expérience accumulée. L'examen des photographies de la suspension colorée au bleu de trypan a montré a posteriori que les cytoplasmes des cellules du RCC4 étaient marrons et granuleux (se référer à la figure supplémentaire 1 de l'article en préparation). Je n'ai pas trouvé d'études ou de sources informelles d'information permettant de mieux comprendre l'aspect des cellules de ccRCCs dissociées sous microscope. Je pense que cette apparence renvoie à un état donné des cellules tumorales, dans lequel une substance est libérée dans le cytoplasme et nuit probablement à l'intégrité de la cellule. Si la qualité de l'échantillon est en cause, cet écueil révèle l'impossibilité d'anticiper jusqu'à la dissociation la qualité d'un échantillon, démontrant la difficulté de travailler avec du tissu humain pour la mise en place de protocoles ambitieux. Aujourd'hui, si une suspension cellulaire présentait le même aspect que ces cellules, je ne procéderai pas à la capture et au séquençage des cellules par précaution.



Ce projet n'est pas finalisé. Ma thèse aura davantage vu sa conception et sa mise en œuvre expérimentale que l'aboutissement de son analyse. Les résultats préliminaires montrent tout de même des sensibilités variables aux traitements, et permettent d'observer des déplacements de populations dans l'espace du UMAP, indiquant des modifications de leur transcriptome en réponse. L'analyse d'expression différentielle donne aussi des résultats préliminaires intéressants puisque chaque traitement induit des différences d'expression spécifiques. Ce projet est encore en cours. Nous avons réalisées des analyses relativement basiques et devons les élargir ; par exemple préciser les réponses au sein des sous-populations tumorales ou intégrer davantage l'information des mécanismes et voies ciblées par les molécules. De façon générale, cette phase d'analyse représente un défi car peu d'outils et aucune approche standard n'existent encore à l'heure actuelle pour prendre en charge des données de traitement à l'échelle de la cellule unique.

Nous avons retiré les cellules immunitaires, qui restaient en proportions importantes après dissociation et occultaient les autres types cellulaires. Cette déplétion est en grand décalage avec la plupart des études scRNA-seq dans le ccRCC qui se concentrent sur la caractérisation du compartiment immunitaire. Elle est justifiée par l'enrichissement des cellules malignes cependant car c'est dans ces cellules qu'agissent principalement les TKIs. Cela pose la question des modèles précliniques à utiliser pour étudier de la même façon les effets des immunothérapies à l'échelle de la cellule unique. Tous les modèles ont des avantages et des inconvénients, cependant, il est possible d'envisager le recours à des transplantations sur des souris humanisées, afin d'avoir un modèle avec une systémique. Il sera alors nécessaire d'y adapter l'utilisation du scRNA-seq, à l'image des mises au point dans ce projet, notamment en ce qui concerne la quantité et la qualité des tissus.

Une des limites de mon protocole est que je peux difficilement prolonger la culture à plus de 48h, et je ne l'ai pas optimisé en ce sens. Il faut noter que j'examine la dissociation tissulaire après culture, au contraire de la majorité des études utilisant la culture organotypique de coupes de tumeur. Dans ces études, les auteurs prolongent parfois la culture jusqu'à 4 ou 5 jours (Meijer et al., 2017), sans dissociation toutefois. Or cette dernière induit un stress cellulaire important, d'autant plus impactant que les coupes ont été fragilisées par la culture *ex vivo*. De façon générale, mes résultats supplémentaires montrent que des paramètres sont encore à améliorer, par exemple pour prendre en charge de façon plus systématique chaque tumeur, le plus indépendamment possible de la qualité du tissu.





## 4.3. UNCOVER : UNE RESSOURCE WEB DE DONNEES GENOMIQUES DEDIEE AUX CANCERS UROLOGIQUES

### 4.3.1. INTRODUCTION

Depuis 2009 et la première étude parue, le nombre de publications référencées dans Pubmed en recherchant « single-cell RNA-sequencing » ne cesse de croître, jusqu'à atteindre plus de 3020 en 2022 Parmi elles, plus du tiers concerne la cancérologie (« single-cell RNA-sequencing AND cancer » ; 1147 résultats). La progression fulgurante de cette technologie entraîne de rapides évolutions dans son état de l'art, à la fois expérimentales et analytiques (Lähnemann *et al.*, 2020; Zappia and Theis, 2021). De façon comparable au début du scRNA-seq, nous assistons à un développement important des technologies de transcriptomique spatiale basées sur les code-barres, à l'image du Visium de 10X Genomics, qui devraient prochainement s'imposer dans tous les laboratoires (Wu *et al.*, 2022). Pour accompagner cet essor important, il devient nécessaire de développer, en parallèle, des bases de données qui répertorient ce type d'études transcriptomiques de nouvelle génération et donne accès à leur analyse. Par ailleurs, étant donné leur nombre grandissant (G. Sun *et al.*, 2021), il est nécessaire d'organiser ces études en thématiques afin de centraliser et organiser les données de façon cohérente, d'après des annotations réalisées par des experts du domaine. L'une des façons de procéder pourrait être de s'adresser aux communautés travaillant sur des types de cancers spécifiques afin d'organiser la connaissance issue de cette multidisciplinarité de façon cohérente. À cet égard, la communauté de l'uro-oncologie exprime une forte demande d'accès aux analyses de ces nouvelles technologies pour envisager de les traduire en bénéfices cliniques. À la fois grande et très dynamique, elle est encline à recourir à des outils tels que les bases de données scRNA-seq et de transcriptomique spatiale. Les données de scRNA-seq sont riches d'informations et permettent de poser des questions variées, comme l'illustre la réutilisation fréquente de jeux de données déjà publiés, par exemple dans le cancer du rein (Fendler *et al.*, 2020; Hu *et al.*, 2020; Liu *et al.*, 2021). En ce qui concerne la transcriptomique spatiale, sa dimension histologique couplée à la transcriptomique haut-débit intéresse beaucoup les cliniciens au titre de son rapprochement avec leur expertise en histopathologie. De façon surprenante, les bases de données de scRNA-seq généralistes (Y. Cao *et al.*, 2017; Franzén, Gan and Björkegren, 2019; Papatheodorou *et al.*, 2020) ou axées sur la thématique des cancers (Yuan *et al.*, 2019; D. Sun *et al.*, 2021; Zeng *et al.*, 2022) contiennent très peu ou pas d'études sur les cancers urologiques, tandis que plus d'une vingtaine a déjà été publiée. Sur la base de travaux précédents (Darde *et*

*al.*, 2015, 2019), des rôles connus de l'HIT dans les cancers urologiques (Gerlinger *et al.*, 2015), ainsi que des expertises et de la demande identifiées dans le domaine, mon équipe d'accueil a entrepris de développer une ressource web ayant la double fonction de répertorier systématiquement les données de scRNA-seq accessibles dans le domaine de l'uro-oncologie et de permettre à un public non expert de les explorer et de les analyser. Afin d'anticiper les futures attentes liées aux progrès technologiques, nous avons aussi intégré un premier jeu de données de transcriptomique spatiale sur le cancer de la prostate. Dans ce contexte, mon rôle était notamment de répertorier les études et les données cliniques, ainsi que de formater les données, en collaboration avec Gwendoline Lecuyer, actuellement en thèse au sein du laboratoire, pour les rendre compatibles avec l'interface de l'outil UncoVer.

**4.3.2. MANUSCRIT EN PREPARATION**



## **Format: Letter**

# **UncoVer: A web-based resource for single-cell and spatial transcriptome exploration and comparison for the urologic oncology community**

Gwendoline Lecuyer<sup>1,✉</sup>, Judikael Saout<sup>1,✉</sup>, Thomas A. Darde<sup>2</sup>, Laurence Noël<sup>1</sup>, Paul Rivaud<sup>1</sup>, Matéo Boudet<sup>3,4</sup>, Simon Léonard<sup>1,5,6</sup>, Olivier Sallou<sup>3</sup>, Nathalie Rioux-Leclercq<sup>1,7</sup>, Olivier Collin<sup>3</sup>, Aurélie Lardenois<sup>1</sup>, Frédéric Chalmel<sup>1,\*</sup>

<sup>1</sup> Univ Rennes, Inserm, EHESP, Irset (Institut de Recherche en Santé, Environnement et Travail), UMR\_S 1085, F-35000 Rennes, France.

<sup>2</sup> SciLicum, 10 rue de la sauvaie, 35200 Rennes, France.

<sup>3</sup> Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA/INRIA) - GenOuest platform, Université de Rennes 1, F-35042 Rennes, France.

<sup>4</sup> IGEPP, Agrocampus Ouest, INRAE, Université de Rennes 1, Le Rheu, France.

<sup>5</sup> LabEx IGO “Immunotherapy, Graft, Oncology”, F-35043 Nantes, France.

<sup>6</sup> Univ Rennes, INSERM, EFS, UMR S1236, F-35000 Rennes, France.

<sup>7</sup> Department of Pathology, University of Rennes 1, Rennes, France.

\* To whom correspondence should be addressed

✉ These authors contributed to the manuscript equally

Correspondence: [frederic.chalmel@inserm.fr](mailto:frederic.chalmel@inserm.fr)



# Abstract

In a context of increased interest in tumor heterogeneity and its role in phenotypes such as therapeutic escape or tumor progression and relapse, single-cell RNA-sequencing has emerged as an essential technology for its ability to describe the variety of cell types within a tumor tissue. While this brings new insights, the resulting data have great potential for clinicians but their analysis requires specific bioinformatic skills. To address this issue, we here report the development of UncoVer (The Uro-oncogenomics Viewer), a web-based resource for the visualization and comparison of published single-cell and spatial transcriptomic datasets in Uro-Oncology. The system currently embeds 24 published studies related to urologic neoplasms. Datasets have been reproducibly processed based on a standard workflow supplemented by expert manual curation and a single-cell meta-analysis of genitourinary malignancy atlas composed of more than 730,000 cells has been assembled. UncoVer has a user-friendly interface allowing one to display single-cell and spatial transcriptomic atlas and to query any given gene expression in multiple types of graphical representations. UncoVer is freely available at <https://uncover.genouest.org/>.

## **Patient summary**

Here we introduce an online resource, UncoVer, that allows researchers and clinicians to explore the cellular complexity and heterogeneity of urologic cancers at a single-cell resolution, for the benefit of translational research. Due to its user-friendly interface, users can easily filter data based on clinical and experimental information but also query specific gene expression patterns.

## **Main text**

Genitourinary malignancies account for over 2.4 million cumulative new cases and more than 760,000 deaths in 2020 [1]. Intra-tumor heterogeneity (ITH) has been reported in main urologic cancers and is thought to be a crucial determinant of treatment failure and tumor progression [2]. Over the last few years ITH has been extensively described in many types of tumors with single-cell transcriptome approaches, which has already generated large amounts of data [3,4]. While single-cell datasets and the emerging spatially resolved transcriptomics technologies have great translational potential and must be made accessible for interpretation by the community at large, processing and analyzing these data require specific bioinformatic skills. Although many resources with either broad spectrum or focusing on cancers (Table S1) give access to large single-cell datasets, the urologic oncology field remains generally under-represented, especially regarding the malignant cells. Other transcriptomic resources in urology exist [5–7] (Table S1) but they are not able to manage single-cell and spatial transcriptome data. To meet the pressing need to gather such data within a centralized database, we developed the Uro-oncogenomics Viewer (UncoVer), an online resource designed to assist researchers and clinicians for the exploration and comparison of single-cell and spatial profiling studies accumulated within the field of uro-oncology.

UncoVer has a user-friendly interface that meets all modern web standards and enables users to display it in a web browser on any device (Fig. 1A). Within the system, information is organized in a two-layer architecture (study > dataset): studies are usually associated with peer-reviewed publications and are divided into datasets describing distinct technologies and sub-analyses, according to the original work. All datasets were analyzed from raw count matrices by using a bioinformatic workflow that meets all current standards followed by an expert manual curation (supplementary materials, Fig. 1A, Table S2). In case the authors provided additional information about their published analysis (such as cell clusters or cell coordinates in low-dimensional spaces), a dedicated dataset is created and titled with the prefix 'Article'. Briefly, the 'Browse available studies' button offers rapid access to the list of studies hosted by the system. Studies and their associated publications are presented in a table that can be filtered or sorted according to several features (e.g. technology, tissue type, species). Once a study is selected, users are then invited to choose one of the related datasets on a web page providing detailed information. The user is then redirected to the "exploration" interface which is composed of four main panels (Fig. 2A-D). Briefly, the first panel (Fig. 2A) corresponds to a filtering component allowing users to filter cells based on any kind of available metadata, such as cell types ('CellType') or cell cycle phase ('CellCyclePhase'). The second panel (Fig. 2B) provides three independent customizable graphical components (ICGCs) (pie chart or bar plot as deemed appropriate by users) describing cell distributions according to selected metadata. The third panel (Fig. 2C) includes two ICGCs (scatter, hexagonal or density plot) which allow users to project cells on available low-dimensional spaces (such as UMAP) and to color cells based on selected metadata. Finally, the fourth panel (Fig. 2D) contains two ICGCs (scatter, hexagonal, density, violin or spot plot) dedicated to the analysis of gene expression. Users can choose genes among top variable features or select their own set of genes. Noteworthy, each ICGC in third and fourth panels can be zoomed in/out and exported as a static image file (PNG format). Regarding spatial transcriptomic datasets, cells correspond to spots and the corresponding hematoxylin-and-eosin-stained tissue sections are used as background images in the ICGCs of the third and fourth panels.

In its current state UncoVer embeds 24 single-cell RNA-seq (scRNA-seq) studies representing 77 datasets for a total of 113 patients and 249 samples (Table S3). Although not its primary target, the resource also includes two massive bulk transcriptomic datasets for a total of 2,727 samples of genitourinary malignancies (including 1,519 renal, 735 prostate, 464 bladder, 5 ureter and 5 testis cancers) from the Expression Project for Oncology (expO, microarray data) and from the Cancer Genome Atlas program (TCGA, RNA-seq data) (Table S1) in which samples are clustered based on their transcriptomic profiles and projected on a low-dimensional space. Anticipating that spatial transcriptomics will be widely spread in laboratories [8], we made UncoVer compatible with such technologies and already included one prostate cancer dataset.

Finally, UncoVer also includes an original study corresponding to a massive pan-cancer meta-analysis composed of 19 10X Genomics' scRNA-seq studies hosted by the system, which have been harmonized to reduce batch-effects (Fig. 2). This unique atlas of genitourinary malignancies is composed of 734,054 cells associated with 12 cell types (Fig. 2C) from kidney (13 studies), prostate (7) and bladder (2) cancers (Fig. 2B). The cell distribution by cancer type shows that the tumor microenvironment subpopulations overlap very well from one cancer type to another. By contrast, even though sharing an epithelial origin, we observed substantial differences in malignant subpopulations across cancer types (Fig. 2C). In our results so far, although the degree of heterogeneity correlates with the number of samples in each study, kidney cancers seem to be associated with the highest heterogeneity, ahead of prostate and bladder cancers (Fig. S2).

Overall, UncoVer is an easy-to-use online resource for exploration and comparison of single-cell and spatial profiling studies that may support fundamental and translational research in uro-oncology. UncoVer has the advantages of being manually curated by experts in genitourinary malignancies. Each dataset has been analyzed using standard procedures and complemented with metadata when provided by the authors. Importantly, UncoVer also gives access to a massive meta-analysis which is the first attempt to centralize and integrate all single-cell datasets accumulated in the field. UncoVer will be

complemented with upcoming studies including spatial profiling datasets, and its range will be extended to other technologies (such as single-cell ATAC-sequencing), model organisms, and types of experimental design (such as cultured explants exposed to treatments). In the future, UncoVer also intends to serve as a warehouse for regression-based or deconvolution-based predictive models which may help users to compare their own datasets with those already hosted by the system.

## **Author contributions**

FC supervised the research and designed the study. JS, GL and FC wrote the manuscript. All authors contributed to the manuscript. GL, PR and FC processed and analyzed the data. SL contributed to the analysis. JS, NRL, AL and FC assisted in the manual curation. TAD, LN, PR, MB, OC and OS contributed to the development and deployment of the website. All authors approved the final version of the manuscript.

## **Acknowledgements**

The Uro-oncogenomics Viewer resource is supported, built and maintained by the Research Institute for Environmental and Occupational Health (IRSET), the French School of Public Health (EHESP) and the GenOuest BioInformatics core facility. We thank the Institut national de la santé et de la recherche médicale (Inserm) and the Université de Rennes 1 for supporting this work.

## **Funding**

This study was funded by La ligue contre le cancer committees 22, 35 and 81 ; the clinical and translational research committee (CORECT) of the Rennes University Hospital Center ; the French National Institute of Health and Medical Research (Inserm); the University of Rennes 1; and the French

School of Public Health (EHESP). S.L. was supported by the chair “Cancer & Innovation” of Rennes 1 Foundation (<https://fondation.univ-rennes1.fr/>) and a specific grant from the LabEx IGO program (n° ANR-11-LABX-0016) funded by the «Investment into the Future» French Government program, managed by the National Research Agency (ANR).

## Conflict of interest

The authors have no conflict of interest to declare.

## Bibliography

- [1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021;71:209–49.
- [2] Gerlinger M, Catto JW, Orntoft TF, Real FX, Zwarthoff EC, Swanton C. Intratumour heterogeneity in urologic cancers: from molecular evidence to clinical implications. *Eur Urol* 2015;67:729–37.
- [3] Suvà ML, Tirosh I. Single-Cell RNA Sequencing in Cancer: Lessons Learned and Emerging Challenges. *Mol Cell* 2019;75:7–12.
- [4] Gonzalez Castro LN, Tirosh I, Suvà ML. Decoding Cancer Biology One Cell at a Time. *Cancer Discov* 2021;11:960–70.
- [5] Klein J, Jupp S, Moulos P, Fernandez M, Buffin-Meyer B, Casemayou A, et al. The KUPKB: a novel Web application to access multiomics data on kidney disease. *FASEB J* 2012;26:2145–53.
- [6] Marzec J, Ross-Adams H, Pirrò S, Wang J, Zhu Y, Mao X, et al. The Transcriptomic Landscape of Prostate Cancer Development and Progression: An Integrative Analysis. *Cancers* 2021;13. <https://doi.org/10.3390/cancers13020345>.
- [7] Zhang Q, Yang B, Chen X, Xu J, Mei C, Mao Z. Renal Gene Expression Database (RGED): a relational database of gene expression profiles in kidney disease. *Database* 2014;2014. <https://doi.org/10.1093/database/bau092>.
- [8] Li Q, Zhang X, Ke R. Spatial Transcriptomics for Tumor Heterogeneity Analysis. *Front Genet* 2022;13:906158.

# Figure Legends

**Figure 1. Workflow of the single-cell data analysis and list of featured studies.** (A) Raw data and metadata were downloaded from public repositories such as GEO and SRA. Raw data could be downloaded either as fastq files or count matrix files (.mtx). In the case of fastq raw data, files were first processed according to the standard CellRanger procedure, resulting in gene expression matrix, barcode and features complementary files (one count matrix file for each sample). Count matrices were merged and normalized by using the Seurat package before undergoing dimensionality reductions (PCA and UMAP). Cells were then clustered using the Louvain community detection method implemented in Seurat. Cell clusters were then associated with cell types based on the expression of well-known gene markers (Table S2) and manual curation by experts. For the meta-analysis, Seurat objects from 10X Genomics experiments were merged and batch-effect corrected by using the Harmony R package. Analytical and clinical metadata provided by the authors were integrated to the Seurat objects. Finally, the resulting Seurat objects were converted into loom files, which were designed to efficiently hold large omics datasets, the input format of the UncoVer resource. (B) UncoVer web interface illustrating the search and filtration of available studies.

**Figure 2. UncoVer's exploration web-page.** The exploration web-page is divided into four 4 panels (from A to D). (A) Filtering component. Purple box indicates the drop-down menu field. (B) Independent customizable graphical components (ICGC) showing the cell distributions across samples in the form of pie charts or barplots. Purple box indicates the drop-down menu field. (C) ICGCs that allow the projection of cells in the form of UMAP or PCA. Purple box indicates the drop-down menu field allowing to color cells according to a given metadata such as the tumor or the cell type. (D) ICGCs that allow to display cells' gene expression levels in the form of scatter, hexagonal, density or violin plot. Purple box indicates the gene query and the display type selection fields.

## Supplementary figure legends

**Figure S1. Web architecture of UncoVer.** Web architecture of UncoVer. Data stored in the PostgreSQL database or in the loom files are both sent in the JSON format via a REST API built on top of the Django framework (version 3.0.11) and rendered via ReactJs components (version 17.0.2). Django is a python web framework used to build scalable and secure website applications while React is a front-end JavaScript framework used for building composable user interfaces. The open-source plotting libraries Chart.js (version 2.9.4) and Plotly.js (version 1.58.5) have been integrated to the react components so as to be able to visualize the data according to multiple chart types such as scatter plot, violin plot, dot plot). The Python Imaging Library (PIL) is also used to provide image processing capabilities to the Python interpreter and to enable the visualization of images of tissue sections in the case of spatially resolved transcriptome datasets. To handle website traffic and to provide data security, scalability, and to facilitate deployment, the different components of UncoVer have been containerized with Docker (version 17.12.0-ce) and a Nginx http and reverse proxy-server (version 1.19.1-alpine) is used to serve the application.

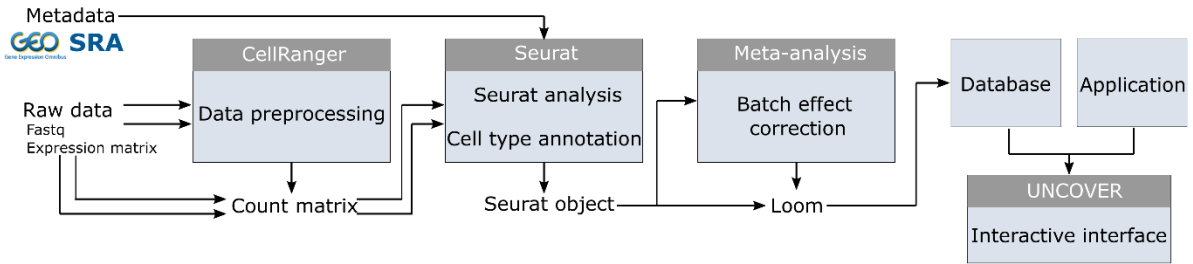
**Figure S2. Malignant and epithelial cells in the meta-analysis.** (A) UMAP plot representing malignant cells, either in primary tumor (“Tumor”), tumoral thrombus in the vena cava of ccRCC (“Thrombus”) or in metastasis (“Metastasis”); and epithelial cells (“Normal”). (B) Distribution of malignant and epithelial cells per organ of origin.



# Figures

Fig. 1

A



B

TITLE	TECHNOLOGY	EXPERIMENTAL DESIGN	BIOMATERIAL TYPE	BIOMATERIAL FROM	BIOMATERIAL ENTITY	TOPIC	NO DATASETS	ARTICLE
Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma	scRNA-Seq	Treatment	Tumor tissue, Metastatic tumor tissue	Kidney, lung, lymph node, Abdominal	Primary site, malignant cell, immune cell	Clear Cell Renal Cell Carcinoma, Papillary Renal Cell Carcinoma	7 datasets	Cancer cell. Bi et al. (2021 May 10) PMID:33711272
Meta-analysis	scRNA-Seq	Investigation	Normal tissue, Tumor tissue, Metastatic tumor tissue	Kidney, urinary bladder, prostate	Primary site, Thrombus	Clear Cell Renal Cell Carcinoma, Chromophobe Renal Cell Carcinoma, Papillary Renal Cell Carcinoma, Adenocarcinoma, Bladder Urothelial Carcinoma, Muscle Invasive Bladder Urothelial Carcinoma, Transitional Cell Carcinoma, Prostatic intraepithelial neoplasia, Small Cell Neuroendocrine Carcinoma, Wilms Tumor	1 dataset	
Identification of a distinct luminal subgroup diagnosing and identifying early stage prostate cancer by tissue-based single-cell RNA sequencing	scRNA-Seq	Investigation	Tumor tissue	prostate	Primary site	Adenocarcinoma	1 dataset	Molecular cancer. Ma et al. (2020 Oct 8) PMID:33032011
Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors	scRNA-Seq	Investigation	Normal tissue, Tumor tissue	Kidney, uretra	Primary site, malignant cell, immune cell	Clear Cell Renal Cell Carcinoma, Papillary Renal Cell Carcinoma, Wilms Tumor	4 datasets	Science (New York, N.Y.). Young et al. (2018 Aug 16) PMID:30093597
Inhibiting WNT and NOTCH in renal cancer stem cells and the implications for human patients	scRNA-Seq	Investigation	Tumor tissue	Kidney		Clear Cell Renal Cell Carcinoma	1 dataset	Nature communications. Fardipour et al. (2020 Feb 17) PMID:32046795
Single-cell sequencing links multiregional immune landscapes and tissue-resident T cells in context to tumor topology and therapy efficacy	scRNA-Seq	Exposure	Normal tissue, Tumor tissue, Metastatic tumor tissue	Kidney, lymph node, PROSTATE	Primary site, immune cell, malignant cell	Clear Cell Renal Cell Carcinoma	5 datasets	Cancer cell. Krishna et al. (2021 May 05) PMID:33861994
TODA	RNA-seq	Investigation	Normal tissue, Tumor tissue	Kidney, urinary bladder, prostate		Clear Cell Renal Cell Carcinoma, Papillary Renal Cell Carcinoma, Adenocarcinoma, Chromophobe Renal Cell Carcinoma, Metastatic Castration-Resistant Prostate Cancer, Bladder Urothelial Carcinoma	1 dataset	

**FILTERS** **RESET**

Title  
All

Exposure  
Investigation  
Treatment

Biomaterial from  
All

Fig. 2

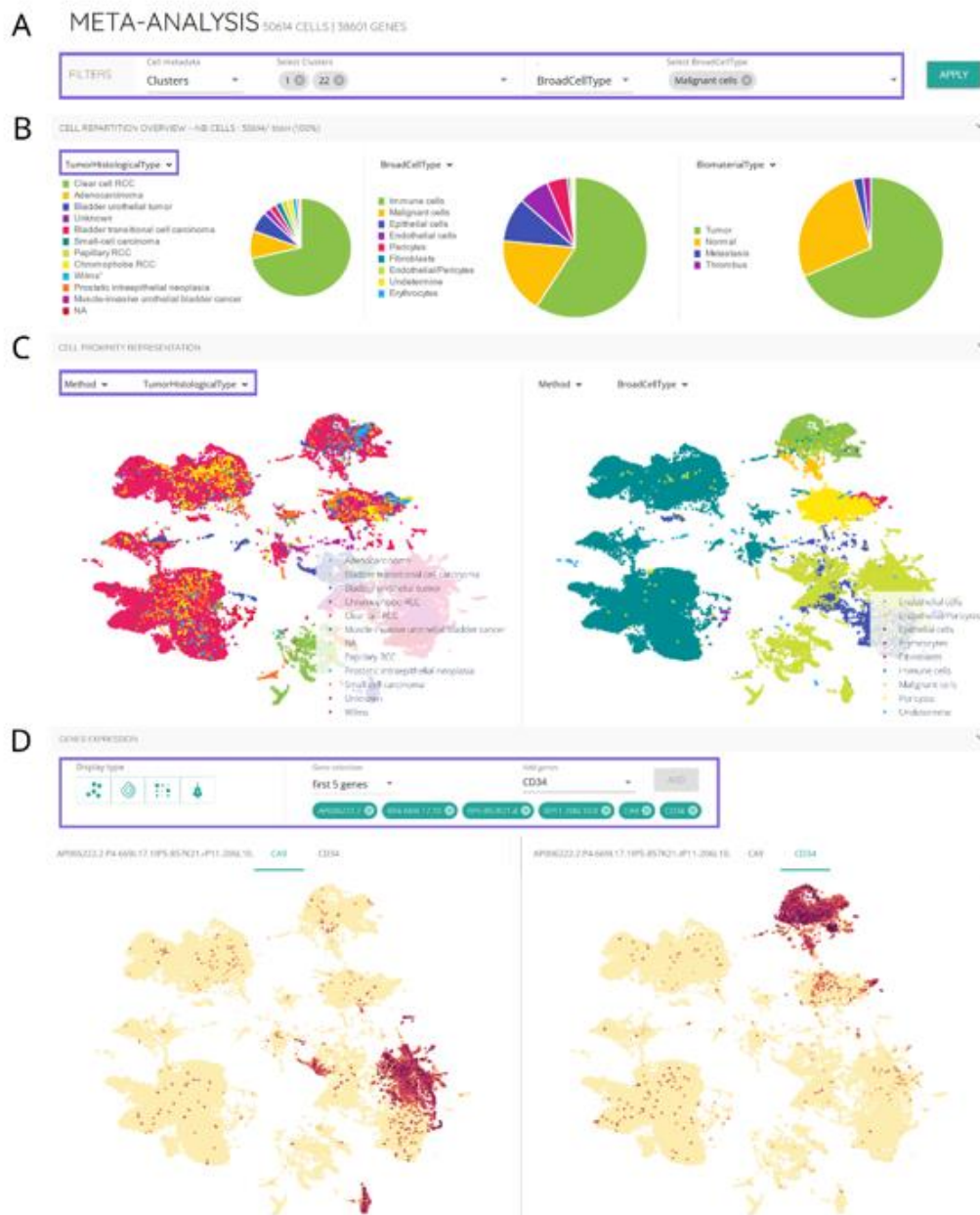


Fig. S1

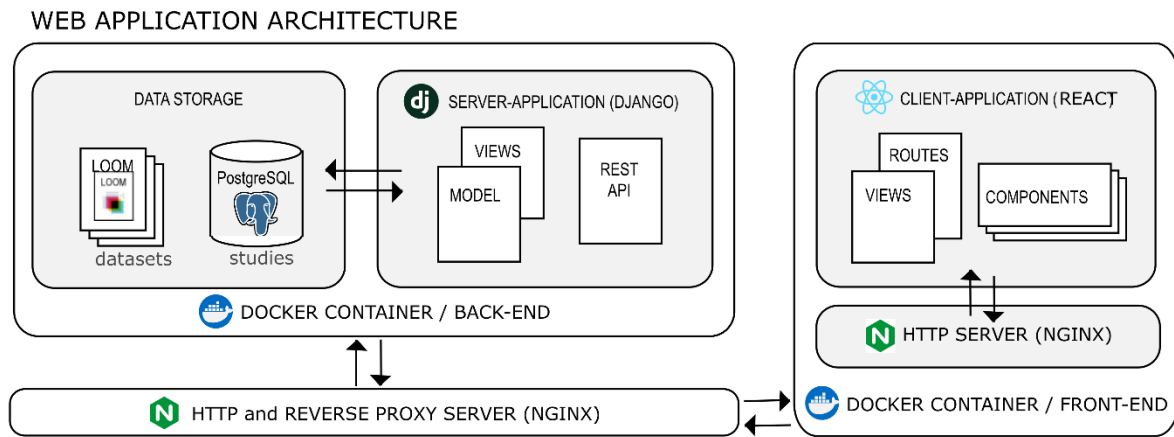
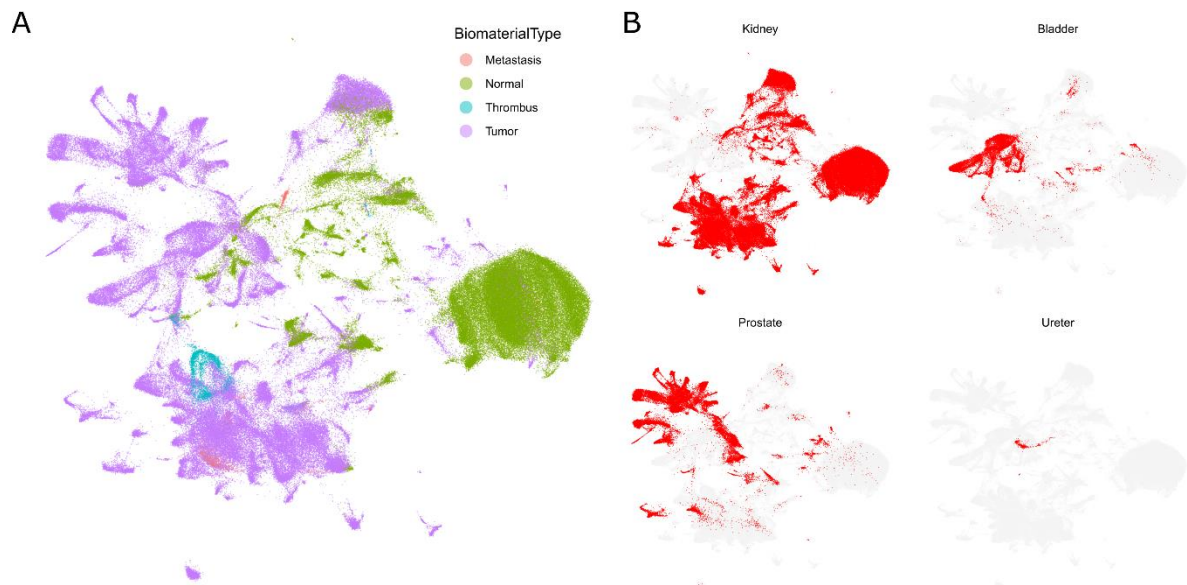


Fig. S2



# Tables

Resource	URL	Pubmed ID
Single cell portal	<a href="https://singlecell.broadinstitute.org">https://singlecell.broadinstitute.org</a>	NA
SCPortalen	<a href="https://single-cell.riken.jp">https://single-cell.riken.jp</a>	PMID: 29045713
Expression Atlas	<a href="https://www.ebi.ac.uk/gxa/home">https://www.ebi.ac.uk/gxa/home</a>	PMID: 31665515
PanglaoDB	<a href="https://panglaodb.se/">https://panglaodb.se/</a>	PMID: 30951143
scRNASeqDB	<a href="https://bioinfo.uth.edu/scrnaseqdb/">https://bioinfo.uth.edu/scrnaseqdb/</a>	PMID: 29206167
CancerSCEM	<a href="https://ngdc.cncb.ac.cn/cancerscem/">https://ngdc.cncb.ac.cn/cancerscem/</a>	PMID: 34643725
CancerSEA	<a href="http://biocc.hrbmu.edu.cn/CancerSEA/home.jsp">http://biocc.hrbmu.edu.cn/CancerSEA/home.jsp</a>	PMID: 30329142
expO	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=gse2109">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=gse2109</a>	NA
TCGA	<a href="https://www.genome.gov/Funded-Programs-Projects/Cancer-Genome-Atlas">https://www.genome.gov/Funded-Programs-Projects/Cancer-Genome-Atlas</a>	NA
TISCH	<a href="http://tisch.comp-genomics.org/">http://tisch.comp-genomics.org/</a>	PMID: 33179754
Descartes	<a href="https://descartes.brotmanbaty.org/">https://descartes.brotmanbaty.org/</a>	PMID: 33184181
RGV	<a href="https://rgv.genouest.org">https://rgv.genouest.org</a>	PMID: 30668675
Nephroseq	<a href="https://www.nephroseq.org">https://www.nephroseq.org</a>	NA
KUPKB	<a href="https://kupkb.org">https://kupkb.org</a>	PMID: 22345404
RGED	<a href="http://rged.wall-eva.net/">http://rged.wall-eva.net/</a>	PMID: 25252782
PIXdb	<a href="https://pixdb.org.uk">https://pixdb.org.uk</a>	PMID: 33477882

Table S1. Single-cell and transcriptomic resources

<b>Cell Type</b>	<b>Gene markers</b>
<b>Epithelial cells</b>	EPCAM, KRT18, ANPEP, CDH2, MME, PAX2, SAA2, SLC16A9, SLC17A3
<b>Tubular epithelial cells</b>	KRT19, ATP6V0D2, CLCNKB
<b>Collecting duct epithelial cells</b>	UMOD
<b>Endothelial cells</b>	TEK, CD34, CDH5, IL3RA, KDR, PLVAP, PTPRB, TSPAN7, ERG
<b>Malignant cells</b>	CA9, NDUFA4L2, CP, ELF3, KLK4, UCA1, CLCA4, PROM1, SLC31A2, ERBB2, HNF1B, PAX8, ABCB1, ALDH1A1, EGFR, FOXM1, HMOX1, MET, SLC16A1, VEGFA
<b>Erythrocytes</b>	AHSP, ALAS2, CA1
<b>Immune cells</b>	LY9, SPN, PTPRC, IGLC3, CD48, CD52, CD53, LAIR1, IL7R, CD4, CD8A, CD7
<b>B cells</b>	CD79A, CD79B, MS4A1
<b>T cells</b>	CD3E, CD3D, CD3G
<b>NK cells</b>	GNLY, NKG7, NCAM1, FCGR3A
<b>Mast cells</b>	TPSAB1, TPSB2, KIT
<b>Myeloid cells</b>	CD14, CD68, AIF1, CSF1R, FUT4
<b>Pericytes</b>	ACTA2, CSPG4, PDGFA, PDGFRB, RGS5, DES, ACAN, ANGPT1, ANGPT2
<b>Fibroblasts</b>	PDGFRA, FAP, LUM, DCN, FBLN1, PRELP, CCDC80, COL1A1

Table S2: Table of genes markers used to identified cell type

<b>Publication</b>	<b>PMID</b>	<b>Tumor organ</b>	<b># datasets</b>	<b>Meta-analysis</b>
<b>Kim et al., 2016</b>	27139883	Kidney	1	No
<b>Young et al., 2018</b>	30093597	Kidney	4	Yes
<b>Vickman et al., 2019</b>	30824526	Prostate	1	No
<b>Vickman et al., 2020</b>	31763714	Prostate	2	Yes
<b>Fendler et al., 2020</b>	32066735	Kidney	1	No
<b>Su et al., 2020</b>	32119131	Prostate	1	No
<b>Lee et al., 2020</b>	32460812	Bladder	4	Yes
<b>Oh et al., 2020</b>	32497499	Bladder	1	Yes
<b>Schnepp et al., 2020</b>	32513898	Prostate	1	No
<b>Ma et al., 2020</b>	33032611	Prostate	4	Yes
<b>Chen et al., 2020</b>	33033240	Bladder	4	Yes
<b>Arguello et al., 2020</b>	33264598	Kidney	4	Yes
<b>Dong et al., 2020</b>	33328604	Prostate	4	Yes
<b>Chen et al., 2021</b>	33420488	Prostate	4	Yes
<b>Borcheding et al., 2021</b>	33504936	Kidney	5	Yes
<b>Bi et al., 2021</b>	33711272	Kidney	7	Yes
<b>Braun et al., 2021</b>	33711273	Kidney	1	Yes
<b>Krishna et al., 2021</b>	33861994	Kidney	5	Yes
<b>Zhang et al., 2021</b>	34099557	Kidney	4	Yes
<b>Yu et al., 2021</b>	34168986	Kidney	4	Yes
<b>Su et al., 2021</b>	34722263	Kidney	4	Yes
<b>Masetti et al., 2022</b>	34919143	Prostate	1	Yes
<b>Zhang et al., 2022</b>	35605062	Kidney	4	Yes
<b>Saout et al, submitted</b>	NA	Kidney	6	Yes
<b>Huang et al., 2021</b>	34813980	Kidney	NA	No
<b>Kourtis et al. , 2022</b>	35668194	Kidney	NA	No
<b>Song et al., 2022</b>	35013146	Prostate	NA	No
<b>M. Alchahin et al., 2022</b>	<a href="https://doi.org/10.1038/s41467-022-33375-w">https://doi.org/10.1038/s41467-022-33375-w</a>	Kidney	NA	No
<b>Wang et al., 2020</b>	32111252	Bladder	NA	No
<b>Yuen et al., 2020</b>	32405063	Bladder, Kidney	NA	No
<b>T. Neal et al., 2018</b>	30550791	Kidney	NA	No
<b>Kfoury et al., 2021</b>	34719426	Prostate	NA	No

Table S3. Studies datasets features in UncoVer

**Supplementary materials:**

**UncoVer: A web-based resource for single-cell and  
spatial transcriptome exploration and comparison  
for the urologic oncology community**

Gwendoline Lecuyer<sup>1,✉</sup>, Judikael Saout<sup>1,✉</sup>, Thomas A. Darde<sup>2</sup>, Laurence Noël<sup>1</sup>, Paul Rivaud<sup>1</sup>, Matéo Boudet<sup>3,4</sup>, Simon Léonard<sup>1,5,6</sup>, Olivier Sallou<sup>3</sup>, Nathalie Rioux-Leclercq<sup>1,7</sup>, Olivier Collin<sup>3</sup>, Aurélie Lardenois<sup>1</sup>, Frédéric Chalmel<sup>1,\*</sup>

<sup>1</sup> Univ Rennes, Inserm, EHESP, Irset (Institut de Recherche en Santé, Environnement et Travail), UMR\_S 1085, F-35000 Rennes, France.

<sup>2</sup> SciLicism, 10 rue de la Sauvaie, 35200 Rennes, France.

<sup>3</sup> Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA/INRIA) - GenOuest platform, Université de Rennes 1, F-35042 Rennes, France.

<sup>4</sup> IGEPP, Agrocampus Ouest, INRAE, Université de Rennes 1, Le Rheu, France.

<sup>5</sup> LabEx IGO “Immunotherapy, Graft, Oncology”, F-35043 Nantes, France.

<sup>6</sup> Univ Rennes, INSERM, EFS, UMR S1236, F-35000 Rennes, France.

<sup>7</sup> Department of Pathology, University of Rennes 1, Rennes, France.

\* To whom correspondence should be addressed

✉ These authors contributed to the manuscript equally

## **Supplementary methods**

### **Data retrieval within the Django/react application**

Data stored in the PostgreSQL database or in the loom files were both sent in the JSON format via a REST API built on top of the Django framework (version 3.0.11) and rendered via ReactJs components (version 17.0.2) (Figure S1). Briefly, Django is a python web framework used to build scalable and secure website applications while ReactJs is a front-end JavaScript framework used for building composable user interfaces. The open-source plotting libraries Chart.js (version 2.9.4) and Plotly.js (version 1.58.5) have been integrated to the ReactJs components so as to be able to visualize the data according to multiple chart types such as scatter plot, violin plot, dot plot). The Python Imaging Library (PIL) was also used to provide image processing capabilities to the Python interpreter and to enable the visualization of images of tissue sections in the case of spatially resolved transcriptome datasets. To handle website traffic and to provide data security, scalability, and to facilitate deployment, the different components of UncoVer have been containerized with Docker (version 17.12.0-ce) and a Nginx http and reverse proxy-server (version 1.19.1-alpine) is used to serve the application.

### **Data storage**

All datasets were stored in the Loom file format. The Loom format is based on HDF5, a standard for storing large numerical datasets, and adapted by the Linnarson Lab to store single cell genomics data and metadata. Briefly, each loom file contains the main gene expression matrix, as well as row and column attributes. Row attributes enable the characterization of genes, whereas column attributes are used to characterize the cells (such as cell cluster, cell type, cell cycle phase). To be accessible in the UncoVer application, a dataset has to be associated with a study, which is described via another set of metadata stored in a PostgreSQL database. Each study was manually curated and annotated so as to



enable end-users to search and filter the resources available via a common set of parameters (such as cancer type, publication, technology).

## **Dataset description**

The vast majority of studies have been downloaded from the NCBI Gene Expression Omnibus Repository [1] (Table S3). The current version of UncoVer hosts datasets generated with either the 10X Genomics, fluidigm C1 or CEL-seq2 platforms. When provided by the authors, metadata such as analytical (cell clusters, cell types and low-dimensional space coordinates) and clinical information were included into the corresponding study/dataset.

## **Data processing and analysis**

**Single-cell RNA-seq data.** UncoVer can take as input either raw fastq files or a count matrix corresponding to a given dataset. In the case of fastq, the sequencing reads were mapped to the GRCh38 human reference genome and quantified on the basis of unique molecular identifiers (UMIs), in order to generate a count matrix. Count matrices were then analyzed with the Seurat R package (v4) [2]. First, cells expressing <500 genes and genes expressed in <10 cells were removed. Subsequently, the scater R package (v1.10.1) was used to remove outlier cells based on their proportions of reads mapping to mitochondrial and ribosomal genes, and median number of genes and UMIs [3]. Doublets were removed with DoubletFinder R package [4]. Data were normalized by using the SCTransform function implemented into the Seurat suite (v4) by regressing out the unwanted variation due to cell cycle (S score and G2M score) and to the proportion of mitochondrial genes. The top-3000 most varying genes were used to perform a principal component analysis (PCA) with the RunPCA function implemented in Seurat. Cells were then clustered by using the Louvain community detection method (FindNeighbors and FindClusters functions) on the top-30 principal components, with default parameters. Finally, we

used the Uniform Manifold Approximation and Projection (UMAP) method (RunUMAP function) to project cells in a 2D space. Cell clusters were annotated using a set of known marker genes (Table S2).

**Meta-analysis of single-cell data.** UncoVer includes an original study corresponding to a massive meta-analysis of 19 of its studies that use the chromium 10X technology. Datasets were merged and batch-effect corrected with the Harmony package [5]. The resulting dataset was next analyzed in Seurat using the same procedure as described in the **Single-cell RNA-seq data** section.

**Bulk RNA-seq data from the TCGA database.** TCGA expression and clinical data from bulk-RNA-seq of kidney, prostate and bladder cancer were downloaded with the TCGABiolinks v3.15 R package [6]. Two outlier patients (TCGA.CJ.4642.01B.01R.1305.07 and TARGET.52.PASXGF.01A.01R) were removed based on their expression distribution. The resulting count matrix was then normalized with the VST method implemented into the DESeq2 R package [7]. Count and normalized data were next integrated into a Seurat object. After scaling, the top-3000 most varying genes were used to perform a principal component analysis (PCA) followed by a projection on a 2D space (UMAP). Samples were clustered by using the Louvain community detection method (FindNeighbors and FindClusters functions) on the top-50 principal components, with default parameters.

**Bulk microarray data from the expO database.** CEL files were downloaded from the GEO repository (GSE2109). The data quality was verified by plotting the surface intensity distribution, 3'-5' RNA degradation and log<sub>2</sub> signal distribution across samples. Data were normalized using the Robust Multi-Array Average (RMA) method implemented into the affy R package [8]. Normalized intensity signal data were next integrated into a Seurat object. After scaling, the top-3000 most varying genes were used to perform a principal component analysis (PCA) followed by a projection on a 2D space (UMAP). Samples were clustered by using the Louvain community detection method (FindNeighbors and FindClusters functions) on the top-50 principal components, with default parameters.

**Spatial transcriptomic data.** The count matrix file and the image of the corresponding tissue section were downloaded from the 10X Genomics website. The analysis was performed in Seurat. Briefly count data were normalized using the SCTransform function. The top-3000 most varying genes were used to perform a principal component analysis (PCA). Samples were clustered by using the Louvain community detection method (FindNeighbors and FindClusters functions) on the top-50 principal components, with default parameters.

## **Study and dataset completion with metadata**

Following data processing, each dataset was associated with a single Seurat object which was complemented with metadata. Prior to their integration, metadata were first manually reported and homogenized into an Excel sheet based on controlled vocabularies. In case the authors provided additional information about their published analysis (such as cell clusters, cell types or cell coordinates in 2D spaces (t-SNE or UMAP), a dedicated dataset with specific metadata was created and titled with the prefix 'Article'.

## **Converting Seurat objects to loom files**

All datasets were stored in the Loom file format. To address this issue Seurat objects were converted into loom files by using the SeuratDisk package (v0.0.0.9019) (<https://github.com/mojaveazure/seurat-disk>).

## **Supplementary results**

### **Description of scRNA-seq studies**

The scRNA-seq studies embedded in UncoVer cover a wide range of topics, including (*i*) the investigation of the intra-tumoral heterogeneity and micro-environment [9–13], with important focuses

of the immune micro-environment [14–17]; (ii) the therapy efficacy in ccRCC [14,16,18,19], in bladder [20] and in prostate [21] cancers; and, (iii) the identification of new specific cell markers [22]. Other studies have specific focuses such as the characterization of sporadic bilateral ccRCC [23], cancer associated fibroblasts (CAFs) [13] or tumor associated macrophages (TAMs) [24].

## Supplementary references

- [1] Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets-update. *Nucleic Acids Res* 2013;41:D991–5.
- [2] Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;36:411–20.
- [3] McCarthy DJ, Campbell KR, Lun ATL, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* 2017:btw777. <https://doi.org/10.1093/bioinformatics/btw777>.
- [4] McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* 2019;8:329–37.e4.
- [5] Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* 2019;16:1289–96.
- [6] Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* 2016;44:e71.
- [7] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
- [8] Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003;31:e15.
- [9] Zhang Y, Narayanan SP, Mannan R, Raskind G, Wang X, Vats P, et al. Single-cell analyses of renal cell cancers reveal insights into tumor microenvironment, cell of origin, and therapy response. *Proc Natl Acad Sci U S A* 2021;118. <https://doi.org/10.1073/pnas.2103240118>.
- [10] Zhang M, Zhai W, Miao J, Cheng X, Luo W, Song W, et al. Single cell analysis reveals intra-tumour heterogeneity, microenvironment and potential diagnosis markers for clear cell renal cell carcinoma. *Clin Transl Med* 2022;12:e713.
- [11] Young MD, Mitchell TJ, Vieira Braga FA, Tran MGB, Stewart BJ, Ferdinand JR, et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* 2018;361:594–9.
- [12] Chen Z, Zhou L, Liu L, Hou Y, Xiong M, Yang Y, et al. Single-cell RNA sequencing highlights the role of inflammatory cancer-associated fibroblasts in bladder urothelial carcinoma. *Nat Commun* 2020;11:5077.
- [13] Vickman RE, Broman MM, Lanman NA, Franco OE, Sudyanti PAG, Ni Y, et al. Heterogeneity of human prostate carcinoma-associated fibroblasts implicates a role for subpopulations in myeloid cell recruitment. *Prostate* 2020;80:173–85.
- [14] Krishna C, DiNatale RG, Kuo F, Srivastava RM, Vuong L, Chowell D, et al. Single-cell sequencing links multiregional immune landscapes and tissue-resident T cells in ccRCC to tumor topology and therapy efficacy. *Cancer Cell* 2021;39:662–77.e6.
- [15] Braun DA, Street K, Burke KP, Cookmeyer DL, Denize T, Pedersen CB, et al. Progressive immune dysfunction with advancing disease stage in renal cell carcinoma. *Cancer Cell* 2021;39:632–48.e8.
- [16] Bi K, He MX, Bakouny Z, Kanodia A, Napolitano S, Wu J, et al. Tumor and immune

- reprogramming during immunotherapy in advanced renal cell carcinoma. *Cancer Cell* 2021;39:649–61.e5.
- [17] Oh DY, Kwek SS, Raju SS, Li T, McCarthy E, Chow E, et al. Intratumoral CD4 T Cells Mediate Anti-tumor Cytotoxicity in Human Bladder Cancer. *Cell* 2020;181:1612–25.e13.
- [18] Borcherding N, Vishwakarma A, Voigt AP, Bellizzi A, Kaplan J, Nepple K, et al. Mapping the immune environment in clear cell renal carcinoma by single-cell genomics. *Commun Biol* 2021;4:122.
- [19] Kim K-T, Lee HW, Lee H-O, Song HJ, Jeong DE, Shin S, et al. Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma. *Genome Biol* 2016;17:80.
- [20] Lee HW, Chung W, Lee H-O, Jeong DE, Jo A, Lim JE, et al. Single-cell RNA sequencing reveals the tumor microenvironment and facilitates strategic choices to circumvent treatment failure in a chemorefractory bladder cancer patient. *Genome Med* 2020;12:47.
- [21] Schnepf PM, Shelley G, Dai J, Wakim N, Jiang H, Mizokami A, et al. Single-Cell Transcriptomics Analysis Identifies Nuclear Protein 1 as a Regulator of Docetaxel Resistance in Prostate Cancer Cells. *Mol Cancer Res* 2020;18:1290–301.
- [22] Su C, Lv Y, Lu W, Yu Z, Ye Y, Guo B, et al. Single-Cell RNA Sequencing in Multiple Pathologic Types of Renal Cell Carcinoma Revealed Novel Potential Tumor-Specific Markers. *Front Oncol* 2021;11:719564.
- [23] Yu Z, Lu W, Su C, Lv Y, Ye Y, Guo B, et al. Single-Cell RNA-seq Identification of the Cellular Molecular Characteristics of Sporadic Bilateral Clear Cell Renal Cell Carcinoma. *Front Oncol* 2021;11:659251.
- [24] Masetti M, Carriero R, Portale F, Marelli G, Morina N, Pandini M, et al. Lipid-loaded tumor-associated macrophages sustain tumor growth and invasiveness in prostate cancer. *J Exp Med* 2022;219. <https://doi.org/10.1084/jem.20210564>.

### 4.3.3. DISCUSSION ET PERSPECTIVES

UncoVer est la première base de données dédiée spécifiquement aux données de scRNA-seq dans les cancers urologiques. Le principal atout de cet outil selon moi est son interactivité, qui permet de parfaitement répondre à de nombreuses questions que nous avons eues de la part de collaborateurs, concernant par exemple l'expression de gènes d'intérêt dans certaines populations ou en fonction de données cliniques. Nous avons identifié ce besoin assez tôt dans le cadre de nos projets, ce qui m'a permis de prendre part au recensement et au formatage des données de scRNA-seq pour UncoVer. UncoVer a pour but de permettre de nouvelles approches d'analyse des données existantes. Premièrement, il inclut une méta-analyse intégrant les données de scRNA-seq de 19 études. Les principaux cancers urologiques sont des carcinomes, et ont en commun leur origine épithéliale. Leur analyse conjointe à l'échelle de la cellule unique permet donc de comparer de façon inédite les grandes populations cellulaires qu'ils ont en commun. Il est intéressant par exemple de chercher les programmes d'expression partagés ou spécifiques des cellules malignes à travers les différents cancers. Toutefois, comme c'est souvent le cas dans les études scRNA-seq, notre analyse montre une hétérogénéité inter-cancer importante dans les cellules malignes, malgré les expressions partagées de marqueurs épithéliaux tels qu'EPCAM, KRT18 ou les gènes de la famille SLC. Par ailleurs, le microenvironnement tumoral est bien superposé à travers les différents cancers, de façon plus attendue, ce qui suggère que la différence observée entre cellules malignes relève quant à elle davantage d'une réalité biologique que d'éventuels biais techniques. En l'état, il est difficile de déterminer si cette hétérogénéité tient à des processus de carcinogenèse fondamentalement différents, ou si cela tient plutôt au fond transcriptomique initial des cellules épithéliales à partir desquelles ont évolué les cellules malignes, spécialisées dans les fonctions de leur organe. Pour répondre à cette question, il faudrait approfondir la sous-analyse des cellules malignes que nous avons réalisées. En projetant les cellules malignes dans un nouvel espace en deux dimensions retranscrivant mieux leurs différences, elle permet en première approche d'identifier celles qui sont proches transcriptionnellement, bien qu'appartenant à différents cancers. Il sera peut-être nécessaire d'identifier celles qui sont particulièrement différenciées afin de faire ressortir les programmes communs à la carcinogenèse. Identifier de tels programmes pourrait permettre, du point de vue fondamental, de mieux comprendre le Cancer et, du point de vue translationnel, de concevoir de nouvelles molécules par exemple multi-cibles. La problématique des programmes d'expression communs entre cancers et a

fortiori, entres carcinomes, est transversale (Suvà and Tirosh, 2019). Même si les organes du tractus uro-génital ont une même origine développementale dans le mésoderme intermédiaire (Libretti and Aeddula, 2020), les programmes d'expression communs des cancers sont pour l'essentiel indépendants de cette origine et ont plus trait aux fameuses *hallmarks* du cancer (Hanahan and Weinberg, 2011). Deux programmes d'expression récurrents ont été identifiés dans une revue de la littérature sur les leçons du scNRA-seq dans différents cancers ; ceux liés au cycle cellulaire et ceux liés à la réponse au stress (Suvà and Tirosh, 2019). Ainsi, bien que la problématique soit plus générale, il reste intéressant de chercher ces programmes d'expression dans les données de notre méta-analyse, et de voir si d'autres programmes ne ressortent pas, qu'ils soient communs aux carcinomes, restreints aux carcinomes de la sphère urologiques, ou spécifiques d'un type de cancer.

UncoVer comprend un premier jeu de transcriptomique spatiale, une technologie qui est actuellement en train de se développer considérablement en oncologie (Wu *et al.*, 2022). Je pense que cette résolution d'analyse est très prometteuse et suscitera un vif intérêt des cliniciens et experts en anatomopathologie, étant donné sa dimension histologique. Plus encore, il est probable que les cliniciens prennent la main sur ce genre d'analyse étant donné leur expertise, ce qui justifie plus avant la nécessité de partage de données opéré par UncoVer. De façon générale, la transcriptomique spatiale.

En plus d'une réanalyse systématique des différents jeux de données par un « *workflow* » standardisé, UncoVer offre également la possibilité de retrouver des analyses fidèles à la publication d'origine, ce qui en fait une force. Cependant, ce processus est dépendant des auteurs et tous n'ont pas répondu à la requête de métadonnées cliniques ou d'analyses. C'est un des défauts de l'approche. UncoVer devra gagner en notoriété, notamment auprès de la communauté de l'uro-oncologie, pour que ce problème ne devienne pas trop systématique.

Enfin, il faudra à l'avenir palier à un problème actuel d'UncoVer : l'affichage des cellules. Il est pour le moment impossible d'afficher des millions de points sur un navigateur classique en raison des ressources mémoires nécessaires. Pour l'instant, nous avons recours à un échantillonnage aléatoire de 100 000 à 150 000 cellules sur les jeux de données trop conséquents, qui permet de conserver les proportions de chaque étude et chaque type cellulaire. Peut-être qu'il faudra passer par d'autres solutions d'affichage. UncoVer fait encore l'objet de développements. Il sera actualisé à mesure que des études paraîtront et inclura plus tard des études liées au cancer du testicule (encore aucune étude à ce jour pour ce cancer urologique) ou aux perturbations expérimentales de tumeurs, par exemple par des nouvelles thérapies.

# 5.DISCUSSION





Dans ce travail de thèse, j'ai utilisé la technologie de séquençage ARN sur cellules uniques (« single-cell RNA-sequencing », ou scRNA-seq) pour étudier l'hétérogénéité intratumorale (HIT) dans le carcinome rénal à cellules claires (ccRCC). J'ai en particulier cherché à comprendre comment cette information pouvait être utilisée pour améliorer la prise en charge des patients, notamment des points de vue pronostiques et thérapeutiques. Pour cela, j'ai contribué à mettre en place deux approches complémentaires ; l'une descriptive qui visait à établir un atlas du ccRCC à l'échelle de la cellule unique ; et l'autre, appliquée, qui visait à étudier l'action de thérapies ciblées sur cette HIT. Enfin, pour mettre en valeur ces travaux, j'ai participé au développement d'une base de données interactive répertoriant et permettant l'analyse de données de transcriptomique dans les cancers urologiques.

Ces travaux participent d'un effort collectif plus vaste de caractérisation de l'hétérogénéité des tissus humains, incarné notamment par le consortium international du « *Human Cell Atlas* » (HCA) (Regev *et al.*, 2017). Depuis sa création en 2016, le HCA réunit plus de 1000 instituts de recherches dans le monde et a pour but de cartographier chaque type cellulaire dans le corps humain, des stades développementaux jusqu'au stade adulte, afin de mieux appréhender diverses maladies, dont le cancer. Ce projet, que leurs acteurs comparent au premier séquençage complet du génome humain (« *The human genome project* »), a émergé avec l'avènement de la technologie du scRNA-seq. Il recense actuellement plus de 130 études et 13,5 millions de cellules. Dans le même temps en oncologie, le scRNA-seq a naturellement trouvé une place très importante, à la suite du séquençage ARN de type *bulk*, pour permettre d'accéder à l'échelle de la cellule unique. En effet, la cellule génétiquement et épigénétiquement innovante est l'unité fondamentale opérationnelle du cancer (Baslan and Hicks, 2017) ; c'est en son sein qu'adviennent les événements moléculaires à l'origine du cancer. Qu'elle soit en prolifération ou quiescente, au sein de la tumeur primaire ou d'une métastase, c'est la cellule qui régit les paramètres dictant la biologie du cancer. Cette résolution d'analyse inédite ouvre donc un large éventail d'applications et de questions permettant d'approfondir notre compréhension fondamentale de cette maladie, et suscite actuellement un vif intérêt dans tous les laboratoires. Tel un préalable à ses perspectives, l'une des premières applications du scRNA-seq, à l'image du HCA, fut de cartographier les cellules dans chaque type de cancer (G. Sun *et al.*, 2021). Les premiers atlas « *single-cell* » ont été publiés dans des revues majeures telles que Nature ou Science, permettant à certains laboratoires de s'imposer comme nouveaux référents dans ce domaine, à l'image de l'équipe du docteur Itay Tirosh au Weizman Institute, qui a eu la primeur de publications d'atlas du glioblastome, le mélanome,

ou encore le cancer de la tête et du cou (Patel *et al.*, 2014; Tirosh *et al.*, 2016; Puram *et al.*, 2017). Aujourd'hui, la majorité des cancers a fait l'objet de telles publications. Cependant, ces atlas varient souvent dans la taille, la qualité ou l'origine technologique des données. Par ailleurs, au contraire d'un organe cartographié dans le HCA, les atlas de cancers sont moins figés, l'état des cellules cancéreuses étant par nature dynamique (Barkley *et al.*, 2021). Des atlas dits d'états cellulaires (les « *cancer single-cell state atlases* ») ont d'ailleurs suivis les premiers pour prendre en compte cette dimension (Yuan *et al.*, 2019). Étant donné la variabilité entre les atlas, il reste des zones d'ombre et des compartiments cellulaires moins explorés que d'autres. L'HIT par exemple a été décrite pour ce qui concerne la tumeur dans son ensemble, incluant le microenvironnement tumoral, mais moins en ce qui concerne spécifiquement les cellules malignes qui sont, par essence, hétérogènes.

Dans ce contexte, l'un des objectifs de ma thèse de doctorat était de bâtir un atlas de cellules uniques dans le ccRCC, avec une attention particulière portée aux cellules malignes. Après être parvenu à les enrichir, j'ai pu observer plusieurs degrés d'hétérogénéité. Premièrement, les cellules malignes dans chaque tumeur ont un programme d'expression propre, qui les distingue significativement d'une tumeur à l'autre. Les cellules du microenvironnement tumoral (« tumor microenvironment » ou TME) sont quant à elles davantage homogènes, à la fois entre et au sein des tumeurs. La littérature fait état des mêmes observations dans la plupart des cancers solides (Suvà and Tirosh, 2019). Cependant, alors que quelques programmes d'expression communs ont été identifiés tels que ceux du cycle cellulaire et de la réponse au stress, le constat d'une hétérogénéité inter-tumorale si fréquente dans les études reste à éclaircir. Il est probable qu'elle soit essentiellement le fruit d'histoires évolutives distinctes et relève d'une réalité biologique que l'analyse scRNA-seq permet de retranscrire. Cela suggère que la taille des cohortes doit être augmentée pour faire converger les atlas et approcher une image complète de la diversité cellulaire dans les ccRCCs. À cet égard, la méta-analyse réalisée et mise à disposition dans UncoVer réunit la majorité des données disponibles dans le cancer du rein actuellement. En première approche, elle permet de penser que les cellules malignes dans le ccRCC se superposent globalement. Cependant, le fait qu'elles soient analysées en même temps que les cellules des cancers de la prostate et de la vessie réduit leurs différences lors des étapes de réduction de dimensions. En effet, l'analyse en composantes principales utilisée pour réduire les dimensions retient une trentaine de composantes permettant de différencier les cellules, ce qui est trop faible pour distinguer convenablement le type de cancer. En conséquence, les cellules sont artificiellement rapprochées dans le UMAP. Pour

obtenir une image plus fiable de l'hétérogénéité dans le ccRCC, il faudrait donc analyser à part l'ensemble des cellules provenant de ce sous-type de carcinome. Ceci fait actuellement l'objet de la suite de notre méta-analyse. Toutefois, cette méta-analyse comme celle de l'ensemble des types de tumeurs peut être limitée en dernière instance par des biais techniques d'une étude à l'autre, liés aux manipulations d'échantillons ou à des plans expérimentaux très variables, qui impactent le profil transcriptomique des cellules et induisent des différences non-biologiques. La correction de ces effets par des outils bio-informatiques peut palier à ce problème, mais il n'est pas à exclure que ces outils appliquent une sur-correction qui, dans le sens inverse, aura tendance à homogénéiser excessivement les populations cellulaires. Cette balance illustre l'équilibre qu'il est nécessaire de trouver dans l'analyse (et surtout la méta-analyse) de données de single-cell RNA-seq. Bien que ces pistes soient à explorer, la description plus complète de l'hétérogénéité dans le ccRCC, pour dépasser ces limites, passera sans doute plutôt par la mise en place de cohortes de patients conséquentes, dont les échantillons seront traités de façon homogène.

Alors que l'acquisition des données de scRNA-seq dans l'axe 1 de cette thèse a coûté près de 21 000€ pour 7 échantillons, il est possible aujourd'hui de séquencer aux alentours de 16-20 échantillons pour le même prix, en utilisant par exemple le multiplexage couplé à la technique de single-nuclei RNA-seq (snRNA-seq). En effet, le scRNA-seq requiert du tissu frais et en bon état afin d'obtenir des cellules dissociées viables, le seuil minimum de viabilité recommandé par 10X Genomics étant de 70%. Cela m'a contraint aux prises en charge séquentielle des échantillons dans mes projets de thèse, chaque jour d'opération, et à l'utilisation d'un kit de réactifs pour chacun. Quant au snRNA-seq, il a l'avantage d'être adapté aux échantillons congelés car leur fragilité, souvent problématique pour obtenir des cellules viables en scRNA-seq, n'est pas une limite à l'extraction des noyaux (Slyper *et al.*, 2020). D'autre part, le *nuclei hashing* (Gaublomme *et al.*, 2019), variation du *cell hashing*, permet de multiplexer les noyaux provenant de différents échantillons dans une expérience de Chromium 10X et ainsi d'économiser des coûts importants de réactifs. Cette approche permet donc de collecter et congeler de multiples échantillons, de les traiter simultanément, puis de les multiplexer dans une même librairie de séquençage, ce qui réduirait considérablement les coûts et biais techniques. Elle permettrait selon moi d'établir des cohortes très conséquentes et de bâtir des atlas suffisamment importants et représentatifs. De façon plus générale, l'analyse d'échantillons cryoconservés connaît une forte progression, car il rend possible l'analyse d'énormes banques de tissus congelés dans les hôpitaux et plateformes de stockage. Le snRNA-seq a également d'autres avantages par rapport au scRNA-seq. En scRNA-seq, le processus de

dissociation peut nuire à certaines cellules sensibles, telles que les cellules de forme allongée ou les cellules stromales dans la matrice collagénique (Wu *et al.*, 2019), rendant leur obtention et séquençage séquençage difficiles. Le snRNA-seq est beaucoup moins stringent quant à lui, car il requiert uniquement l'isolation du noyau et non de la cellule entière. Par ailleurs, contrairement au scRNA-seq, il repose sur des protocoles rapides et légers qui permettent d'éviter certains biais liés par exemple à la digestion par les protéases ou à l'expression génique trompeuse causée par les ribosomes cytoplasmiques. Dans mes données de thèse, les gènes ribosomiaux sont par exemple très représentés dans les cellules, de la même façon que dans la majorité des études de scRNA-seq. Le snRNA-seq est cependant loin d'être dénué de limitations qui inciteront plutôt les chercheurs à considérer l'une ou l'autre technique selon leurs objectifs. Premièrement, il faut prendre en compte que le transcriptome nucléaire n'est pas le même que le transcriptome cytoplasmique, car les ARNm ne sont pas matures. Deuxièmement, la quantité d'ARNs de départ est globalement moins importante dans le noyau par rapport au cytoplasme. Il faut généralement s'attendre à un nombre médian de gènes par noyau inférieur à celui obtenu avec des cellules entières (Ding *et al.*, 2020). Or, le nombre de gènes est un paramètre important qui affecte la qualité des données et la classification des cellules. Toutefois, des protocoles semblent permettre de pallier ce problème en diminuant les volumes des gouttelettes de capture, ce qui augmente la concentration relative en ARN (Davey *et al.*, 2021). Ces limitations peuvent en partie expliquer pourquoi le snRNA-seq ne parvient pas à identifier des états cellulaires fins là où le scRNA-seq le permet, comme c'est le cas avec les cellules de la microglie (Thrupp *et al.*, 2020).

Selon les problématiques et objectifs de recherche, il faut plutôt opter pour l'une ou l'autre technique. Si ma thèse était à refaire dans l'état actuel de la littérature, je considérerais davantage le snRNA-seq pour plusieurs raisons. La première est l'économie de coût précédemment évoquée, qui permettrait de cumuler plus d'échantillons, moyennant la mise au point de la technique de nuclei hashing. Cela aurait mécaniquement augmenté la puissance statistique et le poids de nos observations, au détriment d'une diminution du rapport signal/bruit toutefois. Deuxièmement, nous avons eu assez tôt la volonté de conserver des échantillons congelés afin de permettre leur analyse a posteriori dans les cas où nous obtiendrions les métastases associées par exemple. Cela a longtemps pris la forme de cellules de ccRCC dissociées et congelées. Désormais, la nouvelle orientation de l'équipe en matière de prise en charge des tumeurs consiste à congeler le tissu entier en vue d'employer le snRNA-seq dans un second temps. D'un point de vue pratique, cela revient à ne pas dissocier systématiquement chaque échantillon mais plutôt à choisir ultérieurement lesquels exploiter et ainsi réaliser

l'ensemble des expériences en même temps diminuant par la même les biais techniques liés aux manipulations. Ce principe, associé à un protocole plus aisé pour l'isolation des noyaux, représente un gain de temps considérable. Dans cette logique, nous avons ainsi récemment généré des données de snRNA-seq sur des couples de tissus primaires et métastatiques qui avaient été congelés en prévision d'une étude de lignage tumorale et du processus métastatique. Les résultats préliminaires montrent que tous les types cellulaires sont bien représentés, en comparaison de nos projets de scRNA-seq, mais que le nombre de lectures est globalement inférieur, de façon cohérente avec les limitations évoquées précédemment. Dans ce projet d'étude du lignage cellulaire et de la progression métastatique, ce paramètre sera à prendre en compte, et l'analyse révélera son impact. Concernant l'enjeu du stockage d'échantillons congelés, j'ai cherché à démontrer dans l'axe 1 de ma thèse que notre protocole était compatible avec le séquençage de cellules uniques après dissociation de tissu de ccRCC congelé (échantillon RCC4). J'ai démontré que le transcriptome des cellules dissociées et congelées était très similaire à celui des cellules fraîchement dissociées, ouvrant la voie au type d'approches précédemment citées. Cela ne constituait qu'une preuve de concept avec un échantillon cependant, or je sais pour avoir essuyé des échecs qu'une partie significative des cellules décongelées est trop abîmée pour être analysée. Bien que j'aie dans un premier temps considéré ces résultats caducs après avoir pris connaissance des possibilités du snRNA-seq, je considère désormais que cette preuve de faisabilité reste intéressante, car elle pourrait bien permettre d'envisager plus tard des études d'échantillons asynchrones sur des questions pour lesquelles le snRNA-seq est limité, en raison des désavantages décrits plus haut.

Une autre raison importante pour laquelle l'usage du snRNA-seq aurait été intéressant selon moi est que nous avons eu du mal à identifier des fibroblastes. Cette population, participant de façon cruciale à la biologie tumorale, à la composition du TME et au processus métastatique (Asif *et al.*, 2021), est sous-représentée dans nos données et dans la littérature en général, probablement en raison de sa plus grande fragilité (Lendahl, Muhl and Betsholtz, 2022). Or le snRNA-seq semble plus à même de capturer ces populations fragiles, comme évoqué précédemment. Une étude l'a d'ailleurs comparé au scRNA-seq dans un contexte de fibrose rénale – relativement similaire au contexte du ccRCC –, et mis en évidence sa capacité accrue à détecter des populations discrètes telles que les cellules endothéliales, mésangiales et les podocytes glomérulaires (Wu *et al.*, 2019), suggérant qu'elle serait adaptée à l'étude des fibroblastes. Cependant, nos données préliminaires de snRNA-seq dans le projet d'étude de la progression métastatique ne montrent pas de proportions supérieures de CAFs (« *Cancer associated fibroblasts* », à l'encontre de cette théorie. Il faudrait donc étudier plus attentivement

la question. Il est possible que les CAFs soient une population d'emblée moins fréquente dans certaines tumeurs de ccRCC, ou que l'hétérogénéité tissulaire et les contraintes de prélèvement en dehors des zones trop remaniées empêchent de les capturer. À la lumière de ces observations, je pense toutefois que les projets s'intéressant aux rôles des CAFs dans le ccRCC devront prendre en compte les possibilités offertes par le snRNA-seq et établir des plans d'expérience adaptés.

Dans l'objectif de bâtir des cohortes importantes, l'utilisation du sciRNA-seq (« *Single cell combinatorial indexing RNA sequencing* ») est une autre approche possible. Le protocole est plus accessible, car il ne nécessite pas d'équipement particulier. Il s'affranchit notamment de la technologie microfluidique et de la capture en gouttelette au cœur de la technologie de 10X Genomics. De ce fait, le sciRNA-seq permet de générer, par marquages successifs avec des code-barres nucléotidiques, des données sur des millions de cellules à un coût moins important puisqu'il ne nécessite pas l'achat des kits commerciaux très coûteux. Mon équipe d'accueil a commencé à implémenter cette technique en collaboration avec une autre équipe sur le site rennais. Le seul coût conséquent est celui des index nécessaires à marquer les cellules individuelles. Ces index permettent d'identifier les cellules mais également d'identifier les échantillons d'origine. L'idée de cette collaboration est donc de mettre en commun l'achat des index entre plusieurs laboratoires et de multiplexer les cellules de plusieurs expériences pour réaliser du scRNA-seq haut-débit et moins coûteux, en routine. Par ailleurs, ce protocole a des avantages intéressants quelques soient les développements de nos thématiques. Il est très souple car il permet aussi bien d'analyser plus d'un million de cellule à partir d'un échantillon que 10 000 cellules pour 96 échantillons. Il bénéficie également d'un développement « open-source » et d'une multitude de protocoles alternatifs permettant par exemple de réaliser de la multiomique ou d'étudier l'accessibilité à la chromatine (scATAC-seq) (Cusanovich *et al.*, 2015; Vitak *et al.*, 2017).

Globalement, cette méthode pourrait faciliter à l'avenir l'établissement de cohortes conséquentes à même de répondre à des problématiques dédiées (atlas complets). Cependant nous manquons de recul sur la technique et devons encore faire des mises au point avant d'apprécier son potentiel réel.

Les locutions « hétérogénéité tumorale » et « hétérogénéité intra-tumorale » sont parfois confondues et sont devenues courantes à travers différents contextes de recherche. De ce fait, elles recouvrent plusieurs significations. Dans son sens original l'hétérogénéité tumorale était bien utilisée pour désigner la diversité observée dans les cellules malignes elles-mêmes

(Heppner, 1984), ce qui est le plus souvent désigné par hétérogénéité intra-tumorale aujourd'hui. Mais l'HIT désigne aussi parfois la diversité des cellules malignes et normales comprises dans une tumeur cancéreuse, y compris dans des atlas descriptifs à l'échelle de la cellule unique. Selon l'emploi qui est fait du terme, je considère que les études scRNA-seq ne dressent pas toujours le constat d'une HIT « maligne ». De ce point de vue-là, l'une des hypothèses de mon travail de thèse postulant que l'HIT serait importante dans les ccRCCs ne semble pas complètement vérifiée. Il est possible que les technologies actuelles de scRNA-seq ne soient en réalité pas encore suffisamment sensibles pour ce type de discrimination. En leur état et en dépit de nos stratégies de prélèvements multiples, l'HIT n'est pas la règle dans nos résultats ; chaque tumeur possède plutôt un profil particulier, certaines présentant une importante diversité clonale, et d'autres, un clone malin majoritaire et uniforme. Notre capacité à capturer cette hétérogénéité dans certaines tumeurs suggère que les prélèvements multiples y ont peut-être contribué. La stratégie de prélèvements multiples a déjà été utilisée avec du RNA-seq bulk dans une étude de référence sur la question de l'hétérogénéité dans le ccRCC, et a montré une diversité de clones répartis spatialement en des zones macroscopiques distinctes des tumeurs (Gerlinger *et al.*, 2012). Cependant, à l'inverse de cette étude, nous n'apportons pas la preuve du rôle de ces prélèvements multiples car leur information spatiale n'a pas été conservée, faute de moyens. Ainsi, bien que ce soit peu probable, les cinq clones malins identifiés dans une tumeur, dans l'axe 1, pourraient théoriquement provenir d'un seul prélèvement. C'est là un des points d'amélioration de mes travaux. Si je devais recommencer ces approches dans le contexte actuel, j'utiliserais l'une des solutions de multiplexage des cellules évoquées précédemment pour marquer les cellules correspondant à chaque prélèvement. Cela aurait permis d'intégrer une dimension spatiale à notre description de l'HIT. Nous aurions alors pu mettre les profils transcriptomiques en relation avec les examens macro- et microscopiques dans les zones échantillonnées, et appuyer le constat d'hétérogénéité.

Bien qu'intéressantes, ces approches ont une limitation infranchissable ; les cellules analysées sont dissociées et il est impossible d'appréhender leur organisation spatiale. Pour cela, les technologies de transcriptomique spatiale sont toutes indiquées. Ce domaine connaît une forte progression depuis quelques années avec la technique des codes-barres (la « *Spatial barcoding-based transcriptomics* »), inspirée des codes-barres cellulaires en scRNA-seq, et dont le Visium de chez *10X Genomics* est le plus utilisé. Son principe est de collecter les ARN sur une lame spéciale au niveau de points (ou « *spots* ») organisés en quadrillage, et de superposer informatiquement ce quadrillage sur les lames d'histologie. Actuellement, le Visium a une résolution de 55µm, soit 1 à 5 cellules, selon leur taille. Au départ de 100µm, cette



résolution continue d'augmenter et devrait très bientôt permettre le séquençage de cellules uniques dans leur contexte tissulaire. La technique est actuellement développée dans notre équipe, qui possède une véritable expertise en histopathologie du ccRCC. L'identification précise, la caractérisation transcriptomique puis la comparaison de l'organisation spatiale des cellules des différentes zones permettra de décrire l'HIT comme jamais avant. Au-delà de l'hétérogénéité cellulaire, la transcriptomique spatiale est susceptible d'apporter des éléments de compréhension nouveaux en lien avec l'expertise histopathologique. Elle pourra par exemple aider à mieux caractériser le microenvironnement tumoral en général, par exemple en ce qui concerne les cellules immunitaires. Bien que je les aie retirées expérimentalement pour les besoins de mes projets, ces cellules font l'objet d'une très grande attention dans de nombreux laboratoires étant donné leur rôle dans l'inflammation et dans le fonctionnement des immunothérapies. L'une des questions que je me suis posée dans mes projets concerne la nature infiltrante ou circulante des cellules immunitaires. Après dissociation des ccRCCs, j'obtenais des très grandes proportions de cellules immunitaires dans les suspensions cellulaires (près de 80%). En plus de poser problème pour l'étude des cellules malignes, je me demandais si elles avaient infiltré le tissu tumoral et étaient susceptibles d'en être affectées, ou si elles provenaient du sang périphérique souvent très abondant dans les échantillons après opération. Dans le premier cas, j'estimais que les cellules avaient davantage d'intérêt pour élucider la réponse et l'échappement immunitaire par exemple et, dans le second cas, que les cellules immunitaires n'étaient pas spécifiques du contexte tumoral et moins intéressantes à ce titre. Pour mieux comprendre leur nature, j'ai essayé de comparer les proportions immunitaires dans les données de cytométrie en flux avec celles estimées avec un marquage immunohistochimique, sur la base de comptages automatisés sur le logiciel ImageJ. Les différences que j'ai observées, allant d'un facteur 6 à 22, suggéraient que la majorité des cellules immunitaires en suspension n'étaient pas infiltrantes (n=6). Avec du recul, j'aurais pu essayer de laver abondamment les échantillons au PBS pour en retirer le sang et les cellules circulantes, bien que cela revête un caractère quelque peu aléatoire. Ici, la transcriptomique spatiale a un rôle important à jouer selon moi puisqu'elle permettra normalement d'identifier histologiquement et de caractériser par transcriptomique les cellules immunitaires infiltrantes. Cependant, en l'état actuel, la sensibilité de la technologie reste une limitation, aussi bien pour caractériser les cellules immunitaires que l'hétérogénéité intra-tumorale. Actuellement, ce sont près de 10 000 transcrits par point (~55µm), soit entre 2000 et 10 000 par cellule, là où le scRNA-seq situe les standards à environ 25 000 transcrits par cellules. Or, la sensibilité de l'analyse détermine directement la capacité à identifier et classifier des sous-populations cellulaires. Cela permettra normalement d'identifier

les principaux types cellulaires sans difficulté. Cet exemple d'application illustre les possibilités de la transcriptomique spatiale pour caractériser le microenvironnement tumoral, avec les limites évoquées toutefois. De la même façon, elle pourrait identifier et aider à caractériser les CAFs, en s'affranchissant ici aussi de leur dissociation délétère. Au-delà de l'HIT et du microenvironnement, la transcriptomique spatiale pourrait également aider à comprendre le front d'invasion de la tumeur dans le tissu environnant, que les experts en histopathologie identifient bien sur lames histologiques. L'invasion des tissus est une étape clef du processus métastatique. Peut-être qu'à terme, il sera possible d'identifier grâce à cette technologie les cellules invasives et leur profil transcriptomique. Enfin, d'autres pistes peuvent être à développer comme l'étude de la néovascularisation et son contexte cellulaire ou la caractérisation transcriptomique liée aux grades nucléolaires de l'ISUP, qui est un marqueur direct et graduel de la malignité d'une cellule. Je pense qu'il faut à l'avenir considérer l'idée de coupler les technologies de *single-cell*, dont el scRNA-seq, avec la transcriptomique spatiale, à partir d'un même échantillon, afin de superposer et compléter les données. Dans tous les cas, cette technologie porte de belles promesses qui prolongent naturellement plusieurs thématiques abordées dans ce travail de thèse.

J'avais pour objectif dans le premier axe de ma thèse d'évaluer la pertinence clinique de l'HIT, à savoir de déterminer si cette information pouvait être utilisée au bénéfice du patient. À partir d'une tumeur particulièrement hétérogène constituée de plusieurs clones tumoraux, un modèle de déconvolution a été utilisé permettant de s'appuyer sur les profils transcriptomiques de nos populations cellulaires pour inférer leur représentativité dans les données RNA-seq *bulk* (à l'échelle du tissu entier) chez les patients de la base de données du TCGA-KIRC. Nous avons ensuite démontré que certaines populations cellulaires inférées, en particulier certains clones tumoraux, étaient associées à la survie des patients. C'est notamment le cas pour une catégorie de patients dits à « bas-risque » (« *low-risk tumor* ») dont la tumeur progresse après résection chirurgicale malgré des éléments de diagnostic clinique favorables. L'enjeu n'est pas seulement d'améliorer la communication avec le patient mais bien d'améliorer sa prise en charge, en accentuant la surveillance, voire en décidant d'une stratégie thérapeutique plus adaptée, éventuellement sous forme de traitements adjuvants. Ces résultats incitent à reproduire cette stratégie à plus grande échelle et à constituer des cohortes de validation ; projet en cours dans notre laboratoire sous la direction des Dr. Aurélie Lardenois, Dr. Frédéric Chalmel et du Pr Nathalie Rioux-Leclercq, en collaboration avec la plateforme du CRB (Centre de ressources biologiques humaines) Santé du CHU de Rennes. Le CRB Santé possède des centaines

d'échantillons congelés sur lesquels seront effectuées des analyses BRB-seq, une technique de séquençage ARN *bulk* à bas coût. Au sein de ces données seront reproduites et améliorées les approches de déconvolution à des fins pronostiques. Cette approche fait actuellement l'objet d'une étude de brevetabilité par Inserm Transfert. Il y a encore peu, une revue de référence anticipait un recours important à la déconvolution dans les bases de données transcriptomiques *bulk*, en raison des trop grands coûts et investissements nécessaires à l'établissement de large cohorte de scRNA-seq (Suvà and Tirosh, 2019). Bien que ces cohortes doivent arriver dans les prochaines années selon moi, contrebalançant le point de vue des auteurs, ces approches de déconvolution restent pertinentes. Les travaux menés dans mon équipe d'accueil incitent à l'utilisation du BRB-seq en clinique. Cela revêt un enjeu crucial selon moi, car il est nécessaire de développer de nouveaux marqueurs et que beaucoup, tels que les signatures géniques pronostiques, sont transcriptomiques. Alors que le RNA-seq tarde à s'imposer en clinique, le BRB-seq, technologie beaucoup moins onéreuse, est facile à externaliser via des plateformes dédiées. Considérant les banques de données existantes, d'autres initiatives utiliseront probablement l'analyse conjointe de données de type RNA-seq *bulk* et *single-cell* RNA-seq avec des approches de déconvolution pour étudier les associations entre la proportion prédite de certains types cellulaires et l'issue de la maladie ou encore la réponse au traitement.

Dans le deuxième axe de mes travaux de recherche, j'ai développé une approche combinant culture organotypique et scRNA-seq. J'ai démontré la faisabilité de ce protocole en obtenant des suspensions de cellules individuelles et des données de cytométrie de bonne qualité. L'analyse montre que les tranches de ccRCC cultivées récapitulent globalement les mêmes populations cellulaires que dans l'atlas précédemment établi. Bien que les analyses statistiques et fonctionnelles des données ne soient pas achevées, l'approche développée est suffisamment sensible pour permettre d'observer une réponse différentielle aux traitements dans la grande majorité des populations et sous-populations cellulaires, incluant les cellules malignes. En tant que preuve de faisabilité, l'étude comportait initialement quatre tumeurs, et l'analyse, après écartement d'un échantillon, ne comporte que trois tumeurs soit 18 échantillons (6 conditions de traitement par tumeur). Cela représente une faible puissance statistique, d'autant que, à l'image de nos premières données, l'hétérogénéité inter-tumorale dans les cellules malignes est importante. Cette approche nécessitera à l'avenir une augmentation du nombre de réplicas biologiques. Ici aussi, les évolutions technologiques précédemment citées pourront permettre des études futures plus conséquentes. Par ailleurs, d'autres modèles d'étude de traitements pourraient s'avérer plus adéquats pour tester un grand nombre de molécules et

obtenir des résultats statistiquement significatifs. À cet égard, mon équipe d'accueil a entamé une collaboration avec l'équipe du Pr. Fred Saad, au centre hospitalier de l'université de Montréal, qui a développé des dispositifs microfluidiques (autrement appelé « *lab-on-chip* ») permettant la culture en trois dimensions d'explants *ex vivo* dans des chambres et le contrôle des flux d'échanges de gaz et de milieu (Dorrigiv *et al.*, 2021). Ils utilisent notamment ces dispositifs en oncologie pour étudier le cancer de la prostate sous traitements. Les intérêts majeurs justifiant de se porter sur cette technologie à l'avenir résident dans le nombre d'explants testés, entre 20 et 70 par dispositif, dans le contrôle des flux dans le système de culture (oxygène, nutriments) et, dans les temps de culture plus importants (jusqu'à deux semaines) (Dorrigiv *et al.*, 2021). L'application de scRNA-seq sur des tissus ainsi cultivés reste à valider cependant. Aux vues des mises au point nécessaires dans mon projet, je pense qu'une limitation éventuelle sera la quantité de tissu disponible, étant donné la taille réduite des explants dans ce dispositif. Mais elle pourra probablement être contournée en rassemblant plusieurs explants traités de façon homogène. Le snRNA-seq pourrait par ailleurs palier à cela car il nécessite moins de quantité de matériel biologique, en raison du meilleur rendement d'obtention des noyaux. De façon très intéressante ces mêmes collaborateurs optimisent actuellement leurs techniques pour permettre la culture de tissus préalablement congelés, dans la double perspective d'exploiter de grandes banques de tissus cancéreux dans les hôpitaux, pour lesquelles les données cliniques du patient sont cumulées avec du recul, et de permettre des plans expérimentaux originaux nécessitant de passer par la congélation. La spécialité de nos collaborateurs pour le cancer de la prostate permettra par ailleurs de renforcer les thématiques naissantes envisagées dans notre laboratoire concernant la caractérisation par scRNA-seq des effets de l'apalutamide, un antagoniste des récepteurs aux androgènes, à l'échelle de la cellule unique.

Ces travaux représentent un apport méthodologique significatif selon moi, en particulier dans le domaine du cancer du rein. En effet, à ma connaissance une seule étude a réalisé une approche similaire sur culture organotypique de coupes, dans le glioblastome (Zhao *et al.*, 2021), et aucune approche n'a combiné le scRNA-seq à un modèle 3D d'étude des traitements dans le cancer du rein. Ces stratégies combinant approches single-cell et modèles de culture 3D vont très probablement se développer considérablement à l'avenir. Elles seront complémentaires des études sur cohortes humaines traitées qui, bien qu'elles représentent un modèle *in vivo* plus adapté à l'étude des thérapies ciblées, requièrent de gros investissements et une bonne coordination dans des centres hospitaliers d'envergure. Le modèle *in vivo* ne permet

pas par ailleurs le même contrôle du modèle et l'homogénéisation des conditions de traitements. Les deux approches, comme la plupart des modèles, restent complémentaires.

De façon générale, nos approches devrait bénéficier de l'attention particulière portée à la technologie de scRNA-seq lors de la 10<sup>e</sup> édition du symposium international sur le cancer du rein (« International Kidney Cancer Symposium » ou IKCS), pendant laquelle David Braun a présenté d'importants travaux sur la caractérisation du microenvironnement immunitaire à l'échelle de la cellule unique et son évolution avec l'avancement de la maladie (Braun *et al.*, 2021). Il a présenté le plus grand atlas dans le ccRCC et l'un des plus grands en général, comprenant plus de 160 000 cellules immunitaires variées. Il a notamment démontré que les lymphocytes T CD8+ épuisés en phase terminale sont enrichis dans les métastases et que les macrophages pro-inflammatoires diminuaient avec l'avancée de la maladie. Ces travaux ont été très bien reçus, ce qui promet d'alimenter un vif intérêt de la communauté du cancer du rein pour les approches de type -omiques à l'échelle de la cellule unique.

Bien que la technologie 10X Genomics soit la plus utilisée aujourd'hui, elle ne permet le séquençage que de la partie 3' des transcrits, sur moins d'une centaine de nucléotides, et prive les chercheurs de la séquence complète du transcrit. La technique scCOLOR-seq, qui s'appuie sur la technologie Nanopore (Philpott *et al.*, 2021), permet de dépasser cette limite et de réaliser du séquençage transcriptomique *long-read* sur cellules uniques, donnant accès à la séquence complète et aux mutations éventuelles sur les transcrits, ainsi qu'à certaines isoformes de gènes. Cela permet de façon générale de caractériser les cellules de façon plus complète. Dans le ccRCC, la carcinogénèse reste encore peu comprise (voir la section **Mécanismes et pathophysiologie**), mais les mutations, telles que celles des gènes VHL, BAP1, PBRM1 ou SETD2, sont mieux connues (Hsieh *et al.*, 2017) et constituent des repères moléculaires importants. Ces mutations ont été utilisées pour étudier l'évolution tumorale et établir des arbres phylogénétiques dans le ccRCC (Gerlinger *et al.*, 2012) décrivant la dynamique de l'hétérogénéité intra-tumorale dans plusieurs tumeurs. Je pense qu'il est possible de s'inspirer de cette approche et combiner par exemple le scCOLOR-seq avec la construction d'arbre phylogénétiques basés sur les mutations, afin de caractériser, dans chaque sous-population, le fonds transcriptomique associés aux mutations les plus fréquentes et d'établir les profils fins chaque cellule en début et en fin de lignage. Cela pourra être comparé et complété avec des approches de lignage basées sur l'expression génique, comme celles proposées par Monocle3 (Cao *et al.*, 2019). Globalement, cela permettra de mieux comprendre la carcinogénèse et

l'évolution tumorale dans le ccRCC. Une autre technique adaptée de la technologie 10X, appelée GoT (« *Genotyping of Transcriptome* ») permet de mesurer les SNV (« *Single Nucleotid Variant* ») et mutations en même temps que l'expression génique dans les cellules individuelles (Nam *et al.*, 2019). Elle pourrait théoriquement permettre les mêmes approches, bien qu'elles ne donnent pas accès au transcrit entier. Le GoT et bien d'autres techniques font partie des technologies dites de « multiomique » à l'échelle single-cell. Ces méthodes se développent très rapidement en écho à l'essor du scRNA-seq, et permettent de détailler simultanément des processus tels que la méthylation, la variation génétique, l'abondance des protéines ou l'accessibilité de la chromatine (Perkel, 2021). Le CITE-seq, par exemple, permet d'établir les profils transcriptomiques et protéiques dans une cellule (Stoeckius *et al.*, 2017). Une autre technique, le scNMT-seq, combine la mesure du transcriptome, de la méthylation et de l'accessibilité à la chromatine, au sein des cellules uniques (Clark *et al.*, 2018). Le G&T-seq quant à lui capture à la fois l'ADN génomique et l'ARN (Macaulay *et al.*, 2015). Ces méthodes multiomiques visent à obtenir des informations complexes qui ne pourraient pas être atteintes avec une seule technique. Cela permettra à terme d'obtenir des profils particulièrement détaillés des populations cellulaires constituant l'hétérogénéité des tumeurs. De façon générale, l'avenir de ce domaine, en oncologie comme en général, tend vers la caractérisation complète de la cellule dans toutes ses dimensions. Cela représentera un véritable défi du point de vue de l'analyse bio-informatique cependant, et nécessitera en parallèle le développement d'outils intégrés permettant l'interprétation de telles données.

D'un point de vue personnel et disciplinaire, cette thèse m'a placé à l'interface entre plusieurs domaines, depuis l'anatomopathologie clinique jusqu'aux analyses bio-informatiques, en passant par les expérimentations de biologie et de culture cellulaire. Je pense que ce type de profils multidisciplinaires est nécessaire pour mener les approches décrites dans cette thèse. A l'image de cette multidisciplinarité, je pense que les axes d'étude de ma thèse articulaient des aspects fondamentaux et d'autres plus appliqués (développement d'un modèle d'étude de l'impact des traitements et d'un site web interactif) qui se complétaient bien de ce point de vue. Ces travaux contribuent selon moi aux apports et à la traduction clinique d'une technologie innovante, le scRNA-seq, dans le domaine du cancer du rein. L'information du degré d'HIT réel d'une tumeur, la prévalence d'un clone tumoral agressif, ou la mise en place de routine clinique incluant le test de molécules à des fins de décisions thérapeutique sont autant d'apports significatifs à la recherche et au bénéfice des patients. Je pense que dans les cinq

prochaines années, nous assisterons aux premières applications cliniques du scRNA-seq et qu'à terme, la technologie se démocratisera, devenant routinière et simple d'utilisation et d'analyse. Des investissements importants seront nécessaires d'ici là pour libérer tout son potentiel en recherche, découverte de médicaments et diagnostic. Cependant, un grand élan est en cours et, une fois les défis résolus, il ne fait aucun doute que le scRNA-seq révolutionnera la médecine personnalisée.

# Bibliographie





- Aceto, N. *et al.* (2014) ‘Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis’, *Cell*, 158(5). doi: 10.1016/j.cell.2014.07.013.
- Adams, M. D. *et al.* (1991) ‘Complementary DNA sequencing: Expressed sequence tags and human genome project’, *Science*, 252(5013). doi: 10.1126/science.2047873.
- Ahmed, N. *et al.* (2018) ‘Tumour microenvironment and metabolic plasticity in cancer and cancer stem cells: Perspectives on metabolic and immune regulatory signatures in chemoresistant ovarian cancer stem cells’, *Seminars in Cancer Biology*. doi: 10.1016/j.semcancer.2018.10.002.
- Akbani, R. *et al.* (2015) ‘Genomic Classification of Cutaneous Melanoma’, *Cell*, 161(7). doi: 10.1016/j.cell.2015.05.044.
- Akino, T. *et al.* (2009) ‘Cytogenetic abnormalities of tumor-associated endothelial cells in human malignant tumors’, *American Journal of Pathology*, 175(6). doi: 10.2353/ajpath.2009.090202.
- Albiges, L. *et al.* (2011) ‘Vascular endothelial growth factor-targeted therapies in advanced renal cell carcinoma’, *Hematology/Oncology Clinics of North America*. doi: 10.1016/j.hoc.2011.04.006.
- Alpern, D. *et al.* (2019) ‘BRB-seq: Ultra-affordable high-throughput transcriptomics enabled by bulk RNA barcoding and sequencing’, *Genome Biology*, 20(1). doi: 10.1186/s13059-019-1671-x.
- Amezquita, R. A. *et al.* (2020) ‘Orchestrating single-cell analysis with Bioconductor’, *Nature Methods*, 17(2). doi: 10.1038/s41592-019-0654-x.
- Amin, M. B. *et al.* (2017) ‘The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more “personalized” approach to cancer staging’, *CA: A Cancer Journal for Clinicians*, 67(2). doi: 10.3322/caac.21388.
- Anagnostou, V. *et al.* (2017) ‘Evolution of neoantigen landscape during immune checkpoint blockade in non-small cell lung cancer’, *Cancer Discovery*, 7(3). doi: 10.1158/2159-8290.CD-16-0828.
- Andrews, T. S. *et al.* (2021) ‘Tutorial: guidelines for the computational analysis of single-cell RNA sequencing data’, *Nature Protocols*. doi: 10.1038/s41596-020-00409-w.
- Arai, E. *et al.* (2008) ‘Genetic clustering of clear cell renal cell carcinoma based on array-comparative genomic hybridization: Its association with DNA methylation alteration and patient outcome’, *Clinical Cancer Research*, 14(17). doi: 10.1158/1078-0432.CCR-08-0443.
- Argüello, R. J. *et al.* (2020) ‘SCENITH: A Flow Cytometry-Based Method to Functionally Profile Energy Metabolism with Single-Cell Resolution’, *Cell Metabolism*, 32(6). doi: 10.1016/j.cmet.2020.11.007.
- Asif, P. J. *et al.* (2021) ‘The role of cancer-associated fibroblasts in cancer invasion and metastasis’, *Cancers*. doi: 10.3390/cancers13184720.
- Asp, M. *et al.* (2019) ‘A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart’, *Cell*, 179(7). doi: 10.1016/j.cell.2019.11.025.
- Aubry, M. *et al.* (2015) ‘From the core to beyond the margin: A genomic picture of glioblastoma intratumor heterogeneity’, *Oncotarget*, 6(14). doi: 10.18632/oncotarget.3297.
- Ayoub, N. M., Al-Shami, K. M. and Yaghan, R. J. (2019) ‘Immunotherapy for HER2-positive breast cancer: Recent advances and combination therapeutic approaches’, *Breast Cancer: Targets and Therapy*. doi: 10.2147/BCTT.S175360.
- Azizi, E. *et al.* (2018) ‘Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment’, *Cell*, 174(5). doi: 10.1016/j.cell.2018.05.060.
- Bacher, R. *et al.* (2017) ‘SCnorm: Robust normalization of single-cell RNA-seq data’, *Nature Methods*, 14(6). doi: 10.1038/nmeth.4263.
- Bakhtyar, N. *et al.* (2013) ‘Clear cell renal cell carcinoma induces fibroblast-mediated production of stromal periostin’, *European Journal of Cancer*, 49(16). doi:

10.1016/j.ejca.2013.06.032.

Balkwill, F. and Mantovani, A. (2001) 'Inflammation and cancer: Back to Virchow?', *Lancet*. doi: 10.1016/S0140-6736(00)04046-0.

Baran-Gale, J., Chandra, T. and Kirschner, K. (2018) 'Experimental design for single-cell RNA sequencing', *Briefings in Functional Genomics*, 17(4). doi: 10.1093/bfgp/elx035.

Barkley, D. *et al.* (2021) 'Cancer cell states and emergent properties of the dynamic tumor system', *Genome Research*. doi: 10.1101/gr.275308.121.

Baslan, T. and Hicks, J. (2017) 'Unravelling biology and shifting paradigms in cancer with single-cell sequencing', *Nature Reviews Cancer*. doi: 10.1038/nrc.2017.58.

Bast-Habersbrunner, A. *et al.* (2021) 'LncRNA Ctflos orchestrates transcription and alternative splicing in thermogenic adipogenesis', *EMBO reports*, 22(7). doi: 10.15252/embr.202051289.

Batai, K. *et al.* (2018) 'Whole-transcriptome sequencing identified gene expression signatures associated with aggressive clear cell renal cell carcinoma', *Genes and Cancer*, 9(5–6). doi: 10.18632/genesandcancer.183.

Baumans, V. (2004) 'Use of animals in experimental research: An ethical dilemma?', *Gene Therapy*. doi: 10.1038/sj.gt.3302371.

Bejarano, L., Jordão, M. J. C. and Joyce, J. A. (2021) 'Therapeutic targeting of the tumor microenvironment', *Cancer Discovery*. doi: 10.1158/2159-8290.CD-20-1808.

Bensalah, K. *et al.* (2006) 'Prognostic value of thrombocytosis in renal cell carcinoma', *Journal of Urology*, 175(3). doi: 10.1016/S0022-5347(05)00526-4.

Bensalah, K. *et al.* (2020) 'Recommandations françaises du Comité de cancérologie de l'AFU – actualisation 2020–2022 : prise en charge du cancer du rein', *Progrès en Urologie*, 30(12). doi: 10.1016/s1166-7087(20)30749-1.

Bentley, D. R. *et al.* (2008) 'Accurate whole human genome sequencing using reversible terminator chemistry', *Nature*, 456(7218). doi: 10.1038/nature07517.

Van Den Berg, A. *et al.* (2019) 'Personalised organs-on-chips: functional testing for precision medicine', *Lab on a Chip*, 19(2). doi: 10.1039/c8lc00827b.

Bergamaschi, A. *et al.* (2009) 'Molecular profiling and characterization of luminal-like and basal-like in vivo breast cancer xenograft models', *Molecular Oncology*, 3(5–6). doi: 10.1016/j.molonc.2009.07.003.

Van den Berge, K. *et al.* (2018) 'Observation weights unlock bulk RNA-seq tools for zero inflation and single-cell applications', *Genome Biology*, 19(1). doi: 10.1186/s13059-018-1406-4.

Van den Berge, K. *et al.* (2020) 'Trajectory-based differential expression analysis for single-cell sequencing data', *Nature Communications*, 11(1). doi: 10.1038/s41467-020-14766-3.

Berglund, E. *et al.* (2018) 'Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity', *Nature Communications*, 9(1). doi: 10.1038/s41467-018-04724-5.

Bergström, A. *et al.* (2001) 'Obesity and renal cell cancer – a quantitative review', *British Journal of Cancer*, 85(7). doi: 10.1054/bjoc.2001.2040.

Beroukhir, R. *et al.* (2009) 'Patterns of gene expression and copy-number alterations in von-Hippel Lindau disease-associated and sporadic clear cell carcinoma of the kidney', *Cancer Research*, 69(11). doi: 10.1158/0008-5472.CAN-09-0146.

Bertrand, E. *et al.* (2020) 'Human Kidney Tumor Dissociation for single-cell genomics', pp. 1–11.

Bērziņa, S. *et al.* (2021) 'Technological advances in tumor-on-chip technology: From bench to bedside', *Cancers*. doi: 10.3390/cancers13164192.

Beutner, U. *et al.* (2015) 'Survival after renal cell carcinoma metastasis to the thyroid:

Single center experience and systematic review of the literature', *Thyroid*. doi: 10.1089/thy.2014.0498.

Beyers, R. P. J. *et al.* (2019) 'Mitochondrial haplotypes affect metabolic phenotypes in the Drosophila Genetic Reference Panel', *Nature Metabolism*, 1(12). doi: 10.1038/s42255-019-0147-3.

Bex, A. *et al.* (2016) 'The Role of Cytoreductive Nephrectomy: European Association of Urology Recommendations in 2016', *European Urology*, 70(6). doi: 10.1016/j.eururo.2016.07.005.

Bi, K. *et al.* (2021) 'Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma', *Cancer Cell*, pp. 1–13. doi: 10.1016/j.ccell.2021.02.015.

Bianchi, M. *et al.* (2012) 'Distribution of metastatic sites in renal cell carcinoma: A population-based analysis', *Annals of Oncology*, 23(4). doi: 10.1093/annonc/mdr362.

Bissell, M. J., Hall, H. G. and Parry, G. (1982) 'How does the extracellular matrix direct gene expression?', *Journal of Theoretical Biology*, 99(1). doi: 10.1016/0022-5193(82)90388-5.

Bissell, M. J. and Hines, W. C. (2011) 'Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression', *Nature Medicine*. doi: 10.1038/nm.2328.

Blackmur, J. P. *et al.* (2021) 'Leibovich score is the optimal clinico-pathological system associated with recurrence of non-metastatic clear cell renal cell carcinoma', *Urologic Oncology: Seminars and Original Investigations*, 39(7). doi: 10.1016/j.urolonc.2021.04.007.

Blondel, V. D. *et al.* (2008) 'Fast unfolding of communities in large networks', *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10). doi: 10.1088/1742-5468/2008/10/P10008.

Boj, S. F. *et al.* (2015) 'Organoid models of human and mouse ductal pancreatic cancer', *Cell*, 160(1–2). doi: 10.1016/j.cell.2014.12.021.

Borcherding, N. *et al.* (2021) 'Mapping the immune environment in clear cell renal carcinoma by single-cell genomics', *Communications Biology*, 4(1). doi: 10.1038/s42003-020-01625-6.

Boutros, P. C. *et al.* (2015) 'Spatial genomic heterogeneity within localized, multifocal prostate cancer', *Nature Genetics*, 47(7). doi: 10.1038/ng.3315.

Boveri, T. (2008) 'Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris', *Journal of Cell Science*, 121(SUPPL. 1). doi: 10.1242/jcs.025742.

Brabletz, T. *et al.* (2018) 'EMT in cancer', *Nature Reviews Cancer*. doi: 10.1038/nrc.2017.118.

Braun, D. A. *et al.* (2021) 'Progressive immune dysfunction with advancing disease stage in renal cell carcinoma', *Cancer Cell*, 39(5). doi: 10.1016/j.ccell.2021.02.013.

Brennan, C. W. *et al.* (2013) 'The somatic genomic landscape of glioblastoma', *Cell*, 155(2). doi: 10.1016/j.cell.2013.09.034.

Brodaczewska, K. K. *et al.* (2016) 'Choosing the right cell line for renal cell cancer research', *Molecular Cancer*. doi: 10.1186/s12943-016-0565-8.

Brokaw, C. J. (2000) 'Aneuploidy, the somatic mutation that makes cancer a species of its own', *Cell Motility and the Cytoskeleton*. doi: 10.1002/1097-0169(200010)47:2<81::aid-cml1>3.0.co;2-%23.

De Bruin, E. C. *et al.* (2014) 'Spatial and temporal diversity in genomic instability processes defines lung cancer evolution', *Science*, 346(6206). doi: 10.1126/science.1253462.

Bulyk, M. L. (2006) 'DNA microarray technologies for measuring protein-DNA interactions', *Current Opinion in Biotechnology*. doi: 10.1016/j.copbio.2006.06.015.

Burrell, R. A. *et al.* (2013) 'The causes and consequences of genetic heterogeneity in

cancer evolution', *Nature*. doi: 10.1038/nature12625.

Butler, A. *et al.* (2018) 'Integrating single-cell transcriptomic data across different conditions, technologies, and species.', *Nature biotechnology*. doi: 10.1038/nbt.4096.

Campbell, P. J. *et al.* (2008) 'Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing', *Proceedings of the National Academy of Sciences of the United States of America*, 105(35). doi: 10.1073/pnas.0801523105.

Campbell, S. C. *et al.* (2009) 'Guideline for Management of the Clinical T1 Renal Mass', *Journal of Urology*, 182(4 SUPPL.). doi: 10.1016/j.juro.2009.07.004.

Cannoodt, R. *et al.* (2021) 'Spearheading future omics analyses using dyngen, a multi-modal simulator of single cells', *Nature Communications*, 12(1). doi: 10.1038/s41467-021-24152-2.

Cao, J. *et al.* (2017) 'Comprehensive single-cell transcriptional profiling of a multicellular organism', *Science*, 357(6352). doi: 10.1126/science.aam8940.

Cao, J. *et al.* (2019) 'The single-cell transcriptional landscape of mammalian organogenesis', *Nature*, 566(7745). doi: 10.1038/s41586-019-0969-x.

Cao, J. *et al.* (2020) 'A human cell atlas of fetal gene expression', *Science*, 370(6518). doi: 10.1126/science.aba7721.

Cao, Y. *et al.* (2017) 'scRNAseqDB: A database for RNA-seq based gene expression profiles in human single cells', *Genes*, 8(12). doi: 10.3390/genes8120368.

Capitano, U. *et al.* (2019) 'Epidemiology of Renal Cell Carcinoma [Figure presented]', *European Urology*. doi: 10.1016/j.eururo.2018.08.036.

Carranza-Torres, I. E. *et al.* (2015) 'Organotypic culture of breast tumor explants as a multicellular system for the screening of natural compounds with antineoplastic potential', *BioMed Research International*, 2015. doi: 10.1155/2015/618021.

Castro, L. N. G., Tirosh, I. and Suvà, M. L. (2021) 'Decoding cancer biology one cell at a time', *Cancer Discovery*. doi: 10.1158/2159-8290.CD-20-1376.

Casuscelli, J. *et al.* (2019) 'Characterization and Impact of TERT Promoter Region Mutations on Clinical Outcome in Renal Cell Carcinoma', *European Urology Focus*, 5(4). doi: 10.1016/j.euf.2017.09.008.

Centenera, M. M. *et al.* (2013) 'Ex vivo culture of human prostate tissue and drug development', *Nature Reviews Urology*. doi: 10.1038/nrurol.2013.126.

Chalmel, F. and Primig, M. (2008) 'The Annotation, Mapping, Expression and Network (AMEN) suite of tools for molecular systems biology', *BMC Bioinformatics*, 9. doi: 10.1186/1471-2105-9-86.

Chantal Dreyer, Eric Raymond and Sandrine Faivre (2009) 'La voie de signalisation PI3K/AKT/mTOR', *Cancéro digest*, 1(3). doi: 10.4267/2042/28434.

Chen, B. *et al.* (2018) 'Profiling tumor infiltrating immune cells with CIBERSORT', in *Methods in Molecular Biology*. doi: 10.1007/978-1-4939-7493-1\_12.

Chen, F. *et al.* (2016) 'Multilevel Genomics-Based Taxonomy of Renal Cell Carcinoma', *Cell Reports*, 14(10). doi: 10.1016/j.celrep.2016.02.024.

Chen, G., Ning, B. and Shi, T. (2019) 'Single-cell RNA-seq technologies and related computational data analysis', *Frontiers in Genetics*. doi: 10.3389/fgene.2019.00317.

Chen, J. *et al.* (2017) 'Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq', *Nature Protocols*, 12(3). doi: 10.1038/nprot.2017.003.

Chen, M. *et al.* (2009) 'Genome-wide profiling of chromosomal alterations in renal cell carcinoma using high-density single nucleotide polymorphism arrays', *International Journal of Cancer*, 125(10). doi: 10.1002/ijc.24642.

Chen, P. *et al.* (2018) 'MicroRNA-191-5p exerts a tumor suppressive role in renal cell carcinoma', *Experimental and Therapeutic Medicine*, 15(2). doi: 10.3892/etm.2017.5581.

Chen, Q. *et al.* (2017) 'Growth-induced stress enhances epithelial-mesenchymal

transition induced by il-6 in clear cell renal cell carcinoma via the akt/gsk-3 $\beta$ / $\beta$ -catenin signaling pathway', *Oncogenesis*, 6(8). doi: 10.1038/oncsis.2017.74.

Chen, Q. *et al.* (2019) 'Cancer Immunotherapies and Humanized Mouse Drug Testing Platforms', *Translational Oncology*. doi: 10.1016/j.tranon.2019.04.020.

Chen, W. *et al.* (2022) 'Live-seq enables temporal transcriptomic recording of single cells', *Nature*. Springer US, 608(August). doi: 10.1038/s41586-022-05046-9.

Cheng, Y. H. *et al.* (2019) 'Hydro-Seq enables contamination-free high-throughput single-cell RNA-sequencing for circulating tumor cells', *Nature Communications*, 10(1). doi: 10.1038/s41467-019-10122-2.

Chevrier, S. *et al.* (2017) 'An Immune Atlas of Clear Cell Renal Cell Carcinoma', *Cell*, 169(4). doi: 10.1016/j.cell.2017.04.016.

Chi, Y. *et al.* (2020) 'Cancer cells deploy lipocalin-2 to collect limiting iron in leptomeningeal metastasis', *Science*, 369(6501). doi: 10.1126/science.aaz2193.

Chien, C. C. *et al.* (2017) 'Epidemiology of cancer in end-stage renal disease dialysis patients: A national cohort study in Taiwan', *Journal of Cancer*, 8(1). doi: 10.7150/jca.16550.

Cho, D. S. *et al.* (2016) 'Prognostic significance of modified Glasgow Prognostic Score in patients with non-metastatic clear cell renal cell carcinoma', *Scandinavian Journal of Urology*, 50(3). doi: 10.3109/21681805.2015.1136677.

Chuanyu, S. *et al.* (2017) 'Periostin promotes migration and invasion of renal cell carcinoma through the integrin/focal adhesion kinase/c-Jun N-terminal kinase pathway', *Tumor Biology*, 39(4). doi: 10.1177/1010428317694549.

Clark, S. J. *et al.* (2018) 'ScNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells e', *Nature Communications*, 9(1). doi: 10.1038/s41467-018-03149-4.

Clevers, H. (2016) 'Modeling Development and Disease with Organoids', *Cell*. doi: 10.1016/j.cell.2016.05.082.

Corrò, C. and Moch, H. (2018) 'Biomarker discovery for renal cancer stem cells', *Journal of Pathology: Clinical Research*. doi: 10.1002/cjp2.91.

Creighton, C. J. *et al.* (2013) 'Comprehensive molecular characterization of clear cell renal cell carcinoma', *Nature*, 499(7456), pp. 43–49. doi: 10.1038/nature12222.

Crusz, S. M. *et al.* (2016) 'Heterogeneous response and progression patterns reveal phenotypic heterogeneity of tyrosine kinase inhibitor response in metastatic renal cell carcinoma', *BMC Medicine*. doi: 10.1186/s12916-016-0729-9.

Cumberbatch, M. G. *et al.* (2016) 'The Role of Tobacco Smoke in Bladder and Kidney Carcinogenesis: A Comparison of Exposures and Meta-analysis of Incidence and Mortality Risks', *European Urology*. doi: 10.1016/j.eururo.2015.06.042.

Cusanovich, D. A. *et al.* (2015) 'Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing', *Science*, 348(6237). doi: 10.1126/science.aab1601.

Dagher, J. *et al.* (2017) 'Clear cell renal cell carcinoma: validation of World Health Organization/International Society of Urological Pathology grading', *Histopathology*, 71(6). doi: 10.1111/his.13311.

Dagher, J. *et al.* (2021) 'Multiple metastatic clones assessed by an integrative multiomics strategy in clear cell renal carcinoma: A case study', *Journal of Clinical Pathology*. doi: 10.1136/jclinpath-2020-207326.

Dangles-Marie, V. *et al.* (2007) 'Establishment of human colon cancer cell lines from fresh tumors versus xenografts: Comparison of success rate and cell line features', *Cancer Research*, 67(1). doi: 10.1158/0008-5472.CAN-06-0594.

Darde, T. A. *et al.* (2015) 'The ReproGenomics Viewer: An integrative cross-species toolbox for the reproductive science community', *Nucleic Acids Research*. doi:

10.1093/nar/gkv345.

Darde, T. A. *et al.* (2019) ‘The ReproGenomics Viewer: a multi-omics and cross-species resource compatible with single-cell studies for the reproductive science community’, *Bioinformatics*. doi: 10.1093/bioinformatics/btz047.

Darwin, C. (1860) ‘The Origin of Species, by Means of Natural Selection, or the Preservation of Favored Races in the Struggle for Life’, *The Crayon*, 7(5). doi: 10.2307/25528056.

Davey, K. *et al.* (2021) ‘A flexible microfluidic system for single-cell transcriptome profiling elucidates phased transcriptional regulators of cell cycle’, *Scientific Reports*, 11(1). doi: 10.1038/s41598-021-86070-z.

Davies, E. J. *et al.* (2015) ‘Capturing complex tumour biology in vitro: Histological and molecular characterisation of precision cut slices’, *Scientific Reports*, 5. doi: 10.1038/srep17187.

Davis, A., Gao, R. and Navin, N. (2017) ‘Tumor evolution: Linear, branching, neutral or punctuated?’, *Biochimica et Biophysica Acta - Reviews on Cancer*, 1867(2). doi: 10.1016/j.bbcan.2017.01.003.

Delahunt, B. *et al.* (2011) ‘Grading of clear cell renal cell carcinoma should be based on nucleolar prominence’, *American Journal of Surgical Pathology*, 35(8). doi: 10.1097/PAS.0b013e318220697f.

Delahunt, B. *et al.* (2013) ‘The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters’, *American Journal of Surgical Pathology*, 37(10). doi: 10.1097/PAS.0b013e318299f0fb.

Delahunt, B. and Eble, J. N. (2005) ‘History of the development of the classification of renal cell neoplasia’, *Clinics in Laboratory Medicine*. doi: 10.1016/j.cll.2005.01.007.

Dexter, D. L. *et al.* (1978) ‘Heterogeneity of Tumor Cells from a Single Mouse Mammary Tumor’, *Cancer Research*, 38(10).

Díaz-Montero, C. M., Rini, B. I. and Finke, J. H. (2020) ‘The immunology of renal cell carcinoma’, *Nature Reviews Nephrology*. doi: 10.1038/s41581-020-0316-3.

Ding, J. *et al.* (2020) ‘Systematic comparison of single-cell and single-nucleus RNA-sequencing methods’, *Nature Biotechnology*, 38(6). doi: 10.1038/s41587-020-0465-8.

Ding, L. *et al.* (2012) ‘Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing’, *Nature*, 481(7382). doi: 10.1038/nature10738.

Dong, R. *et al.* (2020) ‘Single-Cell Characterization of Malignant Phenotypes and Developmental Trajectories of Adrenal Neuroblastoma’, *Cancer Cell*, 38(5). doi: 10.1016/j.ccell.2020.08.014.

Dorriv, D. *et al.* (2021) ‘Microdissected tissue vs tissue slices—a comparative study of tumor explant models cultured on-chip and off-chip’, *Cancers*, 13(16). doi: 10.3390/cancers13164208.

Dumanskiy, Y. V. *et al.* (2013) ‘Markers of epithelial-mesenchymal transition in renal cell carcinoma’, *Experimental Oncology*, 35(4).

Dutcher, J. P. *et al.* (2020) ‘On the Shoulders of Giants: The Evolution of Renal Cell Carcinoma Treatment—Cytokines, Targeted Therapy, and Immunotherapy’, *American Society of Clinical Oncology Educational Book*, (40). doi: 10.1200/edbk\_280817.

Ediriwickrema, A. *et al.* (2020) ‘Single-cell mutational profiling enhances the clinical evaluation of AML MRD’, *Blood Advances*, 4(5). doi: 10.1182/bloodadvances.2019001181.

Eldredge, N. and Gould, S. J. (1972) ‘Punctuated equilibria: an alternative to phyletic gradualism’, *Models in paleobiology*, 82.

Erdoğan, F., Demirel, A. and Polat, Ö. (2004) ‘Prognostic significance of morphologic parameters in renal cell carcinoma’, *International Journal of Clinical Practice*, 58(4). doi: 10.1111/j.1368-5031.2004.00008.x.

Errarte, P., Larrinaga, G. and López, J. I. (2020) 'The role of cancer-associated fibroblasts in renal cell carcinoma. An example of tumor modulation through tumor/non-tumor cell interactions', *Journal of Advanced Research*. doi: 10.1016/j.jare.2019.09.004.

Esch, E. W., Bahinski, A. and Huh, D. (2015) 'Organs-on-chips at the frontiers of drug discovery', *Nature Reviews Drug Discovery*. doi: 10.1038/nrd4539.

Escudier, B. *et al.* (2014) 'Renal cell carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up', *Annals of Oncology*, 25(August), pp. iii49–iii56. doi: 10.1093/annonc/mdw328.

European Association of Urology (2021) 'EAU Guidelines. Edn. Presented at the EAU Annual Congress Milan 2021', in. Milan. Available at: <https://uroweb.org/guideline/renal-cell-carcinoma/>.

Falzone, L., Salomone, S. and Libra, M. (2018) 'Evolution of cancer pharmacological treatments at the turn of the third millennium', *Frontiers in Pharmacology*. doi: 10.3389/fphar.2018.01300.

Fan, D. *et al.* (2020) 'Prognostic significance of PI3K/AKT/ mTOR signaling pathway members in clear cell renal cell carcinoma', *PeerJ*, 2020(6). doi: 10.7717/peerj.9261.

Fang, Y. *et al.* (2013) 'Protein Expression of ZEB2 in Renal Cell Carcinoma and Its Prognostic Significance in Patient Survival', *PLoS ONE*, 8(5). doi: 10.1371/journal.pone.0062558.

Farace, F. *et al.* (2011) 'Levels of circulating CD45 dim CD34 VEGFR2 progenitor cells correlate with outcome in metastatic renal cell carcinoma patients treated with tyrosine kinase inhibitors', *British Journal of Cancer*, 104(7). doi: 10.1038/bjc.2011.72.

Fearon R. Eric and Bert, V. (1990) 'A Genetic Model for Colorectal Tumorigenesis', *Cell*, 61(2).

Fendler, A. *et al.* (2020) 'Inhibiting WNT and NOTCH in renal cancer stem cells and the implications for human patients', *Nature Communications*, 11(1). doi: 10.1038/s41467-020-14700-7.

Feng, X. *et al.* (2019) 'Frequency, incidence and survival outcomes of clear cell renal cell carcinoma in the United States from 1973 to 2014 A SEER-based analysis', *Medicine (United States)*, 98(31). doi: 10.1097/MD.00000000000016684.

Ferlay, J. *et al.* (2015) 'Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012', *International Journal of Cancer*, 136(5). doi: 10.1002/ijc.29210.

Fialkow, P. J. (1974) 'Human genetic markers as tracers of tumour histogenesis', *Journal of Clinical Pathology*, 27(SUPPL. 7). doi: 10.1136/jcp.27.Suppl\_7.11.

Ficarra, V. *et al.* (2005) 'Original and reviewed nuclear grading according to the Fuhrman system: A multivariate analysis of 388 patients with conventional renal cell carcinoma', *Cancer*. doi: 10.1002/cncr.20749.

Filbin, M. G. *et al.* (2018) 'Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq', *Science*, 360(6386). doi: 10.1126/science.aao4750.

Finke, J. H. *et al.* (2013) 'Modification of the tumor microenvironment as a novel target of renal cell carcinoma therapeutics', *Cancer Journal (United States)*. doi: 10.1097/PPO.0b013e31829da0ae.

Finnberg, N. K. *et al.* (2017) 'Application of 3D tumoroid systems to define immune and cytotoxic therapeutic responses based on tumoroid and tissue slice culture molecular signatures', *Oncotarget*, 8(40). doi: 10.18632/oncotarget.19965.

Fleck, C. *et al.* (2000) 'Ex vivo stimulation of renal tubular p-aminohippurate transport by dexamethasone triiodothyronine in human renal cell carcinoma', *Urological Research*, 28(6). doi: 10.1007/s002400000142.

Flecknell, P. (2002) 'Replacement, reduction and refinement.', *ALTEX: Alternativen zu*



*Tierexperimenten*, 19(2).

Franzén, O., Gan, L.-M. and Björkegren, J. L. M. (2019) 'PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data', *Database*, 2019, pp. 1–9. doi: 10.1093/database/baz046.

Freese, E. (1962) 'On the evolution of the base composition of DNA', *Journal of Theoretical Biology*, 3(1). doi: 10.1016/S0022-5193(62)80005-8.

Fukuda, S. *et al.* (2021) 'Impact of C-reactive protein flare-response on oncological outcomes in patients with metastatic renal cell carcinoma treated with nivolumab', *Journal for ImmunoTherapy of Cancer*, 9(2). doi: 10.1136/jitc-2020-001564.

Gago-Dominguez, M. *et al.* (2002) 'Lipid peroxidation: A novel and unifying concept of the etiology of renal cell carcinoma (United States)', *Cancer Causes and Control*, 13(3). doi: 10.1023/A:1015044518505.

Gao, R. *et al.* (2016) 'Punctuated copy number evolution and clonal stasis in triple-negative breast cancer', *Nature Genetics*, 48(10). doi: 10.1038/ng.3641.

Galalde, D. R. *et al.* (2018) 'Highly parallel direct RNA sequencing on an array of nanopores', *Nature Methods*, 15(3). doi: 10.1038/nmeth.4577.

García-Figueiras, R. *et al.* (2019) 'How clinical imaging can assess cancer biology', *Insights into Imaging*. doi: 10.1186/s13244-019-0703-0.

Gati, A. *et al.* (2014) 'Obesity and renal cancer: Role of adipokines in the tumor-immune system conflict', *OncImmunology*. doi: 10.4161/onci.27810.

Gaublomme, J. T. *et al.* (2019) 'Nuclei multiplexing with barcoded antibodies for single-nucleus genomics', *Nature Communications*, 10(1). doi: 10.1038/s41467-019-10756-2.

Gautier Defossez Sandra Le Guyader-Peyrou Zoé Uhry Pascale Grosclaude Marc Colonna Emmanuelle Dantony Patricia Delafosse Florence Molinié Anne-Sophie Woronoff Anne-Marie Bouvier Laurent Remontet Nadine Bossard Alain Monnereau, A. (2019) 'Estimations nationales de l'incidence et de la mortalité par cancer en Étude à partir des registres des cancers du réseau Francim Synthèse', p. 2. Available at: <http://invs.santepubliquefrance.fr/Dossiers-thematiques/>.

Gawad, C., Koh, W. and Quake, S. R. (2014) 'Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics', *Proceedings of the National Academy of Sciences of the United States of America*, 111(50). doi: 10.1073/pnas.1420822111.

Gayed, B. A. *et al.* (2014) 'Ki67 is an independent predictor of oncological outcomes in patients with localized clear-cell renal cell carcinoma', *BJU International*, 113(4). doi: 10.1111/bju.12263.

Gejman, R. S. *et al.* (2018) 'Rejection of immunogenic tumor clones is limited by clonal fraction', *eLife*, 7. doi: 10.7554/eLife.41090.

Gerlach, M. M. *et al.* (2014) 'Slice cultures from head and neck squamous cell carcinoma: A novel test system for drug susceptibility and mechanisms of resistance', *British Journal of Cancer*, 110(2). doi: 10.1038/bjc.2013.700.

Gerlinger, M. *et al.* (2012) 'Intratumor heterogeneity and branched evolution revealed by multiregion sequencing.', *The New England journal of medicine*, 366(10), pp. 883–892. doi: 10.1056/NEJMoal113205.

Gerlinger, M. *et al.* (2014) 'Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing', *Nature Genetics*, 46(3), pp. 225–233. doi: 10.1038/ng.2891.

Gerlinger, M. *et al.* (2015) 'Intratour heterogeneity in urologic cancers: From molecular evidence to clinical implications', *European Urology*. doi: 10.1016/j.eururo.2014.04.014.

Giacosa, S. *et al.* (2021) 'Cooperative blockade of CK2 and ATM kinases drives apoptosis in VHL-deficient renal carcinoma cells through ROS overproduction', *Cancers*,

13(3). doi: 10.3390/cancers13030576.

Gilbert, S. F. and Barresi, M. J. F. (2017) 'DEVELOPMENTAL BIOLOGY, 11TH EDITION 2016', *American Journal of Medical Genetics Part A*, 173(5). doi: 10.1002/ajmg.a.38166.

Gild, P., Ehdai, B. and Kluth, L. A. (2017) 'Effect of obesity on bladder cancer and renal cell carcinoma incidence and survival', *Current Opinion in Urology*. doi: 10.1097/MOU.0000000000000425.

Gill, B. J. *et al.* (2014) 'MRI-localized biopsies reveal subtype-specific differences in molecular and cellular composition at the margins of glioblastoma', *Proceedings of the National Academy of Sciences of the United States of America*, 111(34). doi: 10.1073/pnas.1405839111.

Gnarra, J. R. *et al.* (1994) 'Mutations of the VHL tumour suppressor gene in renal carcinoma', *Nature Genetics*, 7(1). doi: 10.1038/ng0594-85.

Gojo, J. *et al.* (2020) 'Single-Cell RNA-Seq Reveals Cellular Hierarchies and Impaired Developmental Trajectories in Pediatric Ependymoma', *Cancer Cell*, 38(1). doi: 10.1016/j.ccell.2020.06.004.

Gomez, R. A., Norwood, V. F. and Tufro-Mcreddie, A. (1997) 'Development of the kidney vasculature', *Microscopy Research and Technique*. doi: 10.1002/(SICI)1097-0029(19971101)39:3<254::AID-JEMT5>3.0.CO;2-K.

Grassi, L. *et al.* (2019) 'Organoids as a new model for improving regenerative medicine and cancer personalized therapy in renal diseases', *Cell Death and Disease*, 10(3). doi: 10.1038/s41419-019-1453-0.

Grael, A. L. *et al.* (2020) 'TGF $\beta$ -blockade uncovers stromal plasticity in tumors by revealing the existence of a subset of interferon-licensed fibroblasts', *Nature Communications*, 11(1). doi: 10.1038/s41467-020-19920-5.

Greaves, M. and Maley, C. C. (2012) 'Clonal evolution in cancer', *Nature*. doi: 10.1038/nature10762.

Griffiths, J. A., Scialdone, A. and Marioni, J. C. (2018) 'Using single-cell genomics to understand developmental processes and cell fate decisions', *Molecular Systems Biology*, 14(4). doi: 10.15252/msb.20178046.

Grisanzio, C. *et al.* (2011) 'Orthotopic xenografts of RCC retain histological, immunophenotypic and genetic features of tumours in patients', *Journal of Pathology*, 225(2). doi: 10.1002/path.2929.

Guo, H. *et al.* (2015) 'The PI3K/AKT Pathway and Renal Cell Carcinoma', *Journal of Genetics and Genomics*. doi: 10.1016/j.jgg.2015.03.003.

Guo, X. *et al.* (2018) 'Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing', *Nature Medicine*, 24(7). doi: 10.1038/s41591-018-0045-3.

Guyot, C. *et al.* (2007) 'Specific activation of the different fibrogenic cells in rat cultured liver slices mimicking in vivo situations', *Virchows Archiv*, 450(5). doi: 10.1007/s00428-007-0390-y.

Hafemeister, C. and Satija, R. (2019) 'Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression', *Genome Biology*, 20(1). doi: 10.1186/s13059-019-1874-1.

Haghverdi, L. *et al.* (2018) 'Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors', *Nature Biotechnology*, 36(5). doi: 10.1038/nbt.4091.

Hajdu, S. I. (2004) 'A note from history: the first tumor pathologist.', *Annals of clinical and laboratory science*, 34(3).

Hakimi, A. A., Ostrovnaya, I., *et al.* (2013) 'Adverse outcomes in clear cell renal cell carcinoma with mutations of 3p21 epigenetic regulators BAP1 and SETD2: A report by

MSKCC and the KIRC TCGA research network', *Clinical Cancer Research*, 19(12). doi: 10.1158/1078-0432.CCR-12-3886.

Hakimi, A. A., Chen, Y. B., *et al.* (2013) 'Clinical and pathologic impact of select chromatin-modulating tumor suppressors in clear cell renal cell carcinoma', *European Urology*, 63(5). doi: 10.1016/j.eururo.2012.09.005.

Hakimi, A. A. *et al.* (2016) 'An Integrated Metabolic Atlas of Clear Cell Renal Cell Carcinoma', *Cancer Cell*, 29(1). doi: 10.1016/j.ccell.2015.12.004.

Hakimi, A. A. *et al.* (2019) 'Transcriptomic profiling of the tumor microenvironment reveals distinct subgroups of clear cell renal cell cancer: Data from a randomized phase III trial', *Cancer Discovery*, 9(4). doi: 10.1158/2159-8290.CD-18-0957.

Hakimi, A. A., Pham, C. G. and Hsieh, J. J. (2013) 'A clear picture of renal cell carcinoma', *Nature Genetics*, 45(8). doi: 10.1038/ng.2708.

Hamburger, A. W. and Salmon, S. E. (1977) 'Primary bioassay of human tumor stem cells', *Science*, 197(4302). doi: 10.1126/science.560061.

Hanahan, D. and Weinberg, R. A. (2011) 'Hallmarks of cancer: The next generation', *Cell*. doi: 10.1016/j.cell.2011.02.013.

Harada, K. I. *et al.* (2012) 'Expression of epithelial-mesenchymal transition markers in renal cell carcinoma: Impact on prognostic outcomes in patients undergoing radical nephrectomy', *BJU International*, 110(11 C). doi: 10.1111/j.1464-410X.2012.11297.x.

Harrison, R. G. *et al.* (1907) 'Observations of the living developing nerve fiber', *The Anatomical Record*, 1(5). doi: 10.1002/ar.1090010503.

Harrison, R. G. (1910) 'The outgrowth of the nerve fiber as a mode of protoplasmic movement', *Journal of Experimental Zoology*, 9(4). doi: 10.1002/jez.1400090405.

Hartigan, J. A. and Wong, M. A. (1979) 'Algorithm AS 136: A K-Means Clustering Algorithm', *Applied Statistics*, 28(1). doi: 10.2307/2346830.

Hashimshony, T. *et al.* (2012) 'CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification', *Cell Reports*, 2(3). doi: 10.1016/j.celrep.2012.08.003.

Hashimshony, T. *et al.* (2016) 'CEL-Seq2: Sensitive highly-multiplexed single-cell RNA-Seq', *Genome Biology*, 17(1). doi: 10.1186/s13059-016-0938-8.

Havel, J. J., Chowell, D. and Chan, T. A. (2019) 'The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy', *Nature Reviews Cancer*. doi: 10.1038/s41568-019-0116-x.

He, H. and Magi-Galluzzi, C. (2014) 'Epithelial-to-mesenchymal transition in renal neoplasms', *Advances in Anatomic Pathology*. doi: 10.1097/PAP.0000000000000018.

Heidegger, I., Pircher, A. and Pichler, R. (2019) 'Targeting the tumor microenvironment in renal cell cancer biology and therapy', *Frontiers in Oncology*, 9(JUN). doi: 10.3389/fonc.2019.00490.

Hemminki, O. *et al.* (2020) 'Treatment of Advanced Renal Cell Carcinoma: Immunotherapies Have Demonstrated Overall Survival Benefits While Targeted Therapies Have Not', *European Urology Open Science*. doi: 10.1016/j.euro.2020.11.003.

Heppner, G. H. (1984) 'Tumor Heterogeneity', *Cancer Research*, 44(6), pp. 2259–2265. doi: 10.1001/jamanetworkopen.2020.3677.

Hicks, S. C. *et al.* (2018) 'Missing data and technical variability in single-cell RNA-sequencing experiments', *Biostatistics*, 19(4). doi: 10.1093/biostatistics/kxx053.

Hida, K. *et al.* (2018) 'Contribution of tumor endothelial cells in cancer progression', *International Journal of Molecular Sciences*. doi: 10.3390/ijms19051272.

Hidalgo, M. *et al.* (2014) 'Patient-derived Xenograft models: An emerging platform for translational cancer research', *Cancer Discovery*, 4(9). doi: 10.1158/2159-8290.CD-14-0001.

Hidayat, K. *et al.* (2017) 'Blood pressure and kidney cancer risk: Meta-analysis of prospective studies', *Journal of Hypertension*. doi: 10.1097/HJH.0000000000001286.

- Ho, M. Y. *et al.* (2012) ‘TNF- $\alpha$  induces epithelial-mesenchymal transition of renal cell carcinoma cells via a GSK3 $\beta$ -dependent mechanism’, *Molecular Cancer Research*, 10(8). doi: 10.1158/1541-7786.MCR-12-0160.
- Hoffman, L. M. *et al.* (2016) ‘Spatial genomic heterogeneity in diffuse intrinsic pontine and midline high-grade glioma: implications for diagnostic biopsy and targeted therapeutics’, *Acta neuropathologica communications*, 4. doi: 10.1186/s40478-015-0269-0.
- Hong, S. P. *et al.* (2019) ‘Single-cell transcriptomics reveals multi-step adaptations to endocrine therapy’, *Nature Communications*, 10(1). doi: 10.1038/s41467-019-11721-9.
- Hou, Y. *et al.* (2012) ‘Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm’, *Cell*, 148(5). doi: 10.1016/j.cell.2012.02.028.
- Hovestadt, V. *et al.* (2019) ‘Resolving medulloblastoma cellular architecture by single-cell genomics’, *Nature*, 572(7767). doi: 10.1038/s41586-019-1434-6.
- Hsieh, J. J. *et al.* (2017) ‘Renal cell carcinoma 1 2’, *Nature Reviews Disease Primers*.
- Hu, F., Zeng, W. and Liu, X. (2019) ‘A gene signature of survival prediction for kidney renal cell carcinoma by multi-omic data analysis’, *International Journal of Molecular Sciences*, 20(22). doi: 10.3390/ijms20225720.
- Hu, J. *et al.* (2020) ‘Single-Cell Transcriptome Analysis Reveals Intratumoral Heterogeneity in ccRCC, which Results in Different Clinical Outcomes’, *Molecular Therapy*. Elsevier Ltd., 28(7), pp. 1658–1672. doi: 10.1016/j.ymthe.2020.04.023.
- Ilicic, T. *et al.* (2016) ‘Classification of low quality cells from single-cell RNA-seq data’, *Genome Biology*, 17(1). doi: 10.1186/s13059-016-0888-1.
- Iliopoulos, O. *et al.* (1995) ‘Tumour suppression by the human von hippel-lindau gene product’, *Nature Medicine*, 1(8). doi: 10.1038/nm0895-822.
- Iliopoulos, O. *et al.* (1996) ‘Negative regulation of hypoxia-inducible genes by the von Hippel-Lindau protein’, *Proceedings of the National Academy of Sciences of the United States of America*, 93(20). doi: 10.1073/pnas.93.20.10595.
- ‘Invention of the Microscope’ (1858) *Scientific American*, 13(41). doi: 10.1038/scientificamerican06191858-321a.
- Ireland, A. S. *et al.* (2020) ‘MYC Drives Temporal Evolution of Small Cell Lung Cancer Subtypes by Reprogramming Neuroendocrine Fate’, *Cancer Cell*, 38(1). doi: 10.1016/j.ccell.2020.05.001.
- Islam, S. *et al.* (2014) ‘Quantitative single-cell RNA-seq with unique molecular identifiers’, *Nature Methods*, 11(2). doi: 10.1038/nmeth.2772.
- Ito, R. *et al.* (2017) ‘The impact of obesity and adiponectin signaling in patients with renal cell carcinoma: A potential mechanism for the “obesity paradox”’, *PLoS ONE*, 12(2). doi: 10.1371/journal.pone.0171615.
- Ivanov, S. *et al.* (2001) ‘Expression of hypoxia-inducible cell-surface transmembrane carbonic anhydrases in human cancer’, *American Journal of Pathology*, 158(3). doi: 10.1016/S0002-9440(10)64038-2.
- Jeong, B. J. *et al.* (2012) ‘CD44 is associated with tumor recurrence and is an independent poor prognostic factor for patients with localized clear cell renal cell carcinoma after nephrectomy’, *Experimental and Therapeutic Medicine*, 3(5). doi: 10.3892/etm.2012.505.
- Jerby-Arnon, L. *et al.* (2018) ‘A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade’, *Cell*, 175(4). doi: 10.1016/j.cell.2018.09.006.
- Jessa, S. *et al.* (2019) ‘Stalled developmental programs at the root of pediatric brain tumors’, *Nature Genetics*, 51(12). doi: 10.1038/s41588-019-0531-7.
- Ji, A. L. *et al.* (2020) ‘Multimodal Analysis of Composition and Spatial Architecture in Human Squamous Cell Carcinoma’, *Journal of Cleaner Production*. doi: 10.1016/j.cell.2020.05.039.
- Ji, Z. and Ji, H. (2016) ‘TSCAN: Pseudo-time reconstruction and evaluation in single-

cell RNA-seq analysis', *Nucleic Acids Research*, 44(13). doi: 10.1093/nar/gkw430.

Jian, Y. *et al.* (2021) 'Current Advance of Immune Evasion Mechanisms and Emerging Immunotherapies in Renal Cell Carcinoma', *Frontiers in Immunology*. doi: 10.3389/fimmu.2021.639636.

Jiang, K. *et al.* (2019) 'miR-33a inhibits cell growth in renal cancer by downregulation of MDM4 expression', *Molecular Genetics and Genomic Medicine*, 7(8). doi: 10.1002/mgg3.833.

Jiang, X. *et al.* (2017) 'Long-lived pancreatic ductal adenocarcinoma slice cultures enable precise study of the immune microenvironment', *OncImmunity*. doi: 10.1080/2162402X.2017.1333210.

Jin, L. *et al.* (2017) 'MiR-15a-5p acts as an oncogene in renal cell carcinoma', *Molecular Medicine Reports*, 15(3). doi: 10.3892/mmr.2017.6121.

Jing, X. *et al.* (2019) 'Role of hypoxia in cancer therapy by regulating the tumor microenvironment', *Molecular Cancer*. doi: 10.1186/s12943-019-1089-9.

Jordan, N. V. *et al.* (2016) 'HER2 expression identifies dynamic functional states within circulating breast cancer cells', *Nature*, 537(7618). doi: 10.1038/nature19328.

Kaelin, W. G. (2007) 'Von Hippel-Lindau disease', *Annual Review of Pathology*, 2, pp. 145–173. doi: 10.1146/annurev.pathol.2.010506.092049.

Kalbasi, A. and Ribas, A. (2020) 'Tumour-intrinsic resistance to immune checkpoint blockade', *Nature Reviews Immunology*. doi: 10.1038/s41577-019-0218-4.

Kameya, T., Ohsawa, N. and Nomura, T. (1976) 'Human Gastric Choriocarcinoma Serially Transplanted in Nude Mice', *Journal of the National Cancer Institute*, 56(2). doi: 10.1093/jnci/56.2.325.

Kamli, H., Li, L. and Gobe, G. C. (2019) 'Limitations to the therapeutic potential of tyrosine kinase inhibitors and alternative therapies for kidney cancer', *Ochsner Journal*, 19(2). doi: 10.31486/toj.18.0015.

Kapitsinou, P. P. and Haase, V. H. (2008) 'The VHL tumor suppressor and HIF: Insights from genetic studies in mice', *Cell Death and Differentiation*. doi: 10.1038/sj.cdd.4402313.

Karakiewicz, P. I. *et al.* (2007) 'Multi-institutional validation of a new renal cancer-specific survival nomogram', *Journal of Clinical Oncology*, 25(11). doi: 10.1200/JCO.2006.06.1218.

Karekla, E. *et al.* (2017) 'Primary Tumor Responses To Anticancer Therapy', *Cancer research*. doi: 10.1158/0008-5472.CAN-16-1121.

Karthaus, W. R. *et al.* (2014) 'Identification of multipotent luminal progenitor cells in human prostate organoid cultures', *Cell*, 159(1). doi: 10.1016/j.cell.2014.08.017.

Keith, B., Johnson, R. S. and Simon, M. C. (2012) 'HIF1  $\alpha$  and HIF2  $\alpha$ : sibling rivalry in hypoxic tumour growth and progression', *Nature Reviews Cancer*. doi: 10.1038/nrc3183.

Kenerson, H. L. *et al.* (2020) 'Tumor slice culture as a biologic surrogate of human cancer', *Annals of Translational Medicine*, 8(4). doi: 10.21037/atm.2019.12.88.

Khan, M. I. *et al.* (2016) 'Gene set enrichment analysis and ingenuity pathway analysis of metastatic clear cell renal cell carcinoma cell line', *American Journal of Physiology - Renal Physiology*, 311(2). doi: 10.1152/ajprenal.00138.2016.

Khawam, K. *et al.* (2009) 'Human renal cancer cells express a novel membrane-bound interleukin-15 that induces, in response to the soluble interleukin-15 receptor  $\alpha$  chain, epithelial-to-mesenchymal transition', *Cancer Research*, 69(4). doi: 10.1158/0008-5472.CAN-08-3198.

Kim, H. L. *et al.* (2003) 'Paraneoplastic signs and symptoms of renal cell carcinoma: Implications for prognosis', *Journal of Urology*, 170(5). doi: 10.1097/01.ju.0000092764.81308.6a.

Kim, H. L. *et al.* (2004) 'Cachexia-like, symptoms predict a worse prognosis in localized

T1 renal cell carcinoma', *Journal of Urology*, 171(5). doi: 10.1097/01.ju.0000121440.82581.d3.

Kim, K.-T. *et al.* (2016) 'Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma.', *Genome biology*, 17(1), p. 80. doi: 10.1186/s13059-016-0945-9.

Kim, T. M. *et al.* (2015) 'Subclonal genomic architectures of primary and metastatic colorectal cancer based on intratumoral genetic heterogeneity', *Clinical Cancer Research*, 21(19). doi: 10.1158/1078-0432.CCR-14-2413.

Kim, Y. S. *et al.* (2014) 'Role of the epithelial-mesenchymal transition and its effects on embryonic stem cells', *Experimental and Molecular Medicine*. doi: 10.1038/emm.2014.44.

Kimura, M. (1983) 'Rare variant alleles in the light of the neutral theory', *Molecular Biology and Evolution*, 1(1). doi: 10.1093/oxfordjournals.molbev.a040305.

Kivioja, T. *et al.* (2012) 'Counting absolute numbers of molecules using unique molecular identifiers', *Nature Methods*, 9(1). doi: 10.1038/nmeth.1778.

Klinghammer, K., Walther, W. and Hoffmann, J. (2017) 'Choosing wisely – Preclinical test models in the era of precision medicine', *Cancer Treatment Reviews*. doi: 10.1016/j.ctrv.2017.02.009.

Ko, J. J. *et al.* (2015) 'The international metastatic renal cell carcinoma database consortium model as a prognostic tool in patients with metastatic renal cell carcinoma previously treated with first-line targeted therapy: A population-based study', *The Lancet Oncology*, 16(3). doi: 10.1016/S1470-2045(14)71222-7.

Koboldt, D. C. *et al.* (2012) 'Comprehensive molecular portraits of human breast tumours', *Nature*, 490(7418). doi: 10.1038/nature11412.

Koerfer, J. *et al.* (2016) 'Organotypic slice cultures of human gastric and esophagogastric junction cancer', *Cancer Medicine*, 5(7). doi: 10.1002/cam4.720.

Kondo, K. *et al.* (2002) 'Inhibition of HIF is necessary for tumor suppression by the von Hippel-Lindau protein', *Cancer Cell*, 1(3). doi: 10.1016/S1535-6108(02)00043-0.

Kondo, K. I. *et al.* (2001) 'PTEN/MMAC1/TEP1 mutations in human primary renal-cell carcinomas and renal carcinoma cell lines', *International Journal of Cancer*, 91(2). doi: 10.1002/1097-0215(200002)9999:9999<::AID-IJC1034>3.0.CO;2-S.

Korsunsky, I. *et al.* (2019) 'Fast, sensitive and accurate integration of single-cell data with Harmony', *Nature Methods*, 16(12). doi: 10.1038/s41592-019-0619-0.

Kraan, J. *et al.* (2012) 'Clinical value of circulating endothelial cell detection in oncology', *Drug Discovery Today*. doi: 10.1016/j.drudis.2012.01.011.

Kreso, A. and Dick, J. E. (2014) 'Evolution of the cancer stem cell model', *Cell Stem Cell*. doi: 10.1016/j.stem.2014.02.006.

Krishna, C. *et al.* (2021) 'Single-cell sequencing links multiregional immune landscapes and tissue-resident T cells in ccRCC to tumor topology and therapy efficacy', *Cancer Cell*, 39(5). doi: 10.1016/j.ccell.2021.03.007.

Kumar, A. *et al.* (2018) 'Renal Cell Carcinoma: Molecular Aspects', *Indian Journal of Clinical Biochemistry*. doi: 10.1007/s12291-017-0713-y.

Kurtenbach, S. *et al.* (2021) 'Uphyloplot2: visualizing phylogenetic trees from single-cell RNA-seq data', *BMC Genomics*, 22(1). doi: 10.1186/s12864-021-07739-3.

Lacraz, G. P. A. *et al.* (2017) 'Tomo-Seq Identifies SOX9 as a Key Regulator of Cardiac Fibrosis during Ischemic Injury', *Circulation*, 136(15). doi: 10.1161/CIRCULATIONAHA.117.027832.

LaFramboise, T. (2009) 'Single nucleotide polymorphism arrays: A decade of biological, computational and technological advances', *Nucleic Acids Research*. doi: 10.1093/nar/gkp552.

Lähnemann, D. *et al.* (2020) 'Eleven grand challenges in single-cell data science',

*Genome Biology*. doi: 10.1186/s13059-020-1926-6.

Lai, Y. *et al.* (2017) 'Current status and perspectives of patient-derived xenograft models in cancer research', *Journal of Hematology and Oncology*. doi: 10.1186/s13045-017-0470-7.

Lai, Yulin *et al.* (2018) 'MicroRNA-181a-5p functions as an oncogene in Renal cell carcinoma', *Molecular Medicine Reports*, 17(6). doi: 10.3892/mmr.2018.8899.

Lambrechts, D. *et al.* (2018) 'Phenotype molding of stromal cells in the lung tumor microenvironment', *Nature Medicine*, 24(8). doi: 10.1038/s41591-018-0096-5.

Lancaster, M. A. *et al.* (2013) 'Cerebral organoids model human brain development and microcephaly', *Nature*, 501(7467). doi: 10.1038/nature12517.

Landolt, L. *et al.* (2017) 'Clear Cell Renal Cell Carcinoma is linked to Epithelial-to-Mesenchymal Transition and to Fibrosis', *Physiological Reports*, 5(11). doi: 10.14814/phy2.13305.

Lapidot, T. *et al.* (1994) 'A cell initiating human acute myeloid leukaemia after transplantation into SCID mice', *Nature*, 367(6464). doi: 10.1038/367645a0.

Lareau, C. A. *et al.* (2020) 'Inference and effects of barcode multiplets in droplet-based single-cell assays', *Nature Communications*, 11(1). doi: 10.1038/s41467-020-14667-5.

Larsson, L., Frisén, J. and Lundeberg, J. (2021) 'Spatially resolved transcriptomics adds a new dimension to genomics', *Nature Methods*. doi: 10.1038/s41592-020-01038-7.

Lathia, J. D. *et al.* (2015) 'Cancer stem cells in glioblastoma', *Genes and Development*. doi: 10.1101/gad.261982.115.

Laver, T. *et al.* (2015) 'Assessing the performance of the Oxford Nanopore Technologies MinION', *Biomolecular Detection and Quantification*, 3. doi: 10.1016/j.bdq.2015.02.001.

Lawson, D. A. *et al.* (2015) 'Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells', *Nature*, 526(7571). doi: 10.1038/nature15260.

Lê, S., Josse, J. and Husson, F. (2008) 'FactoMineR: An R package for multivariate analysis', *Journal of Statistical Software*, 25(1). doi: 10.18637/jss.v025.i01.

Lee, C.-H. *et al.* (2014) 'Src pathway activation in RCC and the correlation with grade and survival and the development of a rational new target in RCC.', *Journal of Clinical Oncology*, 32(4\_suppl). doi: 10.1200/jco.2014.32.4\_suppl.453.

Lee, C. H. *et al.* (2016) 'Bevacizumab Monotherapy as Salvage Therapy for Advanced Clear Cell Renal Cell Carcinoma Pretreated with Targeted Drugs', *Clinical Genitourinary Cancer*, 14(1). doi: 10.1016/j.clgc.2015.07.010.

Lee, C. H. and Motzer, R. J. (2017) 'Kidney cancer in 2016: The evolution of anti-Angiogenic therapy for kidney cancer', *Nature Reviews Nephrology*. doi: 10.1038/nrneph.2016.194.

Lee, H. J. *et al.* (2017) 'Combination of immunohistochemistry, FISH and RT-PCR shows high incidence of Xp11 translocation RCC: Comparison of three different diagnostic methods', *Oncotarget*, 8(19). doi: 10.18632/oncotarget.16481.

Lee, Z. *et al.* (2020) 'Local Recurrence Following Resection of Intermediate-High Risk Nonmetastatic Renal Cell Carcinoma: An Anatomical Classification and Analysis of the ASSURE (ECOG-ACRIN E2805) Adjuvant Trial', *The Journal of urology*, 203(4). doi: 10.1097/JU.0000000000000588.

Leibovich, B. C. *et al.* (2005) 'A scoring algorithm to predict survival for patients with metastatic clear cell renal cell carcinoma: A stratification tool for prospective clinical trials', *Journal of Urology*, 174(5). doi: 10.1097/01.ju.0000177487.64651.3a.

Lendahl, U., Muhl, L. and Betsholtz, C. (2022) 'Identification, discrimination and heterogeneity of fibroblasts', *Nature Communications*, 13(1), pp. 1–14. doi: 10.1038/s41467-022-30633-9.

- Lewandowska, A. M. *et al.* (2019) 'Environmental risk factors for cancer - review paper', *Annals of Agricultural and Environmental Medicine*. doi: 10.26444/aaem/94299.
- Li, B. *et al.* (2017) 'Tumor-initiating cells contribute to radiation resistance in primary human renal clear cell carcinomas by activating the dna damage checkpoint response', *Oncology Letters*, 14(3). doi: 10.3892/ol.2017.6504.
- Li, Y. *et al.* (2012) 'Single-cell sequencing analysis characterizes common and cell-lineage-specific mutations in a muscle-invasive bladder cancer', *GigaScience*, 61(1). doi: 10.1186/2047-217X-1-12.
- Liao, L. M. *et al.* (2017) 'Circulating levels of obesity-related markers and risk of renal cell carcinoma in the PLCO cancer screening trial', *Cancer Causes and Control*, 28(7). doi: 10.1007/s10552-017-0901-3.
- Libretti, S. and Aeddula, N. R. (2020) *Embryology, Genitourinary, StatPearls*.
- Lin, H. and Zelterman, D. (2002) 'Modeling Survival Data: Extending the Cox Model', *Technometrics*, 44(1). doi: 10.1198/tech.2002.s656.
- Linder, D. and Gartler, S. M. (1965) 'Glucose-6-phosphate dehydrogenase mosaicism: Utilization as a cell marker in the study of leiomyomas', *Science*, 150(3692). doi: 10.1126/science.150.3692.67.
- Linehan, W. M., Srinivasan, R. and Schmidt, L. S. (2010) 'The genetic basis of kidney cancer: A metabolic disease', *Nature Reviews Urology*. doi: 10.1038/nrurol.2010.47.
- Ling, S. *et al.* (2015) 'Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution', *Proceedings of the National Academy of Sciences of the United States of America*, 112(47). doi: 10.1073/pnas.1519556112.
- Litovchenko, M. *et al.* (2021) 'Extensive tissue-specific expression variation and novel regulators underlying circadian behavior', *Science Advances*, 7(5). doi: 10.1126/SCIADV.ABC3781.
- Liu, C. *et al.* (2019) 'Comprehensive molecular characterization of clear cell renal cell carcinoma with caval tumour thrombus', *European Urology Supplements*, 18(1). doi: 10.1016/s1569-9056(19)31522-2.
- Liu, J. *et al.* (2020) 'Inter-embryo gene expression variability recapitulates the hourglass pattern of evo-devo', *BMC Biology*, 18(1). doi: 10.1186/s12915-020-00842-z.
- Liu, K. *et al.* (2021) 'Single-cell analysis reveals metastatic cell heterogeneity in clear cell renal cell carcinoma', *Journal of Cellular and Molecular Medicine*, 25(9). doi: 10.1111/jcmm.16479.
- Liu, N. W. *et al.* (2013) 'Impact of ischemia and procurement conditions on gene expression in renal cell carcinoma', *Clinical Cancer Research*, 19(1). doi: 10.1158/1078-0432.CCR-12-2606.
- Ljungberg, B. *et al.* (2019) 'European Association of Urology Guidelines on Renal Cell Carcinoma: The 2019 Update', *European Urology*, 75(5). doi: 10.1016/j.eururo.2019.02.011.
- Ljungberg, B. *et al.* (2020) 'EAU Guidelines. Edn. presented at the EAU Annual Congress Amsterdam 2020.', *European Urology*, 67(5).
- Lobo, I. (2008) 'Chromosome Abnormalities and Cancer Cytogenetics', *Nature Education*, 1(1).
- Lomakin, A. *et al.* (2021) 'Spatial genomics maps the structure, character and evolution of cancer clones', *bioRxiv*.
- López, J. I. and Angulo, J. C. (2018) 'Pathological Bases and Clinical Impact of Intratumor Heterogeneity in Clear Cell Renal Cell Carcinoma', *Current Urology Reports*. doi: 10.1007/s11934-018-0754-7.
- Lorusso, G. and Rüegg, C. (2008) 'The tumor microenvironment and its contribution to tumor evolution toward metastasis', *Histochemistry and Cell Biology*. doi: 10.1007/s00418-008-0530-8.



- Lotan, Y. *et al.* (2016) ‘Renal-cell carcinoma risk estimates based on participants in the prostate, lung, colorectal, and ovarian cancer screening trial and national lung screening trial’, *Urologic Oncology: Seminars and Original Investigations*, 34(4). doi: 10.1016/j.urolonc.2015.10.011.
- Lovisolò, J. A. J. *et al.* (2006) ‘Gene expression profiling of renal cell carcinoma: A DNA microarray analysis’, *BJU International*, 98(1). doi: 10.1111/j.1464-410X.2006.06209.x.
- Lovly, C. M. and Shaw, A. T. (2014) ‘Molecular pathways: Resistance to kinase inhibitors and implications for therapeutic strategies’, *Clinical Cancer Research*, 20(9). doi: 10.1158/1078-0432.CCR-13-1610.
- Lowe, R. *et al.* (2017) ‘Transcriptomics technologies’, *PLoS Computational Biology*, 13(5). doi: 10.1371/journal.pcbi.1005457.
- Lu, H., Giordano, F. and Ning, Z. (2016) ‘Oxford Nanopore MinION Sequencing and Genome Assembly’, *Genomics, Proteomics and Bioinformatics*. doi: 10.1016/j.gpb.2016.05.004.
- Luecken, M. D. and Theis, F. J. (2019) ‘Current best practices in single-cell RNA-seq analysis: a tutorial’, *Molecular Systems Biology*, 15(6). doi: 10.15252/msb.20188746.
- Lummertz Da Rocha, E. *et al.* (2018) ‘Reconstruction of complex single-cell trajectories using CellRouter’, *Nature Communications*, 9(1). doi: 10.1038/s41467-018-03214-y.
- Lun, A. T. L. *et al.* (2017) ‘Assessing the reliability of spike-in normalization for analyses of single-cell RNA sequencing data’, *Genome Research*, 27(11). doi: 10.1101/gr.222877.117.
- Lun, A. T. L. *et al.* (2019) ‘EmptyDrops: Distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data’, *Genome Biology*, 20(1). doi: 10.1186/s13059-019-1662-y.
- Lun, A. T. L., Bach, K. and Marioni, J. C. (2016) ‘Pooling across cells to normalize single-cell RNA sequencing data with many zero counts’, *Genome Biology*, 17(1). doi: 10.1186/s13059-016-0947-7.
- Van Der Maaten, L. and Hinton, G. (2008) ‘Visualizing data using t-SNE’, *Journal of Machine Learning Research*, 9.
- Macaulay, I. C. *et al.* (2015) ‘G&T-seq: Parallel sequencing of single-cell genomes and transcriptomes’, *Nature Methods*, 12(6). doi: 10.1038/nmeth.3370.
- MacLeod, L. C. *et al.* (2013) ‘Risk factors for renal cell carcinoma in the VITAL study’, *Journal of Urology*, 190(5). doi: 10.1016/j.juro.2013.04.130.
- Macosko, E. Z. *et al.* (2015) ‘Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets’, *Cell*, 161(5). doi: 10.1016/j.cell.2015.05.002.
- Maeda, H. and Khatami, M. (2018) ‘Analyses of repeated failures in cancer therapy for solid tumors: poor tumor-selective drug delivery, low therapeutic efficacy and unsustainable costs’, *Clinical and Translational Medicine*, 7(1). doi: 10.1186/s40169-018-0185-6.
- Mahadevan, V. (2019) ‘Anatomy of the kidney and ureter’, *Surgery (United Kingdom)*. doi: 10.1016/j.mpsur.2019.04.005.
- Majmundar, A. J., Wong, W. J. and Simon, M. C. (2010) ‘Hypoxia-Inducible Factors and the Response to Hypoxic Stress’, *Molecular Cell*. doi: 10.1016/j.molcel.2010.09.022.
- Mandal, A., Littler, Y. and Libertiny, G. (2012) ‘Asymptomatic renal cell carcinoma with metastasis to the skin and duodenum: A case report and review of the literature’, *BMJ Case Reports*. doi: 10.1136/bcr.02.2012.5764.
- La Manno, G. *et al.* (2018) ‘RNA velocity of single cells’, *Nature*, 560(7719). doi: 10.1038/s41586-018-0414-6.
- Marinov, G. K. *et al.* (2014) ‘From single-cell to cell-pool transcriptomes: Stochasticity in gene expression and RNA splicing’, *Genome Research*, 24(3). doi: 10.1101/gr.161034.113.
- Martin, B. K. *et al.* (2021) *An optimized protocol for single cell transcriptional profiling*

by *combinatorial indexing*. Available at:  
<https://arxiv.org/ftp/arxiv/papers/2110/2110.15400.pdf>.

Marusyk, A., Janiszewska, M. and Polyak, K. (2020) 'Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance', *Cancer Cell*. doi: 10.1016/j.ccell.2020.03.007.

Masson, N. and Ratcliffe, P. J. (2014) 'Hypoxia signaling pathways in cancer metabolism: the importance of co-selecting interconnected physiological pathways', *Cancer & Metabolism*, 2(1). doi: 10.1186/2049-3002-2-3.

Maxam, A. M. and Gilbert, W. (1977) 'A new method for sequencing DNA', *Proceedings of the National Academy of Sciences of the United States of America*, 74(2). doi: 10.1073/pnas.74.2.560.

Maynard, A. *et al.* (2020) 'Therapy-Induced Evolution of Human Lung Cancer Revealed by Single-Cell RNA Sequencing', *Cell*, 182(5). doi: 10.1016/j.cell.2020.07.017.

Mazzoleni, G., Di Lorenzo, D. and Steimberg, N. (2009) 'Modelling tissues in 3D: The next future of pharmaco-toxicology and food research?', *Genes and Nutrition*. doi: 10.1007/s12263-008-0107-0.

Mazzucotelli, V. *et al.* (2017) 'De novo cancer in patients on dialysis and after renal transplantation: north-western Italy, 1997–2012', *Journal of Nephrology*, 30(6). doi: 10.1007/s40620-017-0385-y.

McCarthy, D. J. *et al.* (2017) 'Scater: Pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R', *Bioinformatics*, 33(8), pp. 1179–1186. doi: 10.1093/bioinformatics/btw777.

McGinnis, C. S., Murrow, L. M. and Gartner, Z. J. (2019) 'DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors', *Cell Systems*, 8(4). doi: 10.1016/j.cels.2019.03.003.

McGranahan, N. *et al.* (2016) 'Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade', *Science*, 351(6280). doi: 10.1126/science.aaf1490.

McGranahan, N. and Swanton, C. (2015) 'Biological and Therapeutic Impact of Intratumor Heterogeneity in Cancer Evolution', *Cancer Cell*, 27(1), pp. 15–26. doi: 10.1016/j.ccell.2014.12.001.

McInnes, L. *et al.* (2018) 'UMAP: Uniform Manifold Approximation and Projection', *Journal of Open Source Software*, 3(29). doi: 10.21105/joss.00861.

Meacham, C. E. and Morrison, S. J. (2013) 'Tumour heterogeneity and cancer cell plasticity', *Nature*. doi: 10.1038/nature12624.

Meijer, T. G., Jager, A. and Gent, D. C. Van (2017) 'Review-Ex vivo tumor culture systems for functional drug testing and therapy response prediction-Meijer-2017', 3. doi: 10.4155/fsoa-2017-0003.

Menko, F. H. and Maher, E. R. (2016) 'Diagnosis and management of hereditary renal cell cancer', *Recent Results in Cancer Research*, 205. doi: 10.1007/978-3-319-29998-3\_6.

Van De Merbel, A. F. *et al.* (2018) 'An ex vivo tissue culture model for the assessment of individualized drug responses in prostate and bladder cancer', *Frontiers in Oncology*, 8(OCT). doi: 10.3389/fonc.2018.00400.

van de Merbel, A. F., van der Horst, G. and van der Pluijm, G. (2021) 'Patient-derived tumour models for personalized therapeutics in urological cancers', *Nature Reviews Urology*. doi: 10.1038/s41585-020-00389-2.

Mercer, T. R., Dinger, M. E. and Mattick, J. S. (2009) 'Long non-coding RNAs: Insights into functions', *Nature Reviews Genetics*. doi: 10.1038/nrg2521.

Merlo, L. M. F. *et al.* (2006) 'Cancer as an evolutionary and ecological process', *Nature Reviews Cancer*. doi: 10.1038/nrc2013.

Milo, I. *et al.* (2018) 'The immune system profoundly restricts intratumor genetic heterogeneity', *Science Immunology*, 3(29). doi: 10.1126/sciimmunol.aat1435.

- Misra, S. *et al.* (2019) 'Ex vivo organotypic culture system of precision-cut slices of human pancreatic ductal adenocarcinoma', *Scientific Reports*, 9(1). doi: 10.1038/s41598-019-38603-w.
- Miyamoto, D. T. *et al.* (2015) 'RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance', *Science*, 349(6254). doi: 10.1126/science.aab0917.
- Moch, H. *et al.* (2016) 'The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs—Part A: Renal, Penile, and Testicular Tumours', *European Urology*, 70(1). doi: 10.1016/j.eururo.2016.02.029.
- Mockler, T. C. and Ecker, J. R. (2005) 'Applications of DNA tiling arrays for whole-genome analysis', *Genomics*. doi: 10.1016/j.ygeno.2004.10.005.
- Molina, A. M. *et al.* (2014) 'Sunitinib objective response in metastatic renal cell carcinoma: Analysis of 1059 patients treated on clinical trials', *European Journal of Cancer*. doi: 10.1016/j.ejca.2013.08.021.
- Moncada, R. *et al.* (2020) 'Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas', *Nature Biotechnology*, 38(3). doi: 10.1038/s41587-019-0392-8.
- Moor, A. E. *et al.* (2018) 'Spatial Reconstruction of Single Enterocytes Uncovers Broad Zonation along the Intestinal Villus Axis', *Cell*, 175(4). doi: 10.1016/j.cell.2018.08.063.
- Moore, L. E. *et al.* (2012) 'Genomic copy number alterations in clear cell renal carcinoma: Associations with case characteristics and mechanisms of VHL gene inactivation', *Oncogenesis*, 1(6). doi: 10.1038/oncsis.2012.14.
- Morgan, K. M. *et al.* (2017) 'Patient-derived xenograft models of non-small cell lung cancer and their potential utility in personalized medicine', *Frontiers in Oncology*, 7(JAN). doi: 10.3389/fonc.2017.00002.
- Mortazavi, A. *et al.* (2008) 'Mapping and quantifying mammalian transcriptomes by RNA-Seq', *Nature Methods*, 5(7). doi: 10.1038/nmeth.1226.
- Motzer, R. J. *et al.* (2007) 'Sunitinib versus interferon alfa in metastatic renal-cell carcinoma', *New England Journal of Medicine*. doi: 10.1056/NEJMoa065044.
- Motzer, R. J. *et al.* (2015) 'Kidney Cancer, Version 3.2015', *Journal of the National Comprehensive Cancer Network*, 13(2). doi: 10.6004/jnccn.2015.0022.
- Muglia, V. F. and Prando, A. (2015) 'Renal cell carcinoma: histological classification and correlation with imaging findings', *Radiologia Brasileira*, 48(3). doi: 10.1590/0100-3984.2013.1927.
- Muzny, D. M. *et al.* (2012) 'Comprehensive molecular characterization of human colon and rectal cancer', *Nature*, 487(7407). doi: 10.1038/nature11252.
- Mylka, V. *et al.* (2022) 'Comparative analysis of antibody- and lipid-based multiplexing methods for single-cell RNA-seq', *Genome Biology*, 23(1). doi: 10.1186/s13059-022-02628-8.
- N. Dudley, PhD, W., Wickham, PhD, RN, AOCN, R. and Coombs, MS, N. (2016) 'An Introduction to Survival Statistics: Kaplan-Meier Analysis', *Journal of the Advanced Practitioner in Oncology*, 7(1). doi: 10.6004/jadpro.2016.7.1.8.
- Nam, A. S. *et al.* (2019) 'Somatic mutations and cell identity linked by Genotyping of Transcriptomes', *Nature*, 571(7765). doi: 10.1038/s41586-019-1367-0.
- Nam, A. S., Chaligne, R. and Landau, D. A. (2021) 'Integrating genetic and non-genetic determinants of cancer evolution by single-cell multi-omics', *Nature Reviews Genetics*. doi: 10.1038/s41576-020-0265-5.
- Navin, N. *et al.* (2011) 'Tumour evolution inferred by single-cell sequencing', *Nature*, 472(7341). doi: 10.1038/nature09807.
- Neal, J. T. *et al.* (2018) 'Organoid Modeling of the Tumor Immune Microenvironment', *Cell*, 175(7). doi: 10.1016/j.cell.2018.11.021.

- Neftel, C. *et al.* (2019) ‘An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma’, *Cell*, 178(4). doi: 10.1016/j.cell.2019.06.024.
- Nelson, N. J. (2001) ‘Microarrays have arrived: Gene expression tool matures’, *Journal of the National Cancer Institute*. doi: 10.1093/jnci/93.7.492.
- Nerich, V. *et al.* (2014) ‘Clinical impact of targeted therapies in patients with metastatic clear-cell renal cell carcinoma’, *OncoTargets and Therapy*, 7. doi: 10.2147/OTT.S56370.
- Neuzillet, Y. *et al.* (2011) ‘Renal cell carcinoma (RCC) in patients with end-stage renal disease exhibits many favourable clinical, pathologic, and outcome features compared with RCC in the general population’, *European Urology*, 60(2). doi: 10.1016/j.eururo.2011.02.035.
- Nichterwitz, S. *et al.* (2016) ‘Laser capture microscopy coupled with Smart-seq2 for precise spatial transcriptomic profiling’, *Nature Communications*, 7. doi: 10.1038/ncomms12139.
- Nichterwitz, S. *et al.* (2018) ‘LCM-seq: A method for spatial transcriptomic profiling using laser capture microdissection coupled with PolyA-based RNA sequencing’, in *Methods in Molecular Biology*. doi: 10.1007/978-1-4939-7213-5\_6.
- Nik-Zainal, S. *et al.* (2012) ‘The life history of 21 breast cancers’, *Cell*, 149(5). doi: 10.1016/j.cell.2012.04.023.
- Nowell, P. C. (1976) ‘The clonal evolution of tumor cell populations’, *Science*, 194(4260). doi: 10.1126/science.959840.
- Nowell, P. C. (2007) ‘Discovery of the Philadelphia chromosome: A personal perspective’, *Journal of Clinical Investigation*. doi: 10.1172/JCI31771.
- Öhlund, D., Elyada, E. and Tuveson, D. (2014) ‘Fibroblast heterogeneity in the cancer wound’, *Journal of Experimental Medicine*. doi: 10.1084/jem.20140692.
- Okegawa, T. *et al.* (2017) ‘Intratumor Heterogeneity in Primary Kidney Cancer Revealed by Metabolic Profiling of Multiple Spatially Separated Samples within Tumors’, *EBioMedicine*, 19. doi: 10.1016/j.ebiom.2017.04.009.
- Ortiz, C. *et al.* (2020) ‘Molecular atlas of the adult mouse brain’, *Science Advances*, 6(26). doi: 10.1126/sciadv.abb3446.
- Pampaloni, F., Reynaud, E. G. and Stelzer, E. H. K. (2007) ‘The third dimension bridges the gap between cell culture and live tissue’, *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm2236.
- Pan, Q. *et al.* (2020) ‘The immune infiltration in clear cell renal cell carcinoma and their clinical implications: A study based on TCGA and GEO databases’, *Journal of Cancer*, 11(11). doi: 10.7150/jca.37285.
- Papatheodorou, I. *et al.* (2020) ‘Expression Atlas update: From tissues to single cells’, *Nucleic Acids Research*, 48(D1). doi: 10.1093/nar/gkz947.
- Parasramka, M. *et al.* (2016) ‘Validation of Gene Expression Signatures to Identify Low-risk Clear-cell Renal Cell Carcinoma Patients at Higher Risk for Disease-related Death’, *European Urology Focus*, 2(6). doi: 10.1016/j.euf.2016.03.008.
- Parekh, S. *et al.* (2016) ‘The impact of amplification on differential expression analyses by RNA-seq’, *Scientific Reports*, 6. doi: 10.1038/srep25533.
- Park, H. (2017) ‘Hypoxia suffocates histone demethylases to change gene expression: A metabolic control of histone methylation’, *BMB Reports*, 50(11). doi: 10.5483/BMBRep.2017.50.11.200.
- Patard, J. J. *et al.* (2004) ‘Multi-institutional validation of a symptom based classification for renal cell carcinoma’, *Journal of Urology*, 172(3). doi: 10.1097/01.ju.0000135837.64840.55.
- Patel, A. *et al.* (2020) ‘Neutrophil-to-Lymphocyte Ratio as a Prognostic Factor of Disease-free Survival in Postnephrectomy High-risk Locoregional Renal Cell Carcinoma: Analysis of the S-TRAC Trial’, *Clinical Cancer Research*, 26(18). doi: 10.1158/1078-

0432.CCR-20-0704.

Patel, A. P. *et al.* (2014) ‘Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma’, *Science*, 344(6190), pp. 1396–1401. doi: 10.1126/science.1254257.

Paul, J. (1980) ‘Methods in enzymology, Vol. LVIII, cell culture’, *Trends in Biochemical Sciences*, 5(4). doi: 10.1016/0968-0004(80)90267-4.

Pavía-Jiménez, A., Tcheuyap, V. T. and Brugarolas, J. (2014) ‘Establishing a human renal cell carcinoma tumorgraft platform for preclinical drug testing’, *Nature Protocols*, 9(8). doi: 10.1038/nprot.2014.108.

Peired, A. J. *et al.* (2021) ‘Sex and gender differences in kidney cancer: Clinical and experimental evidence’, *Cancers*. doi: 10.3390/cancers13184588.

Peña-Llopis, S. *et al.* (2012) ‘BAP1 loss defines a new class of renal cell carcinoma’, *Nature Genetics*, 44(7). doi: 10.1038/ng.2323.

Peres-Neto, P. R., Jackson, D. A. and Somers, K. M. (2005) ‘How many principal components? stopping rules for determining the number of non-trivial axes revisited’, *Computational Statistics and Data Analysis*, 49(4). doi: 10.1016/j.csda.2004.06.015.

Perkel, J. M. (2021) ‘Single-cell analysis enters the multiomics age’, *Nature*. doi: 10.1038/d41586-021-01994-w.

Petillo, D. *et al.* (2009) ‘MicroRNA profiling of human kidney cancer subtypes’, *International Journal of Oncology*, 35(1). doi: 10.3892/ijo\_00000318.

Peto, R. *et al.* (1977) ‘Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples’, *British Journal of Cancer*, 35(1). doi: 10.1038/bjc.1977.1.

Phillippy, A. M. (2017) ‘New advances in sequence assembly’, *Genome Research*. doi: 10.1101/gr.223057.117.

Philpott, M. *et al.* (2021) ‘Nanopore sequencing of single-cell transcriptomes with scCOLOR-seq’, *Nature Biotechnology*, 39(12). doi: 10.1038/s41587-021-00965-w.

Del Piccolo, N. *et al.* (2021) ‘Tumor-on-chip modeling of organ-specific cancer and metastasis’, *Advanced Drug Delivery Reviews*. doi: 10.1016/j.addr.2021.05.008.

Picelli, S. *et al.* (2013) ‘Smart-seq2 for sensitive full-length transcriptome profiling in single cells’, *Nature Methods*, 10(11). doi: 10.1038/nmeth.2639.

Picelli, S. (2017) ‘Single-cell RNA-sequencing: The future of genome biology is now’, *RNA Biology*, 14(5). doi: 10.1080/15476286.2016.1201618.

Pircher, A. *et al.* (2011) ‘Biomarkers in tumor angiogenesis and anti-angiogenic therapy’, *International Journal of Molecular Sciences*. doi: 10.3390/ijms12107077.

Pircher, A. *et al.* (2017) ‘Synergies of targeting tumor angiogenesis and immune checkpoints in non-small cell lung cancer and renal cell cancer: From basic concepts to clinical reality’, *International Journal of Molecular Sciences*. doi: 10.3390/ijms18112291.

Powles, T. *et al.* (2017) ‘Updated European Association of Urology Guidelines Recommendations for the Treatment of First-line Metastatic Clear Cell Renal Cancer.’, *European urology*, 73(3), pp. 311–315. doi: 10.1016/j.eururo.2017.11.016.

Powley, I. R. *et al.* (2020) ‘Patient-derived explants (PDEs) as a powerful preclinical platform for anti-cancer drug and biomarker discovery’, *British Journal of Cancer*. doi: 10.1038/s41416-019-0672-6.

Pozhitkov, A. E., Tautz, D. and Noble, P. A. (2007) ‘Oligonucleotide microarrays: Widely applied - Poorly understood’, *Briefings in Functional Genomics and Proteomics*, 6(2). doi: 10.1093/bfgp/elm014.

Prasetyanti, P. R. and Medema, J. P. (2017) ‘Intra-tumor heterogeneity from a cancer stem cell perspective’, *Molecular Cancer*. doi: 10.1186/s12943-017-0600-4.

Prochazka, M. *et al.* (1992) ‘The nonobese diabetic scid mouse: Model for spontaneous thymomagenesis associated with immunodeficiency’, *Proceedings of the National Academy of*

*Sciences of the United States of America*, 89(8). doi: 10.1073/pnas.89.8.3290.

Puram, S. V. *et al.* (2017) 'Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer', *Cell*, 171(7). doi: 10.1016/j.cell.2017.10.044.

Putney, S. D., Herlihy, W. C. and Schimmel, P. (1983) 'A new troponin T and cDNA clones for 13 different muscle proteins, found by shotgun sequencing [21]', *Nature*. doi: 10.1038/302718a0.

Qiu, X. *et al.* (2017) 'Single-cell mRNA quantification and differential analysis with Census', *Nature Methods*, (December 2016). doi: 10.1038/nmeth.4150.

Quan, J. *et al.* (2019) 'MiR-23a-3p acts as an oncogene and potential prognostic biomarker by targeting PNR2 in RCC', *Biomedicine and Pharmacotherapy*, 110. doi: 10.1016/j.biopha.2018.11.065.

Rabjerg, M. (2017) 'Identification and validation of novel prognostic markers in renal cell carcinoma', *Danish Medical Journal*, 64(10).

Rabjerg, M. *et al.* (2017) 'Nuclear localization of the CK2 $\alpha$ -subunit correlates with poor prognosis in clear cell renal cell carcinoma', *Oncotarget*, 8(1). doi: 10.18632/oncotarget.13693.

Rachel Schor-Bardach, David C Alsop, Ivan Pedrosa, Stephanie A Solazzo, Xiaoen Wang, Robert P Marquis, Michael B Atkins, Meredith Regan, Sabina Signoretti, Robert E Lenkinski, S. N. G. (2009) 'Does arterial spin-labeling MR imaging-measured tumor perfusion correlate with renal cell cancer response to antiangiogenic therapy in a mouse model? (Radiology (2009) 251, 3, (731-742))', *Radiology*. doi: 10.1148/radiol.2532092534.

Raimondi, A. *et al.* (2020) 'Predictive Biomarkers of Response to Immunotherapy in Metastatic Renal Cell Cancer', *Frontiers in Oncology*. doi: 10.3389/fonc.2020.01644.

Rambow, F. *et al.* (2018) 'Toward Minimal Residual Disease-Directed Therapy in Melanoma', *Cell*, 174(4). doi: 10.1016/j.cell.2018.06.025.

Ramos-Casals, M. *et al.* (2020) 'Immune-related adverse events of checkpoint inhibitors', *Nature Reviews Disease Primers*. doi: 10.1038/s41572-020-0160-6.

Rao, M. *et al.* (2020) 'Comparative single-cell RNA sequencing (scRNA-seq) reveals liver metastasis-specific targets in a patient with small intestinal neuroendocrine cancer', *Cold Spring Harbor Molecular Case Studies*, 6(2). doi: 10.1101/MCS.A004978.

Regev, A. *et al.* (2017) 'The human cell atlas', *eLife*, 6, pp. 1–30. doi: 10.7554/eLife.27041.

Rezapourbehagh, S. *et al.* (2019) 'Targeted Agents and Resistance Mechanism in Renal Cell Cancer', *The Bulletin of Urooncology*, 18(2). doi: 10.4274/uob.galenos.2018.1090.

Rhoads, A. and Au, K. F. (2015) 'PacBio Sequencing and Its Applications', *Genomics, Proteomics and Bioinformatics*. doi: 10.1016/j.gpb.2015.08.002.

Ribas, A. and Wolchok, J. D. (2018) 'Cancer immunotherapy using checkpoint blockade', *Science*. doi: 10.1126/science.aar4060.

Rich, J. T. *et al.* (2010) 'A practical guide to understanding Kaplan-Meier curves', *Otolaryngology - Head and Neck Surgery*, 143(3). doi: 10.1016/j.otohns.2010.05.007.

Ricketts, C. J. *et al.* (2018) 'The Cancer Genome Atlas Comprehensive Molecular Characterization of Renal Cell Carcinoma', *Cell Reports*, 23(1). doi: 10.1016/j.celrep.2018.03.075.

Rini, B. *et al.* (2015) 'A 16-gene assay to predict recurrence after surgery in localised renal cell carcinoma: Development and validation studies', *The Lancet Oncology*, 16(6). doi: 10.1016/S1470-2045(15)70167-1.

Rini, B. I. (2009) 'Metastatic renal cell carcinoma: Many treatment options, one patient', *Journal of Clinical Oncology*. doi: 10.1200/JCO.2008.19.9836.

Rini, B. I. *et al.* (2018) 'Validation of the 16-gene recurrence score in patients with locoregional, high-risk renal cell carcinoma from a phase III trial of adjuvant sunitinib', *Clinical*

*Cancer Research*, 24(18). doi: 10.1158/1078-0432.CCR-18-0323.

Rini, B. I., Campbell, S. C. and Escudier, B. (2009) 'Renal cell carcinoma', *The Lancet*. Elsevier Ltd, 373(9669), pp. 1119–1132. doi: 10.1016/S0140-6736(09)60229-4.

Robb, V. A. *et al.* (2007) 'Activation of the mTOR Signaling Pathway in Renal Clear Cell Carcinoma', *Journal of Urology*, 177(1). doi: 10.1016/j.juro.2006.08.076.

Rodríguez-Esteban, G. *et al.* (2015) 'Digital gene expression approach over multiple RNA-Seq data sets to detect neoblast transcriptional changes in *Schmidtea mediterranea*', *BMC Genomics*, 16(1). doi: 10.1186/s12864-015-1533-1.

Rodrigues, S. G. *et al.* (2019) 'Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution', *Science*, 363(6434). doi: 10.1126/science.aaw1219.

Roelants, C. *et al.* (2018a) 'Combined inhibition of PI3K and Src kinases demonstrates synergistic therapeutic efficacy in clear-cell renal carcinoma', *Oncotarget*. doi: 10.1016/j.chiabu.2004.12.010.

Roelants, C. *et al.* (2018b) 'Combined inhibition of PI3K and Src kinases demonstrates synergistic therapeutic efficacy in clear-cell renal carcinoma', *Oncotarget*. doi: 10.18632/oncotarget.25700.

Roelants, C. *et al.* (2020) 'Ex-vivo treatment of tumor tissue slices as a predictive preclinical method to evaluate targeted therapies for patients with renal carcinoma', *Cancers*, 12(1). doi: 10.3390/cancers12010232.

Rohlenova, K. *et al.* (2018) 'Endothelial Cell Metabolism in Health and Disease', *Trends in Cell Biology*. doi: 10.1016/j.tcb.2017.10.010.

Rojas, J. D. *et al.* (2019) 'Ultrasound Measurement of Vascular Density to Evaluate Response to Anti-Angiogenic Therapy in Renal Cell Carcinoma', *IEEE Transactions on Biomedical Engineering*, 66(3). doi: 10.1109/TBME.2018.2860932.

Rowley, J. D. (1973) 'A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining', *Nature*, 243(5405). doi: 10.1038/243290a0.

Sabatino, M. *et al.* (2009) 'Serum vascular endothelial growth factor and fibronectin predict clinical response to high-dose interleukin-2 therapy', *Journal of Clinical Oncology*, 27(16). doi: 10.1200/JCO.2008.19.1106.

Sachs, N. *et al.* (2018) 'A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity', *Cell*, 172(1–2). doi: 10.1016/j.cell.2017.11.010.

Sade-Feldman, M. *et al.* (2018) 'Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma', *Cell*, 175(4). doi: 10.1016/j.cell.2018.10.038.

Saelens, W. *et al.* (2019) 'A comparison of single-cell trajectory inference methods', *Nature Biotechnology*, 37(5). doi: 10.1038/s41587-019-0071-9.

Samanta, M. P., Tongprasit, W. and Stolc, V. (2007) 'In-depth query of large genomes using tiling arrays', *Methods in Molecular Biology*, 377. doi: 10.1385/1-59745-390-0:163.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) 'DNA sequencing with chain-terminating inhibitors.', *Proceedings of the National Academy of Sciences of the United States of America*, 74(12). doi: 10.1073/pnas.74.12.5463.

Sarkar, A. and Stephens, M. (2021) 'Separating measurement and expression models clarifies confusion in single-cell RNA sequencing analysis', *Nature Genetics*. doi: 10.1038/s41588-021-00873-4.

Sasagawa, Y. *et al.* (2013) 'Quartz-Seq: A highly reproducible and sensitive single-cell RNA sequencing method, reveals nongenetic gene-expression heterogeneity', *Genome Biology*, 14(4). doi: 10.1186/gb-2013-14-4-r31.

Sasagawa, Y. *et al.* (2018) 'Quartz-Seq2: A high-throughput single-cell RNA-sequencing method that effectively uses limited sequence reads', *Genome Biology*, 19(1). doi: 10.1186/s13059-018-1407-3.

- Sato, Y. *et al.* (2013) 'Integrated molecular analysis of clear-cell renal cell carcinoma', *Nature Genetics*, 45(8). doi: 10.1038/ng.2699.
- Sawada, M. *et al.* (1994) 'Clonal analysis of human gynecologic cancers by means of the polymerase chain reaction', *International Journal of Cancer*, 58(4). doi: 10.1002/ijc.2910580406.
- Schatz, M. C. (2017) 'Nanopore sequencing meets epigenetics', *Nature Methods*. doi: 10.1038/NMETH.4240.
- Scheyltjens, I. *et al.* (2022) 'Single-cell RNA and protein profiling of immune cells from the mouse brain and its border tissues', *Nature Protocols*. Springer US. doi: 10.1038/s41596-022-00716-4.
- Scudiero, D. A. *et al.* (1988) 'Evaluation of a Soluble Tetrazolium/Formazan Assay for Cell Growth and Drug Sensitivity in Culture Using Human and Other Tumor Cell Lines', *Cancer Research*, 48(17).
- See, P. *et al.* (2018) 'A single-cell sequencing guide for immunologists', *Frontiers in Immunology*. doi: 10.3389/fimmu.2018.02425.
- Semenza, G. L. (2013) 'HIF-1 mediates metabolic responses to intratumoral hypoxia and oncogenic mutations', *Journal of Clinical Investigation*. doi: 10.1172/JCI67230.
- Şenbabaoğlu, Y. *et al.* (2016) 'Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures', *Genome Biology*, 17(1). doi: 10.1186/s13059-016-1092-z.
- Setty, M. *et al.* (2016) 'Wishbone identifies bifurcating developmental trajectories from single-cell data', *Nature Biotechnology*, 34(6). doi: 10.1038/nbt.3569.
- Shackleton, M. *et al.* (2009) 'Heterogeneity in Cancer: Cancer Stem Cells versus Clonal Evolution', *Cell*. doi: 10.1016/j.cell.2009.08.017.
- Shaffer, S. M. *et al.* (2017) 'Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance', *Nature*, 546(7658). doi: 10.1038/nature22794.
- Shah, S. P. *et al.* (2012) 'The clonal and mutational evolution spectrum of primary triple-negative breast cancers', *Nature*, 486(7403). doi: 10.1038/nature10933.
- Shang, M. *et al.* (2019) 'Microfluidic modelling of the tumor microenvironment for anti-cancer drug development', *Lab on a Chip*. doi: 10.1039/c8lc00970h.
- Sharifi, N. and Farrar, W. L. (2006) 'Perturbations in hypoxia detection: A shared link between hereditary and sporadic tumor formation?', *Medical Hypotheses*, 66(4). doi: 10.1016/j.mehy.2005.11.003.
- Sharma, R. *et al.* (2021) 'Determinants of resistance to VEGF-TKI and immune checkpoint inhibitors in metastatic renal cell carcinoma', *Journal of Experimental and Clinical Cancer Research*. doi: 10.1186/s13046-021-01961-3.
- Shenoy, N. and Pagliaro, L. (2016) 'Sequential pathogenesis of metastatic VHL mutant clear cell renal cell carcinoma: Putting it together with a translational perspective', *Annals of Oncology*, 27(9). doi: 10.1093/annonc/mdw241.
- Shibue, T. and Weinberg, R. A. (2017) 'EMT, CSCs, and drug resistance: The mechanistic link and clinical implications', *Nature Reviews Clinical Oncology*. doi: 10.1038/nrclinonc.2017.44.
- Shin, J. *et al.* (2015) 'Single-Cell RNA-Seq with Waterfall Reveals Molecular Cascades underlying Adult Neurogenesis', *Cell Stem Cell*, 17(3). doi: 10.1016/j.stem.2015.07.013.
- Shinawi, M. and Cheung, S. W. (2008) 'The array CGH and its clinical applications', *Drug Discovery Today*. doi: 10.1016/j.drudis.2008.06.007.
- Shingarev, R. and Jaimes, E. A. (2017) 'Renal cell carcinoma: New insights and challenges for a clinician scientist', *American Journal of Physiology - Renal Physiology*. doi: 10.1152/ajprenal.00480.2016.
- Shinojima, T. *et al.* (2007) 'Renal cancer cells lacking hypoxia inducible factor (HIF)-



1 $\alpha$  expression maintain vascular endothelial growth factor expression through HIF-2 $\alpha$ ', *Carcinogenesis*, 28(3). doi: 10.1093/carcin/bgl143.

Shoemaker, R. H. (2006) 'The NCI60 human tumour cell line anticancer drug screen', in *Nature Reviews Cancer*. doi: 10.1038/nrc1951.

Shrewsbury, A. B. *et al.* (2014) 'Renal cell carcinoma in patients with end-stage renal disease has favorable overall prognosis', *Clinical Transplantation*, 28(2). doi: 10.1111/ctr.12299.

Shuib, S. *et al.* (2011) 'Copy number profiling in von hippel-lindau disease renal cell carcinoma', *Genes Chromosomes and Cancer*, 50(7). doi: 10.1002/gcc.20865.

Singla, N. *et al.* (2020) 'Pancreatic tropism of metastatic renal cell carcinoma', *JCI Insight*, 5(7). doi: 10.1172/JCI.INSIGHT.134564.

Slyper, M. *et al.* (2020) 'A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors', *Nature Medicine*, 26(5). doi: 10.1038/s41591-020-0844-1.

Soneson, C. and Robinson, M. D. (2018) 'Bias, robustness and scalability in single-cell differential expression analysis', *Nature Methods*, 15(4). doi: 10.1038/nmeth.4612.

Sontheimer-Phelps, A., Hassell, B. A. and Ingber, D. E. (2019) 'Modelling cancer in microfluidic human organs-on-chips', *Nature Reviews Cancer*. doi: 10.1038/s41568-018-0104-6.

Sorbellini, M. *et al.* (2005) 'A postoperative prognostic nomogram predicting recurrence for patients with conventional clear cell renal cell carcinoma', *Journal of Urology*, 173(1). doi: 10.1097/01.ju.0000148261.19532.2c.

Ståhl, P. L. *et al.* (2016) 'Visualization and analysis of gene expression in tissue sections by spatial transcriptomics', *Science*. doi: 10.1126/science.aaf2403.

Stengel, B. (2010) 'Chronic kidney disease and cancer: A troubling connection', *Journal of Nephrology*.

Stoeckius, M. *et al.* (2017) 'Simultaneous epitope and transcriptome measurement in single cells', *Nature Methods*. doi: 10.1038/nmeth.4380.

Stoeckius, M. *et al.* (2018) 'Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics', *Genome Biology*, 19(1). doi: 10.1186/s13059-018-1603-1.

Street, K. *et al.* (2018) 'Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics', *BMC Genomics*. doi: 10.1186/s12864-018-4772-0.

Su, C. *et al.* (2021) 'Single-Cell RNA Sequencing in Multiple Pathologic Types of Renal Cell Carcinoma Revealed Novel Potential Tumor-Specific Markers', *Frontiers in Oncology*, 11. doi: 10.3389/fonc.2021.719564.

Su, Y. *et al.* (2017) 'Single-cell analysis resolves the cell state transition and signaling dynamics associated with melanoma drug-induced resistance', *Proceedings of the National Academy of Sciences of the United States of America*, 114(52). doi: 10.1073/pnas.1712064115.

Su, Z. *et al.* (2019) 'Inferring the evolution and progression of small-cell lung cancer by single-cell sequencing of circulating tumor cells', *Clinical Cancer Research*, 25(16). doi: 10.1158/1078-0432.CCR-18-3571.

Sun, D. *et al.* (2021) 'TISCH: A comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment', *Nucleic Acids Research*, 49(D1). doi: 10.1093/nar/gkaa1020.

Sun, G. *et al.* (2021) 'Single-cell RNA sequencing in cancer: Applications, advances, and emerging challenges', *Molecular Therapy - Oncolytics*. doi: 10.1016/j.omto.2021.04.001.

Sun, M. *et al.* (2011) 'Prognostic factors and predictive models in renal cell carcinoma: A contemporary review', *European Urology*. doi: 10.1016/j.eururo.2011.06.041.

Sun, S. *et al.* (2019) 'Accuracy, robustness and scalability of dimensionality reduction methods for single-cell RNA-seq analysis', *Genome Biology*, 20(1). doi: 10.1186/s13059-019-

1898-6.

Sun, W. *et al.* (2019) ‘Organ-on-a-Chip for Cancer and Immune Organs Modeling’, *Advanced Healthcare Materials*. doi: 10.1002/adhm.201801363.

Sung, H. *et al.* (2021) ‘Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries’, *CA: A Cancer Journal for Clinicians*, 71(3). doi: 10.3322/caac.21660.

Sung, W. W. *et al.* (2021) ‘Trends in the kidney cancer mortality-to-incidence ratios according to health care expenditures of 56 countries’, *Scientific Reports*, 11(1). doi: 10.1038/s41598-020-79367-y.

Suvà, M. L. and Tirosh, I. (2019) ‘Single-Cell RNA Sequencing in Cancer: Lessons Learned and Emerging Challenges’, *Molecular Cell*. doi: 10.1016/j.molcel.2019.05.003.

Svensson, V. *et al.* (2017) ‘Power analysis of single-cell mRNA-sequencing experiments’, *Nature Methods*, 14(4). doi: 10.1038/nmeth.4220.

Svensson, V., Vento-Tormo, R. and Teichmann, S. A. (2018) ‘Exponential scaling of single-cell RNA-seq in the past decade’, *Nature Protocols*. doi: 10.1038/nprot.2017.149.

Swanton, C. (2012) ‘Intratumor heterogeneity: Evolution through space and time’, *Cancer Research*. doi: 10.1158/0008-5472.CAN-12-2217.

Syed, J. S. *et al.* (2020) ‘Transcriptomics in RCC’, *Urologic Oncology: Seminars and Original Investigations*. doi: 10.1016/j.urolonc.2019.12.003.

Tan, M. H. *et al.* (2004) ‘Gene expression profiling of renal cell carcinoma’, *Clinical Cancer Research*, 10(18 II), pp. 6315–6322. doi: 10.1158/1078-0432.CCR-050002.

Tan, M. H. *et al.* (2010) ‘Genomic expression and single-nucleotide polymorphism profiling discriminates chromophobe renal cell carcinoma and oncocytoma’, *BMC Cancer*, 10. doi: 10.1186/1471-2407-10-196.

Tang, F. *et al.* (2009) ‘mRNA-Seq whole-transcriptome analysis of a single cell’, *Nature Methods*, 6(5). doi: 10.1038/nmeth.1315.

Tate, J. G. *et al.* (2019) ‘COSMIC: The Catalogue Of Somatic Mutations In Cancer’, *Nucleic Acids Research*, 47(D1). doi: 10.1093/nar/gky1015.

Team, R. C. (2021) ‘R: A Language and Environment for Statistical Computing’, *R Foundation for Statistical Computing*.

Thrane, K. *et al.* (2018) ‘Spatially resolved transcriptomics enables dissection of genetic heterogeneity in stage III cutaneous malignant melanoma’, *Cancer Research*, 78(20). doi: 10.1158/0008-5472.CAN-18-0747.

Thrupp, N. *et al.* (2020) ‘Single-Nucleus RNA-Seq Is Not Suitable for Detection of Microglial Activation Genes in Humans’, *Cell Reports*, 32(13). doi: 10.1016/j.celrep.2020.108189.

Ting, D. T. *et al.* (2014) ‘Single-cell RNA sequencing identifies extracellular matrix gene expression by pancreatic circulating tumor cells’, *Cell Reports*, 8(6). doi: 10.1016/j.celrep.2014.08.029.

Tirosh, I. *et al.* (2016) ‘Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq’, *Science*, 352(6282), pp. 189–196. doi: 10.1126/science.aad0501.

Toma, M. I. *et al.* (2008) ‘Loss of heterozygosity and copy number abnormality in clear cell renal cell carcinoma discovered by high-density Affymetrix 10K single nucleotide polymorphism mapping array’, *Neoplasia*, 10(7). doi: 10.1593/neo.08160.

Tomczak, K., Czerwińska, P. and Wiznerowicz, M. (2015) ‘The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge’, *Wspolczesna Onkologia*. doi: 10.5114/wo.2014.47136.

Trapnell, C. *et al.* (2014) ‘The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells’, *Nature Biotechnology*, 32(4). doi: 10.1038/nbt.2859.

- Trapnell, C. and Cacchiarelli, D. (2014) ‘Monocle: Differential expression and time-series analysis for single-cell RNA-Seq and qPCR experiments’, *Bioconductor.Fmrp.Usp.Br.*
- Trujillo-de Santiago, G. *et al.* (2018) ‘Gut-microbiota-on-a-chip: an enabling field for physiological research’, *Microphysiological Systems*, 1. doi: 10.21037/mps.2018.09.01.
- Tsai, H. F. *et al.* (2017) ‘Tumour-on-a-chip: Microfluidic models of tumour morphology, growth and microenvironment’, *Journal of the Royal Society Interface*. doi: 10.1098/rsif.2017.0137.
- Tung, P. Y. *et al.* (2017) ‘Batch effects and the effective design of single-cell gene expression studies’, *Scientific Reports*, 7. doi: 10.1038/srep39921.
- Ueda, K. *et al.* (2020) ‘The prognostic value of systemic inflammatory markers in advanced renal cell carcinoma patients treated with molecular targeted therapies’, *Anticancer Research*, 40(3). doi: 10.21873/anticancer.14127.
- Vaira, V. *et al.* (2010) ‘Preclinical model of organotypic culture for pharmacodynamic profiling of human tumors’, *Proceedings of the National Academy of Sciences*, 107(18), pp. 8352–8356. doi: 10.1073/pnas.0907676107.
- Velculescu, V. E. *et al.* (1995) ‘Serial analysis of gene expression’, *Science*, 270(5235). doi: 10.1126/science.270.5235.484.
- Vander Velde, R. *et al.* (2020) ‘Resistance to targeted therapies as a multifactorial, gradual adaptation to inhibitor specific selective pressures’, *Nature Communications*, 11(1). doi: 10.1038/s41467-020-16212-w.
- Vendramin, R., Litchfield, K. and Swanton, C. (2021) ‘Cancer evolution: Darwin and beyond’, *The EMBO Journal*, 40(18). doi: 10.15252/embj.2021108389.
- Verhaak, R. G. W. *et al.* (2010) ‘Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1’, *Cancer Cell*, 17(1). doi: 10.1016/j.ccr.2009.12.020.
- Vickers, A. E. m. *et al.* (2004) ‘Kidney Slices of Human and Rat to Characterize Cisplatin-Induced Injury on Cellular Pathways and Morphology’, *Toxicologic Pathology*, 32(5). doi: 10.1080/01926230490508821.
- Vickman, R. E. *et al.* (2020) ‘Heterogeneity of human prostate carcinoma-associated fibroblasts implicates a role for subpopulations in myeloid cell recruitment’, *Prostate*, 80(2). doi: 10.1002/pros.23929.
- Vieth, B. *et al.* (2019) ‘A systematic evaluation of single cell RNA-seq analysis pipelines’, *Nature Communications*, 10(1). doi: 10.1038/s41467-019-12266-7.
- Vitak, S. A. *et al.* (2017) ‘Sequencing thousands of single-cell genomes with combinatorial indexing’, *Nature Methods*, 14(3). doi: 10.1038/nmeth.4154.
- Vuong, L. *et al.* (2019) ‘Tumor microenvironment dynamics in clear-cell renal cell carcinoma’, *Cancer Discovery*. doi: 10.1158/2159-8290.CD-19-0499.
- Wallace, M. A. (1998) ‘Anatomy and physiology of the kidney.’, *AORN journal*. doi: 10.1016/s0001-2092(06)62377-6.
- Wang, C. C. *et al.* (2014) ‘A time- and matrix-dependent TGFBR3-JUND-KRT5 regulatory circuit in single breast epithelial cells and basal-like premalignancies’, *Nature Cell Biology*, 16(4). doi: 10.1038/ncb2930.
- Wang, H. *et al.* (2014) ‘Downregulation of FAP suppresses cell proliferation and metastasis through PTEN/PI3K/AKT and Ras-ERK signaling in oral squamous cell carcinoma’, *Cell Death and Disease*, 5(4). doi: 10.1038/cddis.2014.122.
- Wang, J. *et al.* (2018) ‘Bioinformatics and functional analyses of key genes and pathways in human clear cell renal cell carcinoma’, *Oncology Letters*, 15(6). doi: 10.3892/ol.2018.8473.
- Wang, M. *et al.* (2017) ‘Role of tumor microenvironment in tumorigenesis’, *Journal of Cancer*. doi: 10.7150/jca.17648.

- Wang, X. *et al.* (2014) ‘The role of angiopoietins as potential therapeutic targets in renal cell carcinoma’, *Translational Oncology*, 7(2). doi: 10.1016/j.tranon.2014.02.003.
- Wang, Y. *et al.* (2014) ‘Clonal evolution in breast cancer revealed by single nucleus genome sequencing’, *Nature*, 512(7513). doi: 10.1038/nature13600.
- Wang, Yuhao *et al.* (2019) ‘Extent and characteristics of immune infiltration in clear cell renal cell carcinoma and the prognostic value’, *Translational Andrology and Urology*, 8(6). doi: 10.21037/tau.2019.10.19.
- Wang, Yejinpeng *et al.* (2019) ‘Fifteen hub genes associated with progression and prognosis of clear cell renal cell carcinoma identified by coexpression analysis’, *Journal of Cellular Physiology*, 234(7). doi: 10.1002/jcp.27692.
- Ward, C. *et al.* (2021) ‘Preclinical Organotypic Models for the Assessment of Novel Cancer Therapeutics and Treatment’, in *Current Topics in Microbiology and Immunology*. doi: 10.1007/82\_2019\_159.
- Wei, E. Y. and Hsieh, J. J. (2015) ‘A river model to map convergent cancer evolution and guide therapy in RCC’, *Nature Reviews Urology*. doi: 10.1038/nrurol.2015.260.
- Weikert, S. *et al.* (2008) ‘Blood pressure and risk of renal cell carcinoma in the European prospective investigation into cancer and nutrition’, *American Journal of Epidemiology*, 167(4). doi: 10.1093/aje/kwm321.
- Van De Wetering, M. *et al.* (2015) ‘Prospective derivation of a living organoid biobank of colorectal cancer patients’, *Cell*, 161(4). doi: 10.1016/j.cell.2015.03.053.
- Weygant, N. *et al.* (2015) ‘DCLK1 is a broadly dysregulated target against epithelial-mesenchymal transition, focal adhesion, and stemness in clear cell renal carcinoma’, *Oncotarget*, 6(4). doi: 10.18632/oncotarget.3059.
- White, E. *et al.* (2004) ‘VITamins and Lifestyle Cohort Study: Study Design and Characteristics of Supplement Users’, *American Journal of Epidemiology*, 159(1). doi: 10.1093/aje/kwh010.
- Wilhelm, M. *et al.* (2002) ‘Array-based comparative genomic hybridization for the differential diagnosis of renal cell cancer’, *Cancer Research*, 62(4).
- Williams, M. J. *et al.* (2016) ‘Identification of neutral tumor evolution across cancer types’, *Nature Genetics*, 48(3). doi: 10.1038/ng.3489.
- Wolf, F. A. *et al.* (2019) ‘PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells’, *Genome Biology*, 20(1). doi: 10.1186/s13059-019-1663-x.
- Wolf, F. A., Angerer, P. and Theis, F. J. (2018) ‘SCANPY: Large-scale single-cell gene expression data analysis’, *Genome Biology*, 19(1). doi: 10.1186/s13059-017-1382-0.
- Wolf, M. M., Kimryn Rathmell, W. and Beckermann, K. E. (2020) ‘Modeling clear cell renal cell carcinoma and therapeutic implications’, *Oncogene*. doi: 10.1038/s41388-020-1234-3.
- Wolf, Y. *et al.* (2019) ‘UVB-Induced Tumor Heterogeneity Diminishes Immune Response in Melanoma’, *Cell*, 179(1). doi: 10.1016/j.cell.2019.08.032.
- Wolock, S. L., Lopez, R. and Klein, A. M. (2019) ‘Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data’, *Cell Systems*, 8(4). doi: 10.1016/j.cels.2018.11.005.
- Woodward, E. R. *et al.* (2000) ‘Familial clear cell renal cell carcinoma (FCRC): Clinical features and mutation analysis of the VHL, MET, and CUL2 candidate genes’, *Journal of Medical Genetics*, 37(5). doi: 10.1136/jmg.37.5.348.
- Wu, H. *et al.* (2018) ‘Single-cell RNA sequencing reveals diverse intratumoral heterogeneities and gene signatures of two types of esophageal cancers’, *Cancer Letters*. Elsevier, 438, pp. 133–143. doi: 10.1016/j.canlet.2018.09.017.
- Wu, H. *et al.* (2019) ‘Advantages of single-nucleus over single-cell RNA sequencing of

adult kidney: Rare cell types and novel cell states revealed in fibrosis', *Journal of the American Society of Nephrology*, 30(1). doi: 10.1681/ASN.2018090912.

Wu, Y. *et al.* (2022) 'Spatial omics: Navigating to the golden era of cancer research', *Clinical and Translational Medicine*. doi: 10.1002/ctm2.696.

Xia, Z. N. *et al.* (2022) 'Identification of a differentiation-related prognostic nomogram based on single-cell RNA sequencing in clear cell renal cell carcinoma', *Scientific Reports*. Nature Publishing Group UK, 12(1), pp. 1–14. doi: 10.1038/s41598-022-15206-6.

Xu, J. *et al.* (2016) 'Mechanistically distinct cancer-associated mTOR activation clusters predict sensitivity to rapamycin', *Journal of Clinical Investigation*, 126(9). doi: 10.1172/JCI86120.

Xu, J. *et al.* (2019) 'Single-cell lineage tracing by endogenous mutations enriched in transposase accessible mitochondrial DNA', *eLife*, 8. doi: 10.7554/eLife.45105.

Yachida, S. *et al.* (2010) 'Distant metastasis occurs late during the genetic evolution of pancreatic cancer', *Nature*, 467(7319). doi: 10.1038/nature09515.

Yamamoto, D. *et al.* (2009) 'Metastatic gastric tumor from renal cell carcinoma', *Gastric Cancer*, 12(3). doi: 10.1007/s10120-009-0519-6.

Yang, X. F. *et al.* (2018) 'Twist2 and CD24 expression alters renal microenvironment in obesity associated kidney cancer', *European Review for Medical and Pharmacological Sciences*, 22(2). doi: 10.26355/eurrev\_201801\_14180.

Yao, T. *et al.* (2016) 'Identification of genes associated with renal cell carcinoma using gene expression profiling analysis', *Oncology Letters*, 12(1). doi: 10.3892/ol.2016.4573.

Yates, L. R. *et al.* (2015) 'Subclonal diversification of primary breast cancer revealed by multiregion sequencing', *Nature Medicine*, 21(7). doi: 10.1038/nm.3886.

Yip, S. H., Sham, P. C. and Wang, J. (2018) 'Evaluation of tools for highly variable gene discovery from single-cell RNA-seq data', *Briefings in Bioinformatics*, 20(4). doi: 10.1093/bib/bby011.

Yoshihara, K. *et al.* (2013) 'Inferring tumour purity and stromal and immune cell admixture from expression data', *Nature Communications*, 4. doi: 10.1038/ncomms3612.

Young, M. D. *et al.* (2018) 'Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors', *Science*, 361(6402). doi: 10.1126/science.aat1699.

Young, M. D. *et al.* (2021) 'Single cell derived mRNA signals across human kidney tumors', *Nature Communications*, 12(1). doi: 10.1038/s41467-021-23949-5.

Yu, C. *et al.* (2014) 'Discovery of biclonal origin and a novel oncogene SLC12A5 in colon cancer by single-cell sequencing', *Cell Research*, 24(6). doi: 10.1038/cr.2014.43.

Yu, S. *et al.* (2017) 'Immortalized cancer-associated fibroblasts promote prostate cancer carcinogenesis, proliferation and invasion', in *Anticancer Research*. doi: 10.21873/anticancer.11824.

Yuan, H. *et al.* (2019) 'CancerSEA: A cancer single-cell state atlas', *Nucleic Acids Research*, 47(D1). doi: 10.1093/nar/gky939.

Zappia, L. and Theis, F. J. (2021) 'Over 1000 tools reveal trends in the single-cell RNA-seq analysis landscape', *Genome Biology*. doi: 10.1186/s13059-021-02519-4.

Zeng, J. *et al.* (2022) 'CancerSCEM: A database of single-cell expression map across various human cancers', *Nucleic Acids Research*, 50(D1). doi: 10.1093/nar/gkab905.

Zhang, C. *et al.* (2019) 'Development and validation of a metastasis-associated prognostic signature based on single-cell RNA-seq in clear cell renal cell carcinoma', *Aging*, 11(22). doi: 10.18632/aging.102434.

Zhang, F. *et al.* (2021) 'Discovery and construction of prognostic model for clear cell renal cell carcinoma based on single-cell and bulk transcriptome analysis', *Translational Andrology and Urology*, 10(9). doi: 10.21037/tau-21-581.

Zhang, J. M., Kamath, G. M. and Tse, D. N. (2019) 'Valid Post-clustering Differential

- Analysis for Single-Cell RNA-Seq', *Cell Systems*, 9(4). doi: 10.1016/j.cels.2019.07.012.
- Zhang, Jianjun *et al.* (2014) 'Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing', *Science*, 346(6206). doi: 10.1126/science.1256930.
- Zhang, L. *et al.* (2018) 'Tumor necrosis as a prognostic variable for the clinical outcome in patients with renal cell carcinoma: A systematic review and meta-analysis', *BMC Cancer*, 18(1). doi: 10.1186/s12885-018-4773-z.
- Zhang, L. *et al.* (2020) 'Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer', *Cell*, 181(2). doi: 10.1016/j.cell.2020.03.048.
- Zhang, M. *et al.* (2022) 'IDEAS: individual level differential expression analysis for single-cell RNA-seq data', *Genome Biology*, 23(1). doi: 10.1186/s13059-022-02605-1.
- Zhang, N. *et al.* (2018) 'Multiple metastases of clear-cell renal cell carcinoma to different region of the nasal cavity and paranasal sinus 3 times successively', *Medicine (United States)*. doi: 10.1097/MD.00000000000010286.
- Zhang, X. *et al.* (2019) 'Comparative Analysis of Droplet-Based Ultra-High-Throughput Single-Cell RNA-Seq Systems', *Molecular Cell*, 73(1). doi: 10.1016/j.molcel.2018.10.020.
- Zhang, Z. *et al.* (2020) 'Construction of a novel gene-based model for prognosis prediction of clear cell renal cell carcinoma', *Cancer Cell International*, 20(1). doi: 10.1186/s12935-020-1113-6.
- Zhao, C. X., Luo, C. L. and Wu, X. H. (2015) 'Hypoxia promotes 786-O cells invasiveness and resistance to sorafenib via HIF-2 $\alpha$ /COX-2', *Medical Oncology*, 32(1). doi: 10.1007/s12032-014-0419-4.
- Zhao, W. *et al.* (2021) 'Deconvolution of cell type-specific drug responses in human tumor tissue with single-cell RNA-seq', *Genome Medicine*, 13(1). doi: 10.1186/s13073-021-00894-y.
- Zheng, C. *et al.* (2017) 'Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing', *Cell*, 169(7). doi: 10.1016/j.cell.2017.05.035.
- Zheng, F. *et al.* (2016) 'Organ-on-a-Chip Systems: Microengineering to Biomimic Living Systems', *Small*. doi: 10.1002/smll.201503208.
- Zheng, G. X. Y. *et al.* (2017) 'Massively parallel digital transcriptional profiling of single cells', *Nature Communications*, 8. doi: 10.1038/ncomms14049.
- Zhou, F. *et al.* (2017) 'Circulating endothelial cells and their subsets: Novel biomarkers for cancer', *Biomarkers in Medicine*. doi: 10.2217/bmm-2017-0143.
- Zhou, N. *et al.* (2016) 'IL-8 induces the epithelial-mesenchymal transition of renal cell carcinoma cells through the activation of AKT signaling', *Oncology Letters*, 12(3). doi: 10.3892/ol.2016.4900.
- Ziegenhain, C. *et al.* (2017) 'Comparative Analysis of Single-Cell RNA Sequencing Methods', *Molecular Cell*, 65(4). doi: 10.1016/j.molcel.2017.01.023.
- Zigeuner, R. *et al.* (2010) 'External Validation of the Mayo Clinic Stage, Size, Grade, and Necrosis (SSIGN) Score for Clear-Cell Renal Cell Carcinoma in a Single European Centre Applying Routine Pathology', *European Urology*, 57(1). doi: 10.1016/j.eururo.2008.11.033.
- Zuo, S. *et al.* (2019) 'Identification of a universal 6-lncRNA prognostic signature for three pathologic subtypes of renal cell carcinoma', *Journal of Cellular Biochemistry*, 120(5). doi: 10.1002/jcb.28012.
- Zustovich, F. *et al.* (2008) 'Cardiac metastasis from renal cell carcinoma without inferior vena involvement: A review of the literature based on a case report. Two different patterns of spread?', *International Journal of Clinical Oncology*, 13(3). doi: 10.1007/s10147-007-0730-6.



# Annexes





## **5.1. REVUE DE LA LITTERATURE - 1**

Targeting the PD-1/PD-L1 Pathway in Renal cell carcinoma





Review

# Targeting the PD-1/PD-L1 Pathway in Renal Cell Carcinoma

Solène-Florence Kammerer-Jacquet <sup>1,2,\*</sup>, Antoine Deleuze <sup>3</sup>, Judikaël Saout <sup>2</sup>,  
Romain Mathieu <sup>2,4</sup>, Brigitte Laguerre <sup>3</sup>, Gregory Verhoest <sup>4</sup>, Frédéric Dugay <sup>2,5</sup>,  
Marc-Antoine Belaud-Rotureau <sup>2,5</sup>, Karim Bensalah <sup>4</sup> and Nathalie Rioux-Leclercq <sup>1,2</sup>

<sup>1</sup> Department of Pathology, University Hospital, 35000 Rennes, France; nathalie.rioux-leclercq@chu-rennes.fr

<sup>2</sup> Université Rennes, Inserm, EHESP (Ecole des Hautes Etudes en Santé Publique), Irset (Institut de recherche en santé, environnement et travail) - UMR\_S 1085, F-35000 Rennes, France; judikael.saout@univ-rennes1.fr (J.S.); romain.mathieu@chu-rennes.fr (R.M.); frederic.dugay@chu-rennes.fr (F.D.); marc-antoine.belaud-rotureau@chu-rennes.fr (M.-A.B.-R.)

<sup>3</sup> Department of Medical Oncology, Centre Eugene Marquis, 35000 Rennes, France; acjdeleuze@gmail.com (A.D.); b.laguerre@rennes.unicancer.fr (B.L.)

<sup>4</sup> Department of Urology, University Hospital, 35000 Rennes, France; gregory.verhoest@chu-rennes.fr (G.V.); karim.bensalah@chu-rennes.fr (K.B.)

<sup>5</sup> Department of Cytogenetics, University Hospital, F-35000 Rennes, France

\* Correspondence: jacquet.sf@gmail.com; Tel.: +33-299-28-42-79; Fax: +33-2-99-28-42-84

Received: 17 March 2019; Accepted: 27 March 2019; Published: 4 April 2019



**Abstract:** Renal cell carcinoma encompass distinct diseases with different pathologic features and distinct molecular pathways. Immune checkpoint inhibitors targeting the programmed death receptor ligand 1 (PD-L1)/programmed death receptor 1 (PD-1) pathway alone or in combination have greatly changed clinical management of metastatic renal cell carcinoma, now competing with antiangiogenic drugs in monotherapy for first-line treatment. However, long-term response rates are low, and biomarkers are needed to predict treatment response. Quantification of PD-L1 expression by immunohistochemistry was developed as a promising biomarker in clinical trials, but with many limitations (different antibodies, tumour heterogeneity, specimens, and different thresholds of positivity). Other biomarkers, including tumour mutational burden and molecular signatures, are also developed and discussed in this review.

**Keywords:** immunotherapy; immune checkpoint inhibition; renal cell carcinoma; PD-1; PD-L1; biomarker; tumour mutational burden

## 1. Introduction

Renal cell carcinoma (RCC) represents 2% to 3% of all cancers, corresponding to 338,000 new cases diagnosed each year worldwide [1]. Around 30% of patients will present metastatic disease at the time of diagnosis and metastases are found in 30% to 40% of those initially treated in curative intent. If the incidence and mortality tend to decrease in Western Europe [2], recent data from the United States show a continued increase in incidence [3].

Clear cell renal cell carcinomas account for about 75% of RCC [1]. The other histologies are numerous according to the last WHO classification and mainly include the following entities: Papillary (20%) and chromophobe RCC (5%), which are the most frequent non-clear cell RCC (nccRCC), collecting duct carcinoma; medullar RCC, and translocation-associated RCC, with distinct pathologic features and different molecular alterations [4]. Among them, collecting duct carcinoma has been identified as an immunogenic tumour characterized by the upregulation of genes involved in the immune response [5].

The management and treatment of metastatic renal cell cancer have undergone major improvements over the last 20 years. The development of targeted therapies, including vascular endothelial growth factor (VEGF) pathway inhibitors, mammalian target of rapamycin (mTOR) inhibitors, and, more recently, the emergence of immune checkpoint inhibitors (ICI), such as anti-programmed death receptor 1 (PD-1), anti-programmed death receptor ligand 1 (PD-L1), and anti-cytotoxic T lymphocytes antigen 4 (CTLA-4), have led to a wide treatment landscape remodelling, targeting the tumour micro-environment (Table 1).

**Table 1.** Checkpoint inhibitors and corresponding IHC antibodies used in clinical trials for renal cell carcinoma.

Checkpoint Inhibitor	Target	Antibody for PD-L1 IHC	PD-L1 Assessment
Ipilimumab	CTLA-4	Not applicable	Not applicable
Nivolumab	PD-1	Dako 28-8	% tumour cells
Pembrolizumab	PD-1	Dako 22C3	% tumour cells
Atezolizumab	PD-L1	Ventana SP142	% immune cells
Durvalumab	PD-L1	Ventana SP263	Not available
Avelumab	PD-L1	Dako 7310	% immune cells

Abbreviations: CTLA-4, cytotoxic T-lymphocyte associated protein 4 antibody; PD-1, programmed cell death protein 1; PD-L1, programmed death-ligand 1

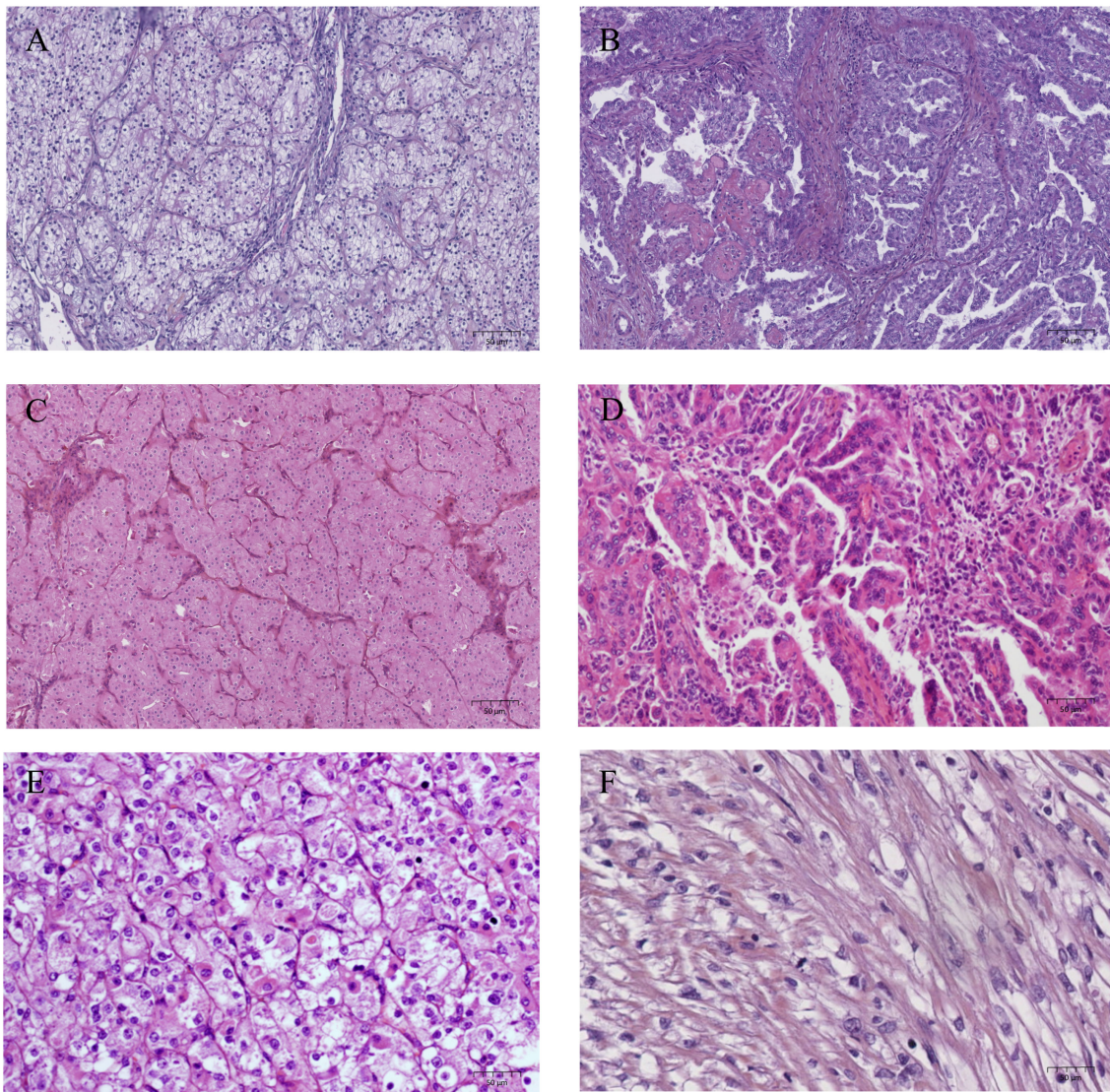
The advent of targeted therapies highlights the central role of the tumour micro-environment. The role of angiogenesis was already well known in ccRCC, with the VHL inactivation pathway leading to angiogenesis. The immune response with ICI is a more recent strategy. There is also a biological rationale to use them in association.

In this comprehensive review, we will discuss: (1) the histologic subtypes and the molecular pathways related, (2) the anti-tumour immune response and its interaction with angiogenesis, (3) the current landscape of targeted immunotherapy in RCC, and (4) the predictive biomarkers in RCC, including their limitations and the perspectives.

## 2. Histologic Subtypes and Related Molecular Pathways

### 2.1. Clear Cell Renal Cell Carcinoma

Clear cell RCC (ccRCC) is the most frequent histologic entity derived from the proximal convoluted tubule with a morphologic heterogeneity, and is composed of cells with clear or eosinophilic cytoplasm (Figure 1A). These tumours are driven by inactivation of the von Hippel Lindau (*VHL*) tumour suppressor gene [6]. The VHL protein is a multifunctional protein insuring variable regulatory functions, such as remodelling of the extra cellular matrix and control of the cell cycle, but also apoptosis and endocytosis. The function that has been most thoroughly studied is the stability regulation of hypoxia inducible factor (HIF). The VHL protein is a component of an E3 ubiquitin-ligase which targets proteins, including the transcription factor HIF, leading to their ubiquitination and degradation by proteasomes. The consequence of VHL protein impairment is a stabilization of HIF and an increased level of the latter. Such a situation can lead to the transcription of genes regulated by HIF, such as vascular endothelial growth factor (*VEGF*) or carbonic anhydrase IX (*CAIX*) [5]. Consequently, *VHL* acts as one of the main triggers of the angiogenic processes in ccRCC. Other pathways, such as mTOR and MAP kinases pathways, have been described, which also lead to the expression of angiogenic genes [7]. Recently, other common molecular drivers have been identified with mutations commonly involving chromatin-remodelling genes (*PBRM1*, *KDM5C*, *SETD2*, and *BAP1*) [8]. A metabolic shift was also described in aggressive tumours, involving, in particular, the downregulation of genes of the citric acid cycle.



**Figure 1.** Histological features of renal cell carcinoma, HES staining, x200. (A) clear cell renal cell carcinoma, (B) papillary renal cell carcinoma, (C) chromophobe renal cell carcinoma, (D) TFE3 translocation renal cell carcinoma, (E) TFE3 translocation renal cell carcinoma, (F) sarcomatoid component.

## 2.2. Non-Clear Cell Renal Cell Carcinoma

Papillary renal cell carcinoma (pRCC) is derived from the renal tubular epithelium and represents a heterogeneous disease, including tumours with an indolent outcome (type 1) and tumours with a more aggressive phenotype (type 2) (Figure 1B) [4]. Type 1 papillary renal cell carcinomas are enriched in *MET* alterations, mainly mutations and gains in chromosome 7 involving the *MET* locus. However, *MET* copy number gain is also found in up to 50% of type 2 pRCC. Recurrent alterations of *SETD2*, *EGFR*, *CDKN2A*, *NF2*, *TERT*, and *FH* were also described and occurred more frequently in type 2 pRCC, suggesting deregulation of chromatin remodelling and activation of the cell cycle and MAP kinases pathway [9].

Chromophobe renal cell carcinoma (cRCC) is an eosinophilic tumour (Figure 1C) derived from the distal nephron, characterized by mitochondrial alterations with increased mitochondria count, expression of genes involved in the citric acid cycle, and mutations in mitochondrial DNA copies [10]. They display frequent chromosome loss and less somatic mutations than in other types of RCC. The most frequently mutated gene was *TP53* and the most frequent oncogenic pathway alterations involved mTOR.



Translocation renal cell carcinoma (tRCC) is characterized by gene fusions involving two members of the Microphthalmia-associated transcription factor (MiTF) family: *TFE3* and *TFEB* located on Xp11.2 and 6p21.1 respectively [4,11]. *TFE3* tRCC displays a papillary architecture with epithelioid clear cells and psammoma bodies, as shown in Figure 1D,E. *TFEB* tRCC are biphasic neoplasms with epithelioid cells and smaller cell clusters. *TFE3* and *TFEB* have multiple translocation partners, mainly involved in messenger RNA splicing [12]. In addition, *TFE3* tRCC have genome-wide expression alterations that are not found in tumours with *TFEB* gene fusions, suggesting different oncogenic mechanisms between tRCC [13].

Collecting duct carcinoma (CDC) arises from the principal cells of the renal collecting ducts of Bellini and was initially thought to share similarities with upper urinary tract urothelial carcinoma [14,15]. However, molecular analysis revealed that CDCs had distinct copy number alterations, as well as a specific gene expression signature closer to other RCC. They have been identified as immunogenic tumours characterized by the upregulation of genes involved in the immune response, including T-cell activation and proliferation and a high rate of lymphocyte infiltration [5]. Additional reports of CDC genomic alterations highlight the prevalence of the metabolic shift and cell cycle deregulations. Mutations of *NF2*, *SETD2*, and *SMARCB1* were also reported [16].

Renal medullary carcinoma (RMC) is a rare tumour centred on the renal medulla whose current understanding is limited [17]. Several reports highlight a consistent loss of *SMARCB1/INI1*, a core component of the SWI/SNF chromatin remodelling complex [18].

Sarcomatoid features can occur in each histologic subtype (Figure 1F) and in the sarcomatoid component, demonstrate an increased mutational burden along with a higher frequency of *TP53*, *CDKN2A*, and *NF2* mutations and chromatin remodelling genes, *BAP1* and *ARID1A* [19].

### 3. Anti-Tumour Immune Response and Interaction with Angiogenesis

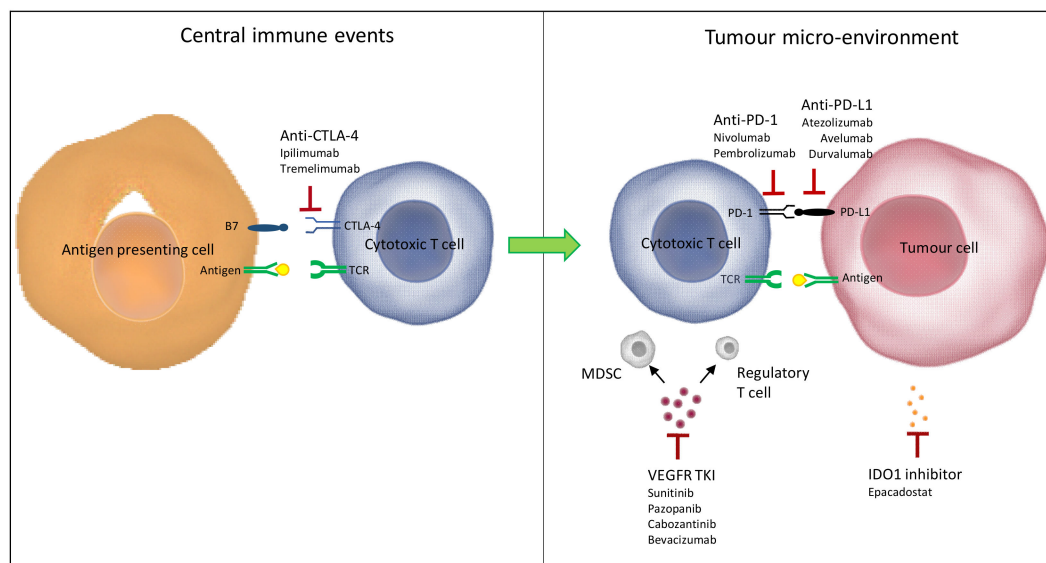
#### 3.1. The Anti-Tumour Immune Response

The anti-tumour immune response is based on the cooperation between innate and adaptive immune cells [20]. The first step of the cancer-immunity cycle described by Chen et al. [21] is the release of tumour cell antigens and their capture by dendritic cells (innate immune cells) (Figure 2). Then, in the lymph node, dendritic cells present the captured antigens on MHC I and MHC II to T cells (adaptive immune cells), leading to the activation of effector T cell responses against tumour cells antigens. Through blood vessels, the activated effector T cells migrate to tumours and specifically recognize and bind to tumour cells through the interaction between its T cell receptor (TCR) and its cognate antigen bound to MHC, thus killing antigen-bearing tumour cells.

Unfortunately, in patients, the anti-tumour immune response is not optimal. For instance, tumour antigens may not be detected [22]. Immune cells may consider antigens as self rather than foreign, leading to T regulatory cell responses. In addition, T cells may be inhibited from infiltrating tumour cells [23]. Most importantly, factors in the tumour micro-environment might suppress those effector T cells that are produced.

The activity of T cells is modulated by immune checkpoints expressed on their surface that can generate inhibitory signals. In lymph node sites, the activation of CTLA-4 on T cells when they binds to B7 ligands on antigen-presenting cells decreases T cells and their activity. In tumour sites, PD-1 expressed on activated T cells and upon binding to its ligand PD-L1 on tumour cells, leads to T cell exhaustion. PD-L1 is not only expressed by tumour cells, but also by immune cells, such as macrophages [24].

Biological rationale encouraged the synergy of CTLA-4 inhibition, which favours the development of an active immune response at the level of T-cell proliferation, with PD-1 inhibition, which modulates the immune response at the level of the tumour micro-environment.



**Figure 2.** Rationale for combining immune checkpoint inhibitors with antiangiogenic drugs, anti-tumour immune response modulation adapted from Amin et al. *Frontiers in Oncology*. 2019. TCR, T cell receptor; MDSC, myeloid derived suppressor cells; VEGF, vascular endothelial growth factor; PD-1, programmed death 1; PD-L1, programmed death ligand 1; CTLA 4, cytotoxic T lymphocyte antigen 4.

### 3.2. Interaction between Angiogenesis and the Anti-Tumour Immune Response

In ccRCC, the inactivation of *VHL* results in an increase of growth factors, such as VEGF, that favour the proliferation and migration of endothelial cells. Increased levels of VEGF in the tumour can induce suppression of both innate and adaptive immune responses. For instance, VEGF have been demonstrated to directly inhibit dendritic cell maturation, decreasing antigen presentation and T cells development, and thus leading to their exhaustion [25–27]. In addition, VEGF induce more immune suppression in the tumour micro-environment by increasing the expression of inhibitory checkpoints and by the recruitment of myeloid-derived suppressor cells (MDSC) and T-regulatory cells (Figure 2) [28–30]. VEGF also results in neo-angiogenesis with tortuous vasculature and abnormal vessel formation that alters the quality and quantity of infiltration of immune cells in the tumour immune micro-environment.

Interestingly, VEGF inhibitors are able to reverse these effects by improving dendritic cell function and antigen presentation, vasculature normalization with greater trafficking of immune cells, increased cytotoxic T cell infiltration, and decreased MDSC and T-regulatory cells that could potentially reduce the immunosuppressive effect in the tumour micro-environment [31,32].

## 4. Clinical Trials Targeting the PD-L1/PD-1 Pathway

### 4.1. Immune Checkpoint Inhibitors Alone or in Combination

Since the use of interleukin-2 and interferon-alpha in metastatic renal cell carcinoma at the end of the 1990s, the understanding of tumour immunity has significantly improved [33,34]. Nivolumab, a human IgG4 anti-PD-1 antibody, showed its efficacy in metastatic ccRCC in patients previously treated by antiangiogenics, over everolimus (mTOR inhibitor), in a phase III trial, CheckMate 025 (NCT01668784) (Table 2) [35]. The study met its primary endpoint with a significantly longer median overall survival (OS): 25 months with nivolumab versus 19.6 months with everolimus (HR = 0.73; 98.5% CI, 0.57 to 0.93;  $p = 0.0018$ ). The median progression-free survival (PFS) was 4.6 months vs. 4.4 months. Nivolumab also showed benefits in secondary outcomes with a higher objective response rate of 25% versus 5% and a better safety profile compared to everolimus. The expression of PD-L1 was associated with a poorer survival, but was not associated to its response to nivolumab.



**Table 2.** Clinical trials evaluating immune checkpoint inhibitors alone or in combination.

NCT Number/Study Name	Targeting Agents	Comparison	Phase	Histology	Primary Endpoint	Therapy Setting	Status
NCT01668784/CheckMate 025	nivolumab	everolimus	III	ccRCC	OS	at least second line	published
NCT02231749/CheckMate 214	Nivolumab + ipilimumab	sunitinib	III	ccRCC	PFS, OS, ORR	first line	published
NCT03260894	Pembrolizumab + epacadostat	sunitinib/pazopanib	III	ccRCC	ORR	first line	active not recruiting
NCT02853344/Keynote-427	pembrolizumab	-	II	cc/nccRCC	ORR	all lines	active not recruiting
NCT02964078	pembrolizumab+interleukin-2	-	II	ccRCC	ORR	all lines	active not recruiting
NCT02960906/BIONIKK	nivolumab/ipilimumab/VEGFR-TKI	-	II	ccRCC	ORR	first line	recruiting
NCT03469713/NIVES	Nivolumab + SBRT	-	II	ccRCC	ORR	at least second line	recruiting
NCT02446860/ADAPTeR	nivolumab	-	II	ccRCC	Safety	pre and post-operative	recruiting
NCT02819596/Calypso	Durvalumab +/tremelimumab/savolitinib	-	II	cc/pRCC	DLT, OR	at least second line	recruiting
NCT03308396	Durvalumab + guadecitabine	-	I/II	ccRCC	safe dose/ORR	at least second line	recruiting
NCT02989714	Nivolumab + interleukin-2	-	I/II	ccRCC	safety	third line	recruiting
NCT03024996/IMmotion010	atezolizumab	placebo	III	ccRCC	DFS	adjuvant	recruiting
NCT03055013	nivolumab	observation	III	ccRCC	DFS	adjuvant	recruiting
NCT03138512/CheckMate 914	Nivolumab + ipilimumab	placebo	III	ccRCC	DFS	adjuvant	recruiting
NCT03142334/Keynote-564	pembrolizumab	placebo	III	ccRCC	DFS	adjuvant	recruiting
NCT02575222	nivolumab	-	I	ccRCC	safety	neo-adjuvant	active not recruiting
NCT03177239/ANZUP1602	Nivolumab + ipilimumab	-	II	nccRCC	ORR	all lines	recruiting
NCT03075423/SUNIFORECAST	Nivolumab + ipilimumab	sunitinib	II	nccRCC	OS	first line	recruiting

Abbreviations: ccRCC, clear cell renal cell carcinoma; nccRCC, non clear cell renal cell carcinoma; OS, overall survival; OR, overall response; PFS, progression-free survival; ORR, objective response rate; DLT, Dose-Limiting Toxicities; -, absence of comparison group.

Nivolumab was then associated with another immune checkpoint inhibitor: Ipilimumab (anti-CTLA-4) (Table 3) [36–38]. In CheckMate 214 (NCT02231749), a phase III study by Motzer et al. confirmed the benefit of the nivolumab and ipilimumab association versus sunitinib in first line metastatic RCC [39]. Patients with intermediate and poor risk according to the International Metastatic Database Consortium (IMDC) score benefitted the most from the association of ipilimumab and nivolumab with a median OS not reached with nivolumab plus ipilimumab versus 26 months with sunitinib (HR = 0.63; 99.8% CI, 0.44 to 0.89;  $p < 0.001$ ). The objective response rate was 42% versus 27%, and the complete response rate was 9% versus 1%. By contrast, the response rate was significantly higher with sunitinib in favourable-risk patients. In an exploratory analysis, a longer PFS was observed in patients with PD-L1 expression. However, the OS and objective response rate (ORR) improvements were demonstrated irrespective of PD-L1 expression. The frequency of treatment-related adverse events was similar in the two arms with less grade 3 or 4 events in the nivolumab plus ipilimumab arm. Nevertheless, treatment-related adverse events leading to discontinuation occurred more frequently in the nivolumab plus ipilimumab group.

A phase 3 trial (NCT03260894) originally proposed the association of epacadostat to pembrolizumab compared to sunitinib. Epacadostat targets and binds to IDO1, an enzyme responsible for the oxidation of tryptophan into kynurenine (Figure 2). By inhibiting IDO1 and decreasing kynurenine in tumour cells, epacadostat restores and increases the proliferation and activation of immune cells, including dendritic cells, natural killer (NK) cells, and T-lymphocytes, as well as interferon production, and a reduction in tumour-associated regulatory T cells.

The effect of immunotherapy could be enhanced by radiotherapy. Tumour destruction by radiotherapy induced the immune response by releasing tumour antigens corresponding to the abscopal effect [40]. The addition of radiotherapy is currently being evaluated in a phase II trial (NCT03469713).

Similar trials in metastatic nccRCC are currently recruiting. In particular, the nivolumab plus ipilimumab association is being evaluated in a first line versus sunitinib in a phase II trial (NCT03075423). Pembrolizumab is also being evaluated in the cohort B of the Keynote 427 phase II trial. In pRCC, durvalumab is being evaluated in combination with savolitinib, a highly selective MET tyrosine kinase inhibitor, in the Calypso phase II trial (NCT02819596). Interestingly, in a retrospective study including metastatic RCC ( $n = 35$ ) with several histologies, nivolumab in monotherapy had an objective response rate of 20%, a median PFS of 3.5 months, and was well tolerated [41]. Out of the four collecting duct carcinomas, one achieved partial response.

It is not surprising that the important results achieved in these trials have encouraged the exploration of immunotherapy in other adjuvant/neo-adjuvant settings to reduce the incidence of progression. Nivolumab was investigated in pre- and post-operative settings in two phase I/II trials and in association with ipilimumab in CheckMate 914 (NCT03138512), a phase III trial, including randomized versus placebo patients with localized renal cell carcinoma who underwent radical or partial nephrectomy with high risk of recurrence. Pembrolizumab and atezolizumab are currently being evaluated in the phase III trials, IMmotion010 (NCT03024996) and Keynote 564 (NCT03142334), respectively.

**Table 3.** Comparison of pivotal phase III clinical trials with available results evaluating immune checkpoints inhibitors.

Study Name	Comparator Arms	N ITT	Median FU	OS ITT	HR (95% CI)	PFS ITT	HR (95% CI)	ORR (%)	CR (%)	Grade $\geq$ 3 (%)
CheckMate 214	Ipilumab + nivolumab sunitinib	550	25.2	Nr *	0.63 (0.44–0.89)	11.6 *	0.82 (0.64–1.05)	42 *	9 *	46
		546		26 *		8.4 *		27 *	1 *	35
JAVELIN Renal 101	Avelumab + axitinib sunitinib	442	11.6	na		13.8	0.69 (0.56–0.84)	51.4	3.4	71.2
		446	10.7	na		8.4		25.7	1.8	71.5
Keynote 426	Pembrolizumab + axitinib sunitinib	432	12.8	na		15.1	0.69 (0.57–0.84)	59.3	5.8	75.8
		429		na		11.1		35.7	1.9	70.6
IMmotion 151	Atezolizumab + bevacizumab sunitinib	454	15	na		11.2	0.83 (0.70–0.97)	37	5	40
		461		na		8.4		33	2	54

Abbreviations: ITT, intention-to-treat; OS, overall survival; FU, follow-up; HR, Hazard Ratio; PFS, progression-free survival; ORR, objective response rate; CR, complete response; nr, not reached; \* in intermediate and poor-risk patients according to IMDC score; nr, not reached; na, not available.

#### 4.2. Association of Immune Checkpoint Inhibitors and Tyrosine Kinase Inhibitors

Combining an immune checkpoint inhibitor and an antiangiogenic drug is a new strategy, with several ongoing trials and published results of phase III trials (Tables 3 and 4). The difficulty is in finding an association with acceptable toxicity. The Checkmate 016 (NCT01472081), a phase I study, evaluated the safety of the intuitive association of nivolumab plus sunitinib. It resulted in a high frequency of high-grade toxicities, limiting its use in future trials [42].

The phase III study, IMmotion 151 (NCT02420821), explored the after phase II study, IMmotion 150, atezolizumab in association with bevacizumab versus sunitinib [43]. Patients were stratified by PD-L1 status (<1% vs.  $\geq$ 1% PD-L1 expression on tumour-infiltrating immune cells). The coprimary endpoints were PFS in PD-L1+ patients and OS in intention-to-treat (ITT) patients. The secondary endpoints included PFS in ITT patients, the objective response rate (ORR), and duration of response (DOR). The first results were recently released and even if the OS data are not yet mature (median survival follow-up was 15 months), the PFS in ITT patients with PD-L1+ tumour was 11.2 months versus 8.4 months (HR = 0.74; 98.5% CI, 0.57 to 0.96;  $p = 0.02$ ), which is comparable with PFS in IIT patients irrespective of PD-L1 expression (11.2 months versus 8.4 months), with acceptable safety.

Already used as second-line treatment in metastatic RCC, axitinib, with less hepatotoxicity than sunitinib and pazopanib, could represent a better alternative in association with ICI [44]. The combination of avelumab and axitinib showed encouraging results in phase I and II trials [45]. The phase III trial, JAVELIN Renal 101 (NCT02684006), randomized patients in two arms: Avelumab plus axitinib versus sunitinib [46]. Median PFS was significantly longer in the avelumab plus axitinib group than in the sunitinib group: 13.8 months versus 8.4 months, respectively (HR = 0.69; 95% CI, 0.56 to 0.84;  $p < 0.001$ ). PFS and ORR benefits were seen in all subgroups irrespective of IMDC prognostic risk groups, PD-L1 status, and prior nephrectomy status. However, with a median follow-up of only 12 months, OS data are still pending.

The phase III trial, Keynote 426 (NCT02853331), compared pembrolizumab plus axitinib combination versus sunitinib in first line treatment and showed similar results [47]. The median PFS was longer in the pembrolizumab plus axitinib arm with 15.1 months versus 11.1 months (HR = 0.69; 95% CI, 0.57 to 0.84;  $p < 0.001$ ). The safety profile was comparable to the results of the JAVELIN Renal 101 trial. Interestingly, the benefit of pembrolizumab plus axitinib for OS, PFS, and ORR was observed in the entire population irrespective of the prognostic group and PD-L1 tumour expression.

Similar to the IMmotion 151 trial (NCT02853331), a phase II trial (NCT02724878) is evaluating the combination of atezolizumab plus bevacizumab in metastatic nccRCC. Nivolumab plus cabozantinib was also originally tested in this population (NCT03635892). Taking into account the particularity of TFE/translocation RCC, a phase II trial (NCT03595124) with the combination of nivolumab plus axitinib is currently recruiting.

**Table 4.** Clinical trials evaluating immune checkpoint inhibitors in association with antiangiogenic drugs .

NCT Number/Study Name	Targeting Agents	Comparison	Phase	Histology	Primary Endpoint	Therapy Setting	Status
NCT02684006/Javelin Renal 101	Avelumab + axitinib	sunitinib	III	ccRCC	PFS, OS	first line	published
NCT02853331/Keynote 426	Pembrolizumab + axitinib	sunitinib	III	ccRCC	PFS, OS	first line	published
NCT02420821/IMmotion 151	Atezolizumab + bevacizumab	sunitinib	III	ccRCC	PFS, OS	first line	active not recruiting
NCT02811861/CLEAR	Lenvatinib + everolimus/pembrolizumab	sunitinib	III	ccRCC	PFS	first line	recruiting
NCT03141177/CheckMate 9ER	Nivolumab + cabozantinib	sunitinib	III	ccRCC	PFS	first line	recruiting
NCT01984242/IMmotion 150	Atezolizumab +/- bevacizumab	sunitinib	II	ccRCC	PFS	first line	Published
NCT02014636/Keynote-018	Pembrolizumab + pazopanib	pazopanib/pembrolizumab	II	ccRCC	safety, efficacy	first line	active not recruiting
NCT02348008	Pembrolizumab + bevacizumab	-	I/II	ccRCC	safe dose, efficacy	at least second line	active not recruiting
NCT02348008	Pembrolizumab + bevacizumab	-	I/II	ccRCC	safe dose, efficacy	at least second line	active not recruiting
NCT03024437	Atezolizumab + bevacizumab + entinostat	-	I/II	ccRCC	safe dose, ORR	at least second line	recruiting
NCT03172754	nivolumab + axitinib	-	I/II	ccRCC	safety	at least second line	recruiting
NCT02493751/Javelin Renal 100	Avelumab + axitinib	-	I/II	ccRCC	DLT	first line	Published
NCT02501096	Pembrolizumab + lenvatinib	-	I/II	ccRCC	DLT, ORR	all lines	recruiting
NCT03200587	Avelumab + cabozantinib	-	Ib	ccRCC	safe dose	all lines	recruiting
NCT01472081/CheckMate 016	Nivolumab + sunitinib/pazopanib/ipilimumab	nivolumab	I	ccRCC	safety, tolerability	at least second line	Published
NCT02133742	Pembrolizumab + axitinib	-	I	ccRCC	DLT	first line	Published
NCT02724878	Atezolizumab + bevacizumab	-	II	nccRCC	ORR	all lines	recruiting
NCT03635892	Nivolumab + cabozantinib	-	II	nccRCC	ORR	all lines	recruiting
NCT03595124	Nivolumab + axitinib	-	II	tRCC	PFS	all lines	recruiting

Abbreviations: ccRCC = clear cell renal cell carcinoma, nccRCC = non clear cell renal cell carcinoma, pRCC = papillary renal cell carcinoma, tRCC = translocation renal cell carcinoma, OS = overall survival, PFS = progression-free survival, ORR = objective response rate, DLT = Dose-Limiting Toxicities, - absence of comparison group.

## 5. Biomarkers and their Limitations

### 5.1. Response Evaluation for Immunotherapy

The version 1.1 of RECIST (RECISTv1.1) is widely used for assessing treatment response [48]. However, these criteria need to be refined regarding the atypical pattern of response to ICI [49]. Indeed, the appearance of new lesions or the growing of established lesions considered by RECISTv1.1 as progressive disease can be observed before an objective response or stable disease. New criteria, all adapted from RECISTv1.1, such as immune-related response criteria (irRC) and, subsequently, irRECIST, immune-based therapeutics RECIST (iRECIST), and immune-modified RECIST (imRECIST), were therefore proposed to prevent misclassification of atypical response as early progression according to RECISTv1.1 [50–53]

Although it is increasingly accepted that response per immune-related criteria can more accurately assess benefits from immunotherapy, RECISTv1.1 remains the radiological evaluation standard in clinical trials using ICI. The endpoints defined by RECISTv1.1 could potentially impair the identification of predictive biomarkers. For instance, in the CheckMate 010 phase II trial, tumour cell PD-L1 expression was not associated with PFS or ORR defined by RECISTv1.1, but was associated with a longer median irPFS and higher irORR using irRECIST [54]. A high level of T cells with the CD8+ PD-1+ TIM3- LAG3- phenotype, corresponding to T cell activation, was associated with longer median irPFS and higher irORR. Notably, the combination of PD-L1 expression with high levels of T-cell activation identified three groups of patients with significantly different irPFS and irORR.

### 5.2. Immunohistochemical Markers

Predictive biomarker research to select patients eligible for immunotherapy has mainly focused on PD-L1 immunohistochemistry (IHC). Even if PD-L1 was confirmed to be a poor prognostic factor, its use as a predictive marker remains controversial [55]. First, several antibodies are available because a companion antibody corresponds to each drug (Table 1). Studies of the comparability demonstrated variable concordance with interchangeability of the 22C3 and 28-8, but a lower sensitivity of the SP142 assay for evaluating tumour cells [56]. Second, although most trials in RCC considered the percentage of tumour cells, trials evaluating atezolizumab focused on immune cells with different thresholds [43]. Third, PD-L1 is a dynamic biomarker, evolving over time, and its expression is remodelled by the use of antiangiogenic drugs. Fourth, intratumour heterogeneity of PD-L1 and its differential expression between primary RCC and metastases was demonstrated [57,58].

PD-L1 is well known as a ligand localized on the membrane surface. However, there is another isoform of PD-L1; a splice variant generated by splicing out exon 2 encoding an immunoglobulin variable domain (Igv)-like domain was described [59]. However, all other exons are retained without a frameshift. Consequently, the putative translated protein contains all other domains, including the transmembrane region, except for the Igv-like domain. This novel isoform shows a pattern of intracellular membrane that is not able to interact with the PD-1 receptor. Indeed, the anti-PD-L1 antibodies (clone E1L3N, research use only and Dako 28-8) has the ability to detect this non-functional isoform because they target a sequence common to all expressed isoforms within the fourth exon of the *PD-L1* gene [60]. An alternative method of detection of PD-L1, such as flow cytometry, could more accurately characterize the presence of PD-L1 ligand isoforms.

Although most studies used anti-PD-1 drugs, exploratory analyses focused on the ligand PD-L1 expression rather than the receptor PD-1. The expression of PD-1 in tumour-infiltrating lymphocytes was not associated with poor clinical outcomes [61]. If not reported alone as a biomarker in metastatic ccRCC, PD-1 was described in association with other markers to define a T cell activation phenotype correlated with the response to nivolumab [54].

In addition, PD-L2 is another ligand competing to bind to PD-1. While its structure is quite similar to PD-L1, the exact function of PD-L2 is unclear, but it could play a role in regulating the T cell immune

response and immune tolerance. PD-L2 expression was previously described in RCC and could be evaluated as a predictive biomarker as suggested in other malignancies [57,62].

Interestingly, the expression of PD-L1 was previously correlated to c-MET expression [62–64]. This association gives an argument for the combination of c-MET inhibitors and targeted immunotherapy. Indeed, a phase II trial evaluated the safety of the combination of cabozantinib and avelumab and a phase III trial evaluated its association to nivolumab versus sunitinib in first line treatment.

The tumour cells and immune cells of nccRCCs can also express PD-L1. In 101 primary nccRCC, PD-L1 expression on tumour cells and immune cells was observed in pRCC (10% and 60%), cRCC (5.6% and 36%), tRCC (30% and 90%), and CDC (20% and 100%) [65]. In concordance with the results on ccRCC, the expression of PD-L1 on tumour cells was associated with worse outcomes in nccRCC. The high expression of PD-L1 in CDC is favoured by the high lymphocytic infiltration and the increased expression of immune response genes in these tumours.

### 5.3. Gene Expression Signatures Related to ICI

Transcriptomic signatures have been developed for patients with first-line TKI (sunitinib or pazopanib) and four different molecular subgroups associated to response and survival have been identified [66,67].

In IMmotion150, exploratory analyses demonstrated that angiogenesis, T-effector/IFN- $\gamma$  response, and myeloid inflammatory gene expression signatures were associated with clinical outcomes. In particular, atezolizumab monotherapy or therapy combined with the bevacizumab arm was more effective in tumours with the T-effector/IFN- $\gamma$  response gene expression signature [43].

Interestingly, the BIONIKK randomized phase II trial originally stratified patients according to the transcriptomic signatures of ccRCC (ccrcc1, 2, 3, 4) according to Beuselinck et al. [66]. For instance, the ccrcc4 molecular subtype morphologically showed sarcomatoid differentiation with a strong inflammatory, Th1-orientation, which promotes cytotoxic responses, but a suppressive immune micro-environment with high expression of PD-1 and its ligands. In this molecular subgroup, nivolumab alone or combined with ipilimumab was administered.

### 5.4. Tumour Mutational Burden

One emerging biomarker of the response to ICI is the tumour mutational burden (TMB), defined by the total number of mutations per coding area of the tumour genome [68]. Among mutations, insertions/deletions, a rich source of immunogenic neoantigens, are more frequently observed in RCC. One hypothesis is that RCC with the worst clinical prognosis better responds to ICI due to a higher mutational load. Unfortunately, differently than expected, the mutational load does not seem to correlate with prognostic groups (IMDC and MSKCC) [69].

Regarding clinical trials, the exploratory biomarker analyses of IMmotion150 indicated that TMB was not associated with PFS [43]. A phase II trial, NIVES (NCT03469713), currently recruiting, is evaluating nivolumab plus radiotherapy and will assess the TMB in RCC and its impact on the response to therapy.

Regarding nccRCC, median mutation tumour burdens in the pRCC and ccRCC range are comparable between 1 and 10 per megabase. In contrast, cRCC displays a lower tumour burden between 0.1 and 1 mutation per megabase [70].

## 6. Conclusions

Renal cell carcinoma encompasses distinct diseases and molecular alterations with ccRCC the most frequent entity characterized by its angiogenesis and its immune micro-environment targeted by antiangiogenic drugs and ICI.

In metastatic ccRCC, unprecedented OS, PFS, and response rates were reached using a combination of ICI (anti-CTLA-4 and anti-PD1) for intermediate and poor risk patients. The PFS was even improved by the combination of ICI with antiangiogenic drugs, even for favourable risk

patients, with preliminary results of pivotal phase III clinical trials recently released. However, the best complete response rate was reached with ICI combination. Future trials will have to change the control arm to ipilimumab plus nivolumab or ICI plus antiangiogenic drugs instead of sunitinib. In adjuvant settings, ICI are also promising, with ongoing clinical trials.

Even if clinical trials evaluating ICI are also developed in nccRCC, the specificity of each entity is rarely taken into account. Indeed, because the entities are rare, they usually include different entities in the same trial. As proposed for tRCC, we emphasize the need of subtype specific clinical trials, taking into account the distinct molecular alterations in each entity, to allow a personalized approach.

The selection of patients who will respond specifically to ICI remains a challenge. Data from the clinical trials in RCC show that the use of PD-L1 expression alone is insufficient to predict treatment response. The clinical utility of mutational burden has not yet been demonstrated in RCC studies and is still under evaluation and the proportion of insertions and deletions could be more relevant in RCC [71]. The gene expression signatures are promising and currently developed for patients who benefited from ICI.

The future of predictive biomarkers will probably involve a combination of different biologic variables as it is clear that a single biomarker may not be enough to determine treatment response, especially in the context of combined drugs. These composite biomarkers will need to be prospectively validated in the context of therapeutic clinical trials.

**Author Contributions:** Conception and design: S.-F.K.-J., N.R.-L. Collection and assembly of data: S.-F.K.-J., A.D., J.S., B.L., F.D., M.-A.B.-R., R.M., K.B., G.V. Manuscript writing: A.D., J.S., S.-F.K.-J. Final approval of the manuscript: all authors. Accountable for all aspects of the work: all authors.

**Funding:** This research received no external funding.

**Acknowledgments:** The authors would like to acknowledge the Ligue Contre le Cancer, the COmité de la REcherche Clinique et Translationnelle (CORECT), for their financial aid.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D.M.; Forman, D.; Bray, F. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **2015**, *136*, E359–E386. [[CrossRef](#)]
2. Levi, F.; Ferlay, J.; Galeone, C.; Lucchini, F.; Negri, E.; Boyle, P.; La Vecchia, C. The changing pattern of kidney cancer incidence and mortality in Europe. *BJU Int.* **2008**, *101*, 949–958. [[CrossRef](#)]
3. King, S.C.; Pollack, L.A.; Li, J.; King, J.B.; Master, V.A. Continued increase in incidence of renal cell carcinoma, especially in young patients and high grade disease: United States 2001 to 2010. *J. Urol.* **2014**, *191*, 1665–1670. [[CrossRef](#)]
4. Moch, H.; Cubilla, A.L.; Humphrey, P.A.; Reuter, V.E.; Ulbright, T.M. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. *Eur. Urol.* **2016**, *70*, 93–105. [[CrossRef](#)]
5. Malouf, G.G.; Comperat, E.; Yao, H.; Mouawad, R.; Lindner, V.; Rioux-Leclercq, N.; Verkarre, V.; Leroy, X.; Dainese, L.; Classe, M.; et al. Unique Transcriptomic Profile of Collecting Duct Carcinomas Relative to Upper Tract Urothelial Carcinomas and other Kidney Carcinomas. *Sci. Rep.* **2016**, *6*, 30988. [[CrossRef](#)] [[PubMed](#)]
6. Kaelin, W.G., Jr. The von Hippel-Lindau tumor suppressor protein and clear cell renal carcinoma. *Clin. Cancer Res.* **2007**, *13 Pt 2*, 680s–684s. [[CrossRef](#)]
7. Banumathy, G.; Cairns, P. Signaling pathways in renal cell carcinoma. *Cancer Biol.* **2010**, *10*, 658–664. [[CrossRef](#)]
8. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **2013**, *499*, 43–49. [[CrossRef](#)] [[PubMed](#)]
9. Cancer Genome Atlas Research Network; Linehan, W.M.; Spellman, P.T.; Ricketts, C.J.; Creighton, C.J.; Fei, S.S.; Davis, C.; Wheeler, D.A.; Murray, B.A.; Schmidt, L.; et al. Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma. *N. Engl. J. Med.* **2016**, *374*, 135–145.



10. Davis, C.F.; Ricketts, C.J.; Wang, M.; Yang, L.; Cherniack, A.D.; Shen, H.; Buhay, C.; Kang, H.; Kim, S.C.; Fahey, C.C.; et al. The somatic genomic landscape of chromophobe renal cell carcinoma. *Cancer Cell* **2014**, *26*, 319–330. [[CrossRef](#)]
11. Argani, P. MiT family translocation renal cell carcinoma. *Semin. Diagn. Pathol.* **2015**, *32*, 103–113. [[CrossRef](#)] [[PubMed](#)]
12. Malouf, G.G.; Monzon, F.A.; Couturier, J.; Molinie, V.; Escudier, B.; Camparo, P.; Su, X.; Yao, H.; Tamboli, P.; Lopez-Terrada, D.; et al. Genomic heterogeneity of translocation renal cell carcinoma. *Clin. Cancer Res.* **2013**, *19*, 4673–4684. [[CrossRef](#)] [[PubMed](#)]
13. Chen, F.; Zhang, Y.; Senbabaoglu, Y.; Ciriello, G.; Yang, L.; Reznik, E.; Shuch, B.; Micevic, G.; De Velasco, G.; Shinbrot, E.; et al. Multilevel Genomics-Based Taxonomy of Renal Cell Carcinoma. *Cell. Rep.* **2016**, *14*, 2476–2489. [[CrossRef](#)] [[PubMed](#)]
14. Amin, M.B.; Smith, S.C.; Agaimy, A.; Argani, P.; Comperat, E.M.; Delahunt, B.; Epstein, J.I.; Eble, J.N.; Grignon, D.J.; Hartmann, A.; et al. Collecting duct carcinoma versus renal medullary carcinoma: An appeal for nosologic and biological clarity. *Am. J. Surg. Pathol.* **2014**, *38*, 871–874. [[CrossRef](#)] [[PubMed](#)]
15. Srigley, J.R.; Delahunt, B.; Eble, J.N.; Egevad, L.; Epstein, J.I.; Grignon, D.; Hes, O.; Moch, H.; Montironi, R.; Tickoo, S.K.; et al. The International Society of Urological Pathology (ISUP) Vancouver Classification of Renal Neoplasia. *Am. J. Surg. Pathol.* **2013**, *37*, 1469–1489. [[CrossRef](#)] [[PubMed](#)]
16. Pal, S.K.; Choueiri, T.K.; Wang, K.; Khaira, D.; Karam, J.A.; Van Allen, E.; Palma, N.A.; Stein, M.N.; Johnson, A.; Squillace, R.; et al. Characterization of Clinical Cases of Collecting Duct Carcinoma of the Kidney Assessed by Comprehensive Genomic Profiling. *Eur. Urol.* **2016**, *70*, 516–521. [[CrossRef](#)] [[PubMed](#)]
17. Beckermann, K.E.; Sharma, D.; Chaturvedi, S.; Msaouel, P.; Abboud, M.R.; Allory, Y.; Bourdeaut, F.; Calderaro, J.; de Cubas, A.A.; Derebail, V.K.; et al. Renal Medullary Carcinoma: Establishing Standards in Practice. *J. Oncol. Pract.* **2017**, *13*, 414–421. [[CrossRef](#)] [[PubMed](#)]
18. Carlo, M.I.; Chaim, J.; Patil, S.; Kemel, Y.; Schram, A.M.; Woo, K.; Coskey, D.; Nanjangud, G.J.; Voss, M.H.; Feldman, D.R.; et al. Genomic Characterization of Renal Medullary Carcinoma and Treatment Outcomes. *Clin. Genitourin. Cancer* **2017**, *15*, e987–e994. [[CrossRef](#)] [[PubMed](#)]
19. Malouf, G.G.; Ali, S.M.; Wang, K.; Balasubramanian, S.; Ross, J.S.; Miller, V.A.; Stephens, P.J.; Khayat, D.; Pal, S.K.; Su, X.; et al. Genomic Characterization of Renal Cell Carcinoma with Sarcomatoid Dedifferentiation Pinpoints Recurrent Genomic Alterations. *Eur. Urol.* **2016**, *70*, 348–357. [[CrossRef](#)]
20. Gajewski, T.F.; Schreiber, H.; Fu, Y.X. Innate and adaptive immune cells in the tumor microenvironment. *Nat. Immunol.* **2013**, *14*, 1014–1022. [[CrossRef](#)] [[PubMed](#)]
21. Chen, D.S.; Mellman, I. Oncology meets immunology: The cancer-immunity cycle. *Immunity* **2013**, *39*, 1–10. [[CrossRef](#)] [[PubMed](#)]
22. Ferguson, T.A.; Choi, J.; Green, D.R. Armed response: How dying cells influence T-cell functions. *Immunol. Rev.* **2011**, *241*, 77–88. [[CrossRef](#)] [[PubMed](#)]
23. Franciszkiewicz, K.; Boissonnas, A.; Boutet, M.; Combadiere, C.; Mami-Chouaib, F. Role of chemokines and chemokine receptors in shaping the effector phase of the antitumor immune response. *Cancer Res.* **2012**, *72*, 6325–6332. [[CrossRef](#)] [[PubMed](#)]
24. Chen, D.S.; Irving, B.A.; Hodi, F.S. Molecular pathways: Next-generation immunotherapy—inhibiting programmed death-ligand 1 and programmed death-1. *Clin. Cancer Res.* **2012**, *18*, 6580–6587. [[CrossRef](#)] [[PubMed](#)]
25. Gabrilovich, D.I.; Nagaraj, S. Myeloid-derived suppressor cells as regulators of the immune system. *Nat. Rev. Immunol.* **2009**, *9*, 162–174. [[CrossRef](#)] [[PubMed](#)]
26. Gabrilovich, D.I.; Chen, H.L.; Girgis, K.R.; Cunningham, H.T.; Meny, G.M.; Nadaf, S.; Kavanaugh, D.; Carbone, D.P. Production of vascular endothelial growth factor by human tumors inhibits the functional maturation of dendritic cells. *Nat. Med.* **1996**, *2*, 1096–1103. [[CrossRef](#)] [[PubMed](#)]
27. Gabrilovich, D.; Ishida, T.; Oyama, T.; Ran, S.; Kravtsov, V.; Nadaf, S.; Carbone, D.P. Vascular endothelial growth factor inhibits the development of dendritic cells and dramatically affects the differentiation of multiple hematopoietic lineages in vivo. *Blood* **1998**, *92*, 4150–4166. [[PubMed](#)]
28. Henze, A.T.; Mazzone, M. The impact of hypoxia on tumor-associated macrophages. *J. Clin. Investig.* **2016**, *126*, 3672–3679. [[CrossRef](#)]
29. Sica, A.; Massarotti, M. Myeloid suppressor cells in cancer and autoimmunity. *J. Autoimmun.* **2017**, *85*, 117–125. [[CrossRef](#)]

30. Voron, T.; Colussi, O.; Marcheteau, E.; Pernot, S.; Nizard, M.; Pointet, A.L.; Latreche, S.; Bergaya, S.; Benhamouda, N.; Tanchot, C.; et al. VEGF-A modulates expression of inhibitory checkpoints on CD8+ T cells in tumors. *J. Exp. Med.* **2015**, *212*, 139–148. [[CrossRef](#)]
31. Tian, L.; Goldstein, A.; Wang, H.; Ching Lo, H.; Sun Kim, I.; Welte, T.; Sheng, K.; Dobrolecki, L.E.; Zhang, X.; Putluri, N.; et al. Mutual regulation of tumour vessel normalization and immunostimulatory reprogramming. *Nature* **2017**, *544*, 250–254. [[CrossRef](#)] [[PubMed](#)]
32. Thienpont, B.; Lambrechts, D. It's T Time for Normal Blood Vessels. *Dev. Cell* **2017**, *41*, 125–126. [[CrossRef](#)] [[PubMed](#)]
33. Fyfe, G.; Fisher, R.I.; Rosenberg, S.A.; Sznol, M.; Parkinson, D.R.; Louie, A.C. Results of treatment of 255 patients with metastatic renal cell carcinoma who received high-dose recombinant interleukin-2 therapy. *J. Clin. Oncol.* **1995**, *13*, 688–696. [[CrossRef](#)] [[PubMed](#)]
34. Negrier, S.; Escudier, B.; Lasset, C.; Douillard, J.Y.; Savary, J.; Chevreau, C.; Ravaud, A.; Mercatello, A.; Peny, J.; Mousseau, M.; et al. Recombinant human interleukin-2, recombinant human interferon alfa-2a, or both in metastatic renal-cell carcinoma. Groupe Francais d'Immunotherapie. *N. Engl. J. Med.* **1998**, *338*, 1272–1278. [[CrossRef](#)] [[PubMed](#)]
35. Motzer, R.J.; Escudier, B.; McDermott, D.F.; George, S.; Hammers, H.J.; Srinivas, S.; Tykodi, S.S.; Sosman, J.A.; Procopio, G.; Plimack, E.R.; et al. Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *N. Engl. J. Med.* **2015**, *373*, 1803–1813. [[CrossRef](#)] [[PubMed](#)]
36. Hammers, H.J.; Plimack, E.R.; Infante, J.R.; Rini, B.I.; McDermott, D.F.; Lewis, L.D.; Voss, M.H.; Sharma, P.; Pal, S.K.; Razak, A.R.A.; et al. Safety and Efficacy of Nivolumab in Combination With Ipilimumab in Metastatic Renal Cell Carcinoma: The CheckMate 016 Study. *J. Clin. Oncol.* **2017**, *35*, 3851–3858. [[CrossRef](#)] [[PubMed](#)]
37. Antonia, S.J.; Lopez-Martin, J.A.; Bendell, J.; Ott, P.A.; Taylor, M.; Eder, J.P.; Jager, D.; Pietanza, M.C.; Le, D.T.; de Braud, F.; et al. Nivolumab alone and nivolumab plus ipilimumab in recurrent small-cell lung cancer (CheckMate 032): A multicentre, open-label, phase 1/2 trial. *Lancet Oncol.* **2016**, *17*, 883–895. [[CrossRef](#)]
38. Larkin, J.; Chiarion-Sileni, V.; Gonzalez, R.; Grob, J.J.; Cowey, C.L.; Lao, C.D.; Schadendorf, D.; Dummer, R.; Smylie, M.; Rutkowski, P.; et al. Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *N. Engl. J. Med.* **2015**, *373*, 23–34. [[CrossRef](#)]
39. Motzer, R.J.; Tannir, N.M.; McDermott, D.F.; Aren Frontera, O.; Melichar, B.; Choueiri, T.K.; Plimack, E.R.; Barthelemy, P.; Porta, C.; George, S.; et al. Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma. *N. Engl. J. Med.* **2018**, *378*, 1277–1290. [[CrossRef](#)]
40. Xie, G.; Gu, D.; Zhang, L.; Chen, S.; Wu, D. A rapid and systemic complete response to stereotactic body radiation therapy and pembrolizumab in a patient with metastatic renal cell carcinoma. *Cancer Biol.* **2017**, *18*, 547–551. [[CrossRef](#)]
41. Koshkin, V.S.; Barata, P.C.; Zhang, T.; George, D.J.; Atkins, M.B.; Kelly, W.J.; Vogelzang, N.J.; Pal, S.K.; Hsu, J.; Appleman, L.J.; et al. Clinical activity of nivolumab in patients with non-clear cell renal cell carcinoma. *J. Immunother. Cancer* **2018**, *6*, 9. [[CrossRef](#)] [[PubMed](#)]
42. Amin, A.; Plimack, E.R.; Ernststoff, M.S.; Lewis, L.D.; Bauer, T.M.; McDermott, D.F.; Carducci, M.; Kollmannsberger, C.; Rini, B.I.; Heng, D.Y.C.; et al. Safety and efficacy of nivolumab in combination with sunitinib or pazopanib in advanced or metastatic renal cell carcinoma: The CheckMate 016 study. *J. Immunother. Cancer* **2018**, *6*, 109. [[CrossRef](#)] [[PubMed](#)]
43. McDermott, D.F.; Huseni, M.A.; Atkins, M.B.; Motzer, R.J.; Rini, B.I.; Escudier, B.; Fong, L.; Joseph, R.W.; Pal, S.K.; Reeves, J.A.; et al. Clinical activity and molecular correlates of response to atezolizumab alone or in combination with bevacizumab versus sunitinib in renal cell carcinoma. *Nat. Med.* **2018**, *24*, 749–757. [[CrossRef](#)] [[PubMed](#)]
44. Akaza, H.; Fukuyama, T. Axitinib for the treatment of advanced renal cell carcinoma. *Expert Opin. Pharm.* **2014**, *15*, 283–297. [[CrossRef](#)] [[PubMed](#)]
45. Choueiri, T.K.; Larkin, J.; Oya, M.; Thistlethwaite, F.; Martignoni, M.; Nathan, P.; Powles, T.; McDermott, D.; Robbins, P.B.; Chism, D.D.; et al. Preliminary results for avelumab plus axitinib as first-line therapy in patients with advanced clear-cell renal-cell carcinoma (JAVELIN Renal 100): An open-label, dose-finding and dose-expansion, phase 1b trial. *Lancet Oncol.* **2018**, *19*, 451–460. [[CrossRef](#)]
46. Motzer, R.J.; Penkov, K.; Haanen, J.; Rini, B.; Albiges, L.; Campbell, M.T.; Venugopal, B.; Kollmannsberger, C.; Negrier, S.; Uemura, M.; et al. Avelumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. *N. Engl. J. Med.* **2019**, *380*, 1103–1115. [[CrossRef](#)] [[PubMed](#)]

47. Rini, B.I.; Plimack, E.R.; Stus, V.; Gafanov, R.; Hawkins, R.; Nosov, D.; Pouliot, F.; Alekseev, B.; Soulieres, D.; Melichar, B.; et al. Pembrolizumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. *N. Engl. J. Med.* **2019**, *380*, 1116–1127. [[CrossRef](#)]
48. Eisenhauer, E.A.; Therasse, P.; Bogaerts, J.; Schwartz, L.H.; Sargent, D.; Ford, R.; Dancey, J.; Arbuck, S.; Gwyther, S.; Mooney, M.; et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur. J. Cancer* **2009**, *45*, 228–247. [[CrossRef](#)]
49. De Velasco, G.; Krajewski, K.M.; Albiges, L.; Awad, M.M.; Bellmunt, J.; Hodi, F.S.; Choueiri, T.K. Radiologic Heterogeneity in Responses to Anti-PD-1/PD-L1 Therapy in Metastatic Renal Cell Carcinoma. *Cancer Immunol. Res.* **2016**, *4*, 12–17. [[CrossRef](#)]
50. Seymour, L.; Bogaerts, J.; Perrone, A.; Ford, R.; Schwartz, L.H.; Mandrekar, S.; Lin, N.U.; Litiere, S.; Dancey, J.; Chen, A.; et al. iRECIST: Guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol.* **2017**, *18*, e143–e152. [[CrossRef](#)]
51. Hodi, F.S.; Ballinger, M.; Lyons, B.; Soria, J.C.; Nishino, M.; Tabernero, J.; Powles, T.; Smith, D.; Hoos, A.; McKenna, C.; et al. Immune-Modified Response Evaluation Criteria In Solid Tumors (imRECIST): Refining Guidelines to Assess the Clinical Benefit of Cancer Immunotherapy. *J. Clin. Oncol.* **2018**, *36*, 850–858. [[CrossRef](#)] [[PubMed](#)]
52. Le Lay, J.; Jarraya, H.; Lebellec, L.; Penel, N. irRECIST and iRECIST: The devil is in the details. *Ann. Oncol.* **2017**, *28*, 1676–1678. [[CrossRef](#)] [[PubMed](#)]
53. Wolchok, J.D.; Hoos, A.; O'Day, S.; Weber, J.S.; Hamid, O.; Lebbe, C.; Maio, M.; Binder, M.; Bohnsack, O.; Nichol, G.; et al. Guidelines for the evaluation of immune therapy activity in solid tumors: Immune-related response criteria. *Clin. Cancer Res.* **2009**, *15*, 7412–7420. [[CrossRef](#)] [[PubMed](#)]
54. Pignon, J.C.; Jegede, O.; Shukla, S.A.; Braun, D.A.; Horak, C.E.; Wind-Rotolo, M.; Ishii, Y.; Catalano, P.J.; Grosha, J.; Flaifel, A.; et al. irRECIST for the Evaluation of Candidate Biomarkers of Response to Nivolumab in Metastatic Clear Cell Renal Cell Carcinoma: Analysis of a Phase II Prospective Clinical Trial. *Clin. Cancer Res.* **2019**. [[CrossRef](#)] [[PubMed](#)]
55. Erlmeier, F.; Weichert, W.; Schrader, A.J.; Autenrieth, M.; Hartmann, A.; Steffens, S.; Ivanyi, P. Prognostic impact of PD-1 and its ligands in renal cell carcinoma. *Med. Oncol.* **2017**, *34*, 99. [[CrossRef](#)] [[PubMed](#)]
56. Tsao, M.S.; Kerr, K.M.; Kockx, M.; Beasley, M.B.; Borczuk, A.C.; Botling, J.; Bubendorf, L.; Chirieac, L.; Chen, G.; Chou, T.Y.; et al. PD-L1 Immunohistochemistry Comparability Study in Real-Life Clinical Samples: Results of Blueprint Phase 2 Project. *J. Thorac. Oncol.* **2018**, *13*, 1302–1311. [[CrossRef](#)] [[PubMed](#)]
57. Basu, A.; Yearley, J.H.; Annamalai, L.; Pryzbycin, C.; Rini, B. Association of PD-L1, PD-L2, and Immune Response Markers in Matched Renal Clear Cell Carcinoma Primary and Metastatic Tissue Specimens. *Am. J. Clin. Pathol.* **2019**, *151*, 217–225. [[CrossRef](#)] [[PubMed](#)]
58. Lopez, J.I.; Pulido, R.; Cortes, J.M.; Angulo, J.C.; Lawrie, C.H. Potential impact of PD-L1 (SP-142) immunohistochemical heterogeneity in clear cell renal cell carcinoma immunotherapy. *Pathol. Res. Pract.* **2018**, *214*, 1110–1114. [[CrossRef](#)]
59. He, X.H.; Xu, L.H.; Liu, Y. Identification of a novel splice variant of human PD-L1 mRNA encoding an isoform-lacking Igv-like domain. *Acta Pharm. Sin.* **2005**, *26*, 462–468. [[CrossRef](#)]
60. Cogswell, J.; Inzunza, H.D.; Wu, Q.; Feder, J.N.; Mintier, G.; Novotny, J.; Cardona, D.M. An Analytical Comparison of Dako 28-8 PharmDx Assay and an E1L3N Laboratory-Developed Test in the Immunohistochemical Detection of Programmed Death-Ligand 1. *Mol. Diagn.* **2017**, *21*, 85–93. [[CrossRef](#)]
61. Kim, K.S.; Sekar, R.R.; Patil, D.; Dimarco, M.A.; Kissick, H.T.; Bilen, M.A.; Osunkoya, A.O.; Master, V.A. Evaluation of programmed cell death protein 1 (PD-1) expression as a prognostic biomarker in patients with clear cell renal cell carcinoma. *Oncoimmunology* **2018**, *7*, e1413519. [[CrossRef](#)] [[PubMed](#)]
62. Shin, S.J.; Jeon, Y.K.; Kim, P.J.; Cho, Y.M.; Koh, J.; Chung, D.H.; Go, H. Clinicopathologic Analysis of PD-L1 and PD-L2 Expression in Renal Cell Carcinoma: Association with Oncogenic Proteins Status. *Ann. Surg. Oncol.* **2016**, *23*, 694–702. [[CrossRef](#)]
63. Kammerer-Jacquet, S.F.; Medane, S.; Bensalah, K.; Bernhard, J.C.; Yacoub, M.; Dupuis, F.; Ravaud, A.; Verhoest, G.; Mathieu, R.; Peyronnet, B.; et al. Correlation of c-MET Expression with PD-L1 Expression in Metastatic Clear Cell Renal Cell Carcinoma Treated by Sunitinib First-Line Therapy. *Target Oncol.* **2017**, *12*, 487–494. [[CrossRef](#)] [[PubMed](#)]

64. Lalani, A.A.; Gray, K.P.; Albiges, L.; Callea, M.; Pignon, J.C.; Pal, S.; Gupta, M.; Bhatt, R.S.; McDermott, D.F.; Atkins, M.B.; et al. Differential expression of c-Met between primary and metastatic sites in clear-cell renal cell carcinoma and its association with PD-L1 expression. *Oncotarget* **2017**, *8*, 103428–103436. [[CrossRef](#)]
65. Choueiri, T.K.; Fay, A.P.; Gray, K.P.; Callea, M.; Ho, T.H.; Albiges, L.; Bellmunt, J.; Song, J.; Carvo, I.; Lampron, M.; et al. PD-L1 expression in nonclear-cell renal cell carcinoma. *Ann. Oncol.* **2014**, *25*, 2178–2184. [[CrossRef](#)] [[PubMed](#)]
66. Beuselinck, B.; Job, S.; Becht, E.; Karadimou, A.; Verkarre, V.; Couchy, G.; Giraldo, N.; Rioux-Leclercq, N.; Molinie, V.; Sibony, M.; et al. Molecular Subtypes of Clear Cell Renal Cell Carcinoma Are Associated with Sunitinib Response in the Metastatic Setting. *Clin. Cancer Res.* **2015**, *21*, 1329–1339. [[CrossRef](#)] [[PubMed](#)]
67. Hakimi, A.A.; Voss, M.H.; Kuo, F.; Sanchez, A.; Liu, M.; Nixon, B.G.; Vuong, L.; Ostrovskaya, I.; Chen, Y.B.; Reuter, V.; et al. Transcriptomic Profiling of the Tumor Microenvironment Reveals Distinct Subgroups of Clear cell Renal Cell Cancer—Data from a Randomized Phase III Trial. *Cancer Discov.* **2019**. [[CrossRef](#)]
68. Yarchoan, M.; Hopkins, A.; Jaffee, E.M. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *N. Engl. J. Med.* **2017**, *377*, 2500–2501. [[CrossRef](#)]
69. De Velasco, G.; Miao, D.; Voss, M.H.; Hakimi, A.A.; Hsieh, J.J.; Tannir, N.M.; Tamboli, P.; Appleman, L.J.; Rathmell, W.K.; Van Allen, E.M.; et al. Tumor Mutational Load and Immune Parameters across Metastatic Renal Cell Carcinoma Risk Groups. *Cancer Immunol. Res.* **2016**, *4*, 820–822. [[CrossRef](#)]
70. Alexandrov, L.B.; Nik-Zainal, S.; Wedge, D.C.; Aparicio, S.A.; Behjati, S.; Biankin, A.V.; Bignell, G.R.; Bolli, N.; Borg, A.; Borresen-Dale, A.L.; et al. Signatures of mutational processes in human cancer. *Nature* **2013**, *500*, 415–421. [[CrossRef](#)]
71. Turajlic, S.; Litchfield, K.; Xu, H.; Rosenthal, R.; McGranahan, N.; Reading, J.L.; Wong, Y.N.S.; Rowan, A.; Kanu, N.; Al Bakir, M.; et al. Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: A pan-cancer analysis. *Lancet Oncol.* **2017**, *18*, 1009–1021. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



## **5.2. REVUE DE LA LITTERATURE -2**

Immunotherapy in Renal Cell Carcinoma The Future is now





Review

# Immunotherapy in Renal Cell Carcinoma: The Future Is Now

Antoine Deleuze <sup>1,2</sup>, Judikaël Saout <sup>1</sup>, Frédéric Dugay <sup>1,3</sup>, Benoit Peyronnet <sup>4</sup>, Romain Mathieu <sup>1,4</sup>, Gregory Verhoest <sup>4</sup>, Karim Bensalah <sup>4</sup>, Laurence Crouzet <sup>2</sup>, Brigitte Laguerre <sup>2</sup>, Marc-Antoine Belaud-Rotureau <sup>1,3</sup>, Nathalie Rioux-Leclercq <sup>1,5</sup> and Solène-Florence Kammerer-Jacquet <sup>1,5,\*</sup>

- <sup>1</sup> Université Rennes, Inserm, EHESP (Ecole des Hautes Etudes en Santé Publique), IRSET (Institut de recherche en santé, environnement et travail), UMR 1085, 35000 Rennes, France; antoine.deleuze@chu-rennes.fr (A.D.); judikael.saout.1@univ-rennes1.fr (J.S.); frederic.dugay@chu-rennes.fr (F.D.); romain.mathieu@chu-rennes.fr (R.M.); marc-antoine.belaud-rotureau@chu-rennes.fr (M.-A.B.-R.); nathalie.rioux-leclercq@chu-rennes.fr (N.R.-L.)
- <sup>2</sup> Department of Medical Oncology, Centre Eugene Marquis, 35000 Rennes, France; l.crouzet@rennes.unicancer.fr (L.C.); b.laguerre@rennes.unicancer.fr (B.L.)
- <sup>3</sup> Department of Cytogenetics, University Hospital, 35000 Rennes, France
- <sup>4</sup> Department of Urology, University Hospital, 35000 Rennes, France; benoit.peyronnet@chu-rennes.fr (B.P.); gregory.verhoest@chu-rennes.fr (G.V.); karim.bensalah@chu-rennes.fr (K.B.)
- <sup>5</sup> Department of Pathology, University Hospital, 35000 Rennes, France
- \* Correspondence: jacquet.sf@gmail.com; Tel.: +33-2-99-28-42-79; Fax: +33-2-99-28-42-84

Received: 2 March 2020; Accepted: 2 April 2020; Published: 5 April 2020



**Abstract:** Renal cell carcinoma is the third type of urologic cancer and has a poor prognosis with 30% of metastatic patients at diagnosis. The antiangiogenics and targeted immunotherapies led to treatment remodeling emphasizing the role of the tumour microenvironment. However, long-term responses are rare with a high rate of resistance. New strategies are emerging to improve the efficacy and the emerging drugs are under evaluation in ongoing trials. With the different treatment options, there is an urgent need to identify biomarkers in order to predict the efficacy of drugs and to better stratify patients. Owing to the limitations of programmed death-ligand 1 (PD-L1), the most studied immunohistochemistry biomarkers, and of the tumor mutational burden, the identification of more reliable markers is an unmet need. New technologies could help in this purpose.

**Keywords:** immunotherapy; immune checkpoint inhibitors; renal cell carcinoma; PD-1; PD-L1; ongoing trials; biomarkers; emerging drugs

## 1. Introduction

Renal cell carcinoma is the third urological cancer, representing 3% of all cancers in women and 5% in men with an incidence of around 400,000 cases worldwide [1]. The prognosis is poor: 30% of patients are metastatic at diagnosis and almost 30% of the remaining patients will develop metastases detected during the follow-up [2].

Clear cell renal cell carcinomas (ccRCC) represent the vast majority of RCC (around 75%). The other histologies mainly encompass papillary (20%) and chromophobe RCC (5%). The other entities are very rare including translocation-associated RCC, medullar RCC and collecting duct carcinoma. The histological entities based on distinct pathological features present different molecular alterations. Indeed, ccRCC are hallmarked by a frequent alteration of the *VHL* gene, a tumour suppressor gene, leading to angiogenesis through the transcription of genes regulated by *HIF* such as *VEGF* [3–7].



Non ccRCC (nccRCC) represent a heterogeneous group with papillary, chromophobe RCC and translocation RCC, the most frequent entities. Papillary RCC (pRCC) include tumours with indolent outcome (type 1) and more aggressive tumours (type 2) [8]. Type 1 and type 2 pRCC commonly harbor frequent *MET* alterations. However, alterations of *SETD2*, *CDKN2A*, *EGFR*, *NF2* and *TERT* have been described in type 2 and suggest the activation of MAP kinases pathway, cell cycle and deregulation of chromatin remodeling [9]. Chromophobe RCC (cRCC) are rarely metastatic and characterize mitochondrial alterations, frequently mutated *p53* and activation of mTOR pathway [9]. Translocation RCC (tRCC) harbor gene fusions involve *TFE3* and *TFEB*, members of the MiTF family [10]. These transcription factors have multiple partners, mainly involving messenger RNA splicing [11].

The other entities constitute less than 2% of renal tumours. The collecting duct carcinoma have been described as immunogenic tumours with high lymphocyte infiltration resulting from the upregulation of genes involved in T-cell activation and proliferation [12]. Renal medullary carcinoma present a frequent loss *IN11* (*SMARCB1*) implicated in the chromatin remodeling complex [13]. Among familial RCC syndromes, patients with hereditary leiomyomatosis and renal cell carcinoma (HLRCC) syndrome harbor *Fumarate Hydratase* (FH) germline mutation and develop clinical aggressive tumours [14]. The FH mutation by inactivating the enzyme alters the function of the Krebs cycle.

The sarcomatoid component can be found in all the histologic subtypes and demonstrates an increased tumour mutation burden (TMB) with high frequency of *p53*, *CDKN2A* and *NF2* mutations and also genes involved the chromatin remodeling such as *ARID1A* and *BAP1* [15].

The treatment and management of metastatic RCC have radically changed over the past 20 years [16]. Initially, first-generation immunotherapy with cytokines: interleukins or interferon represented standard approaches but with poor results [17,18]. The development of tyrosine kinase inhibitors, mainly vascular endothelial growth factor (VEGF) receptor inhibitors, largely improved the prognosis of both progression free survival (PFS) and overall survival (OS) [19].

The emergence of immune checkpoint inhibitors (ICI) alone or in combination (anti-cytotoxic T-lymphocyte antigen-4 (CTLA4) and anti-programmed death 1 (PD-1)) showed interesting results [20,21]. Targeted immunotherapy is an alternative to antiangiogenics because ccRCC is also considered an immunogenic tumour with high numbers of immune cells such as tumour-infiltrating lymphocytes (TIL) [22–24]. Recent trials proposed antiangiogenics in association with targeted immunotherapy to overcome resistance emphasizing the role of the tumour microenvironment (TME) and this strategy is currently an option in first line treatment [25,26].

Mechanisms of resistance with ICI can be primary or innate and secondary or acquired [27]. They encompass neo-antigen loss, defect of antigen presentation, alternative immune checkpoints and defective interferon signalling. Interferon- $\gamma$  is a major mechanism of resistance by enhancing programmed death-ligand 1 PD-L1 expression and inducing the expression of immune inhibitory molecules [28]. Other immune checkpoints such as TIM-3, LAG-3 and TIGIT play a role in the resistance by inhibiting antitumour immune response [29]. Novel therapeutic approaches try to overcome these mechanisms of resistance and are under evaluation in ongoing trials.

Identifying biomarkers is the key to better select treatments, reduce costs and improve survival in patients with metastatic kidney cancer. However, the limitations of the most studied biomarkers: PD-L1 immunohistochemistry and TMB make necessary the identification of robust markers. New technologies could help in this purpose.

In this comprehensive review, we will discuss: 1. the specificities of the TME in RCC, 2. the treatment update with the results of recent trials, 3. the emerging drugs used in ongoing trials, 4. the predictive biomarkers and 5. the novel technologies.

## 2. Specificities of the Tumour Microenvironment in Renal Cell Carcinoma

### 2.1. Vascular Component

Angiogenesis has been described to play an important role in the progression of RCC and leads to the recruitment of endothelial cells. Recent data suggest that endothelial cells in TME differ from normal endothelial cells [30]. Akino et al. identified aneuploidy in one third of endothelial cells freshly dissociated from RCC [31]. This warrants further investigation as it could impact the response to antiangiogenic drugs. Moreover, Edeline et al. demonstrated two different angiogenic phenotype described as mature and immature. These two patterns could coexist within the same tumour demonstrating the heterogeneity of the vascular component [32]. Dufies et al. demonstrated in experimental tumours that sunitinib stimulated the development of lymphatic vessels. Indeed, these vessels are crucial for the recruitment of immune cells [33].

### 2.2. Immune Component

The expression of PD-L1 is widely represented in RCC suggesting the important role of PD-1/PD-L1 checkpoint with the aberrant expression of tumour cells. Indeed, PD-L1 expression was reported in 23% of ccRCC, 10% of pRCC, 5.6% of cRCC, 30% of t RCC and 20% in collecting duct carcinoma [34,35]. An amplification of 9p24.1, locus of PD-L1, was recently identify in RCC with sarcomatoid component leading to PD-L1 constitutive expression [36].

The immune compartment mainly include T cells, NK cells, B cells, macrophages and dendritic cells with complex interactions. Recently, Chevrier and colleagues used mass cytometry to compile an atlas of immune cells from 73 RCC identifying 22 T cell and 17 tumour-associated macrophage phenotypes with distinct immune composition between tumours [37]. Some macrophages such as M-11 or M-13 were associated with a worse prognosis and could constitute new targets. The phenotype of T cells (CD8+) regarding the expression of immune checkpoints (PD-1, LAG-3, Tim-3) identified the immune-regulated profile as more aggressive in a cohort of 40 RCC [38]. The role of B cells in the TME is unclear with both anti- and pro-tumoral effects. In ccRCC, the density of B cells identified by immunohistochemistry was associated with poor prognosis [39]. Moreover, the B cell signature was correlated with poor prognosis in The Cancer Genome Atlas (TCGA) ccRCC cohort. B cells could exert pro-tumoral functions through different mechanisms such as secretion of immune-regulatory cytokines that affect T cells and macrophages.

To understand the immunomodulation of nivolumab, Chouieri et al. explored the morphological and molecular changes of mRCC ( $n = 91$ ) at screening and on treatment [40]. Immunohistochemical analysis revealed an increase in CD3+, CD8+ and CD4+ lymphocytes. No consistent change was observe in the expression of PD-L1 in tumour cells. Transcriptional analysis identified the up-regulation of genes stimulated by interferon  $\gamma$  with high levels of related chemokine in peripheral blood. A study case reported a histological complete response after nivolumab in metastatic ccRCC [41]. Only fibrotic changes and CD8+ lymphocytes were detected after treatment.

## 3. Treatment Update in Renal Cell Carcinoma

### 3.1. First-Line Treatment

The ICI are the new backbone in the therapeutic landscape of renal cancer alongside tyrosine kinase inhibitors, Figure 1. Innovative combinations of ICI or ICI with TKI are now part of the treatment strategy and based on the results of recently published phase III trials in first line setting, Table 1 [20,25,26,42].

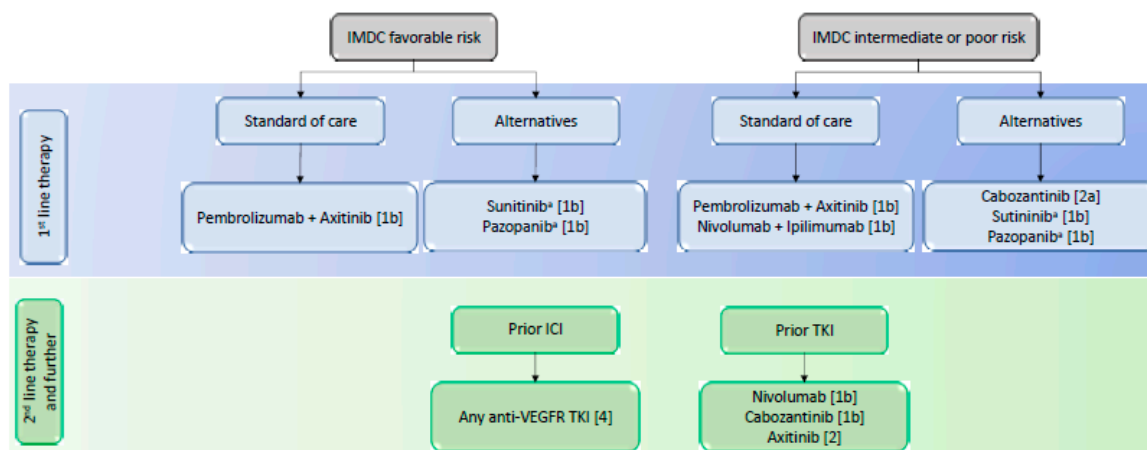
The Checkmate 214 trial confirmed the benefit of the nivolumab and ipilimumab association versus sunitinib in first line metastatic ccRCC among patients with intermediate or poor prognostic risk according to the International Metastatic Database Consortium (IMDC) risk model [20]. The objective response rate (ORR) was 42% versus 27%, and the complete response rate (CRR) was 11% versus 1% ( $p < 0.001$ ). No differences were observed in terms of PFS. A recent update showed an interesting

increase in OS after 30 months of follow-up in favor of nivolumab (anti-PD-1) plus ipilimumab (anti-CTLA4) combination (60% vs. 47%; HR: 0.66 ; CI 0.54 to 0.80 ;  $p < 0.0001$ ) [43].

Recent updates also highlighted the benefit of ICI and TKI in combination. In Keynote 426, Pembrolizumab (anti-PD-1) plus axitinib showed a benefit in terms of OS at 12 month follow-up (90% vs. 78% ; HR: 0.53 ; CI: 0.38–0.74 ;  $p < 0.0001$ ) leading to Food and Drug Administration (FDA) approval of the association in first line [26]. The Javelin 101 Renal comparing avelumab (anti-PD-L1) plus axitinib versus sunitinib in the PD-L1 positive population, defined as  $> 1\%$  of positive immune cells staining within the tumour area, demonstrated a longer PFS, 13.8 months versus 8.4 months (HR = 0.69; 95% CI, 0.56 to 0.84;  $p < 0.001$ ) and an improvement of the ORR (55.2% vs. 25%) [25]. After a 12 month follow-up, OS was not significantly different between the two arms. Similarly, the IMmotion 151 trial explored the atezolizumab (anti-PD-L1) and bevacizumab association regarding PD-L1 expression by immunohistochemistry (IHC) with a 1% cut-off on tumour-infiltrating immune cells [42]. The PFS was 11.2 months versus 8.4 months (HR = 0.74; 98.5% CI, 0.57 to 0.96;  $p = 0.02$ ), but there was no statistical difference in OS at 24 months follow-up.

Among the recent phase III trials released, the distribution of patients in IMDC differed [20,25,26,42]. For instance, Checkmate 214 trial focused on intermediate- and poor-risk populations. The geography of recruitment may have an effect as the Keynote 426 trial had a large proportion of patients from outside the USA and Western Europe. Finally, RR and PFS are not strictly comparable. Moreover, follow-up remains short and it is difficult to formally conclude at this stage.

Considering the differences among trial populations and the lack of mature OS, the choice of first line therapy can be challenging. To date, guidelines for treatment decision rely on the IMDC risk model to stratify patients with untreated mRCC or previously treated with first-line targeted therapies [16,44]. However, the interest of this classification remains unclear with the advent of ICI combination therapies and patient characteristics stay major indicators in the absence of prospectively validated biomarkers. Indeed, a significant number of comorbidities could influence the treatment choice and should be confronted by the safety profile of ICI and TKI.



**Figure 1.** European Association of Urology Guidelines on Renal Cell Carcinoma. IMDC = International Metastatic Renal Cell Carcinoma Database Consortium; OS = overall survival; Oxford level of evidence: [1b] = based on one randomised controlled phase 3 trial; [2a] = based on one randomised controlled phase 2 trial; [2b] = subgroup analysis of a randomised controlled phase 3 trial; [4] = expert opinion. a = No OS benefit proven.

**Table 1.** Comparison of pivotal phase III clinical trials with available results evaluating immune checkpoints inhibitors.

Study Name	Tested Drugs	Comparison	Phase	Histology	Therapy Setting	OS (HR, 95% CI, <i>p</i> )	Median PFS (HR, 95% CI)	ORR (%)	CR (%)	Grade 3 (%)
Javelin Renal 101	avelumab + axitinib	sunitinib	III	ccRCC	First line	12-mo: 86% vs. 83% (0.78; 0.55–1.08; <i>p</i> = 0.14)	13.8 vs. 7.2 mo (0.61)	55.2 vs. 25.5	3.4 vs. 1.8	71.2 vs. 71.5
Keynote 426	pembrolizumab + axitinib	sunitinib	III	ccRCC	First line	12-mo: 90% vs. 78% (0.53; 0.38–0.74; <i>p</i> < 0.0001)	15.1 vs. 11.1 mo (0.69; 0.57–0.84)	59.3 vs. 35.7	5.8 vs. 1.9	75.8 vs. 70.6
CheckMate 214	nivolumab + ipilimumab	sunitinib	III	ccRCC	First line	30-mo : 60% vs, 47% (0.66; 0.54–0.80; <i>p</i> < 0.0001)	11.6 vs. 8.4 mo (0.82; 0.64–1,05)	42 vs. 29	9 vs. 1	47 vs. 64
Immotion 151	atezolizumab + bevacizumab	sunitinib	III	ccRCC	First line	24 mo : 63% vs. 60% (0.93; 0.76–1.14; <i>p</i> = 0.4751)	11.2 vs. 7.7 mo (0.74; 0.57–0.96)	43 vs. 25	9 vs. 4	40 vs. 54

Abbreviations: OS = overall survival; HR = Hazard Ratio; PFS = progression-free survival; ORR = objective response rate; CR = complete response; ccRCC: clear cell renal cell carcinoma, mo = months.

### 3.2. Second-Line Treatment

The choice of treatment in the second-line setting and beyond depends on the therapy previously received. In case of progression after antiangiogenics, the Checkmate 025 trial comparing nivolumab and everolimus showed an improvement of median OS in the nivolumab arm (25 vs. 19.6 months; HR = 0.73; 98.5% CI, 0.57 to 0.93;  $p = 0.0018$ ) [45]. Notably, PD-L1 expression by IHC was associated with a poorer survival, but was not predictive of nivolumab efficacy. After first line ICI, any anti-angiogenic drug was recommended by the European Association of Urology with a low level of relevance [16]. However, due to the release of recent results, there is no prospectively validated data on the best therapeutic option after ICI and TKI combination in the front-line setting.

### 3.3. Adjuvant treatment

Finally, several ongoing phase III trials are studying the efficacy of ICI alone (atezolizumab, Pembrolizumab, nivolumab and tremelimumab) or in combination (nivolumab + ipilimumab) in the challenging peri-operative and adjuvant settings. Unsurprisingly, there is no combination of ICI plus TKI under evaluation as S-TRAC trial was the only one to show a benefit of adjuvant sunitinib in locally advanced high-risk ccRCC [46].

### 3.4. Non-Clear Cell Renal Cell Carcinoma

Unlike ccRCC, the management of nccRCC remains unclear and clinical trials are preferred [47]. Indeed, few prospective data for nccRCC treatment are available and trials have shown a lower efficacy of anti-angiogenic therapy compared to ccRCC with an impact on the prognosis of both PFS and OS [48]. In a recent retrospective trial including 41 patients, nivolumab seemed to demonstrate its efficacy with a favourable safety profile [49]. Indeed, the ORR was 20% and the median PFS was 3.5 months (95% CI; 1.9–5.0 months). Median OS was not reached, and the overall survival at the 10-month time point from the start of nivolumab treatment was 68%. Behind ccRCC, several trials evaluating ICI targeted drugs alone or in combination with TKI are currently recruiting in nccRCC without distinction of the entities, as shown in Tables 2 and 3. Interestingly, the combination of durvalumab and savolitinib (MET TKI) is specifically evaluated in pRCC. Likewise, nivolumab plus axitinib is studied in tRCC. Of note, ongoing trials using emerging drugs or vaccinal strategies include de facto nccRCC under the non-restrictive RCC inclusion criteria, Table 4.

## 4. Emerging Drugs in Ongoing Trials Include Renal Cell Carcinoma

### 4.1. Inhibitory Immune Checkpoints

The anergy of T cell is commonly understood as the result of a sequential transduction of signals implicating immune checkpoints in the immune synapse, Figure 2. If CTLA4 and PD-1 are the most common and well-studied, others are emerging and could be implied in resistance to traditional ICI and are evaluated in clinical trials including RCC, as shown in Tables 2 and 3.

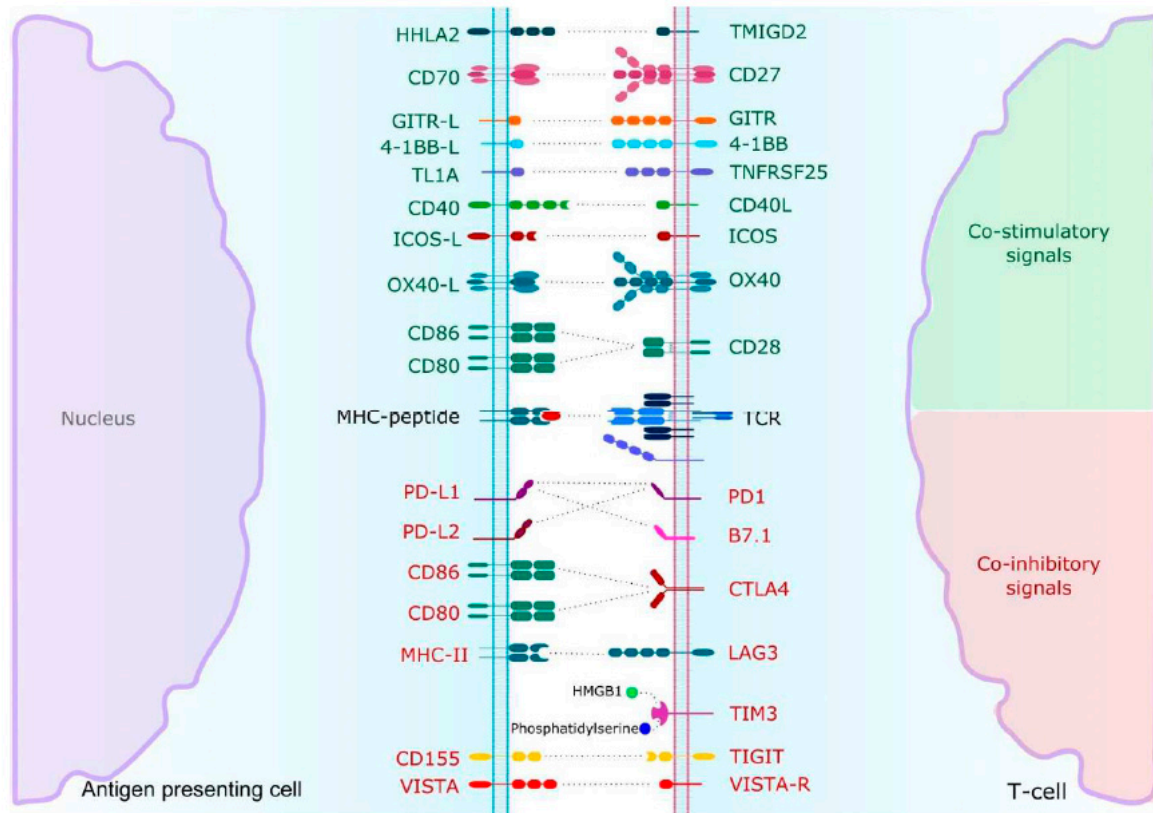
Lymphocyte-associated gene 3 (LAG3) is a transmembrane protein mainly expressed in activated T and natural killer (NK) cells, as shown in Figure 2 [29]. LAG3 is located on T cells. It shares a structural homology with CD4 and binds its ligand, the major histocompatibility complex class II (MHCII) with higher affinity. Moreover, the LAG3 blockade leads to an increased production of interferon gamma (INF $\gamma$ ), tumour necrosis factor alpha (TNF $\alpha$ ) and pro-inflammatory interleukins [50].

Other ICI are emerging. Among them, T cell immunoglobulin and mucin domain3 (TIM3) expressed on a wide variety of immune cells. TIM3 contributes to immune tolerance by inhibiting T cells activation, mostly by upregulation of apoptosis [51]. Interestingly, there could be a synergistic effect with PD-1-PD-L1 blockade, reversing T cell exhaustion and improving anti-tumoral immune response [52].

T cell immunoglobulin and ITIM domain (TIGIT) is mainly found on TIL and is immunosuppressive through disrupting interleukins production and antigen-presenting cell (APC) maturation [53]. The two

ligands Nectin-2 and CD155 are expressed in various cellular types from tumour to immune cells. Similar results in terms of T cells exhaustion were observed in the B7-H3 pathway [54].

The v-domain immunoglobulin suppressor of T cell activation (VISTA) is predominantly expressed on myeloid-derived suppressive cells (MDSC) and APC and down-regulates T cells activation [55]. Remarkably, VISTA blockade seemed to inhibit regulatory T cell immunosuppressive functions [56].



**Figure 2.** T-cell activation or inhibition is the result of an integrated and sequential intracellular signal cascade after MHC-peptide recognition by TCR. Abbreviations: GITR = Glucocorticoid-Induced TNF Receptor, ICOS = Inducible co-stimulator, OX40 = CD134, TCR = T-Cell Receptor, PD1 = Programmed Death 1, CTLA4 = Cytotoxic T-lymphocyte Antigen-4, LAG3 = Lymphocyte-associated gene 3, TIM3 = Tcell Immunoglobulin and Mucin domain-3, TIGIT = T-cell Immunoglobulin and ITIM domain, VISTA = Vdomain Immunoglobulin Suppressor of T-cell Activation.

**Table 2.** Clinical trials evaluating immune checkpoint inhibitors in association with tyrosine kinase inhibitors.

NCT Number	Targeting Agents	Comparison	Phase	Histology	Primary Endpoint	Therapy Setting	Status
NCT03260894	pembrolizumab + epacadostat	sunitinib or pazopanib	III	ccRCC	ORR	First line	Active not recruiting
NCT02811861	lenvatinib + everolimus or pembrolizumab	sunitinib	III	ccRCC	PFS	First line	Active not recruiting
NCT03141177	nivolumab + cabozantinib	sunitinib	III	ccRCC	PFS	First line	Active not recruiting
NCT03793166	nivolumab, ipilimumab, cabozantinib	nivolumab or nivolumab + cabozantinib	III	ccRCC	OS	First line	Recruiting
NCT03937219	nivolumab + ipilimumab + cabozantinib	nivolumab + ipilimumab + placebo	III	ccRCC	DFS	First line	Recruiting
NCT03680521	sitravatinib + nivolumab	–	II	ccRCC	ORR	First line	Recruiting
NCT02960906	nivolumab, ipilimumab, VEGFR-TKI	–	II	ccRCC	ORR	First line	Recruiting
NCT03736330	axitinib + pembrolizumab + D-CIK	–	II	ccRCC	ORR	At least second line	Recruiting
NCT02819596	savolitinib, durvalumab, tremelimumab	–	II	ccRCC, pRCC	DLT, ORR	At least second line	Unknown
NCT02964078	pembrolizumab + interleukin-2	–	II	ccRCC	ORR	Any	Active not recruiting
NCT03092856	anti-OX40 agonist antibody + axitinib	–	II	ccRCC	PFS	No standard anymore available	Recruiting
NCT02724878	atezolizumab + bevacizumab	–	II	nccRCC	ORR	Any	Active not recruiting
NCT03635892	nivolumab + cabozantinib	–	II	nccRCC	ORR	Any	Recruiting
NCT03595124	nivolumab + axitinib	–	II	tRCC	PFS	Any	Recruiting
NCT02493751	avelumab + axitinib	–	I/II	ccRCC	DLT	First line	Published
NCT02899078	ibrutinib + nivolumab	–	I/II	ccRCC, nccRCC	PFS	At least second line	Recruiting
NCT02348008	pembrolizumab + bevacizumab	–	I/II	ccRCC	Safety, efficacy	At least second line	Active not recruiting
NCT03172754	nivolumab + axitinib	–	I/II	ccRCC	Safety	At least second line	Recruiting
NCT03024437	atezolizumab + entinostat + bevacizumab	atezolizumab + entinostat	I/II	ccRCC	Safety, ORR	At least second line	Recruiting
NCT02501096	pembrolizumab + lenvatinib	–	I/II	ccRCC	DLT, ORR	No standard anymore available	Active not recruiting
NCT01472081	nivolumab + sunitinib or pazopanib TSR-022	nivolumab	I	ccRCC	Safety	At least second line	Published
NCT03307785	(anti-TIM3)/niraparib/TSR-042	–	I	RCC and others	Safety	At least second line	Active not recruiting
NCT03200587	(anti-PD1)/chemotherapy/bevacizumab avelumab + cabozantinib	–	Ib	ccRCC	Safety	Any	Recruiting

Abbreviations: OS = overall survival; HR = Hazard Ratio; PFS = progression-free survival; ORR = objective response rate; CR = complete response; ccRCC = clear cell renal cell carcinoma; nccRCC = non clear cell renal cell carcinoma; DLT = dose-limiting toxicity; DFS = disease-free survival; ICI = immune checkpoint inhibitor; TKI = tyrosine kinase inhibitor; VEGFR: vascular endothelial growth factor receptor; \_ = na; SBRT = stereotactic body radiation.



**Table 3.** Clinical trials evaluating targeted immunotherapies alone or in combination.

NCT Number	Targeting Agents	Comparison	Phase	Histology	Primary Endpoint	Therapy Setting	Status
NCT03729245	NKTR-214 (IL2R agonist) + nivolumab	sunitinib or cabozantinib	III	ccRCC	ORR, OS	First line	Recruiting
NCT03873402	nivolumab + ipilimumab	nivolumab	III	ccRCC	PFS, ORR	First line	Recruiting
NCT01668784	nivolumab	everolimus	III	ccRCC	OS	At least second line	Published
NCT03055013	nivolumab	observation	III	ccRCC	DFS	Peri-operative	Recruiting
NCT03024996	atezolizumab	placebo	III	ccRCC	DFS	Adjuvant	Active not recruiting
NCT03138512	nivolumab + ipilimumab	placebo	III	ccRCC	DFS	Adjuvant	Recruiting
NCT03142334	pembrolizumab	placebo	III	ccRCC	DFS	Adjuvant	Active not recruiting
NCT03288532	nivolumab, tremelimumab	–	III	ccRCC	DFS	Adjuvant	Recruiting
NCT02996110	BMS-986205 (IDO1 oral inhibitor) +/- nivolumab +/- ipilimumab	nivolumab +/- ipilimumab	II	ccRCC, nccRCC	ORR, PFS	First line	Recruiting
NCT03552380	nivolumab + ipilimumab + entinostat	–	II	ccRCC	Safety	At least second line	Recruiting
NCT03501381	entinostat + IL2	IL2	II	ccRCC	PFS	At least second line	Recruiting
NCT03469713	nivolumab + SBRT	–	II	ccRCC	ORR	At least second line	Recruiting
NCT03177239	nivolumab + ipilimumab	–	II	nccRCC	ORR	Any	Active, not recruiting
NCT04262375	oleclumab (anti-CD73 antagonist mAb) + durvalumab	–	II	RCC and others	ORR, PFS	Any	Not yet recruiting
NCT03207867	NIR178 (A2aR antagonist) + spartalizumab (anti-PD1)	–	II	RCC and others	ORR	At least second line	Recruiting
NCT03693612	GSK3359609 (anti-ICOS) + tremelimumab	chemotherapies	II	RCC and others	DLT	No standard anymore available	Recruiting
NCT03693612	anti-ICOS + tremelimumab	–	II	RCC and others	Safety, DLT	No standard anymore available	Recruiting
NCT01038778	entinostat + aldesleukin	–	I/II	ccRCC	Dose, ORR	At least second line	Active, not recruiting
NCT03308396	durvalumab + guadecitabine	–	I/II	ccRCC	Safe dose/ORR	At least second line	Recruiting
NCT02989714	nivolumab + interleukin-2	–	I/II	ccRCC	Safety	Third line	Active not recruiting
NCT02460224	LAG525 (anti-LAG3) + spartalizumab (anti-PD1)	–	I/II	RCC and others	DLT, ORR	At least second line	Active, not recruiting
NCT03652077	anti-TIM3	–	I/II	RCC and others	Safety	No standard anymore available	Recruiting
NCT02608268	MBG453 (anti-TIM3) + spartalizumab (anti-PD1)	spartalizumab anti-PD1	I/II	RCC and others	Safety, ORR, DLT	No standard anymore available	Recruiting
NCT01968109	relatlimab (anti-LAG3) + nivolumab	relatlimab	I/II	RCC and others	Safety, ORR, DLT	No standard anymore available	Recruiting
NCT03126110	INCAGN01876 (anti-GITR) + nivolumab + ipilimumab	–	I/II	RCC and others	Safety, tolerability	No standard anymore available	Recruiting
NCT02335918	varlilumab (anti-CD27) + nivolumab	–	I/II	RCC and others	DLT, ORR	At least second line	Completed



Table 3. Cont.

NCT Number	Targeting Agents	Comparison	Phase	Histology	Primary Endpoint	Therapy Setting	Status
NCT02718066	HBI-8000 (HDACi) + nivolumab	–	I/II	RCC and others	RP2D	At least second line	Recruiting
NCT02890069	spartalizumab + LCL16 (IAP inhibitor) + everolimus + panobinostat	–	I/II	RCC and others	DLT	At least second line	Recruiting
NCT02771626	CB-839 (glutaminase inhibitor) + nivolumab	–	I/II	RCC and others	Safety, efficacy	At least second line	Recruiting
NCT02817633	TSR-022 (anti-TIM3)/TSR-042 (anti-PD1)/TSR-033 (anti-LAG3)	–	I	RCC and others	Safety, tolerability	No standard anymore available	Recruiting
NCT03119428	OMP-31M32 (anti-TIGIT) + nivolumab	–	I	RCC and others	DLT	No standard anymore available	Terminated
NCT00351949	IMP321 (anti-LAG3)	–	I	RCC and others	Safety, tolerability	No standard anymore available	Completed
NCT02386111	varlilumab (anti-CD27)	–	I	RCC and others	Safety, tolerability	At least second line	Terminated
NCT03343613	LY3300054 (IDO1 inhibitor) + PD-L1 inhibitor	–	I	RCC and others	DLT	No standard anymore available	Recruiting
NCT02812875	CA-170 (VISTA antagonist)	–	I	RCC and others	DLT	No standard anymore available	Active not recruiting
NCT02655822	ciforadenant (A2aR antagonist) +/- atezolizumab	–	I	RCC and others	DLT	At least second line	Recruiting
NCT04198766	anti-OX40 + pembrolizumab	–	I	RCC and others	Safety	No standard anymore available	Recruiting

Abbreviations: OS = overall survival; HR = Hazard Ratio; PFS = progression-free survival; ORR = objective response rate; CR = complete response; ccRCC = clear cell renal cell carcinoma; nccRCC = non clear cell renal cell carcinoma; DLT = dose-limiting toxicity; DFS = disease-free survival.

**Table 4.** Vaccinal strategies and CAR T-cells trials.

NCT Number	Targeting Agents	Comparison	Phase	Histology	Primary Endpoint	Therapy Setting	Status
NCT00458536	dendritic cell tumor fusion vaccine + GM-CSF	–	I/II	ccRCC, nccRCC	Safety	Any	Active not recruiting
NCT03633110	GEN-009 Adjuvanted Vaccine + nivolumab + pembrolizumab	–	I/II	RCC and others	Safety	At least second line	Recruiting
NCT00722228	Autologous or Allogeneic tumor cells	–	I/II	RCC and others	Safety, efficacy	At least second line	Recruiting
NCT03393936	anti-ROR2 CAR-T or anti AXL CART-T	–	I/II	ccRCC, nccRCC	Safety	No standard available	Recruiting
NCT02830724	anti-CD70 CAR-T	–	I/II	RCC and others	Safety	At least second line	Recruiting
NCT01218867	anti-VEGFR2 CAR-T	–	I/II	RCC and others	ORR	At least second line	Terminated
NCT03638206	anti-C-MET CAR-T	–	I/II	RCC and others	Safety	At least second line	Recruiting
NCT02950766	neovax vaccine + ipilimumab	–	I	ccRCC	DLT	At least second line	Recruiting
NCT00096629	PSMA DNA vaccine	–	I	ccRCC, nccRCC	Safety	Adjuvant	Completed
NCT03548467	VB10.NEO vaccine +/- bempegaldesleukin (NKTR-214)	–	I	RCC and others	Safety	At least second line	Recruiting
NCT03294083	pexastimogene devacirepvec (Pexa-Vec)	–	I	ccRCC, nccRCC	Safety, efficacy	At least second line	Recruiting
NCT03715985	EVAX-01-CAF09b (peptide-based vaccine) +/- anti-PD1 or anti-PD-L1	–	I	RCC and others	Safety, efficacy	First line	Recruiting

Abbreviations: OS = overall survival; HR = Hazard Ratio; PFS = progression-free survival; ORR = objective response rate; CR = complete response; ccRCC = clear cell renal cell carcinoma; nccRCC = non clear cell renal cell carcinoma; DLT = dose-limiting toxicity; DFS = disease-free survival; ICI = immune checkpoint inhibitor; TKI = tyrosine kinase inhibitor; VEGFR: vascular endothelial growth factor receptor; - = na; SBRT = stereotactic body radiation.

#### 4.2. Co-Activating Immune Checkpoints

Tumour-specific T cell-mediated immune response is balanced by both inhibitory and co-stimulatory signals. If ICI are mainly used to restore immune response, agonist drugs are developed to increase co-stimulatory signals and stimulate immune response, as shown in Figure 2. Like ICI, several co-activating immune checkpoints are used in ongoing trials, Tables 2 and 3.

On T cells, CD28 is known to deliver an activating signal when binding CD80/86 after TCR recognition of major histocompatibility complex (MHC). Inducible co-stimulator (ICOS) mainly located on CD4+ T cells belongs to the immunoglobulin family as well as CD28 and produces inflammatory cytokines. Its intracytoplasmic structure has a strong affinity for phosphoinositide 3-kinase (PI3K) favoring the proliferation signal in lymphocytes [57].

The tumour necrosis factor (TNF) receptor superfamily is represented by a group of both soluble and transmembrane receptors involved in inflammation processes and able to bind a variety of ligands such as TNF $\alpha$ , TNF $\beta$  and OX40 ligand [58]. When binding its ligand, OX40 promotes T cells proliferation and survival, particularly CD4+ and CD8+ T cells, by upregulating pro-inflammatory cytokines and anti-apoptotic molecules. Other member TNF receptors such as CD40, CD27 and 4-1BB contribute to increased cytotoxic T cells mediated response through apoptosis or memory-cell differentiation. The glucocorticoid-induced TNF receptor (GITR), located on CD4+ and CD8+ T cells and predominantly on FoxP3+ regulatory T cells, is known to enhance immunity to tumours through the attenuation of the effector activity of immunosuppressive regulatory T cells [59].

#### 4.3. Metabolic Pathways

Metabolic changes influence the TME by providing immunosuppressive metabolites and favoring tumour growth in hypoxic conditions, as shown in Figure 3 [60]. Multiple enzymes have been identified as key regulators of cytotoxic T cells immune response and are under evaluation in clinical trials, as shown in Tables 2 and 3.

Among them, indoleamine 2,3 dioxygenase 1 (IDO1), an intracellular enzyme, catalyses the conversion of tryptophan into kynurenine [61]. Although physiologically expressed in stromal and dendritic cells, it is overexpressed in MDSC and tumour cells. The tryptophan depletion induced by IDO1 expression leads to T cell exhaustion and apoptosis. Moreover, high concentration of kynurenine promotes immune-tolerant dendritic cells and regulatory T cell proliferation [62].

Adenosine is a purine base known to bind G-protein coupled adenosine receptors, upregulated in activated immune cells [63]. Adenosine 2a receptor (A2aR) triggers upregulation of adenylate cyclase activity leading to increased cyclic adenosine monophosphate (cAMP) concentration. It has a profound immunomodulatory effect on immune cells by various mechanisms including ZAP70 inhibition in TCR signalling, IL2 down-regulation, FoxP3 expression enhancement on regulatory T cells and TGF $\beta$  marked secretion [64].

As previously reported, TME acquires immunosuppressive features by reprogramming metabolic processes. Arginine is an essential amino acid in immune and tumour cells demonstrating a high level of arginase [65]. Arginase inhibitors are currently evaluated in combination with several ICI, Table 3.

#### 4.4. Other Strategies

Various innovative strategies in immunotherapy are applied for RCC. The histone conformation impacts the transcription and is regulated through phosphorylation, sumoylation, ubiquitination, acetylation and deacetylation [66]. Histone deacetylase (HDAC) inhibitors modify the chromatin accessibility and play a crucial role in cycle arrest and apoptosis and could eventually enhance tumour antigens release and indirectly improve antigen presentation by APC and T cells priming [67]. In RCC, entinostat, panobinostat and chidamide are under evaluation in combination with ICI or TKI, as shown in Tables 2 and 3.

Several vaccinal strategies are tested in cancers with the common objective to upregulate tumour neo-antigens exposure to immune system, particularly by T cells priming phase improvement [68]. At this time, three broad vaccine types are under investigation including DNA/RNA-based, peptide-based and cell-based vaccines with encouraging results in RCC. Oncolytic viruses are interesting alternatives, designed to infect tumour cells and hijack cellular machinery to induce transgene expression.

The chimeric Antigen Receptor (CAR) T cells are T cells genetically engineered to produce an artificial T cell receptor that combines both antigen-binding and T cell activating functions. The use of CART cells is hampered in RCC by the tumour heterogeneity. CAR T cells are currently tested in renal cell carcinoma, targeting various antigens such as ROR2, AXL, CD70, VEGFR2, MET or CAIX, as seen in Table 3. Moreover, pre-clinical data suggest the rationale of combining CAR T cells with TKI or radiotherapy [69].

## 5. Predictive Biomarkers in RCC

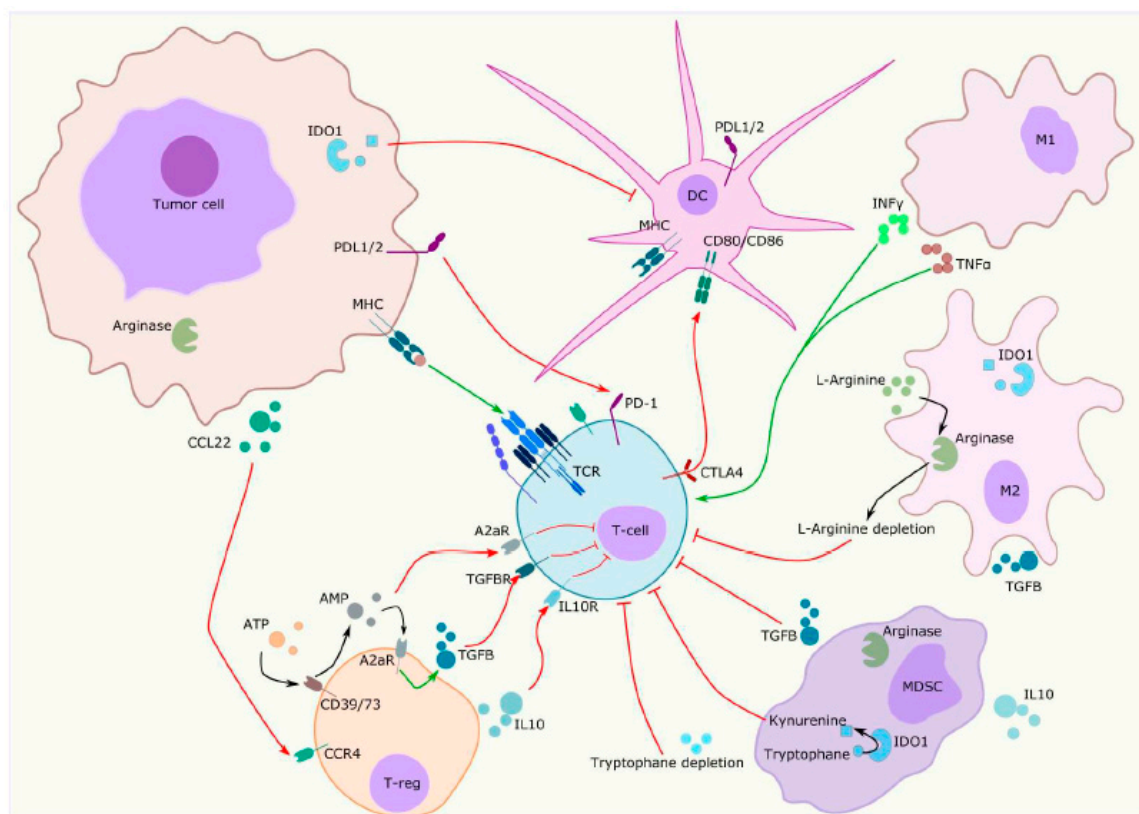
### 5.1. Clinico-Biological Biomarkers

The IMDC risk model includes six clinical and biological variables (poor Karnofsky performance status, less than 1 year between the diagnosis and the treatment, low hemoglobin concentration, high platelet count, high neutrophil count and high serum calcium) and has been validated in ccRCC and nccRCC [70,71]. However, the use of this classification is limited in nccRCC because the group is heterogeneous and samples in each group are small. The guidelines are dichotomized between favorable- and intermediate- or poor-risk disease, as shown in Figure 1 [16]. For instance, the clinical outcome was better in patients with intermediate- or poor-risk disease under nivolumab plus ipilimumab [20]. On the contrary, patients with favorable-risk disease better answered to sunitinib. These results suggest a distinct underlying biology.

### 5.2. Immunohistochemical Biomarkers

The most studied immunohistochemical biomarker PD-L1 failed to demonstrate a predictive capability in metastatic RCC [20]. If it was demonstrated as a poor prognostic factor in both ccRCC and nccRCC; indeed its ability to discriminate between patients the good responders is questionable [35]. Limitations to the use of PD-L1 have been well described and encompass intra-tumoral heterogeneity, variability of cut-offs and heterogeneous expression between primary and metastatic sites among others [72].

The contradictory and unresolved issue about PD-L1 as a biomarker also reveals the complicated interactions between the tumour and immune response, as shown in Figure 3. A comprehensive immune phenotype including other immunosuppressive factors such as TGF $\beta$  or IDO-1 and a better characterization of immune cells could be another step forward in the search of predictive biomarkers. Interestingly, IDO-1 expression was higher in endothelial cells of responders to nivolumab in a cohort of 15 patients with metastatic RCC [73]. This results in a reduced influx of tryptophan in the surrounding tumour tissue leading to a decrease in tumour proliferation.



**Figure 3.** Tumor microenvironment is commonly defined as the co-existence of tumor cells interacting with resident and infiltrating host cells, secreted factors and extracellular matrix proteins. Among them, immunosuppressive cells are recruited in the tumor microenvironment by chemotaxis and are responsible for immunomodulatory cytokines production enhancement and decreased essential amino acids availability, resulting in favorable conditions for tumor growth. MDSC = Myeloid-derived suppressive cells; T-cell = T lymphocyte; T-reg = Regulator lymphocyte; M2 = Type 2 macrophage; DC = dendritic cell; TGFβ = tumor growth factor β; A2aR = adenosine 2a receptor; ATP = adenosine triphosphate; AMP = adenosine monophosphate; IL10 = interleukin-10; MHC = major histocompatibility complex; PD1 = programmed death protein 1; PDL1 = programmed death ligand 1; IDO1 = indoleamine 2,3 dioxygenase 1; CTLA4 = cytotoxic T-lymphocyte antigen-4; CCL22 = C-C motif chemokine Ligand L22; CCR4 = C-C motif chemokine receptor 4; M1 = type 1 macrophage; INFγ = interferon γ; TNF α = tumor necrosis factor α.

### 5.3. Transcriptomic Analysis

Beuselinck and colleagues first performed clustering transcriptomic analysis in patients with metastatic ccRCC ( $n = 53$ ) receiving first-line sunitinib [74,75]. Hakimi et al. performed the same analysis in ccRCC of patients included in COMPARZ trial receiving pazopanib or sunitinib ( $n = 453$ ) and identified four similar clusters [75]. Cluster 3 had the best prognosis with high angiogenic gene expression being associated with a better outcome under antiangiogenic therapy and similar to ccRCC2 cluster reported by Beuselinck et al. [76]. Of note, PBRM1 mutation was frequently associated with angiogenic gene expression. Cluster 4 had a worse prognosis and was similar to ccRCC4 with upregulation of immune pathways. Interestingly, ccRCC4 was enriched in tumours with sarcomatoid differentiation with a frequent expression of PD-L1 [77]. Clusters 1 and 2 were intermediate clusters with a lower expression of angiogenic and immune genes. A consensus classification could emerge from these different studies as proposed in the bladder cancer [78].

McDermott and colleagues explored three gene expression signatures: angiogenesis, T-effector/IFN-γ response, and myeloid inflammatory genes in ccRCC ( $n = 263$ ) from IMmotion 150 trial: a phase 2 trial before IMmotion 151 of atezolizumab (anti-PD-L1) alone or combined with bevacizumab (anti-VEGF)

versus sunitinib [76]. Interestingly, high T effector/IFN $\gamma$  signature was associated with a better outcome to atezolizumab plus bevacizumab and a high myeloid inflammation signature was associated with reduced survival in the atezolizumab alone arm.

#### 5.4. Tumour Mutational Burden and Mismatch Repair Deficiency

The TMB is based on the total number of mutations per coding area of the tumour genome. High mutation burden favors the formation of neo-antigens that enhance the tumour immune response [79]. This biomarker initially failed to identify responders to immunotherapy in ccRCC but is still under evaluation [80,81]. However, Limitations of this marker may be due to technical requirements such as coverage, DNA amount and analysis time and lack of standardization.

Even if RCC are not considered to be in the spectrum of hereditary non-polyposis colon cancer (HNPCC) or Lynch syndrome, loss of mismatch repair (MMR) proteins leading to microsatellite instability is frequently observed [82]. MMR deficient tumours demonstrate a higher rate of mutations but on specific genes. In RCC, data on the response to immunotherapy according to MMR status are also still ongoing.

#### 5.5. Gut Microbiome

The gut microbiome, defined by the environmental conditions and the collection of microorganism and host genomes in an ecosystem, seems to influence the response to ICI. Indeed, the use of antibiotics could unfavorably impact the response. Indeed, Derosa et al. demonstrated a shorter survival and higher rates of progressive disease in patients with mRCC who received antibiotics within the month following the beginning of the treatment [83]. Indeed, antibiotics-related dysbiosis impacted the microbiome and could decrease the activity of ICI. Conversely, they identified that some bacteria such as *B. salyersiae* or *A. muciniphila* could restore the efficacy of ICI.

## 6. Future Directions

### 6.1. Microenvironment Cell Population Counter

The microenvironment cell population counter (MCP counter) aims to assess the proportion of immune and stromal cells in the TME from transcriptomic data [84]. They classified tumours into four molecular TME subgroups: immune infiltration, T and NK lymphocytes, MCH1 expression and fibroblastic infiltration. This classification could be judiciously applied to previous transcriptomic data.

### 6.2. Single-Cell Technologies

The development and proper use of immunotherapies rely on the detailed understanding of tumour composition. However, this purpose is hampered by intratumour heterogeneity, which was demonstrated to be extensive in ccRCC [85]. With recent developments in single-cell technologies though, we can now circumvent and integrate intratumour heterogeneity into analyses in oncology by characterizing each individual cells within tumours [86]. In RCC, Kim and collaborators performed single-cell RNA sequencing to study the intratumour heterogeneity of paired primary RCC and lung metastasis [87]. A considerable variability between the primary and metastatic sites and among tumour cells was demonstrated by the activation of drug target pathways. This study encourages the use of single-cell RNA sequencing to better characterize cell populations and to favor the discovery of new biomarkers.

### 6.3. 3D Culture Models

The 3D culture models of ccRCC could represent good pre-clinical models based on tissue slices preserving the stromal components of the TME [88]. Already used for TKI targeted therapies, these models are increasingly being considered as potential platforms for monitoring immunotherapy responses [89]. Recent studies indeed reported an active immune TME after applying immunomodulatory molecules on

tumour slice organotypic cultures, opening the way for their study on 3D ex vivo tumour models [90]. Interestingly, co-cultures of tumour-derived organoids were enriched with lymphocytes, leading Dijkstra and collaborators to recently demonstrate the generation of tumour-reactive T cells by co-culture of peripheral blood lymphocytes and colorectal cancer-derived organoids [91]. These precursor approaches will allow to learn more about immune escape mechanisms and determine associated predictive biomarkers.

#### 6.4. Imaging

Novel methods of imaging are non-invasive and assess RCC at different time points. Bensch et al. demonstrated that positron emission tomography (PET) imaging was able to localize <sup>89</sup>Zr-labeled atezolizumab to tumours expressing PD-L1 between primary and metastatic sites [92]. The uptake in tumours was heterogeneous and the clinical response better correlated with the pretreatment PET signal than with PD-L1 immunohistochemistry or T-effector gene expression signature. Moreover, CD8+ T cell infiltration was inferred from computed tomography [93].

#### 6.5. Circulating Tumour Cells

The circulating tumour cells (CTC) is a non-invasive method isolating CTC from blood circulation [94]. The challenge is to detect them because of the rare expression of the usual marker epithelial cell adhesion molecule (EpcAM). This is due to the transdifferentiation of tumour cells through the epithelial mesenchymal transition [95]. The antibodies directed against membrane carbonic anhydrase 9 (CA9) and CD147 largely improved the detection of CTC from 17% with EpcAM to 97% of samples with CA9 and CD147 markers [96]. Other methods are based on RT-PCR targeting VHL gene alteration and size-based blood filtration associated with genetic and morphological analyses [97]. The CTC analyses could be used to select patients for clinical trials guided by biomarkers. The phenotype could evolve under treatment and be timely investigated to adapt the therapeutic strategy.

### 7. Conclusions

Renal cell carcinoma include distinct entities with specific molecular alterations. Among them, ccRCC, the most frequent, is particularly characterized by its angiogenic and immunogenic TME with complex interactions between stromal and immune cells. Like tumour cells, intratumour heterogeneity is present in the TME with a variable distribution and phenotype and consequently favours resistance to treatment.

Therapeutic options for patients with mRCC have expanded rapidly over the past decade with targeted immunotherapy being the new corner stone. Several emerging drugs are designed to enhance the antitumour immune response and are tested in ongoing trials.

With an increasing number of treatment options available, improved biomarkers are needed to better stratify patients and define the optimal selection of patients and the sequence of treatment to overcome resistance. Promising biomarkers such as gene expression signatures or gut microbiome are under evaluation. It is likely that the future of predictive biomarkers relies on the combination of different approaches reflecting the complexity of the TME.

We hope that, in the future, new technologies such as single-cell technologies contribute to unravel the intratumour heterogeneity, identify predictive biomarkers and discover new treatment targets.

**Author Contributions:** Conception and design: S.-F.K.-J., N.R.-L.; Collection and assembly of data: S.-F.K.-J., A.D., J.S., L.C., B.L., F.D., M.-A.B.-R., R.M., K.B., G.V.; Manuscript writing: A.D., J.S., S.-F.K.-J.; Final approval of the manuscript: all authors; Accountable for all aspects of the work: all authors. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.



## References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics. *CA Cancer J. Clin.* **2019**, *69*, 7–34. [[CrossRef](#)] [[PubMed](#)]
2. Ljungberg, B.; Campbell, S.C.; Choi, H.Y.; Jacqmin, D.; Lee, J.E.; Weikert, S.; Kiemeny, L.A. The epidemiology of renal cell carcinoma. *Eur. Urol.* **2011**, *60*, 615–621. [[CrossRef](#)] [[PubMed](#)]
3. Gossage, L.; Eisen, T.; Maher, E.R. Vhl, the story of a tumour suppressor gene. *Nat. Rev. Cancer* **2015**, *15*, 55–64. [[CrossRef](#)] [[PubMed](#)]
4. Akaza, H.; Fukuyama, T. Axitinib for the treatment of advanced renal cell carcinoma. *Expert Opin. Pharmacother.* **2014**, *15*, 283–297. [[CrossRef](#)]
5. Motzer, R.J.; Hutson, T.E.; Cella, D.; Reeves, J.; Hawkins, R.; Guo, J.; Nathan, P.; Staehler, M.; de Souza, P.; Merchan, J.R.; et al. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. *N. Engl. J. Med.* **2013**, *369*, 722–731. [[CrossRef](#)]
6. Choueiri, T.K.; Escudier, B.; Powles, T.; Mainwaring, P.N.; Rini, B.I.; Donskov, F.; Hammers, H.; Hutson, T.E.; Lee, J.L.; Peltola, K.; et al. Cabozantinib versus everolimus in advanced renal-cell carcinoma. *N. Engl. J. Med.* **2015**, *373*, 1814–1823. [[CrossRef](#)]
7. Rini, B.I.; Pal, S.K.; Escudier, B.J.; Atkins, M.B.; Hutson, T.E.; Porta, C.; Verzoni, E.; Needle, M.N.; McDermott, D.F. Tivozanib versus sorafenib in patients with advanced renal cell carcinoma (tivo-3): A phase 3, multicentre, randomised, controlled, open-label study. *Lancet Oncol.* **2020**, *21*, 95–104. [[CrossRef](#)]
8. Moch, H.; Cubilla, A.L.; Humphrey, P.A.; Reuter, V.E.; Ulbright, T.M. The 2016 who classification of tumours of the urinary system and male genital organs-part a: Renal, penile, and testicular tumours. *Eur. Urol.* **2016**, *70*, 93–105. [[CrossRef](#)]
9. Cancer Genome Atlas Research, N.; Linehan, W.M.; Spellman, P.T.; Ricketts, C.J.; Creighton, C.J.; Fei, S.S.; Davis, C.; Wheeler, D.A.; Murray, B.A.; Schmidt, L.; et al. Comprehensive molecular characterization of papillary renal-cell carcinoma. *N. Engl. J. Med.* **2016**, *374*, 135–145. [[CrossRef](#)]
10. Argani, P. Mit family translocation renal cell carcinoma. *Semin. Diagn. Pathol.* **2015**, *32*, 103–113. [[CrossRef](#)]
11. Malouf, G.G.; Monzon, F.A.; Couturier, J.; Molinie, V.; Escudier, B.; Camparo, P.; Su, X.; Yao, H.; Tamboli, P.; Lopez-Terrada, D.; et al. Genomic heterogeneity of translocation renal cell carcinoma. *Clin. Cancer Res.* **2013**, *19*, 4673–4684. [[CrossRef](#)] [[PubMed](#)]
12. Malouf, G.G.; Comperat, E.; Yao, H.; Mouawad, R.; Lindner, V.; Rioux-Leclercq, N.; Verkarre, V.; Leroy, X.; Dainese, L.; Classe, M.; et al. Unique transcriptomic profile of collecting duct carcinomas relative to upper tract urothelial carcinomas and other kidney carcinomas. *Sci. Rep.* **2016**, *6*, 30988. [[CrossRef](#)] [[PubMed](#)]
13. Carlo, M.I.; Chaim, J.; Patil, S.; Kemel, Y.; Schram, A.M.; Woo, K.; Coskey, D.; Nanjangud, G.J.; Voss, M.H.; Feldman, D.R.; et al. Genomic characterization of renal medullary carcinoma and treatment outcomes. *Clin. Genitourin. Cancer* **2017**, *15*, e987–e994. [[CrossRef](#)] [[PubMed](#)]
14. Chen, Y.B.; Brannon, A.R.; Toubaji, A.; Dudas, M.E.; Won, H.H.; Al-Ahmadie, H.A.; Fine, S.W.; Gopalan, A.; Frizzell, N.; Voss, M.H.; et al. Hereditary leiomyomatosis and renal cell carcinoma syndrome-associated renal cancer: Recognition of the syndrome by pathologic features and the utility of detecting aberrant succination by immunohistochemistry. *Am. J. Surg. Pathol.* **2014**, *38*, 627–637. [[CrossRef](#)]
15. Malouf, G.G.; Ali, S.M.; Wang, K.; Balasubramanian, S.; Ross, J.S.; Miller, V.A.; Stephens, P.J.; Khayat, D.; Pal, S.K.; Su, X.; et al. Genomic characterization of renal cell carcinoma with sarcomatoid dedifferentiation pinpoints recurrent genomic alterations. *Eur. Urol.* **2016**, *70*, 348–357. [[CrossRef](#)]
16. Albiges, L.; Powles, T.; Staehler, M.; Bensalah, K.; Giles, R.H.; Hora, M.; Kuczyk, M.A.; Lam, T.B.; Ljungberg, B.; Marconi, L.; et al. Updated european association of urology guidelines on renal cell carcinoma: Immune checkpoint inhibition is the new backbone in first-line treatment of metastatic clear-cell renal cell carcinoma. *Eur. Urol.* **2019**, *76*, 151–156. [[CrossRef](#)]
17. Motzer, R.J.; Bacik, J.; Murphy, B.A.; Russo, P.; Mazumdar, M. Interferon-alfa as a comparative treatment for clinical trials of new therapies against advanced renal cell carcinoma. *J. Clin. Oncol.* **2002**, *20*, 289–296. [[CrossRef](#)]
18. Fyfe, G.; Fisher, R.I.; Rosenberg, S.A.; Sznol, M.; Parkinson, D.R.; Louie, A.C. Results of treatment of 255 patients with metastatic renal cell carcinoma who received high-dose recombinant interleukin-2 therapy. *J. Clin. Oncol.* **1995**, *13*, 688–696. [[CrossRef](#)]



19. Motzer, R.J.; Hutson, T.E.; Tomczak, P.; Michaelson, M.D.; Bukowski, R.M.; Rixe, O.; Oudard, S.; Negrier, S.; Szczylik, C.; Kim, S.T.; et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N. Engl. J. Med.* **2007**, *356*, 115–124. [[CrossRef](#)]
20. Motzer, R.J.; Tannir, N.M.; McDermott, D.F.; Aren Frontera, O.; Melichar, B.; Choueiri, T.K.; Plimack, E.R.; Barthelemy, P.; Porta, C.; George, S.; et al. Nivolumab plus ipilimumab versus sunitinib in advanced renal-cell carcinoma. *N. Engl. J. Med.* **2018**, *378*, 1277–1290. [[CrossRef](#)]
21. Motzer, R.J.; Escudier, B.; McDermott, D.F.; George, S.; Hammers, H.J.; Srinivas, S.; Tykodi, S.S.; Sosman, J.A.; Procopio, G.; Plimack, E.R.; et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N. Engl. J. Med.* **2015**, *373*, 1803–1813. [[CrossRef](#)] [[PubMed](#)]
22. Leite, K.R.; Reis, S.T.; Junior, J.P.; Zerati, M.; Gomes Dde, O.; Camara-Lopes, L.H.; Srougi, M. Pd-11 expression in renal cell carcinoma clear cell type is related to unfavorable prognosis. *Diagn. Pathol.* **2015**, *10*, 189. [[CrossRef](#)] [[PubMed](#)]
23. Thompson, R.H.; Dong, H.; Kwon, E.D. Implications of b7-h1 expression in clear cell carcinoma of the kidney for prognostication and therapy. *Clin. Cancer Res.* **2007**, *13*, 709s–715s. [[CrossRef](#)]
24. Thompson, R.H.; Dong, H.; Lohse, C.M.; Leibovich, B.C.; Blute, M.L.; Cheville, J.C.; Kwon, E.D. Pd-1 is expressed by tumor-infiltrating immune cells and is associated with poor outcome for patients with renal cell carcinoma. *Clin. Cancer Res.* **2007**, *13*, 1757–1761. [[CrossRef](#)]
25. Motzer, R.J.; Penkov, K.; Haanen, J.; Rini, B.; Albiges, L.; Campbell, M.T.; Venugopal, B.; Kollmannsberger, C.; Negrier, S.; Uemura, M.; et al. Avelumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. *N. Engl. J. Med.* **2019**, *380*, 1103–1115. [[CrossRef](#)]
26. Rini, B.I.; Plimack, E.R.; Stus, V.; Gafanov, R.; Hawkins, R.; Nosov, D.; Pouliot, F.; Alekseev, B.; Soulieres, D.; Melichar, B.; et al. Pembrolizumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. *N. Engl. J. Med.* **2019**, *380*, 1116–1127. [[CrossRef](#)]
27. Sharma, P.; Hu-Lieskovan, S.; Wargo, J.A.; Ribas, A. Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell* **2017**, *168*, 707–723. [[CrossRef](#)]
28. Dunn, G.P.; Bruce, A.T.; Ikeda, H.; Old, L.J.; Schreiber, R.D. Cancer immunoediting: From immunosurveillance to tumor escape. *Nat. Immunol.* **2002**, *3*, 991–998. [[CrossRef](#)]
29. Anderson, A.C.; Joller, N.; Kuchroo, V.K. Lag-3, tim-3, and tigit: Co-inhibitory receptors with specialized functions in immune regulation. *Immunity* **2016**, *44*, 989–1004. [[CrossRef](#)]
30. Lambrechts, D.; Wauters, E.; Boeckx, B.; Aibar, S.; Nittner, D.; Burton, O.; Bassez, A.; Decaluwe, H.; Pircher, A.; Van den Eynde, K.; et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat. Med.* **2018**, *24*, 1277–1289. [[CrossRef](#)]
31. Akino, T.; Hida, K.; Hida, Y.; Tsuchiya, K.; Freedman, D.; Muraki, C.; Ohga, N.; Matsuda, K.; Akiyama, K.; Harabayashi, T.; et al. Cytogenetic abnormalities of tumor-associated endothelial cells in human malignant tumors. *Am. J. Pathol.* **2009**, *175*, 2657–2667. [[CrossRef](#)] [[PubMed](#)]
32. Edeline, J.; Mottier, S.; Vigneau, C.; Jouan, F.; Perrin, C.; Zerrouki, S.; Fergelot, P.; Patard, J.J.; Rioux-Leclercq, N. Description of 2 angiogenic phenotypes in clear cell renal cell carcinoma. *Hum. Pathol.* **2012**, *43*, 1982–1990. [[CrossRef](#)] [[PubMed](#)]
33. Dufies, M.; Giuliano, S.; Ambrosetti, D.; Claren, A.; Ndiaye, P.D.; Matri, M.; Moghrabi, W.; Cooley, L.S.; Ettaiche, M.; Chamorey, E.; et al. Sunitinib stimulates expression of vegfc by tumor cells and promotes lymphangiogenesis in clear cell renal cell carcinomas. *Cancer Res.* **2017**, *77*, 1212–1226. [[CrossRef](#)] [[PubMed](#)]
34. Thompson, R.H.; Kuntz, S.M.; Leibovich, B.C.; Dong, H.; Lohse, C.M.; Webster, W.S.; Sengupta, S.; Frank, I.; Parker, A.S.; Zincke, H.; et al. Tumor b7-h1 is associated with poor prognosis in renal cell carcinoma patients with long-term follow-up. *Cancer Res.* **2006**, *66*, 3381–3385. [[CrossRef](#)] [[PubMed](#)]
35. Choueiri, T.K.; Fay, A.P.; Gray, K.P.; Callea, M.; Ho, T.H.; Albiges, L.; Bellmunt, J.; Song, J.; Carvo, I.; Lampron, M.; et al. Pd-11 expression in nonclear-cell renal cell carcinoma. *Ann. Oncol.* **2014**, *25*, 2178–2184. [[CrossRef](#)]
36. Gupta, S.; Cheville, J. C.; Jungbluth, A. A.; Zhang, Y.; Zhang, L.; Chen, Y. B.; Tickoo, S. K.; Fine, S. W.; Gopalan, A.; Al-Ahmadie, H. A.; et al. JAK2/PD-L1/PD-L2 (9p24.1) amplifications in renal cell carcinomas with sarcomatoid transformation: Implications for clinical management. *Mod. Pathol.* **2019**, *32*, 1344–1358. [[CrossRef](#)] [[PubMed](#)]
37. Chevrier, S.; Levine, J.H.; Zanotelli, V.R.T.; Silina, K.; Schulz, D.; Bacac, M.; Ries, C.H.; Ailles, L.; Jewett, M.A.S.; Moch, H.; et al. An immune atlas of clear cell renal cell carcinoma. *Cell* **2017**, *169*, 736–749. [[CrossRef](#)]

38. Giraldo, N.A.; Becht, E.; Vano, Y.; Petitprez, F.; Lacroix, L.; Validire, P.; Sanchez-Salas, R.; Ingels, A.; Oudard, S.; Moatti, A.; et al. Tumor-infiltrating and peripheral blood t-cell immunophenotypes predict early relapse in localized clear cell renal cell carcinoma. *Clin. Cancer Res.* **2017**, *23*, 4416–4428. [[CrossRef](#)]
39. Sjoberg, E.; Frodin, M.; Lovrot, J.; Mezheyski, A.; Johansson, M.; Harmenberg, U.; Egevad, L.; Sandstrom, P.; Ostman, A. A minority-group of renal cell cancer patients with high infiltration of cd20 + b-cells is associated with poor prognosis. *Br. J. Cancer* **2018**, *119*, 840–846. [[CrossRef](#)]
40. Choueiri, T.K.; Fishman, M.N.; Escudier, B.; McDermott, D.F.; Drake, C.G.; Kluger, H.; Stadler, W.M.; Perez-Gracia, J.L.; McNeel, D.G.; Curti, B.; et al. Immunomodulatory activity of nivolumab in metastatic renal cell carcinoma. *Clin. Cancer Res.* **2016**, *22*, 5461–5471. [[CrossRef](#)]
41. Shirotake, S.; Kaneko, G.; Nagata, K.; Oyama, M.; Nishimoto, K. Histological complete response with nivolumab for renal cell carcinoma with multiple metastases: A case report. *Mol. Clin. Oncol.* **2019**, *10*, 244–248. [[CrossRef](#)] [[PubMed](#)]
42. Rini, B.I.; Powles, T.; Atkins, M.B.; Escudier, B.; McDermott, D.F.; Suarez, C.; Bracarda, S.; Stadler, W.M.; Donskov, F.; Lee, J.L.; et al. Atezolizumab plus bevacizumab versus sunitinib in patients with previously untreated metastatic renal cell carcinoma (immotion151): A multicentre, open-label, phase 3, randomised controlled trial. *Lancet* **2019**, *393*, 2404–2415. [[CrossRef](#)]
43. Motzer, R.J.; Rini, B.I.; McDermott, D.F.; Aren Frontera, O.; Hammers, H.J.; Carducci, M.A.; Salman, P.; Escudier, B.; Beuselinck, B.; Amin, A.; et al. Nivolumab plus ipilimumab versus sunitinib in first-line treatment for advanced renal cell carcinoma: Extended follow-up of efficacy and safety results from a randomised, controlled, phase 3 trial. *Lancet Oncol.* **2019**, *20*, 1370–1385. [[CrossRef](#)]
44. Lalani, A.A.; McGregor, B.A.; Albiges, L.; Choueiri, T.K.; Motzer, R.; Powles, T.; Wood, C.; Bex, A. Systemic treatment of metastatic clear cell renal cell carcinoma in 2018: Current paradigms, use of immunotherapy, and future directions. *Eur. Urol.* **2019**, *75*, 100–110. [[CrossRef](#)] [[PubMed](#)]
45. Tomita, Y.; Fukasawa, S.; Shinohara, N.; Kitamura, H.; Oya, M.; Eto, M.; Tanabe, K.; Kimura, G.; Yonese, J.; Yao, M.; et al. Nivolumab versus everolimus in advanced renal cell carcinoma: Japanese subgroup analysis from the checkmate 025 study. *Jpn. J. Clin. Oncol.* **2017**, *47*, 639–646. [[CrossRef](#)]
46. Ravaud, A.; Motzer, R.J.; Pandha, H.S.; George, D.J.; Pantuck, A.J.; Patel, A.; Chang, Y.H.; Escudier, B.; Donskov, F.; Magheli, A.; et al. Adjuvant sunitinib in high-risk renal-cell carcinoma after nephrectomy. *N. Engl. J. Med.* **2016**, *375*, 2246–2254. [[CrossRef](#)]
47. Motzer, R.J.; Jonasch, E.; Michaelson, M.D.; Nandagopal, L.; Gore, J.L.; George, S.; Alva, A.; Haas, N.; Harrison, M.R.; Plimack, E.R.; et al. Nccn guidelines insights: Kidney cancer, version 2.2020. *J. Natl. Compr. Cancer Netw.* **2019**, *17*, 1278–1285. [[CrossRef](#)]
48. Vera-Badillo, F.E.; Templeton, A.J.; Duran, I.; Ocana, A.; de Gouveia, P.; Aneja, P.; Knox, J.J.; Tannock, I.F.; Escudier, B.; Amir, E. Systemic therapy for non-clear cell renal cell carcinomas: A systematic review and meta-analysis. *Eur. Urol.* **2015**, *67*, 740–749. [[CrossRef](#)]
49. Koshkin, V.S.; Barata, P.C.; Zhang, T.; George, D.J.; Atkins, M.B.; Kelly, W.J.; Vogelzang, N.J.; Pal, S.K.; Hsu, J.; Appleman, L.J.; et al. Clinical activity of nivolumab in patients with non-clear cell renal cell carcinoma. *J. Immunother. Cancer* **2018**, *6*, 9. [[CrossRef](#)]
50. Benci, J.L.; Xu, B.; Qiu, Y.; Wu, T.J.; Dada, H.; Twyman-Saint Victor, C.; Cucolo, L.; Lee, D.S.M.; Pauken, K.E.; Huang, A.C.; et al. Tumor interferon signaling regulates a multigenic resistance program to immune checkpoint blockade. *Cell* **2016**, *167*, 1540–1554. [[CrossRef](#)]
51. Gao, X.; Zhu, Y.; Li, G.; Huang, H.; Zhang, G.; Wang, F.; Sun, J.; Yang, Q.; Zhang, X.; Lu, B. Tim-3 expression characterizes regulatory t cells in tumor tissues and is associated with lung cancer progression. *PLoS ONE* **2012**, *7*, e30676. [[CrossRef](#)] [[PubMed](#)]
52. Sakuishi, K.; Apetoh, L.; Sullivan, J.M.; Blazar, B.R.; Kuchroo, V.K.; Anderson, A.C. Targeting tim-3 and pd-1 pathways to reverse t cell exhaustion and restore anti-tumor immunity. *J. Exp. Med.* **2010**, *207*, 2187–2194. [[CrossRef](#)] [[PubMed](#)]
53. Casado, J.G.; Pawelec, G.; Morgado, S.; Sanchez-Correa, B.; Delgado, E.; Gayoso, I.; Duran, E.; Solana, R.; Tarazona, R. Expression of adhesion molecules and ligands for activating and costimulatory receptors involved in cell-mediated cytotoxicity in a large panel of human melanoma cell lines. *Cancer Immunol. Immunother.* **2009**, *58*, 1517–1526. [[CrossRef](#)] [[PubMed](#)]

54. Inamura, K.; Amori, G.; Yuasa, T.; Yamamoto, S.; Yonese, J.; Ishikawa, Y. Relationship of b7-h3 expression in tumor cells and tumor vasculature with foxp3+ regulatory t cells in renal cell carcinoma. *Cancer Manag. Res.* **2019**, *11*, 7021–7030. [[CrossRef](#)] [[PubMed](#)]
55. Wang, L.; Rubinstein, R.; Lines, J.L.; Wasiuk, A.; Ahonen, C.; Guo, Y.; Lu, L.F.; Gondek, D.; Wang, Y.; Fava, R.A.; et al. Vista, a novel mouse ig superfamily ligand that negatively regulates t cell responses. *J. Exp. Med.* **2011**, *208*, 577–592. [[CrossRef](#)] [[PubMed](#)]
56. Le Mercier, I.; Chen, W.; Lines, J.L.; Day, M.; Li, J.; Sergent, P.; Noelle, R.J.; Wang, L. Vista regulates the development of protective antitumor immunity. *Cancer Res.* **2014**, *74*, 1933–1944. [[CrossRef](#)]
57. Fos, C.; Salles, A.; Lang, V.; Carrette, F.; Audebert, S.; Pastor, S.; Ghiotto, M.; Olive, D.; Bismuth, G.; Nunes, J.A. Icos ligation recruits the p50alpha pi3k regulatory subunit to the immunological synapse. *J. Immunol.* **2008**, *181*, 1969–1977. [[CrossRef](#)]
58. Aspeslagh, S.; Postel-Vinay, S.; Rusakiewicz, S.; Soria, J.C.; Zitvogel, L.; Marabelle, A. Rationale for anti-ox40 cancer immunotherapy. *Eur. J. Cancer* **2016**, *52*, 50–66. [[CrossRef](#)]
59. Shevach, E.M.; Stephens, G.L. The gitr-gitrl interaction: Co-stimulation or contrasuppression of regulatory activity? *Nat. Rev. Immunol.* **2006**, *6*, 613–618. [[CrossRef](#)]
60. Ramapriyan, R.; Caetano, M.S.; Barsoumian, H.B.; Mafra, A.C.P.; Zambalde, E.P.; Menon, H.; Tsouko, E.; Welsh, J.W.; Cortez, M.A. Altered cancer metabolism in mechanisms of immunotherapy resistance. *Pharmacol. Ther.* **2019**, *195*, 162–171. [[CrossRef](#)]
61. Chen, J.Y.; Li, C.F.; Kuo, C.C.; Tsai, K.K.; Hou, M.F.; Hung, W.C. Cancer/stroma interplay via cyclooxygenase-2 and indoleamine 2,3-dioxygenase promotes breast cancer progression. *Breast Cancer Res.* **2014**, *16*, 410. [[CrossRef](#)] [[PubMed](#)]
62. Cheong, J.E.; Sun, L. Targeting the ido1/tdo2-kyn-ahr pathway for cancer immunotherapy—challenges and opportunities. *Trends Pharmacol. Sci.* **2018**, *39*, 307–325. [[CrossRef](#)] [[PubMed](#)]
63. Leone, R.D.; Emens, L.A. Targeting adenosine for cancer immunotherapy. *J. Immunother. Cancer* **2018**, *6*, 57. [[CrossRef](#)] [[PubMed](#)]
64. Linnemann, C.; Schildberg, F.A.; Schurich, A.; Diehl, L.; Hegenbarth, S.I.; Endl, E.; Lacher, S.; Muller, C.E.; Frey, J.; Simeoni, L.; et al. Adenosine regulates cd8 t-cell priming by inhibition of membrane-proximal t-cell receptor signalling. *Immunology* **2009**, *128*, e728–e737. [[CrossRef](#)]
65. Ananieva, E. Targeting amino acid metabolism in cancer growth and anti-tumor immune response. *World J. Biol. Chem.* **2015**, *6*, 281–289. [[CrossRef](#)]
66. Dawson, M.A.; Kouzarides, T. Cancer epigenetics: From mechanism to therapy. *Cell* **2012**, *150*, 12–27. [[CrossRef](#)]
67. Eckschlager, T.; Plch, J.; Stiborova, M.; Hrabeta, J. Histone deacetylase inhibitors as anticancer drugs. *Int. J. Mol. Sci.* **2017**, *18*. [[CrossRef](#)]
68. Pan, R.Y.; Chung, W.H.; Chu, M.T.; Chen, S.J.; Chen, H.C.; Zheng, L.; Hung, S.I. Recent development and clinical application of cancer vaccine: Targeting neoantigens. *J. Immunol. Res.* **2018**, *2018*, 4325874. [[CrossRef](#)]
69. Li, H.; Ding, J.; Lu, M.; Liu, H.; Miao, Y.; Li, L.; Wang, G.; Zheng, J.; Pei, D.; Zhang, Q. Caix-specific car-t cells and sunitinib show synergistic effects against metastatic renal cancer models. *J. Immunother.* **2020**, *43*, 16–28. [[CrossRef](#)]
70. Heng, D.Y.; Xie, W.; Regan, M.M.; Harshman, L.C.; Bjarnason, G.A.; Vaishampayan, U.N.; Mackenzie, M.; Wood, L.; Donskov, F.; Tan, M.H.; et al. External validation and comparison with other models of the international metastatic renal-cell carcinoma database consortium prognostic model: A population-based study. *Lancet Oncol.* **2013**, *14*, 141–148. [[CrossRef](#)]
71. Kroeger, N.; Klatte, T.; Chamie, K.; Rao, P.N.; Birkhauser, F.D.; Sonn, G.A.; Riss, J.; Kabbinavar, F.F.; Belldegrun, A.S.; Pantuck, A.J. Deletions of chromosomes 3p and 14q molecularly subclassify clear cell renal cell carcinoma. *Cancer* **2013**, *119*, 1547–1554. [[CrossRef](#)] [[PubMed](#)]
72. Kammerer-Jacquet, S.F.; Deleuze, A.; Saout, J.; Mathieu, R.; Laguerre, B.; Verhoest, G.; Dugay, F.; Belaud-Rotureau, M.A.; Bensalah, K.; Rioux-Leclercq, N. Targeting the pd-1/pd-l1 pathway in renal cell carcinoma. *Int. J. Mol. Sci.* **2019**, *20*, 1692. [[CrossRef](#)] [[PubMed](#)]
73. Seeber, A.; Klinglmair, G.; Fritz, J.; Steinkohl, F.; Zimmer, K.C.; Aigner, F.; Horninger, W.; Gastl, G.; Zelger, B.; Brunner, A.; et al. High ido-1 expression in tumor endothelial cells is associated with response to immunotherapy in metastatic renal cell carcinoma. *Cancer Sci.* **2018**, *109*, 1583–1591. [[CrossRef](#)] [[PubMed](#)]

74. Beuselinck, B.; Job, S.; Becht, E.; Karadimou, A.; Verkarre, V.; Couchy, G.; Giraldo, N.; Rioux-Leclercq, N.; Molinie, V.; Sibony, M.; et al. Molecular subtypes of clear cell renal cell carcinoma are associated with sunitinib response in the metastatic setting. *Clin. Cancer Res.* **2015**, *21*, 1329–1339. [[CrossRef](#)] [[PubMed](#)]
75. Hakimi, A.A.; Voss, M.H.; Kuo, F.; Sanchez, A.; Liu, M.; Nixon, B.G.; Vuong, L.; Ostrovskaya, I.; Chen, Y.B.; Reuter, V.; et al. Transcriptomic profiling of the tumor microenvironment reveals distinct subgroups of clear cell renal cell cancer—data from a randomized phase iii trial. *Cancer Discov.* **2019**, *9*, 510–525. [[CrossRef](#)] [[PubMed](#)]
76. McDermott, D.F.; Huseni, M.A.; Atkins, M.B.; Motzer, R.J.; Rini, B.I.; Escudier, B.; Fong, L.; Joseph, R.W.; Pal, S.K.; Reeves, J.A.; et al. Clinical activity and molecular correlates of response to atezolizumab alone or in combination with bevacizumab versus sunitinib in renal cell carcinoma. *Nat. Med.* **2018**, *24*, 749–757. [[CrossRef](#)] [[PubMed](#)]
77. Joseph, R.W.; Millis, S.Z.; Carballido, E.M.; Bryant, D.; Gatalica, Z.; Reddy, S.; Bryce, A.H.; Vogelzang, N.J.; Stanton, M.L.; Castle, E.P.; et al. Pd-1 and pd-11 expression in renal cell carcinoma with sarcomatoid differentiation. *Cancer Immunol. Res.* **2015**, *3*, 1303–1307. [[CrossRef](#)]
78. Kamoun, A.; de Reynies, A.; Allory, Y.; Sjobahl, G.; Robertson, A.G.; Seiler, R.; Hoadley, K.A.; Groeneveld, C.S.; Al-Ahmadie, H.; Choi, W.; et al. A consensus molecular classification of muscle-invasive bladder cancer. *Eur. Urol.* **2019**, *77*, 420–433. [[CrossRef](#)]
79. Chalmers, Z.R.; Connelly, C.F.; Fabrizio, D.; Gay, L.; Ali, S.M.; Ennis, R.; Schrock, A.; Campbell, B.; Shlien, A.; Chmielecki, J.; et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* **2017**, *9*, 34. [[CrossRef](#)]
80. de Velasco, G.; Miao, D.; Voss, M.H.; Hakimi, A.A.; Hsieh, J.J.; Tannir, N.M.; Tamboli, P.; Appleman, L.J.; Rathmell, W.K.; Van Allen, E.M.; et al. Tumor mutational load and immune parameters across metastatic renal cell carcinoma risk groups. *Cancer Immunol. Res.* **2016**, *4*, 820–822. [[CrossRef](#)]
81. Miao, D.; Margolis, C.A.; Gao, W.; Voss, M.H.; Li, W.; Martini, D.J.; Norton, C.; Bosse, D.; Wankowicz, S.M.; Cullen, D.; et al. Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science* **2018**, *359*, 801–806. [[CrossRef](#)] [[PubMed](#)]
82. Dudley, J.C.; Lin, M.T.; Le, D.T.; Eshleman, J.R. Microsatellite instability as a biomarker for pd-1 blockade. *Clin. Cancer Res.* **2016**, *22*, 813–820. [[CrossRef](#)] [[PubMed](#)]
83. Derosa, L.; Hellmann, M.D.; Spaziano, M.; Halpenny, D.; Fidelle, M.; Rizvi, H.; Long, N.; Plodkowski, A.J.; Arbour, K.C.; Chaft, J.E.; et al. Negative association of antibiotics on clinical activity of immune checkpoint inhibitors in patients with advanced renal cell and non-small-cell lung cancer. *Ann. Oncol.* **2018**, *29*, 1437–1444. [[CrossRef](#)] [[PubMed](#)]
84. Becht, E.; Giraldo, N.A.; Lacroix, L.; Buttard, B.; Elarouci, N.; Petitprez, F.; Selves, J.; Laurent-Puig, P.; Sautes-Fridman, C.; Fridman, W.H.; et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **2016**, *17*, 218. [[CrossRef](#)]
85. Gerlinger, M.; Catto, J.W.; Orntoft, T.F.; Real, F.X.; Zwarthoff, E.C.; Swanton, C. Intratumour heterogeneity in urologic cancers: From molecular evidence to clinical implications. *Eur. Urol.* **2015**, *67*, 729–737. [[CrossRef](#)]
86. Suva, M.L.; Tirosch, I. Single-cell rna sequencing in cancer: Lessons learned and emerging challenges. *Mol. Cell* **2019**, *75*, 7–12. [[CrossRef](#)]
87. Kim, K.T.; Lee, H.W.; Lee, H.O.; Song, H.J.; Jeong da, E.; Shin, S.; Kim, H.; Shin, Y.; Nam, D.H.; Jeong, B.C.; et al. Application of single-cell rna sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma. *Genome Biol.* **2016**, *17*, 80. [[CrossRef](#)]
88. Roelants, C.; Pillet, C.; Franquet, Q.; Sarrazin, C.; Peilleron, N.; Giacosa, S.; Guyon, L.; Fontanell, A.; Fiard, G.; Long, J.A.; et al. Ex-vivo treatment of tumor tissue slices as a predictive preclinical method to evaluate targeted therapies for patients with renal carcinoma. *Cancers (Basel)* **2020**, *12*, 232. [[CrossRef](#)]
89. Meijer, T.G.; Naipal, K.A.; Jager, A.; van Gent, D.C. Ex vivo tumor culture systems for functional drug testing and therapy response prediction. *Future Sci. OA* **2017**, *3*, FSO190. [[CrossRef](#)]
90. Seo, Y.D.; Jiang, X.; Sullivan, K.M.; Jalikis, F.G.; Smythe, K.S.; Abbasi, A.; Vignali, M.; Park, J.O.; Daniel, S.K.; Pollack, S.M.; et al. Mobilization of cd8( + ) t cells via cxcr4 blockade facilitates pd-1 checkpoint therapy in human pancreatic cancer. *Clin. Cancer Res.* **2019**, *25*, 3934–3945. [[CrossRef](#)]
91. Dijkstra, K.K.; Cattaneo, C.M.; Weeber, F.; Chalabi, M.; van de Haar, J.; Fanchi, L.F.; Slagter, M.; van der Velden, D.L.; Kaing, S.; Kelderman, S.; et al. Generation of tumor-reactive t cells by co-culture of peripheral blood lymphocytes and tumor organoids. *Cell* **2018**, *174*, 1586–1598. [[CrossRef](#)] [[PubMed](#)]

92. Bensch, F.; van der Veen, E.L.; Lub-de Hooge, M.N.; Jorritsma-Smit, A.; Boellaard, R.; Kok, I.C.; Oosting, S.F.; Schroder, C.P.; Hiltermann, T.J.N.; van der Wekken, A.J.; et al. (89)zr-atezolizumab imaging as a non-invasive approach to assess clinical response to pd-1l blockade in cancer. *Nat. Med.* **2018**, *24*, 1852–1858. [[CrossRef](#)] [[PubMed](#)]
93. Sun, R.; Limkin, E.J.; Vakalopoulou, M.; Derclé, L.; Champiat, S.; Han, S.R.; Verlingue, L.; Brandao, D.; Lancia, A.; Ammari, S.; et al. A radiomics approach to assess tumour-infiltrating cd8 cells and response to anti-pd-1 or anti-pd-1l immunotherapy: An imaging biomarker, retrospective multicohort study. *Lancet Oncol.* **2018**, *19*, 1180–1191. [[CrossRef](#)]
94. Gradilone, A.; Iacovelli, R.; Cortesi, E.; Raimondi, C.; Gianni, W.; Nicolazzo, C.; Petracca, A.; Palazzo, A.; Longo, F.; Frati, L.; et al. Circulating tumor cells and “suspicious objects” evaluated through cellsearch(r) in metastatic renal cell carcinoma. *Anticancer. Res.* **2011**, *31*, 4219–4221. [[PubMed](#)]
95. He, H.; Magi-Galluzzi, C. Epithelial-to-mesenchymal transition in renal neoplasms. *Adv. Anat. Pathol.* **2014**, *21*, 174–180. [[CrossRef](#)] [[PubMed](#)]
96. Liu, S.; Tian, Z.; Zhang, L.; Hou, S.; Hu, S.; Wu, J.; Jing, Y.; Sun, H.; Yu, F.; Zhao, L.; et al. Combined cell surface carbonic anhydrase 9 and cd147 antigens enable high-efficiency capture of circulating tumor cells in clear cell renal cell carcinoma patients. *Oncotarget* **2016**, *7*, 59877–59891. [[CrossRef](#)] [[PubMed](#)]
97. Santoni, M.; Cimadamore, A.; Cheng, L.; Lopez-Beltran, A.; Battelli, N.; Massari, F.; Scarpelli, M.; Galosi, A.B.; Bracarda, S.; Montironi, R. Circulating tumor cells in renal cell carcinoma: Recent findings and future challenges. *Front. Oncol.* **2019**, *9*, 228. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).





**Titre : Étude de l'hétérogénéité intra-tumorale et de l'impact de thérapies ciblées sur le carcinome rénal à cellules claires par une approche transcriptomique sur cellules uniques**

**Mots clés :** Transcriptomique sur cellules uniques, carcinome rénal, hétérogénéité tumorale, traitements

**Résumé :** L'hétérogénéité intra-tumorale (HIT) désigne la diversité des cellules malignes et des cellules normales qui composent le microenvironnement d'une tumeur. Elle constitue un réservoir de cellule innovantes connu pour jouer un rôle dans la progression tumorale et la résistance aux thérapies ciblées dans de nombreux cancers, y compris dans le carcinome rénal à cellules claires (ccRCC). L'un des grands défis de l'oncologie consiste donc à déchiffrer et contourner l'HIT au bénéfice des patients. Dans ce contexte, la technique de séquençage ARN sur cellules uniques (« single-cell RNA-sequencing » ou scRNA-seq) s'est largement développée ces dix dernières années et engage un tournant sans précédent dans le domaine puisqu'elle permet de caractériser le transcriptome de chaque sous-population cellulaire peuplant la tumeur. Dans le premier axe de ce travail de thèse, j'ai utilisé le scRNA-seq afin d'établir un atlas cellulaire du ccRCC révélant différents degrés d'hétérogénéité inter- et intra-tumorale, notamment dans les cellules malignes. En s'appuyant sur de grandes bases de données transcriptomiques, ce travail a démontré des associations significatives

entre la présence de certaines populations malignes et la survie des patients, notamment parmi les tumeurs de bas grade celles qui donnent lieu à une rechute et auraient bénéficié d'un traitement adjuvant. Dans le second axe, davantage appliqué, j'ai établi et validé une stratégie innovante combinant la culture 3D du ccRCC et le scRNA-seq pour étudier l'impact d'agents inhibiteurs de kinases à l'échelle de la cellule unique. L'approche permet notamment d'identifier des populations cellules sensibles et de distinguer des réponses différentielles aux molécules dans le compartiment malin. Enfin, j'ai contribué au développement de UncoVer, une ressource en ligne répertoriant les données de scRNA-seq et transcriptomique spatiale dans le domaine des cancers urologiques. Dans l'ensemble, ces trois axes s'articulent dans le cadre d'une recherche résolument translationnelle et leurs résultats contribuent à mieux caractériser le ccRCC à l'échelle de la cellule unique, ainsi qu'à supporter l'utilisation prochaine de ces technologies innovantes au sein de dispositifs cliniques.

**Title : Study of intra-tumor heterogeneity and anti-cancer therapies in clear cell renal cell carcinoma using single-cell transcriptomics**

**Keywords** Single-cell transcriptomics, renal carcinoma, intra-tumor heterogeneity, treatments

**Abstract :** Intra-tumor heterogeneity (ITH) refers to the diversity of malignant cells and normal cells composing the microenvironment in a tumor. It acts as a pool of innovative cells which is known to play a role in tumor progression and resistance to targeted therapies in many cancers, including clear cell renal cell carcinoma (ccRCC). A major challenge in oncology is therefore to decipher and circumvent ITH for the benefit of patients. In this context, the single-cell RNA-sequencing (scRNA-seq) has been widely developed in the last decade and allows an unprecedented turn in the field since it gives access to the transcriptomic identification of each cellular subpopulation in tumors. In the first axis of this thesis, I used scRNA-seq to establish a cellular atlas of ccRCC revealing different degrees of inter- and intra-tumor heterogeneity, especially in malignant cells. Using large transcriptomic databases, this work contributed to demonstrate significant associations between the

presence of specific malignant clones and patient survival. This enables to better discriminate among low-grade tumors those that relapse and would have benefited from adjuvant treatment. In the second axis, I have established and validated an innovative strategy combining 3D culture of ccRCC and scRNA-seq to study the impact of kinase inhibitors at the single-cell level. This approach allowed us to identify sensitive cell populations and to distinguish differential responses to the drugs tested in the malignant compartment. Finally, I contributed to the development of UncoVer, a web-based resource dedicated to single-cell and spatial profiling data accumulated within the field of urologic oncology. Overall, these results contribute to better characterize the ccRCC at the single-cell resolution, as well as to support the upcoming use of such technologies within clinical settings.