Identification of APOBEC-Associated Frequent Mutations and Characterization of FGFR3-Driven Signaling Pathways in Bladder Cancer

Mingjun Shi

université PARIS-SACLAY

2019SACLS216

Thèse de doctorat

UNIVERSITÉ PARIS SUD
Comprendre le monde, construire l'avenir

Identification of APOBEC-associated hotspot mutations & characterization of FGFR3-driven signaling pathways in bladder cancer

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°582 Cancérologie : biologie – médecine – santé
Spécialité de doctorat: Sciences de la vie et de la santé

Thèse présentée et soutenue à Paris, le 4 septembre 2019, par

Mingjun SHI

Composition du Jury :

Pr. Thierry LEBRET PU-PH, UVSQ, Hôpital Foch, Paris, France	Président
Pr. João Pedro Taborda BARATA Professor, Instituto de Medicina Molecular, Lisbon, Portugal	Rapporteur
Dr. Eric LETOUZE Senior researcher, FuGeST (INSERM U1162), Paris, France	Rapporteur
Pr. Serge ROCHE DR1, CRBM (CNRS), Montpellier, France	Examinateur
Dr. Isabelle BERNARD-PIERROT CR1, Institut Curie (CNRS UMR144), Paris, France	Co-Directeur de thèse
Pr. François RADVANYI DRCE, Institut Curie (CNRS UMR144), Paris, France	Directeur de thèse

Résumé

Le cancer de la vessie (BCa), est une tumeur maligne de l'urothélium, fréquente dans le monde entier, dont le traitement particulièrement coûteux ne permet cependant pas d'éviter les récidives et les progressions. Génétiquement, le BCa présente un grand nombre de mutations, juste après le mélanome et le cancer du poumon. Les mutations somatiques fréquentes sont une conséquence cumulative du processus de mutation endogène et / ou des expositions mutagènes exogènes, en plus d'une sélection fonctionnelle. Chaque exposition / processus de mutation laisse une empreinte caractéristique, appelée signature de mutation, sur le génome du cancer sous la forme de types préférentiels de mutations survenant dans des contextes de séquences spécifiques. La mutagenèse dû à APOBEC, l'une des signatures de ce type induite par l'hyperactivité des APOBEC désaminases, est omniprésente dans plusieurs types de cancer. Elle est en particulier la principale source de mutations dans le BCa. Il est donc utile d'identifier les mutations fréquentes associées à la signature de mutation APOBEC dans le BCa pour en comprendre l'étiologie. Très récemment, il a été rapporté que des patients avec différents types et degrés d'enrichissement de signatures mutationnelles (y compris la mutagenèse d'APOBEC) pouvaient présenter une sensibilité à certains médicaments. Cette découverte étend l'intérêt de l'etude des signatures mutationnelles aux traitements des patients.

FGFR3 est l'un des gènes les plus fréquemment mutés dans le BCa et les cellules tumorales mutées pour ce gène sont dépendantes de son expression pour leur prolifération. Dans la première partie de ma thèse, je me suis intéressé à l'étiologie de la mutation prédominante de *FGFR3*. Nous avons établi le catalogue du spectre des mutations ponctuelles de *FGFR3* dans le BCa et identifié 14 mutations récurrentes (fréquence ≥ 2). Comme déjà décrit, la mutation *FGFR3* S249C est fortement surreprésentée (62% des mutations récurrentes de *FGFR3*). En réalisant une étude approfondie de la signature de mutation, nous avons montré que cette surreprésentation de la mutation *FGFR3* était due à une mutation préférentielle par APOBEC et non à un gain de fonction plus important induit par cette mutation. En plus de *FGFR3* S249C, 44 mutations fréquentes (représentant près de la moitié des mutations fréquentes du BCa) ont été identifiées comme étant associées à la signature mutationnelle APOBEC et la plupart d'entre elles étaient surreprésentées par rapport à d'autres mutations au sein du même gène. Il est intéressant de noter que ces mutations associées à APOBEC incluaient à la fois de nouveaux 'conducteurs' et des 'passagers' fréquents potentiels et qu'elles pouvaient potentiellement prédire la réponse à l'immunothérapie et à un traitement anti-APOBEC.

Dans la deuxième partie de cette thèse, nous nous sommes intéressés aux effets fonctionnels du gène *FGFR3* dans le BCa. En utilisant un modèle de souris transgénique, nous avons apporté la première preuve *in vivo* selon laquelle cette mutation *FGFR3* S249C conférait un pouvoir de transformation maligne. Ce processus était associé à une instabilité accrue du génome et à une angiogenèse accrue, probablement induites par le facteur induisant l'hypoxie (HIF1A). En outre, nous avons caractérisé le réseau de régulation contrôlé par FGFR3 en analysant des données protéomiques obtenues par spectrométrie de masse à partir d'une lignée de cellules cancéreuses du cancer de la vessie portant la mutation *FGFR3* S249C - UMUC14. Plusieurs voies de signalisation bien connues comme étant régulées par FGFR3 ont été identifiées. Nous avons également mis en évidence de nouvelles cascades de signalisation suite à l'activation de FGFR3 pouvant être jouer un rôle dans la progression tumorale, notamment un axe FGFR3 / HIF1A / angiogenèse qui a été validé dans certains modèles de BCa *in vitro* et *in vivo*.

En résumé, cette thèse a permis d'expliquer la raison de la surreprésentation de la mutation S249C de *FGFR3* et a identifié des mutations fréquentes associées à APOBEC. Parallèlement, mes travaux ont démontré le pouvoir oncogénique de la mutation activatrices *FGFR3* S249C dans un modèle de souris transgéniques. Notre étude protéomique a également permis de mettre en évidence l'axe FGFR3 / HIF1A / angiogenèse également identifiée à partir de ce modèle. Les travaux présentés ici apporte un nouvel éclairage sur la valeur clinique potentielle de l'implication d'APOBEC dans la tumorigenèse et améliore les connaissances sur actuelle du FGFR3 en particulier de son rôle dans le BCa.

Abstract

Bladder cancer (BCa) is a worldwide frequent, deadly and costly urothelial malignancy. Genetically, BCa bears considerable tumor mutation burden, just following melanoma and lung cancer. The hotspot somatic mutations are cumulative consequence of the endogenous mutation process and / or exogenous mutagenic exposures in addition to functional selection. Each mutational process/exposure leaves a characteristic fingerprint, termed as a mutational signature, on the cancer genome in the form of preferential types of mutation occurring in specific sequence contexts. One of such mutational signatures, namely APOBEC mutagenesis induced by overactivity of APOBEC deaminases, is ubiquitous across several cancer types and particularly predominant in BCa. Thus, revealing hotspot mutations associated with APOBEC mutational signature is useful to understand the aetiology of BCa. In addition, very recently, it has been reported patients with different type and enrichment degree of mutational signatures (including APOBEC mutagenesis) may be susceptible to certain drugs. Herein, this finding further extends implication of mutational signatures to clinical treatment.

FGFR3 is one of the most frequently mutated genes in BCa and an oncogenic driver. On one hand, we were interested in the aetiology of the over-represented mutation on *FGFR3*. We systematically catalogued the *FGFR3* point mutation spectrum in BCa and identified 14 recurrent residues (frequency ≥ 2). As already reported, one hotspot mutation – *FGFR3* S249C – was strongly over-represented compared to other recurrent *FGFR3* mutations (62% of all recurrent mutations). Based on in-depth investigation of mutational signature, we revealed that this over-representation of *FGFR3* S249C mutation was merely favoured by APOBEC mutagenesis rather than a stronger functional selection compared to other oncodriver mutations on *FGFR3*. Similarly, together with *FGFR3* S249C, 44 hotspot mutations (accounts for nearly half of all hotspot mutations in BCa) were pinpointed to be associated with APOBEC mutational signature and most of them were over-represented compared to other mutations within the same gene. Interestingly, these APOBEC-associated hotspot mutations included both novel potential drivers as well as passengers, and had a potential to predict responders for immunotherapy and future anti-APOBEC treatment.

On the other hand, we were interested in functional effects of FGFR3 activation in BCa. We provided the first *in vivo* evidence that *FGFR3* S249C mutation conferred potency to BCa transformation using a transgenic mouse model. This process was associated with increased genome instability and enhanced angiogenesis probably mediated by hypoxia-inducing factor (*HIF1A*). Further, we tried to characterize FGFR3-driven regulatory network through mass spectrometry based proteomic data generated in a BCa cell line bearing *FGFR3* S249C mutation – UMUC14. As expected, several well-known FGFR3 regulated signaling pathways could be identified. Of note, we also highlighted some novel signaling cascades that may be relevant to FGFR3 activation, including a FGFR3/HIF1A/angiogenesis signaling axis that we validated in several *in vitro* and *in vivo* BCa models.

In summary, this thesis revealed the reason for the over-representation of S249C mutation on *FGFR3* and identified APOBEC-associated hotspot mutations. Meanwhile, my work demonstrated the oncogenic property of *FGFR3* S249C mutation in transgenic mouse model. The FGFR3/HIF1A/angiogenesis regulation identified from this model can also be highlighted by our proteomic study. The work presented here sheds new light on the potential clinical value of APOBEC involvement in tumorigenesis and improves the current understanding of *in vivo* oncogenic property of FGFR3 mutation and signaling pathways induced by altered FGFR3 in BCa.

This thesis is dedicated to my beloved wife, Mme Jie Yang, for her love, company, endless support and encouragement.

Remerciements

Je souhaite tout d'abord remercier mon directeur de thèse, François Radvanyi et ma co-directrice de thèse, Isabelle Bernard-Pierrot, pour m'avoir accueilli au sein de leur laboratoire et m'avoir fait confiance tout au long de cette thèse. Je suis heureux d'avoir été formé à la recherche sérieusement, écouté avec patience et d'avoir pu partager de très nombreuses idées scientifiques.

Je tiens également à adresser mes sincères remerciements à João Barata et Eric Letouzé pour avoir accepté de juger mon travail de thèse en tant que rapporteurs. I would like to thank João Barata for nice discussions and his suggestions to rationalize my project hypotheses. Je remercie beaucoup Serge Roche et Thierry Lebret pour leur participation à ce jury en tant qu'examinateur et président du jury, et leurs lettres de recommandation qui m'ont aidé à être accepté à un cours à HongKong et à obtenir un poste dans un hôpital en Chine. Je voudrais aussi remercier Celio Pouponnot et Christian Auclair pour avoir participé plusieurs fois à mon comité de thèse.

Je remercie chaleureusement toute l'équipe Oncologie Moléculaire : Virginia, qui m'a formé et m'a aidé à réaliser toutes les expériences de protéomique, Xiangyu, pour son amitié et son soutien pour toutes les analyses bioinformatiques, Florent et Hélène, pour le temps passé à me former et à la discussion, Laure, Daniela, Aura et Jacqueline pour l'entraide et pour faire face aux difficultés de la thèse ensemble, Elodie Chapeaublanc, Clémentine et Céline, pour m'assister et leur encouragements, Jing, Linda, Jennifer, Elodie Guyon et Yanish pour partager leur joie de vivre. Je tiens à remercier tous les membres de l'équipe, passés et présents, pour le bout de chemin à mes côtés, pour l'ambiance incroyable, les afterworks et les nombreuses discussions, qu'elles parlent de science ou du reste.

Je souhaite remercier les membres des équipes Glukhova, Chavrier et Johannes, pour les discussions en culture cellulaire ou entre deux portes. Je remercie Fatima pour sa bonne humeur quotidienne. Je remercie Michèle pour son écoute et aide toujours aussi précieuse. Je remercie tous les collègues chinois de Curie pour passer plusieurs 'Chinese Seminar' ensemble. Je remercie aussi le gouvernement chinois pour m'avoir financé pendent quatre ans.

Je remercie enfin tous mes proches, famille et amis, pour leur soutien infaillible, pour avoir cru en moi, pour m'avoir aidé durant toutes ces années, et pour m'avoir permis d'être où je suis aujourd'hui.

Table of Contents

LIST OF ABBREVIATIONS

AID	Activation-Induced cytosine Deaminase
APOBEC	Apolipoprotein B mRNA-Editing enzyme, Catalytic polypeptide-like
ATM	Ataxia-Telangiectasia-Mutated
ATR	Ataxia Telangiectasia and Rad3-related
BAIAP2L1	BAI1 Associated Protein 2 Like 1
BBN	N-butyl-N-(4-hydroxybutyl)-nitrosamine
BCa	Bladder Cancer
BCG	Bacillus Calmette – Guerin
BKV	BK Polyomavirus
BPT	Bladder Preservation Therapy
CA9	Carbonic Anhydrase 9
CaMKII	Calcium/Calmodulin Dependent Protein Kinase II
CBLL1	Cbl Proto-Oncogene Like 1
CCNB1	Cyclin B1
CDC37	Cell Division Cycle 37
CDK1	Cyclin Dependent Kinase 1
CDKN2A	Cyclin Dependent Kinase Inhibitor 2A
CGH	Comparative Genomic Hybridization
CIS	Carcinomas I*n Situ*
CK2	Casein Kinase 2
CME	Clathrin-Mediated Endocytosis
COSMIC	Catalogue of Somatic Mutations in Cancer
CREB1	cAMP-Response Element-Binding protein 1
DUSP6	Dual Specificity Phosphatase 6
EGFR	Epidermal Growth Factor Receptor
ERBB2	Erb-B2 Receptor Tyrosine Kinase 2
FANFT	N-[4-(5-nitro-2-furyl)-2-thiazolyl]-formamide
FDA	Food and Drug Administration
FDR	False Discovery Rates
FGFR 1/2/3/4	Fibroblast Growth Factor Receptor 1/2/3/4
FGFRL1	Fibroblast Growth Factor Receptor Like 1
FRS2	Fibroblast Growth Factor Receptor Substrate 2
GO	Gene Ontology
GRB2	Growth Factor Receptor Bound Protein 2
HIF1A	Hypoxia Inducible Factor 1 Subunit Alpha
HK2	Hexokinase 2
MIBC	Muscle Invasive Bladder Cancer
MNU	N-methyl-N-nitrosourea
MRES	Multiple Regional Epigenetic Silencing
MS	Mass Spectrometry
MYC	MYC Proto-Oncogene, BHLH Transcription Factor
NMIBC	Non-Muscle Invasive Bladder Cancer
$p27^{Kip1}$	Cyclin Dependent Kinase Inhibitor 1B
PAH	Polycyclic Aromatic Hydrocarbons
PDX	Patient-Derived Xenograft
PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
PIK3R1	Phosphoinositide-3-Kinase Regulatory Subunit 1
PKA	Protein Kinase A
PLCG1	Phospholipase C Gamma 1
PP2A	Protein Phosphatase 2 Catalytic Subunit Alpha
PPI	Protein-Protein Interaction
pS	phosphoserine
pT	phosphothreonine
PTB	PhosphoTyrosine-Binding
PTEN	Phosphatase And Tensin Homolog
PTM	Post-Translational Modification
PTP	Protein Tyrosine Phosphatases
PTPN	Protein Tyrosine Phosphatase Non-Receptor Type
PTPRG	Protein Tyrosine Phosphatase Receptor type G
pY	phosphotyrosine
REV1	REV1 DNA Directed Polymerase
RPPA	Reverse Phase Protein Array
RTK	Receptor Tyrosine Kinase
S6K	Ribosomal Protein S6 Kinase B1
SH2	SRC Homology-2
SH2B1	SH2B Adaptor Protein 1
SHC1	SHC Adaptor Protein 1
SILAC	Stable Isotope Labeling with Amino acids in Cell culture
SNP	Single Nucleotide Polymorphism
ssDNA	single-stranded DNA
STAT	Signal Transducer And Activator Of Transcription
STRN	Striatin

HMGB1	High Mobility Group Box 1	*TACC3*	Transforming Acidic Coiled-Coil Containing Protein 3
HPSG	Heparan Sulphate Proteoglycan	TCGA	The Cancer Genome Atlas
HSP90	Heat Shock Protein 90	TF	Transcription Factor
ICAT	Isotope-Coded Affinity Tag	TiO2	Titanium dioxide
Ig	Immunglobin	TKI	Tyrosine Kinase Inhibitor
IGFBP3	Insulin Like Growth Factor Binding Protein 3	TMB	Tumor Mutation Burden
IPA	Ingenuity Pathway Analysis	TMT	Tandem Mass Tag
IP-CoIP	Immunoprecipitation / Co-immunoprecipitation	TURB	Transurethral Resection of Bladder
iTRAQ	isotope Tags for Relative and Absolute Quantification	UII	Uroplakin 2
JAK	Janus Kinase	*UNG*	Uracil DNA Glycosylase
KEA	Kinase Enrichment Analysis	UTI	Urinary Tract Infection
KEGG	Kyoto Encyclopedia of Genes and Genomes	UV	Ultraviolet light
m6A	RNA N^6-methyladenosine modification	*VEGFA*	Vascular Endothelial Growth Factor A
MAPK	Mitogen-Activated Protein Kinase	Y2H	Yeast Two-Hybrid

Abbreviations in italics represent gene names.

List of figures and tables

Figures:

Tables:

Preamble

Cancers arise from various genetic and epigenetic alterations in the genome, including acquired somatic mutations. Regional variation in mutation rates can result from the DNA accessibility to the mutagenic or repair process, which are affected by several factors, like DNA replication timing, chromatin organization and gene expression, etc. Hotspot mutations do not accumulate randomly but are positively selected by two major sources, the endo/exo-genous mutation process itself but also by the functional advantage that these mutations can give to the cells. Each mutation process leaves a characteristic fingerprint on the cancer genome in a DNA sequence context depending manner, termed as a mutational signature. APOBEC mutagenesis, one of such mutational signature induced by overactivity of APOBEC deaminases, is ubiquitous across several cancer types but particularly the predominant source of overall mutations in bladder cancer (BCa), thus might be a reasonable aetiology accounting for some of the hotspot mutations in BCa.

Fibroblast growth factor receptor 3 (*FGFR3*) is one of the most frequently mutated genes in BCa. More than 65% of non-muscle invasive BCa (NMIBC) and 15% of muscle invasive BCa (MIBC) carry an *FGFR3* mutation. For two decades, the strong over-representation of *FGFR3* S249C mutation has been noted in BCa and supposed to be due to a greater functional advantage of this mutation compared to the other recurrent *FGFR3* mutations. However, there is minor difference of malignant transforming potential between S249C and the second frequent *FGFR3* mutation – Y375C despite of their biased mutation rates in BCa.

Receptor tyrosine kinases (RTKs) activated signaling transduction relies on complex protein-protein interactions (PPIs) and hierarchical post-translational protein modifications (PTMs). *FGFR3* mutations result in aberrant activation of FGFR3 and confer uncontrolled growth of BCa. Therapy targeting FGFR3 has been an attractive field and the first pan-FGFR inhibitor has been approved very recently in the US to treat advanced BCa presenting a genetic alteration of the receptor (mutation or translocation). However, optimal treatment regimens need a full understanding of exact signaling pathways to appeal to personalized medical care and to overcome potential acquired drug resistance. Compared to well-studied RTKs, e.g. EGFR, FGFR3-dependent PPIs and PTMs are poorly characterized.

Throughout my PhD, I focused on exploring the association between APOBEC mutagenesis and hotspot mutations (including *FGFR3* S249C) in BCa to understand the responsible

underlying mechanisms. I was also interested in generating genetically engineered mice with *FGFR3* mutation and revealing FGFR3-driven signaling pathways in BCa. The use of large-scale proteomic screening in an endogenous *in vitro* working system, gave me the opportunity to have a more complete picture of FGFR3-driven signaling cascades in BCa.

In order to present the results of my studies, I will first introduce clinical features of BCa from epidemiology to treatments. Next, I will orient to somatic mutations and driving forces accounting for these mutations. Specifically, I will discuss the characteristics of APOBEC mutagenic process and mutation spectra on our gene-of-interest – *FGFR3* in detail. Then I will give a global view of the signaling pathways triggered by RTKs / FGFRs and their trafficking / feedback regulation modes as well as the clinical perspectives by targeting FGFR3 in cancers. In a last introductory section, I will summarize *in vitro* and *in vivo* working models and large-scale strategies for investigating cell signaling driven by RTKs / FGFRs.

Chapter 1 Introduction

The urinary bladder is an organ that accumulates urine via the ureters from the kidneys. Bladder urothelium has specialized sensory and signalling properties that allow the bladder to respond to inflammatory / pathogenic, chemical and mechanical stimuli. Although the urothelium is mitotically quiescent, with a very low constitutive rate of cell turnover (approximately 6 weeks in mice and six months to one year in human), it has a high regenerative capacity and can enter into rapid proliferative state in response to the aforementioned injuries (Kreft et al., 2005; Lewis, 2000).

I. Clinical features of bladder cancer

Epidemiology

Bladder cancer (BCa) is the second frequent urologic malignancy, followed after prostate cancer, with approximately 429,000 new cases and 165,000 deaths per year in the western world (Ferlay et al., 2015). In China, BCa becomes the most prevalent genitourinary cancer with approximately 80,500 new cases and 32,900 deaths yearly (Chen et al., 2016). In recent years, the incidence of BCa is increasing due to improved performance of detection strategies and population aging (Miyazaki and Nishiyama, 2017). Of note, BCa develops in a gender-specific manner and occurs three to four times more frequently in men than in women worldwide (Dobruch et al., 2016), the reasons of which are not fully elucidated. Differences in smoking habits are insufficient to account for this and suggested hormonal effects remain unclear (Hurst et al., 2017).

Risk factors

Tobacco use. Cigarette smoking is the best established independent risk factor for BCar in both men and women, increasing the risk of disease incidence by two to fourfold (Freedman et al., 2011; Zeegers et al., 2004). This risk is attenuated to different degree depending on the duration of smoking cessation and the type of tobacco. Tobacco smoke contains numerous carcinogens that contribute to the initiation and promotion of cancer. The metabolic products of these chemicals are excreted by kidneys, making them directly toxic to the human urinary bladder. As a result of the metabolism of these compounds, DNA-adducts are formed, leading to

permanent genetic mutations that, occurring in oncogenes, tumor suppressor or DNA repair genes, may result in uncontrolled growth and eventually cancer.

Occupational exposure. Following smoking, occupational exposure to carcinogens – like aromatic amines (such as benzidine), polycyclic aromatic hydrocarbons (PAH) and heavy metals – are considered as the second most important factor for increased risk of BCa. Roughly, 20% of all bladder cancers have been suggested to be related to such exposure, mainly in industrial areas processing paint, dye, metal and petroleum products (Burger et al., 2013). Farmers and forestry workers seem to show a relatively lower risk (Al-Zalabani et al., 2016).

Water contaminants. Arsenic is a naturally occurring toxic metalloid prevalent in the earth's crust. It enters drinking-water sources in a dissolved state primarily resulting from the weathering of rocks. Arsenic in drinking water is a public health issue worldwide and has been linked with the risk of urinary BCa depending on dose exposure (Saint-Jacques et al., 2014). Even though the associated-risk is uncertain for low-level exposure, it is commonly agreed that concentrations greater than 150 μg/L could be of risk.

Urinary tract infection (UTI). Chronic inflammation due to urolithiasis or urinary tract infection (caused by bacteria, parasite or virus) has been proposed to be associated with the development of BCa. The most prominent example is the infection by *Schistosoma haematobium*, a parasitic worm that causes urogenital schistosomiasis and squamous cell carcinoma of the bladder in Africa region (Mostafa et al., 1999). Recently, there are also increasing evidence suggesting the oncogenic association between the BK polyomavirus (BKV) reactivated infection, which approaches up to 90% population since their childhood and is latent asymptomatically in human urothelium (Knowles, 2006), and development of BCa (Fioriti et al., 2003; Prado et al., 2018). Beyond all above, the majority of UTI is caused by bacteria (predominantly E. coli) and manifests as cystitis. Several observational studies demonstrate that exposure to UTI favours BCa (Sun et al., 2013; Vermeulen et al., 2015). Some possible explanations to support their hypothesis are: 1) chronic inflammation of the bladder may induce urinary retention and stasis, which can potentially increase exposure and absorption of carcinogens present in the urine; 2) bacterial flora in the urine may contribute to the production of nitrites that are converted to carcinogenic nitrosamines. However, controversial or even inverse results reported by others imply that UTI shows no association or even reduced risk in women patients for BCa due to the corresponding immune response triggered by infection (Jhamb et al., 2007; Jiang et al., 2009). Interestingly, all studies agree that the administration of antibiotics for UTI shows significant protective effect against BCa.

Medical conditions. BCa may arise as a consequence of exposure to ionizing radiation and pharmaceutical agents. An increased incidence rate of BCa following radiotherapy for prostate cancer has been observed (Abern et al., 2013). The long-term use of two pharmacologic agents – Cyclophosphamide (an alkylating agent mainly applied in lymphoma and leukemia) and Pioglitazone (an anti-diabetic drug of the thiazolidinedione class) – has also been found to be associated with a higher incidence of BCa (Knight et al., 2004; Lewis et al., 2011).

Pathology, TNM stage and Grade

Most tumors of the urinary bladder are urothelial carcinomas (90% to 95%), followed by less common histologic types such as sarcoma, signet ring cell carcinoma, squamous cell carcinomas, small cell carcinoma and adenocarcinomas. Urothelial carcinomas is staged using the 2009 consensus Tumour – Node – Metastasis system (TNM system), which describes the extent of invasion (Tis – T4), and it is graded according to their cellular characteristics. Tumors presenting a T2 stage or more are referred to as muscle-invasive bladder cancers (MIBCs) as opposed to superficial or non-muscle-invasive bladder cancers (NMIBCs, carcinomas *in situ* (CIS), Ta and T1 tumors). Grading of a cancer is determined upon microscopic examination of the biopsy sample. BCa was classified in 1973 by the World Health Organisation (WHO) as urothelial papilloma, tumours of grade 1 (G1, well differentiated), grade 2 (G2, moderately differentiated) or grade 3 (G3, poorly differentiated). In 2004, the grading system of BCa was improved and re-classified as classified as urothelial papilloma, papillary urothelial neoplasms of low-malignant potential (PUNLMP), low-grade papillary urothelial carcinoma and high-grade urothelial carcinoma. The details of staging and grading are summarized in Figure 1-1; page 8.

About 35% of PUNLMP and ~70% of low and high grade Ta tumors develop recurrences (Holmang et al., 2004). Moreover, disease progression of PUNLMP was nearly never reported. Low grade Ta tumors rarely progress (approximately 5%) whereas high grade Ta tumors progress in nearly one of four cases (Holmang et al., 2004). All MIBC tumors are systematically high grade cancers. Finally, among NMIBCs, CIS are high grade at neoplasms. CIS have a more frequent progression rate than papillary tumors (40% to 50% *vs.* 5%) and are often associated with MIBC tumors (Eble et al., 2004).

The high recurrence rate of Ta and T1 (80%) tumors requires regular cystoscopy and lifelong surveillance. Therefore, the cost per patient of BCa from diagnosis to death is the highest of all cancers (Botteman et al., 2003). The TNM status has a very significant impact on prognosis,

with a five-year survival rate for NMIBC tumors of ~90%, while dropping from 85% for T2 to 25% for T4 MIBC tumors and a mean survival of six to twelve months for metastatic BCa (Knowles and Hurst, 2015).

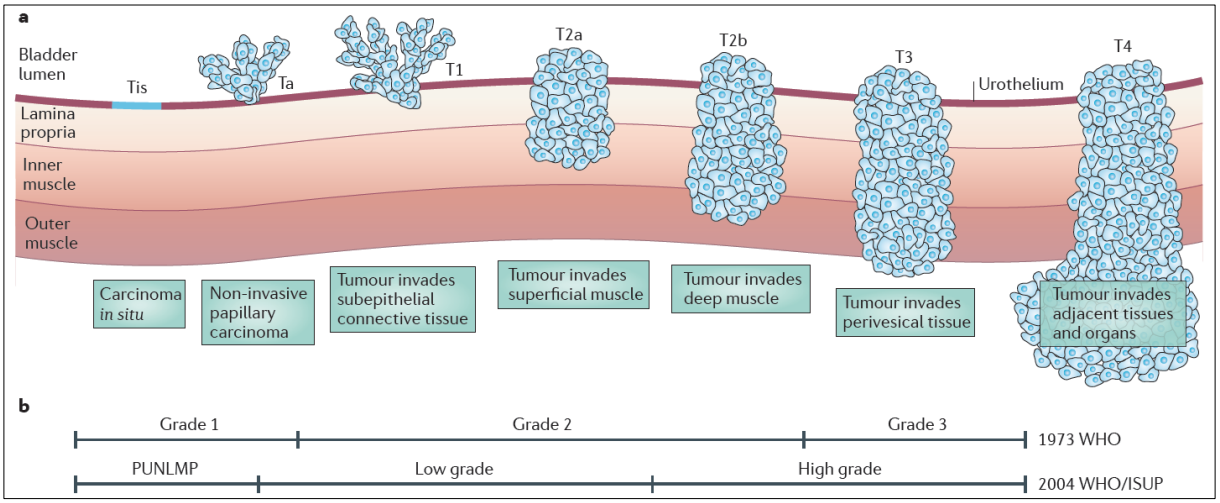


Figure 1-1: *Bladder cancer grading and staging.*

a) *Staging of bladder cancer according to the Tumour – Node – Metastasis (TNM) system.* ***b)*** *Grading according to the 1973 WHO and 2004 WHO (Eble et al., 2004; Mostofi et al., 1974). WHO, World Health Organization. The 1973 grade 1 are mostly reassigned to Papillary Urothelial Neoplasm of Low Malignant Potential (PUNLMP) as well as to low grade tumors. Similarly, 1973 grade 2 are reassigned to low and high grade whereas grade 3 are all assigned to high grade tumors. (Knowles and Hurst, 2015)*

Clinical managements

NMIBC treatments. Transurethral resection of bladder (TURB) is an effective treatment for NMIBC patients. Surgery may be followed by intravesical instillation of a chemotherapy agent (for example, mitomycin C) that has shown promising prospect to delay disease recurrence. Afterward, continual surveillance is conducted by periodic cystoscopy and cytological examination of the urine sediment. Where multiple and/or high-grade tumours or CIS are detected, a course of Bacillus Calmette – Guerin (BCG) intravesical instillation may be used, an approach that is effective in 50-60% of treated patients (Babjuk et al., 2017). Considering the combination of BCG supply shortage and alternative bladder-sparing approaches in patients with BCG-refractory disease, the development of hyperthermia-inducing device-assisted therapies have emerged as new options for NMIBC. Those techniques, such as radiofrequency-induced thermochemotherapeutic effect (RITE), conductive hyperthermic chemotherapy, and

electromotive drug administration (EMDA), have shown promising efficacy (Tan and Kelly, 2018). Very recently, clinical trials focusing on immunotherapy alone (Keynote-057, NCT02625961) or combined with BCG (Keynote-676, NCT03711032) therapies in CIS or BCG-refractory NMIBC patients are ongoing.

MIBC treatments. Organ-confined MIBCs are treated by radical cystectomy. Chemotherapy, mainly cisplatin-based treatment which has been proved as the most effective regimen, may be given before or after cystectomy for eligible patients (Alfred Witjes et al., 2017). Considering high complication rates after radical cystectomy and patients' need for preserving their bladder, clinicians have struggled for many years searching for alternative bladder sparing approaches. Indeed, increasing evidence confirm that the trimodality bladder preservation therapy (BPT), which includes a maximal transurethral resection followed by concurrent chemoradiotherapy, is as efficient as radical cystectomy and even better in selected T2 tumors (Johnson and Yu, 2018). Currently, multiple clinical trials are investigating the role of immunotherapy as either concurrent or neo-adjuvant drug within trimodal-BPT strategy as well (Table 1-1; page 9).

Table 1-1: *Ongoing trials in BPT therapy using immunotherapy.*

BPT, bladder preservation trimodality. (Johnson and Yu, 2018)

Study ID	Phase	Immunotherapy	Treatment	Primary outcome measures
NCT02662062	II	Pembrolizumab	RT + cisplatin + pembrolizumab	Grade 3 or 4 acute toxicities
NCT02560636	I	Pembrolizumab	Hypofractionated RT + pembrolizumab → pembrolizumab	Maximum tolerated dose and toxicity rates
NCT02621151	I/II	Pembrolizumab	Hypofractionated RT + pembrolizumab + gemcitabine	2-year bladder intact disease-free survival
NCT03171025	II	Nivolumab	Concurrent CRT → nivolumab	2-year failure-free survival
NCT03150836	I/II	Durvalumab, tremelimumab	Durvalumab + RT or durvalumab + Tremelimumab + RT	Toxicity and progression-free survival
NCT02891161	I/II	Durvalumab	Durvalumab + RT → durvalumab	Dose limiting toxicity rate, progression-free survival and disease control rate

RT radiation, + with, → followed by, *CRT* chemoradiation

Metastatic cancer. About 50% of MIBC tumors develop metastatic diseases and result in a 5-year survival of ~5% in this setting. Cisplatin-based chemotherapy, the standard first-line treatment, provides overall survival benefit; however, up to two-thirds of patients are ineligible due to impaired performance status or comorbidities (eg. renal dysfunction) (Galsky et al., 2011). Immunotherapy, either approved as second-line treatment for cisplatin failed cases or as first-line treatment for cisplatin-ineligible patients, has shown efficacy and a tolerable safety profile in locally advanced or metastatic urothelial cancer with about 20% objective response rate (Bellmunt et al., 2017; Rosenberg et al., 2016). It is also confirmed that immunotherapy

shows encouraging response as the first-line therapy in cisplatin-ineligible patients with metastatic disease (Balar et al., 2017). Of note, several evidence noted that locally advanced or metastatic urothelial cancer bearing *FGFR3* mutations are less responsive to immunotherapy (Kardos et al., 2016; Sweis et al., 2016). One phase III clinical trial is comparing the efficacy of immunotherapy versus anti-FGFR3 therapy (Erdafitinib) in patients with metastatic urothelial cancer and aberrant FGFRs alterations (THOR, NCT03390504). Very recently, Balversa (Erdafitinib) as a single agent has been approved by the Food and Drug Administration (FDA) to treat locally advanced or metastatic urothelial cancer (https://www.drugs.com/history/balversa.html). To date, this is the first approved targeted therapy based on FGFR inhibition in human cancers.

II. Genetic features of bladder cancer

Cancer as a genetic "train wreck"

In 2011, Hanahan and Weinberg have enumerated a list of hallmarks acquired by cancer cells leading to tumor growth and metastasis (Figure 1-2; page 11) (Hanahan and Weinberg, 2011). Some of these are required from the first steps of tumour growth while others become necessary as the tumour gets bigger (like angiogenesis), or becomes invasive (activating invasion and metastasis). Underlying these acquired capabilities are two enabling characteristics: genome instability, which generates the genetic alterations that facilitates malignant transformation, and inflammation, which promotes multiple hallmark functions. Genetic alterations are modifications conferred in the DNA molecules of a cell, including mutations (eg. point mutations, frameshift mutations, etc) and chromosomal abnormalities (like aneuploidy, copy number variations, translocations, etc). Comprehensive sequence analysis of nearly 1 million tumor samples over the past decade has identified > 2 million coding point mutations, > 6 million noncoding mutations, > 10,000 gene fusions, ~ 61,000 genome rearrangements, ~ 700,000 abnormal copy number segments and > 60 million abnormal expression variants (Forbes et al., 2017). Whole genome sequencing of lung cancer samples in one study showed between 10,000 – 50,000 different single nucleotide variants in tumor cells compared to adjacent normal tissue (Lee et al., 2010). Additionally, epigenetic mechanisms, such as DNA methylation or histone modifications, that play an important role in gene regulation, have

recently been shown to be unstable in cancers too. Accordingly, a multiple regional epigenetic silencing (MRES) phenotype is associated with CIS of bladder (Vallot et al., 2011). In simple terms, cancer is a genetic and epigenetic disease.

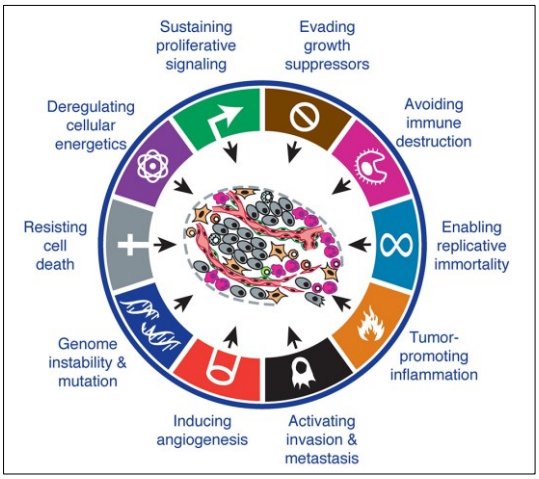


Figure 1-2: *Hallmarks of cancer.*
Illustrated capabilities and characteristics acquired by cancer cells. (Hanahan and Weinberg, 2011)

In my thesis, I am interested in understanding the aetiology of mutations, particularly APOBEC mutational process, and functional effect of point mutations on *FGFR3* gene, one of most frequently gene in BCa. Therefore, the following introduction will focus on these two aspects.

Driving force of somatic mutations

The somatic mutations in cancer genome may be acquired since the fertilized egg and have accumulated over the lifetime of the cancer patient (Figure 1-3; page 12) (Stratton et al., 2009). Overall mutation rate depends on the DNA accessibility to mutagenic or repair process, which is affected by several factors. First, transcription activity is associated with mutation rate. Highly expressed genes shows lower mutation rate, possibly owing to transcription-coupled repair (TCR) process (Lawrence et al., 2013). Second, DNA replication timing is another factor affecting somatic mutation rate where the average mutation rate is higher in the late than earlier replicating regions. Late stages of DNA replication are associated with the slowing down of replication forks due to exhaustion of the dNTP pool with consequent accumulation of single-stranded DNA (ssDNA) regions. Because ssDNA is highly susceptible to endogenous and environmental damage, increased mutation rate would be expected in late replication timing (Stamatoyannopoulos et al., 2009). Third, chromatin organization contributes to regional

variation in mutation rates. Closed chromatins, heterochromatin – a highly condensed form of chromatin and quiescent chromatins – zones markedly depleted of histone modifications, contribute to increased mutation rates may be due to reduced accessibility to DNA repair enzymes (Makova and Hardison, 2015).

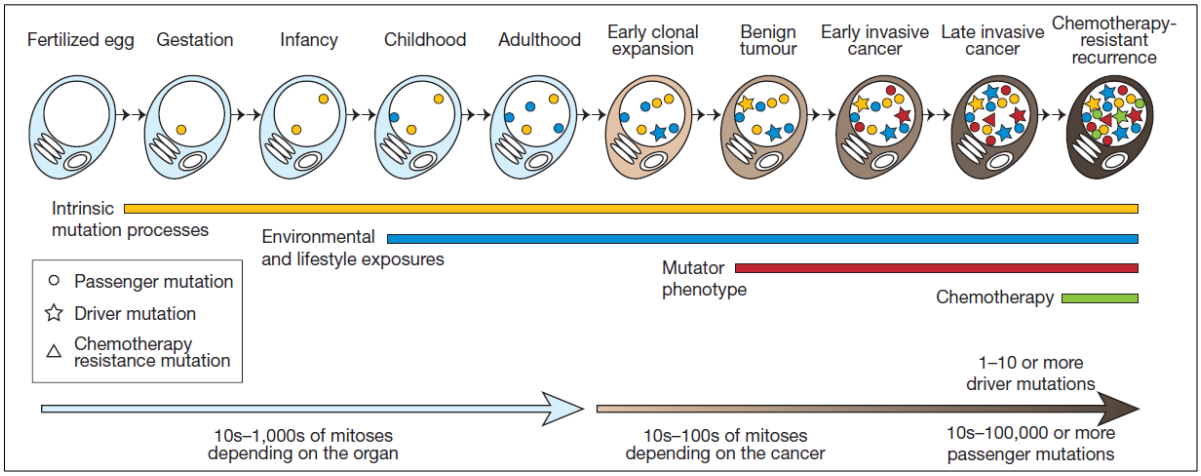


Figure 1-3: *The accumulation of somatic mutations.*
The lineage of mitotic cell divisions from the fertilized egg to a single cell within a cancer showing the timing of the somatic mutations acquired by the cancer cell and the processes that contribute to them. (Stratton et al., 2009)

Further, certain mutation residues can be positively selected by two non-exclusive major sources, functional advantage associated with the mutant and/or DNA sequence dependent affinity to endo- / exogenous mutagenic process.

Functional selection. Somatic mutations providing the fittest advantage for tumor growth are positively selected in a Darwinian fashion, termed 'driver'; conversely, passenger mutations have not been subject to selection and are biologically neutral (Vogelstein et al., 2013). The majority of mutations observed are believed to be passengers and consequently, drivers are always overwhelmed in the sea of passenger mutations in the cancer genome (Greenman et al., 2007). Therefore, it is a formidable challenge to distinguish driver mutations from passenger variants.

DNA sequence dependent selection. DNA sequence dependent forces have been well described to form characteristic mutation patterns (Hodgkinson and Eyre-Walker, 2011), such as frequent C:G → T:A mutations at spontaneously 5-methylated CpG dinucleotide (Cooper and Krawczak, 1989); common CC:GG → TT:AA double nucleotide substitutions in

ultraviolet light (UV)-associated skin cancer (Pfeifer et al., 2005); overwhelming C:G → A:T mutations in tabacoo-associated lung cancer (Rodin and Rodin, 2000); and prevalent T:A → A:T mutation within CTG trinucleotide induced by a herbal plant (aristolochic acid) in urothelial carcinoma (Poon et al., 2015). Coincidentally, aflatoxin exposure leaves a unique fingerprint in *TP53* gene by inducing specifically R249S mutation (C → A substitution) and is associated with liver cancer development (Letouzé et al., 2017; Wogan, 1992).

Mutational signature. Recently, a refined mathematical algorithm has been successfully applied to systematically investigate DNA sequence dependent forces. Thanks to the completion of human genome sequencing and data availability in multiple cancer types, comprehensive catalogues of somatic mutations have been generated by decomposing distinct mutation patterns from a set of cancer samples (Alexandrov et al., 2013a, 2013b; Nik-Zainal et al., 2012; Roberts et al., 2013). In brief, considering the impact of neighbour bases on the rate and type of mutated single nucleotide, this method achieves 96 possible mutated trinucleotides by taking into accout the bases immediately 5' and 3' to each of the six classes of base substitution (C>A, C>G, C>T, T>A, T>C, and T>G), that are used for subsequent fraction in each cancer sample. Throughout the pancancer analysis, 30 clusters have been identified according to fraction profiles, known as mutational signatures (Figure 1-4A; page 14). Cancer related mutational signature is a fast-paced and attractive field, because based on the associations between some of these signatures and either sequence specificity to known endo- / exogenous mutagenic aetiologies or mutated genes relevant to DNA damage / repair processes, they are expected to provide insights into cancer aetiology and the source of mutations (Alexandrov et al., 2013b; Greenman et al., 2007). For instance, signatures similar to previously documented UV or tabacoo-associated mutation patterns (correspond to Catalogue Of Somatic Mutations In Cancer (COSMIC) signature 7 and signature 4, respectively) could be reproduced in skin and lung cancers (Alexandrov et al., 2013b; Forbes et al., 2017) and the associations bewteen various mutational signatures and *TP53* mutations, the most frequently mutated gene in human cancer genome, has been demonstrated comprehensively (Figure 1-4B; page 14) (Giacomelli et al., 2018).

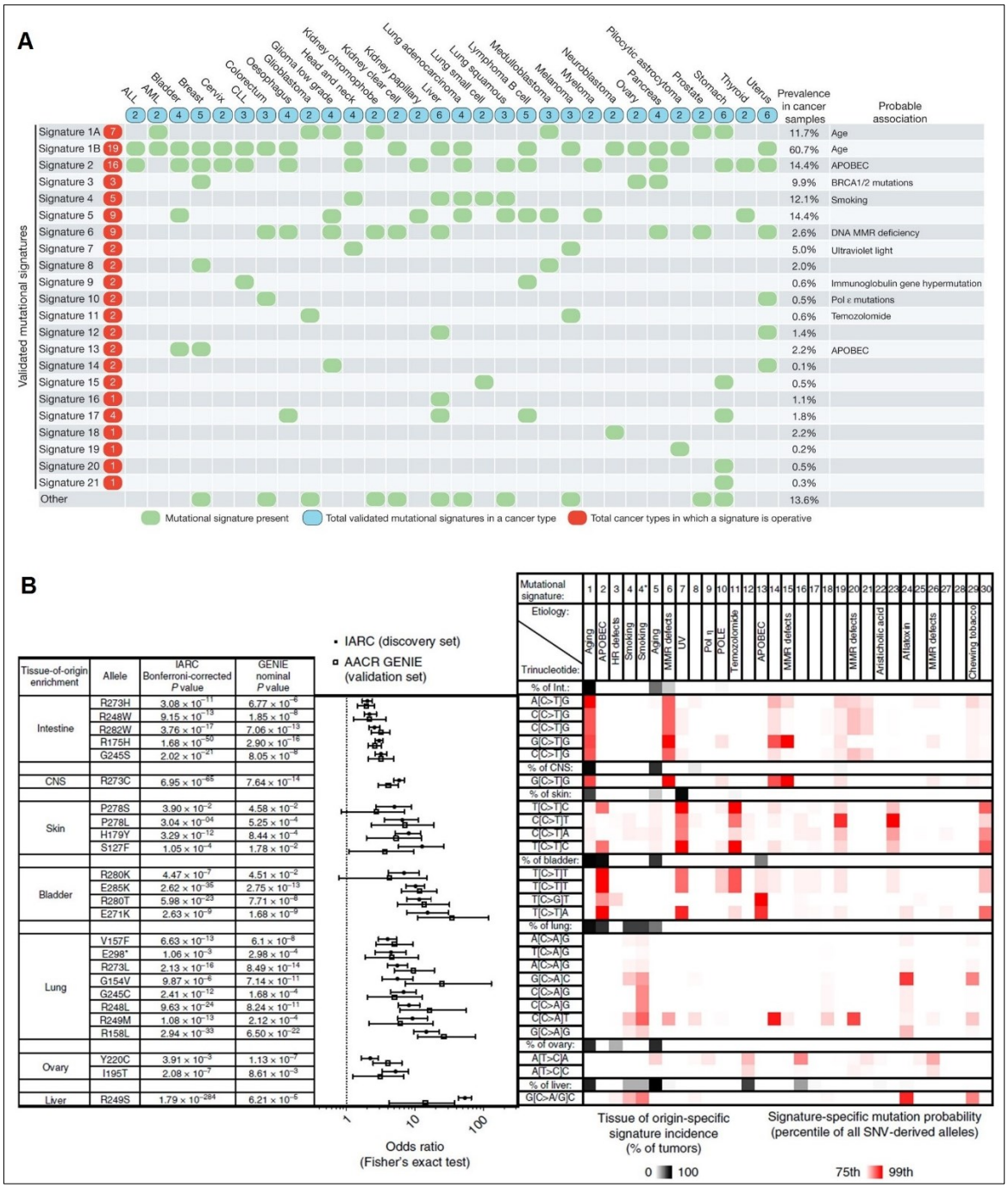


Figure 1-4: *Mutational signatures and specific imprints on cancer gene.*

A) *The presence of mutational signatures across human cancer types.* ***B)*** *Tissue of origin-selective TP53 mutations are linked to specific mutational processes.* (Alexandrov et al., 2013b; Giacomelli et al., 2018)

APOBEC (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like) mutational signature is one of aforementioned signatures and widespread in various cancer types, but significantly enriched in bladder and cervical cancers. This mutagenesis is characterized by

cytosine to uracil (C-to-U) mutation which generally mediated by APOBEC deaminases. In BCa, in addition to APOBEC mutational signature, there are three other mutational signatures have been identified to date from The Cancer Genome Atlas (TCGA) BCa dataset, namely aging-associated signature, POLE-associated signature and ERCC2-associated signature (Robertson et al., 2017). Since APOBEC mutagenesis has been the dominant source of overall mutations in BCa, I will introduce it in particular and its relevance to BCa in more detail below.

APOBEC mutational signature & Bladder cancer

APOBEC deaminases & mutagens. The human cells have the capacity to express a total of 11 distinct APOBEC members including activation-induced cytosine deaminase (AID), *APOBEC1* (A1), *APOBEC2* (A2), *APOBEC3* (A3A – A3H) and *APOBEC4* (A4) (Figure 1-5; page 15). These deaminases normally function as DNA mutators participating in the innate immune system that defends against their targets (retrovirus and retrotransposon) propagation (Petersen-Mahrt et al., 2002; Sheehy et al., 2002). In cancers, except for A2 and A4, all other enzymes are capable of deaminating cytosine to uracil (C-to-U) (Swanton et al., 2015). Of note, subcellular compartmentalization of APOBEC proteins regulates their

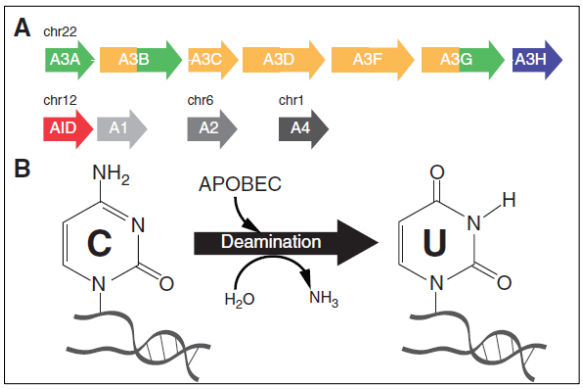


Figure 1-5: *APOBEC family of DNA cytosine deaminases.*
A) *A schematic of the genes encoding the 11-member APOBEC family in humans. The colors represent the different categories of catalytic domains in APOBEC3 enzymes. Green represents a Z1 catalytic domain, orange represents a Z2 catalytic domain, and blue represents a Z3 catalytic domain. AID and A1 are distinctly represented by red and gray, respectively.* ***B)*** *ssDNA cytosine to uracil (C-to-U) deamination is the hallmark biochemical activity of most APOBEC family enzymes. (Swanton et al., 2015)*

DNA-binding activity: although AID, A1, A3A, A3B, A3C and A3H can shuttle between the nuclear and cytoplasmic compartment, AID is preferentially retained in the cytoplasm while A3B predominantly localizes in nucleus; by contrast, A3D, A3F and A3G are exclusively cytoplasmic (Lackey et al., 2013; Salter et al., 2016).

It is difficult to measure APOBEC enzyme activity directly in tumor samples. Multiple studies have shown that the mRNA levels of *APOBEC3A* and *APOBEC3B* are positively correlated with enrichment of APOBEC mutagenesis in various cancer types (Lim et al., 2017). *APOBEC3B* is dominantly associated with cancer cell proliferation and induced by DNA-damaging drugs whereas *APOBEC3A* displays specificity for adaptive immunity and induced by antiviral interferon-stimulated response (Middlebrooks et al., 2016; Ng et al., 2019). Besides, A3A enzyme is over tenfold more proficient than A3B to generate APOBEC mutations (Chan et al., 2015).

Further, other enzymes not belonging to the APOBEC family are also emerging to have an impact on APOBEC mutagenesis, like *REV1* and *UNG* genes (Helleday et al., 2014). In addition, three germline variants have been proposed as independent risk factors associated with APOBEC mutational signature: 1) a 30-kb deletion that eliminates *A3B* and creates an *A3A-A3B* (*A3AB*) chimera (Nik-Zainal et al., 2014); 2) a stable *A3H* haplotypes I (*A3H*-I) in *A3B*-null tumours (Starrett et al., 2016); 3) a three linked single nucleotide polymorphism (SNP) rs1014971 and its two proxies (rs17000526 and rs1004748) (Middlebrooks et al., 2016). All the above variants are associated with APOBEC mutagenesis in breast cancer whereas only SNP rs1014971 is associated with APOBEC mutation pattern in BCa.

Motif specificity & Target ssDNA. COSMIC signature 2 and signature 13 (Figure 1-6; page 17), both characterized by high prevalence of single-nucleotide C → T transition or C → G transversion at stringent 5'-TCW motif (W = A or T), are proposed due to over activity of APOBEC enzymes and emerge as the most popular genetic hallmark of cancer in recent years (Alexandrov et al., 2013b; Roberts et al., 2013). ssDNA are ideal targets of APOBEC enzymes. Even though multiple processes have been known to generate ssDNA, like double-strand breaks (DSBs), R-loop in transcription bubbles and lagging strand template during DNA replication fork, it becomes clear that the main source of ssDNA substrate of APOBEC is related not to DSBs or R-loops but to DNA replication (Adolph et al., 2017; Haradhvala et al., 2016; Seplyarskiy et al., 2016). In addition, the efficiency of APOBEC mutagenesis has also been associated with the propensity of ssDNA to form hairpins, with some APOBEC3 enzymes,

such as APOBEC3A, preferentially targeting loops in the stem-loop structures (Holtz et al., 2013; Nik-Zainal et al., 2016; Sharma and Baysal, 2017).

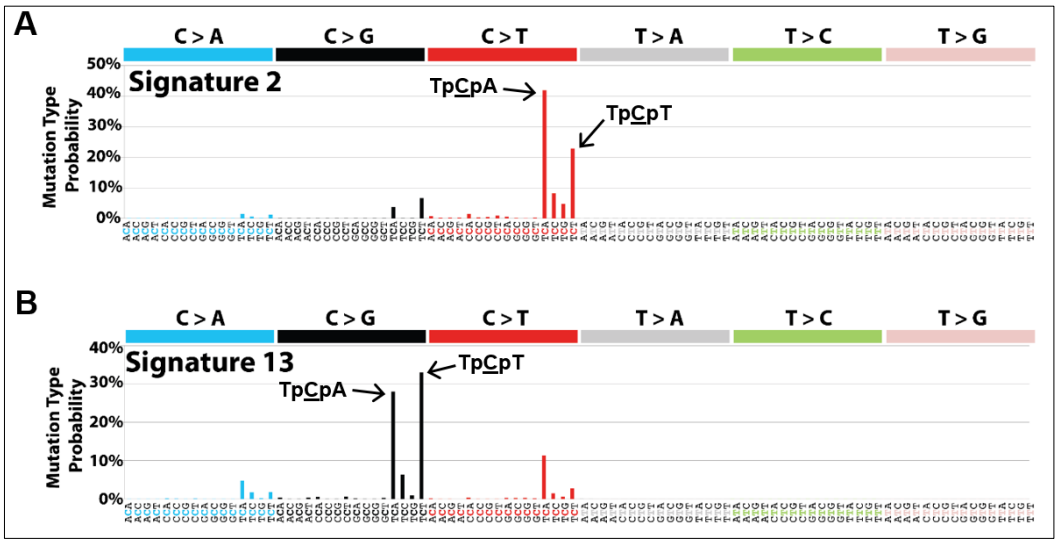


Figure 1-6: *APOBEC mutational signatures show specific patterns.*
A) *COSMIC signature 2 (TCW → TTW, W = A or T, mutated nucleotide underlined).* ***B)*** *COSMIC signature 13 (TCW → TGW, W = A or T, mutated nucleotide underlined). (Forbes et al., 2017)*

Dominance of APOBEC mutagenesis in BCa. Though many of identified mutational signatures are shared across cancer types (eg. age-associated (COSMIC S1) and APOBEC sigantures, etc), several mutational signatures are cancer-type specific due to different extrinsic or intrinsic exposures, such as intensive smoking-associated signature (COSMIC S4) in lung cancer, UV-associated signature (COSMIC S7) in melanoma, BRCA1/2 mutation-associated signature (COSMIC S3) in breast cancer and DNA mismatch repair deficiency or *POLE* mutation-associated signatures (COSMIC S6 and S10) in colorectal and endometrial cancers (Alexandrov et al., 2013b; Forbes et al., 2017). Interestingly, smoking-associated signature has not been identified yet in BCa in spite of considerable contribution of tobacco to bladder tumorigenesis. The reason is still unknown. Of note, following the virus induced cancer (cervical cancer), BCa is the top two cancer type showing high enrichment of APOBEC mutagenesis (Figure 1-7; page 18) (Chen et al., 2017; Roberts et al., 2013). APOBEC mutational signature is the predominant source of overall mutations in BCa, which accounts for ~30% and 65% of all mutations in NMIBC and MIBC, respectively, implying a crucial role of APOBEC mutagenesis in bladder tumorigenesis (Hedegaard et al., 2016; Lamy et al., 2016; Robertson et al., 2017).

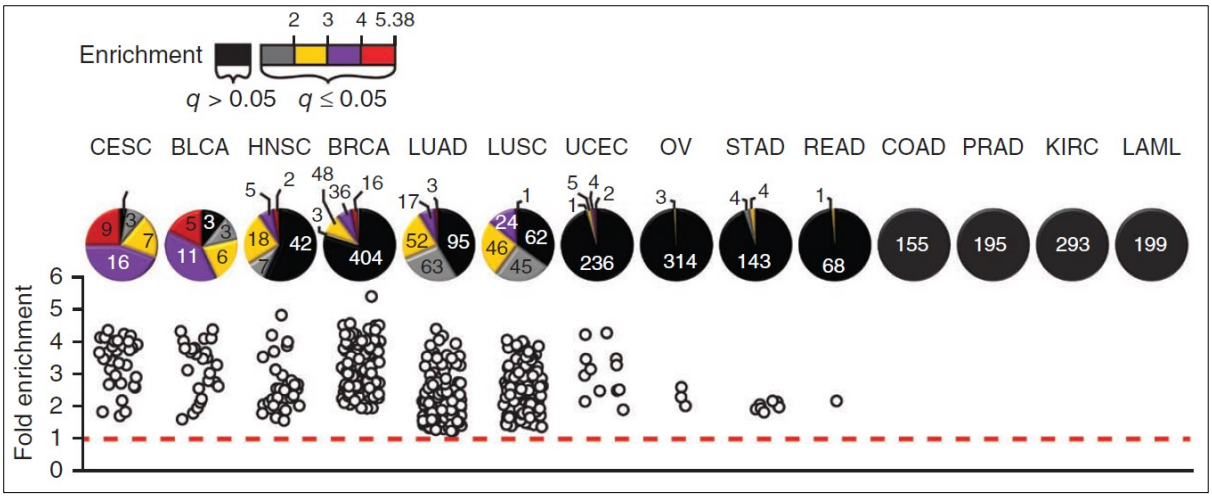


Figure 1-7: *Presence of an APOBEC mutation pattern in various cancer types.*
Cancer types are abbreviated as in TCGA: cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), bladder urothelial carcinoma (BLCA), head and neck squamous cell carcinoma (HNSC), breast invasive carcinoma (BRCA), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), uterine corpus endometrioid carcinoma (UCEC), ovarian serous cystadenocarcinoma (OV), stomach adenocarcinoma (STAD), rectum adenocarcinoma (READ), colon adenocarcinoma (COAD), prostate adenocarcinoma (PRAD), kidney renal clear-cell carcinoma (KIRC) and acute myeloid leukemia (LAML).(Roberts et al., 2013)

Lessons from APOBEC mutagenesis in BCa. BCa ranks the third in terms of bearing tumor mutation burden (TMB), just following melanoma and lung cancer (Figure 1-8A; page 19) (Lawrence et al., 2013). Considering the APOBEC mutagenesis both significantly contributes to mutations in BCa and potentially confers selective pressure, 'driver' mutations are expected to be pinpointed by this mutational process. Indeed, four well-known 'onco-driver' mutations – *PIK3CA* E545K and E542K, *ERBB2* S310F and *TP53* R280T – have been positively associated with APOBEC mutagenesis in BCa. Additionally, the positive association between *PIK3CA* E545K and E542K mutations and APOBEC mutagenesis is also observed in other cancer types presenting an APOBEC mutational signature (cervical cancer, head and neck cancer, breast cancer and lung cancer) (Cannataro et al., 2019; Henderson et al., 2014; Poulos et al., 2018; Temko et al., 2018).

It is well described that numerous of later-arising passenger mutations are responsible for intratumoral heterogeneity (Williams et al., 2016). Interestingly, in lung cancer, APOBEC mutagenesis has been identified as a remarkable contributor to tumor heterogeneity. For

example, APOBEC mutagenesis has been demonstrated as a late event that induces considerable subclonal mutations in operable lung cancer (De Bruin et al., 2014) and as major source of intratumoral heterogeneity in metastatic lung cancer (Roper et al., 2019). BCa is one of the most heterogeneous tumor (Figure 1-8B; page 19) (Raynaud et al., 2018). In spite of considerable contribution to overall mutation by APOBEC mutagenesis and significant intratumoral heterogeneity in BCa, APOBEC mutagenesis seem to be an early and persistant event in BCa (Nordentoft et al., 2014). Hence, the association between APOBEC mutational process with clonal driver mutations can be expected in BCa.

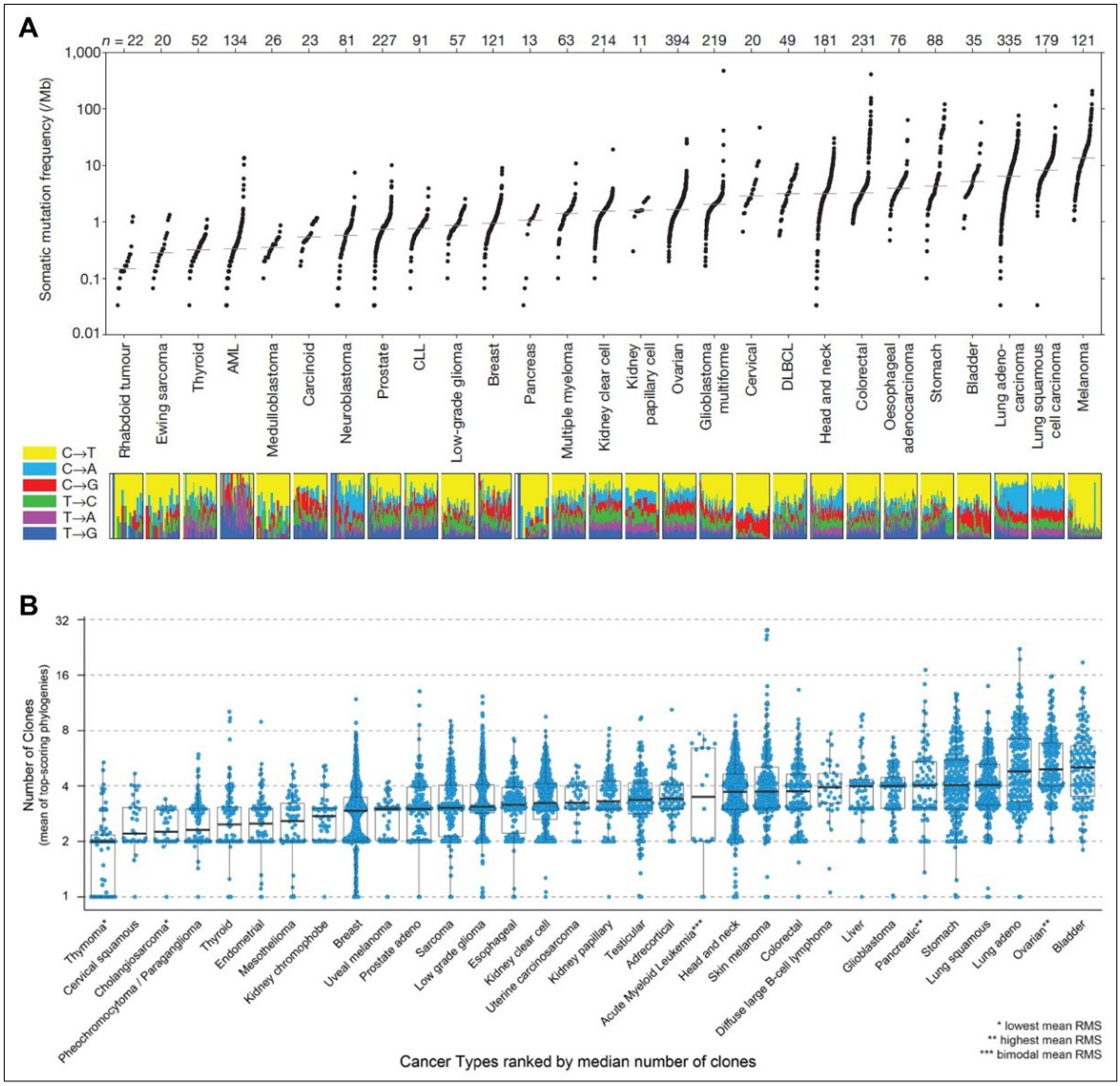


Figure 1-8: *Tumor mutation burden and intratumoral heterogeneity.*

A) *Tumor mutation burden across cancer types (Lawrence et al., 2013). Tumour types are ordered by their median somatic mutation frequency. The bottom panel shows the relative proportions of the six different possible base-pair substitutions, as indicated in the legend on*

the left. ***B)*** *Intratumoral heterogeneity across cancer types (Raynaud et al., 2018). Tumor types are ranked by median number of clones that are estimated in the top scoring PhyloWGS likelihood for each sample. Mean RMS, mean number of reads per mutated sites.*

Last but not least, increasing evidence have highlighted the therapeutic value by perturbation of APOBEC mutagenesis with mainly two strategies: 1) hypo-mutation, by inhibiting APOBEC-dependent tumour to limit intratumoral heterogeneity and potentially suppressing adverse outcomes including recurrence, metastasis, and drug resistance; 2) hyper-mutation, by enhancing the mutagenic effects of APOBEC to make cancer cells suffer catastrophic levels of DNA damage and selectively die (Venkatesan et al., 2018). Indeed, to support the first strategy, specific inhibitors targeting APOBEC enzymes are under investigation (Kvach et al., 2019; Li et al., 2012a). Regarding the second strategy, current studies focus on the interplay between APOBEC activity and DNA damage response (DDR) inhibitors to impair the balance between mutagenic and repair process. In tumors presenting high APOBEC activity, we can expect that the mutagenic process can be accelerated once repressing DNA repair mechanism. Ataxia-telangiectasia-mutated (ATM) and ataxia telangiectasia and Rad3-related (ATR) proteins are crucial DNA damage sensors and key regulators inducing cell cycle arrest and DNA repair (Awasthi et al., 2016). ATM and ATR kinases respond to different types of DNA damages; ATM kinase responds primarily to DNA double-strand breaks (DSBs), whereas ATR kinase responds to a different set of DNA lesions, including those caused by UV light, and to stalled replication forks. In line with the second strategy, several studies have shown that multiple cancer cell lines presenting high APOBEC3A activity are selectively lethal to ATR inhibitors (ATRi), but not to ATM inhibitors (Buisson et al., 2017; Green et al., 2017). Finally, APOBEC-associated tumors exhibited greater likelihood of complete or partial response for immunotherapy in bladder and head and neck cancer in a recent study (Miao et al., 2018). Therefore, there is urgent need to develop method for selecting patients exhibiting high APOBEC mutagenesis and apply optimal clinical treatment for them.

FGFR3 mutations & Bladder cancer

FGFR3 alterations. *FGFR3* (fibroblast growth factor receptor) belongs to a family of structurally related tyrosine kinase receptors (RTKs) encoded by four different genes (*FGFR1-4*). Diverse genetic alterations of *FGFR3* frequently occur in BCa (less common in other cancer types), like point mutations, translocations and copy number variations. *FGFR3* is one of the

most frequently mutated genes in BCa with ~65% of NMIBCs and 15% of MIBCs carry an *FGFR3* mutation (Hedegaard et al., 2016; Robertson et al., 2017). *FGFR3* translocations leading to the production of FGFR3-TACC3 and FGFR3-BAIAP2L1 fusion proteins are identified in 3% of MIBCs (Nakanishi et al., 2015a; Williams et al., 2013; Wu et al., 2013). By contrast, *FGFR3* amplifications is much rare in BCa. The majority of these aberrations identified to date are positively correlated with *FGFR3* gene overexpression and lead to gain-of-function (Mahe et al., 2018; Porębska et al., 2018). Interestingly, overexpression of *FGFR3* is also detectable in a small proportion of *FGFR3* wild-type BCa (Di Martino et al., 2012). The reasons are yet to know.

FGFR3 mutations spectra in human diseases. Most of the somatic mutations of *FGFR3* occur in BCa and in benign skin tumors but much less common in other cancer types. Germline *FGFR3* mutations are also responsible for abnormal bone development. Majority of *FGFR3* mutations occur in three exons (exon 7, exon 10 and exon 15) and show various frequency spectrum among different diseases. In bone disorders – dwarfism and craniosynostosis, many *FGFR3* mutations are frequent and become causal or closely associated to inhibition of chondrocyte proliferation (Foldynova-Trantirkova et al., 2012; Passos-Bueno et al., 2008). All these mutations are germline but identical to somatic mutations in tumors except for frequent P250R, N540K mutation and several nonstop mutations on terminal codon being found only as germline (L'Hôte and Knowles, 2005). Although all *FGFR3*-related dwarfisms present shortening of the long bones, they exhibit a graded phenotypic severity, ranging from relatively milder hypochondroplasia (HCH), severer achondroplasia (ACH) to the neonatal lethal thanatophoric dysplasia I (TD-I) and II (TD-II). This variance is explained by the relative 'activating' potential of a given substitution being wild type < N540K (HCH) < G380R (ACH) << R248C (TDI) = Y373C (TDI) < K650M (TD-I) ≤ K650E (TD-II) (Foldynova-Trantirkova et al., 2012; Krejci et al., 2008; Naski et al., 1996). By contrast, nearly all different kinds of craniosynostosis associated syndromes are exclusively attributed to A391E and P250R mutations. In two common benign skin lesions – seborrhoeic keratosis and epidermal nevus – without malignant potential, activating mutations on two oncogenes (*FGFR3* and *PIK3CA*) have been well documented (Hafner et al., 2007; Logié et al., 2005; Toll and Real, 2008). Even though the mutation residues are similar to those observed in malignant neoplasms, the *FGFR3* mutation spectrum is dominant by R248C mutation in these benign skin tumors (Figure 1-9B, lower; page 22) (Hafner et al., 2010). However, it is still not clear why the same activating mutations do not confer malignant growth in skin as they do in cancers and why they have

opposite effect in bone. In cancers, although *FGFR3* mutations have been identified in multiple tumor types, the highest frequency is observed in BCa (Figure 1-9A; page 22). *FGFR3* mutations were reported in BCa for the first time in 1999 by our team (Cappellen et al., 1999). Since then, multiple following studies confirmed this finding (Al-Ahmadie et al., 2011; Hernández et al., 2006). Of note, the *FGFR3* mutation spectrum is completely different from the one of bone disorders or benign skin lesions. The most frequent mutations in BCa are mutations that create a cysteine residue in the extracellular domain or the first part of the transmembrane domain (S249C being the dominance, Y375C, R248C, G372C), followed by lower frequent S371C, A391E and K650E (Figure 1-9B, upper; page 22).

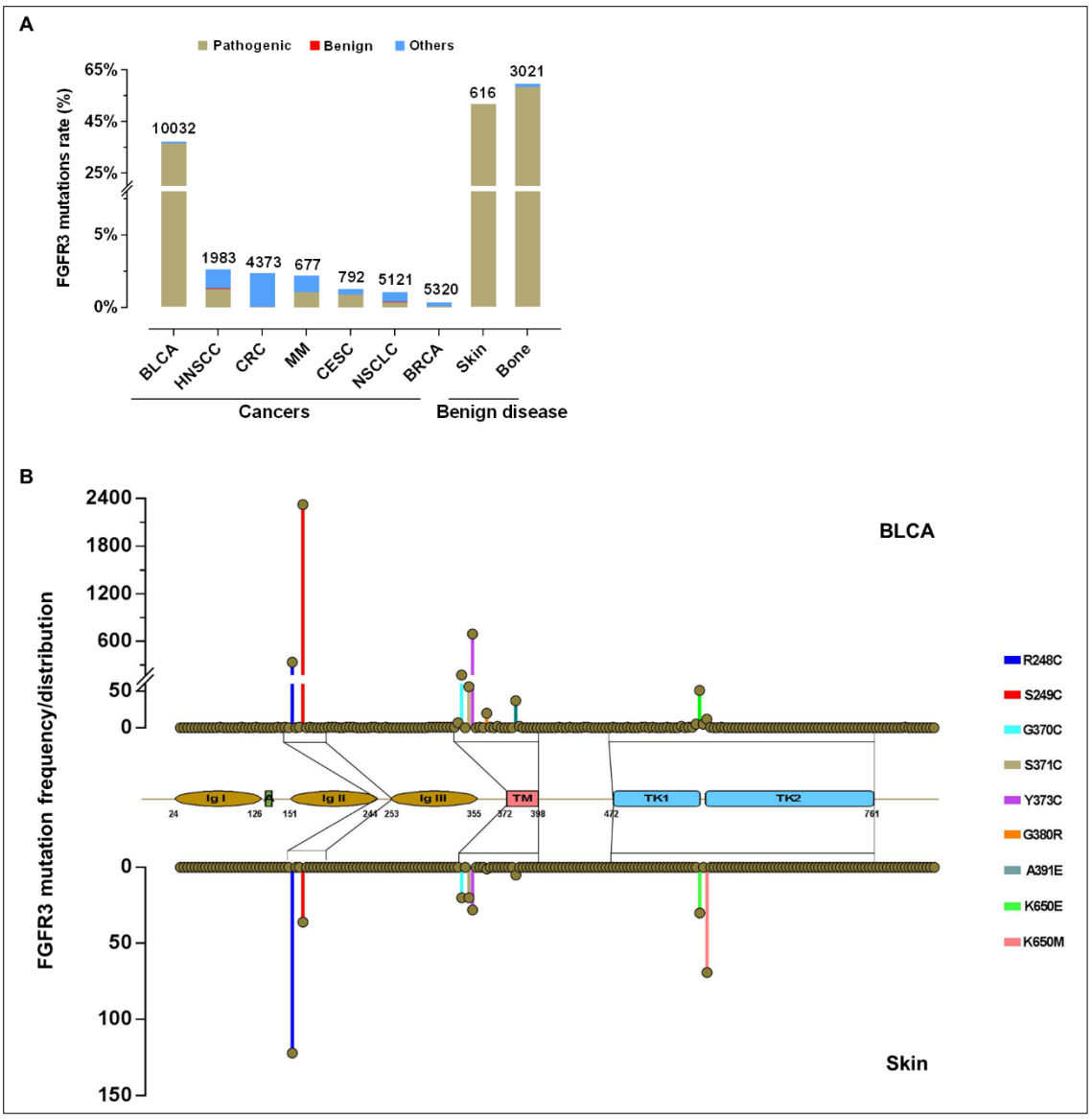


Figure 1-9: *FGFR3 mutation rates and mutation spectra.*

A) *FGFR3 mutation rate across cancer types and benign diseases from 31,935 individuals.* ***B)*** *Examples of mutated amino acid distribution along FGFR3 IIIc sequence in BCa (upper) and*

benign skin lesions (lower). FGFR3 IIIb is the main isoform in cells of epithelial origin and FGFR3 IIIc, the main isoform in chondrocytes. HNSCC, head and neck squamous cell carcinoma; CRC, colorectal cancer; MM, multiple myeloma; NSCLC, non-small cell lung cancer; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; BRCA, breast cancer; Skin, composed of seborrhoeic keratosis and epidermal nevus; Bone, composed of dwarfism and craniosynostosis. FGFR3 mutations found in cancers and benign skin tumors are somatic, those found in bone disorders are germline but identical to the somatic mutations in tumors. TM, transmembrane domain; TK, tyrosine kinase domain. (visualization data based on Shi et al., 2019)

Many of the recurrent *FGFR3* mutations are gain-of-function mutations in cancers. One of the major task of my PhD project is to explore signaling pathways driven by aberrant FGFR3 in BCa. Below, I will introduce general activation mode for FGFRs, what we have known for their downstream cascades and how we can elucidate these signaling pathways.

III. FGFR3-driven signaling pathways in bladder cancer

Divergent pathways in bladder cancer

It is known since decades that two distinct pathways exist in bladder tumorigenesis: the low grade Ta papillary tumor pathway and the carcinoma *in situ* (CIS) pathway (Figure 1-10; page 24) (Billerey et al., 2001; Knowles and Hurst, 2015; Spruck et al., 1994; Wu, 2005). Ta papillary tumors, which account for ~50% of all urothelial tumours, often exhibit low-grade and papillary histology. About 70% of these tumours will recur, but only ~15% will progress. Frequent mutations in the *FGFR3* gene (~70%) are characteristic in this pathway. The homozygous deletion of the tumor suppressor gene – *CDKN2A* – is associated with progression and muscle invasion in this pathway (Rebouissou et al., 2012). CIS pathway (accounting for 20-30% of the urothelial tumours) is often associated with high-grade tumors and invasive potential and characterized by frequent *TP53* mutations and loss of *RB1* function. Additionally, a multiple regional epigenetic silencing (MRES) phenotype is also associated with CIS pathway (Vallot et al., 2011).

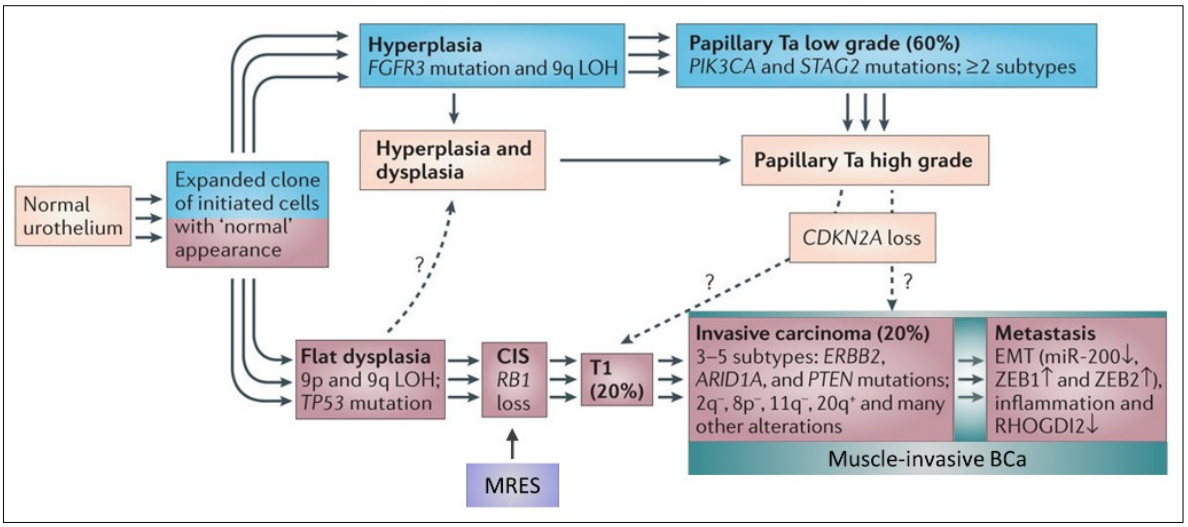


Figure 1-10: *Divergent pathways in bladder tumorigenesis from the normal urothelium. Both pathways are initiated from loss of chromosome 9 (9q–). Ta papillary tumors are much more frequent than tumors arising from the carcinoma in situ (CIS) pathway (70-80% vs. 20-30%, respectively). Ta papillary tumors are characterized by frequent mutations in FGFR3 gene and high recurrence rate (~70%), but rarely progress whereas tumors from the CIS pathway present frequent TP53 mutations, loss of RB1 and MRES phenotype and shows high propensity to invade and metastasize (>50% of tumors of this pathway will give rise to metastasis). MRES, multiple regional epigenetic silencing. (Knowles and Hurst, 2015; Vallot et al., 2011)*

FGFR3 activation

FGFR3 structure. The FGFR3 protein contains ~800 amino acids and is composed of three extracellular Immunoglobulin (Ig)-like domains, an acidic box, a single transmembrane (TM) domain and a cytoplasmic domain containing the catalytic tyrosine kinase (TK) core (Figure 1-11; page 25) (Iyer and Milowsky, 2013). FGFR3 exists as two isoforms, FGFR3 IIIb and FGFR3 IIIc, which are the result of an alternative splicing between exon 8 and exon 9 within Ig-III region. FGFR3 IIIb is the main isoform in the epithelial lineage whereas FGFR3 IIIc is expressed in chondrocytes (L'Hôte and Knowles, 2005). FGFR3 IIIb contains two additional amino acids compared to the FGFR3 IIIc isoform.

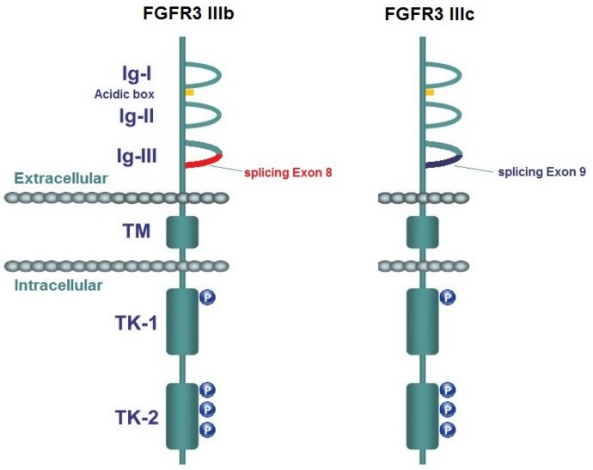


Figure 1-11: *FGFR3 domains and two isoforms owing to alternative splicing. Left, FGFR3 IIIb containing exon 8; right, FGFR3 IIIc containing exon 9. Ig, immunoglobulin; TM, transmembrane domain; TK, tyrosine kinase domain. (Iyer and Milowsky, 2013)*

FGFR3 activation. In physiological condition, the activation of FGFR3 receptor is ligand dependent. The ligand-binding specificity and affinity to distinct FGFR receptors are primarily achieved by alternative splicing event as well as determined by the linker between Ig-like domain II and III (Olsen et al., 2004). Further, the heparan sulphate proteoglycans (HPSGs) or soluble heparin is necessary and increases FGFs binding to FGFRs (Schlessinger et al., 2000). Among the identified 18 FGFs members to date that are capable of binding FGFR1-4, most of them activate a particular subset of FGFRs whereas FGF1 can universally activates each of the seven principal FGFRs (FGFR 1b, 1c, 2b, 2c, 3b, 3c, 4) (Ornitz and Itoh, 2015). Upon the extracellular ligand stimulating signals, a dimerization will be induced between two monomeric receptors. The followed activation of intracellular tyrosine kinase region is provoked by trans-phosphorylation within the dimer. Afterwards, direct phosphorylation on substrates or conformational changes leading to docking sites for adaptor proteins trigger multiple downstream signaling cascades. In cancers, the gain-of-function alterations of *FGFR3*, such as some point mutations or translocations we mentioned previously, will provoke constitutive activation by increased dimerization (like S249C and R248C mutations) or structural changes (like A391E, G380R and K650M/E) in addition to increase the concentration of dimers by gene overexpression (Porębska et al., 2018). Depending on mutation site, this activation sometimes can be independent of ligand stimuli, for instance, *FGFR3* S249C / A391E mutation and FGFR3-TACC3 translocation (Bernard-Pierrot et al., 2006; Chen et al., 2011; Nelson et al., 2016). Of note, *in vitro* evidence shows that *FGFR3* intracellular mutations

(K650N/E/M and X807R) inhibit receptor glycosylation via premature receptor phosphorylation and increase receptor signaling in the Golgi apparatus (Gibbs and Legeai-Mallet, 2007). Interestingly, whereas the dimerization commonly occur between homodimer, heterodimerization has been reported between mutant and wild-type alleles within FGFR3 receptor in bone disorders (Adar et al., 2002; Del Piccolo et al., 2017). In addition, it has been shown heterodimerization can be formed between different RTKs monomers, like heterodimer between EGFR and several RTKs (EGFR-PDGFRβ in BCa and EGFR-MET in lung cancer), ERBB2-ERBB3 heterodimer in esophageal cancer and FGFR1-FGFR2 heterodimer in *in vitro* model (Bellot et al., 1991; Black et al., 2011; Fichter et al., 2014; Tanizaki et al., 2011). Coincidently, Yokote et al reported a formation of heterodimerization between the monomer of FGFR3 and EPHA4 receptors (Yokote et al., 2005).

FGFR3 signaling

Mode of general RTKs signaling pathways. The evolutionary breakthrough regarding tyrosine kinase has been achieved by Hunter et al.'s work in 1980, where he was the first to propose the notion of phosphotyrosine (pY) modification and to prove the v-SRC protein was associated with tyrosine kinase activity (Hunter and Sefton, 1980). The RTK family is part of the TK superfamily. It is clear now that human RTKs are not only crucial for normal cellular response but are also promoters of carcinogenesis when dysregulated by various reasons. All the known 58 RTKs have a similar architecture in terms of intracellular tyrosine kinase region. Their activities switch from an 'inactive' monomeric state to an 'active' dimerization. Autophosphorylation of pY kinase domains of RTKs do not occur randomly, instead, they are mediated by a sequential and precisely controlled three-phase reaction. Taking an example of FGFR1 phosphorylation, the three phases are: 1) the Y653 is autophosphorylated in the activation loop and enables increasing kinase activity by ~10- to 50-fold in the first phase; 2) the following second-phase phosphorylation occurs on three precisely ordered pY sites, sequentially Y583 > Y463 > Y585, that are likely crucial for SH2/PTB domains docking sites to recruit downstream substrates instead of enhancing kinase activity; 3) a further autophosphorylation stimulating Y654 in the third-phase lead to an additionally 10-fold increased kinase activity and reach 100 – 500 times compared to basal levels(Furdui et al., 2006; Lemmon and Schlessinger, 2010).

The SRC homology-2 (SH2) and phosphotyrosine-binding (PTB) domains are the most important elements to be recognized by activated pY sites directly (Pawson, 2004). A number

of docking proteins containing SH2/PTB domains have been identified, some of which are common docking proteins of multiple RTKs (like GRB2, SHC1, PLCG1, and SRC, etc) while others are constrained to specific RTKs (like FGFR adaptor – FRS2α and IGF1R adaptor – IRS-1) (Lew et al., 2009). Mediated by these protein-protein interactions and hierarchical phosphorylation cascades, various signaling pathways are transduced, typically RTKs-GRB2/SOS1-RAS-RAF-MEK-ERK or RTKs-PI3K-AKT-mTOR signaling pathways, to modulate cell proliferation, migration, differentiation, and survival.

Finally, RTKs signaling activation are also under precise control of proper attenuation mechanism involving RTKs autoinhibitory phosphorylation, enhanced phosphatases activity and negative feedback loops that we will discuss in details with an example of FGFR3 in the following sections (Lew et al., 2009).

FGFR3-driven pathways. We can suppose that FGFR3-driven pathways should follow the similar mode of signaling initiation as other FGFRs do. However, our view about the exact signaling pathways induced by FGFR3 is still at a glance of the tip of the iceberg.

First, unlike in FGFR2, a precisely three-phase phosphorylation of pY within FGFR3 has not been discovered yet. Only 5 conserved pY sites of FGFR3 have been described to be required for kinase activity, including Y647 and Y648 in the activation loop as well as Y577, Y724, and Y760 in the non-activation loop residues (Chen et al., 2005; Hart et al., 2001; Webster and Donoghue, 1996).

Second, our knowledge about FGFR3 specific adaptor proteins is still poor. The best identified FGFRs docking proteins is FRS2α which has been shown to be associated with FGFR1 in 2000 (Ong et al., 2000). FRS2α constitutively binds to FGFR1 at the juxtamembrane region (a conserved region among FGFRs family) by its PTB domain and the binding status is

independent of FGF ligand stimuli or receptor activation. FRS2α contains multiple pY sites that are phosphorylated upon FGFR1 activation and recognized by SH2 domains of the docking protein – GRB2. The complex FRS2α:GRB2 further recruits SOS1 substrate to stimulate MAPK pathways (FGFR1-FRS2α:GRB2:SOS1-RAS-RAF-MEK-ERK). To date, among all the pY sites of FGFR3, only Y760 and Y724 are identified to be capable of docking SH2 domains of adaptor proteins. The FGFR3 Y760 (corresponding to Y766 in FGFR1) solely is required for PLCG1 and PIK3R1 (a PI3K regulatory subunit) docking proteins, but both Y760 and Y724 are required for SH2B1 docking protein (Kong et al., 2002; Mohammadi et al., 1991; Salazar et al., 2009). The PLCG1 is essential for PLCG1-PKC signaling while the PIK3R1 protein is crucial to stimulate PI3K pathway (FGFR3-PIK3R1/PI3K-AKT). Interestingly, although PI3K kinase is a well-established substrate of FGFR3, activating mutations of *PIK3CA* often co-occurred with *FGFR3* mutations (Robertson et al., 2017). The reason is unknown. Strikingly, FRS2α is also able to regulate PI3K pathway through an indirect complex intermediated by GRB2 – FRS2α:GRB2:GAB1 as well as PIK3R1 can regulate MAPK pathway, implying an existence of more complicate signaling regulatory network (Ong et al., 2001; Salazar et al., 2009).

Likewise, JAK-STATs signaling are also associated with FGFRs activation. All STAT1/3/5 proteins contain SH2 domains, thus providing possibility to be direct docking proteins. FGFR1 Y677 site is required for STAT3 docking and pY phosphorylation of STAT3, which partially depends on SRC/JAK activation as well (Dudka et al., 2010). Although no clear docking pY sites for STAT1/3 within kinase domain of FGFR3 has been identified yet, both STAT1/3 co-immunoprecipitate with FGFR3, and the phosphorylation of their corresponding activating sites (STAT1 Y701; STAT3 Y705) is FGFR3 dependent (Dudka et al., 2010; Lee et al., 2014). Further, activation of STAT5 is also associated with FGFR3:SH2B1 complex (Kong et al., 2002). Of note, a single pY – Y724 – has been shown to be critical for multiple pathways, like cellular transformation, PI3K activation, MAPK and phosphorylation of STAT1/3 (Hart et al., 2001). A summary of aforementioned FGFRs-driven signaling pathways is illustrated in Figure 1-12; page 29 (Turner and Grose, 2010; Wu, 2005).

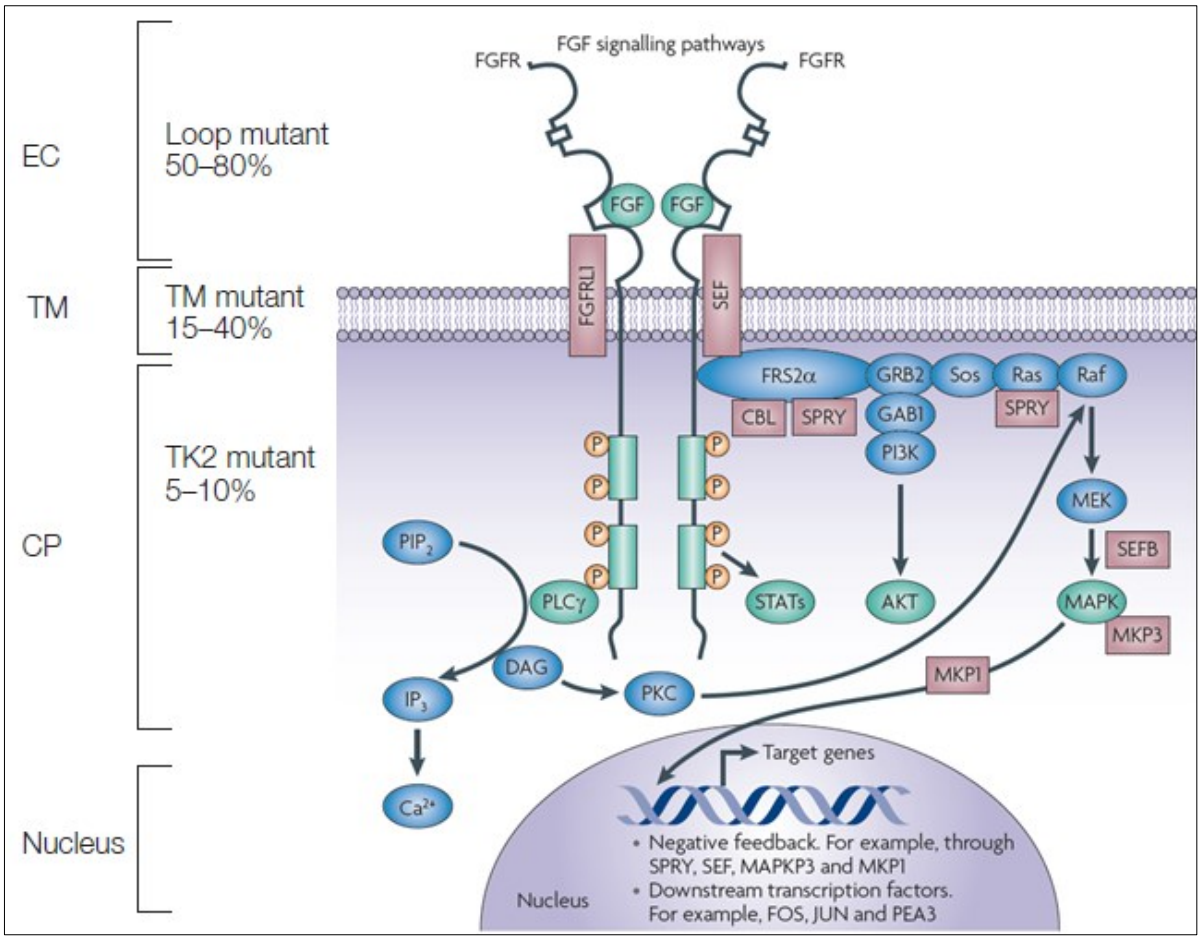


Figure 1-12: *Schema of FGFRs-driven signaling pathways.*
Overall, four key downstream pathways are highlighted: 1) MAPK pathway; 2) PI3K-AKT pathway; 3) PLCG1-PKC pathway and 4) STAT pathways (Turner and Grose, 2010).. Negative regulators are in brown, such as FGFR-like 1 (FGFRL1), SEF, Sprouty (SPRY), CBL, MAPK phosphatase 1 (MKP1) and MKP3. FGFR3 mutation frequency corresponding to different compartments are drawn from (Wu, 2005).

FGFR3 localization and trafficking. All RTKs contain cleavable N-terminal signal peptides that favor receptors locating to the cell surface. Additionally, the juxtamembrane region of FGFR3 also determines the direction to plasma membrane (Bocharov et al., 2013). Whereas RTKs are bona fide plasma membrane proteins, majority of them (including FGFRs) can be localized in various cellular organelles through dynamic trafficking. This subcellular localization of RTKs is a key factor to determine cell fate, for example, nuclear localization of multiple RTKs has been shown to be linked with tumorigenesis (Bergeron et al., 2016; Chen and Hung, 2015; Miaczynska, 2013). Similar to most other RTKs, FGFRs undergo constitutive internalization, but the rate is relatively low and lower than their synthesis and recycling

without ligand stimuli, which facilitates the accumulation of receptors at the cell surface. However, upon ligand stimulation, the rate of internalization is significantly accelerated and hereby provoke endocytosis through several pathways, mainly clathrin-mediated endocytosis (CME) (Porębska et al., 2018). The extended synaptotagmin-2 (ESYT2) protein is the major regulator for CME of FGFRs and it interacts with FGFR1 (Jean et al., 2010). Further, although not fully understandable regarding nuclear transport, FGFRs are detected in the nuclei of multiple cell types (Porębska et al., 2018). Very interestingly, in addition to the typical CME and clathrin-independent nuclear transports of FGFRs, a particular proteolysis mode that cleaves intracellular receptor fragments which are then transported to the nuclei has been identified for FGFR1 and FGFR3 (Chen and Hung, 2015; Chioni and Grose, 2012; Degnin et al., 2011). A general view of FGFRs trafficking is illustrated in Figure 1-13; page 31. Our knowledge about FGFRs internalization is still far from complete.

Negative feedback signals. The signaling pathways need to be exquisitely controlled spatially and temporally to ensure correct cellular behaviour and process (Ullrich and Schlessinger, 1990). Considering that various dysregulation of RTKs are contributing to carcinogenesis, it is of great interest to understand how signals are augmented and attenuated.

First, inhibition of RTKs can be achieved by modulating themselves. Receptor autoinhibition do exist among FGFRs family: 1) the acid box (a linker between Ig-I and Ig-II of about 4~8 amino acids) structure is a conserved region among FGFRs and it can negatively regulate FGFRs activation by competing with FGF and heparin binding (Olsen et al., 2004). Due to different lengths of the acid box within four FGFRs, there exists gradations of autoinhibitory control. For example FGFR1 is likely to be under strong autoinhibition because of the longest acid box whereas FGFR4 is probably under mild autoinhibition as it has the shortest acid box; 2) a ‘molecular brake’ structure, located at the kinase hinge of FGFR2, has shown a critical autoinhibitory function. It is composed of a hydrogen bonding network among four residues – H544, N549, E565 and K641 – within autoinhibited kinases but being disengaged upon kinase activation (Chen et al., 2007). FGFRL1, an analogous protein of true FGFRs but lacks intracellular tyrosine kinase domain, is supposed to be negative regulator by blocking

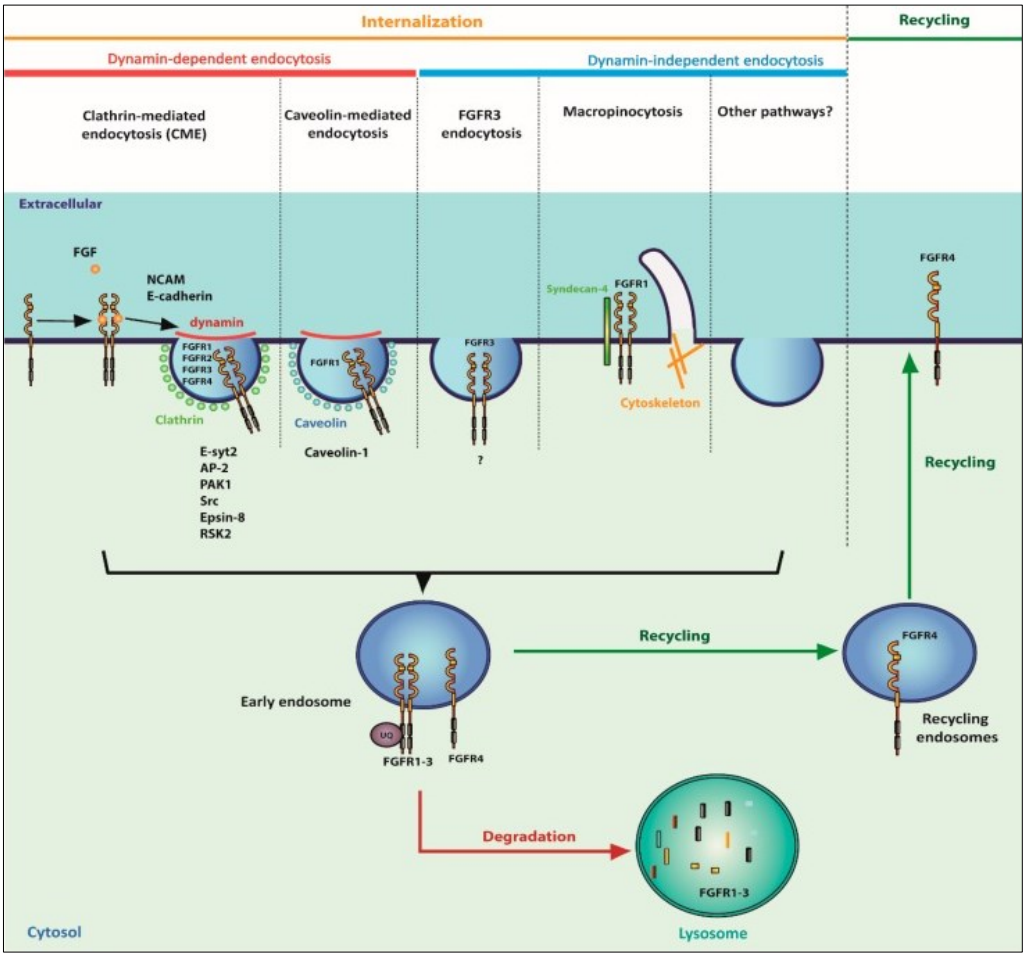


Figure 1-13: *Diverse modes of internalization and cellular endocytosis for FGFR1-4. (Porębska et al., 2018)*

transphosphorylation of FGFRs (Wiedemann and Trueb, 2000). Receptor autophosphorylation can also be attenuated by protein tyrosine phosphatases (PTPs). Compared to other RTKs, much less PTPs have been characterized to mediate FGFRs dephosphorylation. PTPN1/2 are the first identified negative regulators modulating FGFR3 phosphorylation within a consensus motif (D/EYYR/K) in multiple myeloma (St-Germain et al., 2009, 2015). Recently, another PTP – protein tyrosine phosphatase receptor type G (PTPRG) – has been shown to negatively regulate FGFR1 and FGFR4 by dephosphorylation, even though the clear binding motif has not been identified yet (Kostas et al., 2018).

Second, phosphatases targeting adaptor proteins, downstream MAPK or PI3K-AKT signaling play a crucial role as well. PTPN11-mediated dephosphorylation of FGFRs is controlled by the adaptor protein GRB2, which recruits PTPN11 to the activated receptors (Neben et al., 2019). DUSP6 enzyme dephosphorylates ERK1/2 in a FGFRs dependent manner (Nakanishi et al.,

2015b). Sprouty (SPRY) protein can either disrupt SOS1-mediated RAS activation or directly bind to RAF to block FGFRs / EGFR-driven MAPK signaling (Casci et al., 1999; Lao et al., 2007; Ledda and Paratcha, 2007). Moreover, IL17RD (SEF) and PTEN proteins are both engaged in negative feedback regulation of FGFRs-mediated MAPK and PI3K-AKT signaling pathways (Chaffee et al., 2016; Kovalenko et al., 2003; Ledda and Paratcha, 2007).

Targeting FGFR3 against bladder cancer

Of note, therapy strategies targeting FGFR3 is, after immunotherapy, probably one of the most attractive fields in BCa treatment. Though there is not yet FGFR3 specific inhibitor, a number of small molecule tyrosine kinase inhibitors (TKIs) have been developed to target FGFR3, such as multi-targeting TKIs (Dovitinib, Novartis; Lenvatinib, Eisai), selective pan-FGFRs TKIs (Erdafitinib, Janssen; Rogaratinib, BAY 1163877; AZD4547, Astra Zeneca; BGJ398, Novartis; PD173074, Pfizer; and LY2874455, Lilly) and monoclonal antibody (B-701 (vofatamab), BioClin) (Babina and Turner, 2017; Gust et al., 2013). For unknown reason, PD173074 (Pfizer) is not evaluated in clinical trials, even though both *in vitro* and *in vivo* evidence have shown this compound can inhibit FGF-driven neoangiogenesis without obvious toxicity (Dimitroff et al., 1999; Mohammadi et al., 1998; Pardo et al., 2009). In patients with locally advanced or metastatic BCa, phase 1/2 trial evaluating Rogaratinib (BAY 1163877) and phase 2/3 trial evaluating B-701 (BioClin) are ongoing. As we mentioned in the previous section, the first pan-FGFRs inhibitor – Erdafitinib/Balversa – has been recently approved by the FDA (April 2019) for patients with locally advanced or metastatic BCa.

However, the challenges for selecting the eligible patients and the emergence of various drug resistance mechanisms are major constrains in terms of anti-FGFR3 targeted therapy. First, *FGFR3* mutations and translocations account for ~15% and 4% of MIBCs, respectively, but gene overexpression of *FGFR3* is also detected in wild-type tumors. Furthermore, not all patients with mutations or translocations will respond to the treatment. Thus, selecting patients only based on *FGFR3* genetic alterations seems not sufficient to cover the more likely responders to anti-FGFR3 therapy. Second, without a clear knowledge of clonal or sub-clonal status of common *FGFR3* alterations, the intratumoral heterogeneity will be a major source to induce patient selection bias. For example, drug response can be achieved in tumors with clonal *FGFR2* amplification but not in tumors with sub-clonal amplification (Van Cutsem et al., 2017; Pearson et al., 2016). Third, it is crucial to know whether tumors depend on FGFRs signalling for growth and survival, which may predict the effectiveness of anti-FGFR3 treatment. Among

all different *FGFR3* alterations, only oncogenic 'drivers' have such potential to gain tumor growth advantage. For instance, *FGFR3* S249C mutation and FGFR3-TACC3 translocation are indeed proved as oncogenic 'drivers' in BCa (Bernard-Pierrot et al., 2006; Best et al., 2018; Williams et al., 2013). Fourth, FGFRs signaling confer drug resistance through diverse ways: 1) EGFR and ERBB2/3 activation are common mechanisms to mediate acquired resistance of anti-FGFR3 therapy in BCa (Herrera-Abreu et al., 2013; Wang et al., 2015, 2017); 2) although there is not yet such evidence in BCa, acquired resistant mutations to multiple anti-FGFRs inhibitors have also been reported for *FGFR3* V555M mutation in KMS-11 myeloma cell line, *FGFR1* V561M mutation in COS-7 cell line and *FGFR2* N550K mutation in BaF3 cell line (Blencke et al., 2004; Byron et al., 2015; Chell et al., 2013); 3) oppositely, FGFR3 activation, by either overexpression in lung cancer (Lee et al., 2014) or FGFR3-TACC3 fusion in lung and head and neck cancer (Allen et al., 2017; Daly et al., 2017), can be acquired to account for resistance of anti-EGFR blockade as well. Considering these crossed resistance mechanisms between FGFR3 and EGFR and emerging therapeutic value of anti-EGFR in BCa (Rebouissou et al., 2014), drugs combination targeting FGFR3 and EGFR will be promising. Hopefully, with the increasing understanding of FGFR3 binding partners, phosphorylation events as well as FGFR3 trafficking mechanism, many alternative choices of treatment will be elucidated instead of targeting the receptor itself (Mahe et al., 2018). In one sentence, targeting therapies based on FGFR3 or relevant pathways have both challenges and perspectives, and there is still a long way to overcome BCa.

IV. Strategies to decipher aberrant FGFRs/RTK signaling

Working models

In vitro models. Cell lines, including cancer cell lines, are basic tools used in laboratories to understand the cell and molecular biology. Although not perfect, human cancer-derived cancer cells reliably reflect genetic characteristics of primary cancer (Barretina et al., 2012), although some of them were derived from the primary tumor decades ago, such as the first isolated cancer cell line – HeLa – in 1950s (Scherer et al., 1953). Both non-cancer (BaF3, HEK293T or NIH-3T3) and cancer cell lines have been used to investigate FGFRs signaling, mostly through artificial overexpression of FGFRs. However, several BCa cell lines endogenously harbouring activating *FGFR3* alterations are available, like UMUC-14 (*FGFR3* S249C

mutation), MGH-U3 (*FGFR3* Y375C mutation), and RT-112 (FGFR3-TACC3 fusion) cells. Therefore, they represent good models to decipher FGFR3-driven signaling in BCa at endogenous level. Recently, 3D systems (organoids or patient-derived tumoroids) appear as novel *in vitro* models and have been lately applied to evaluate drug response (Lee et al., 2018; Mullenders et al., 2019; Vasyutin et al., 2019). MGH-U3 and RT-112 cells have successfully employed in organoid and spheroid fabrication as representative of low-grade (grade 1) bladder carcinomas (Vasyutin et al., 2019). These 3D techniques mimic *in vivo*-like cellular properties and present evident advantages to 2D systems.

In vivo models. There are two major classes of mouse models in BCa: autochthonous and non-autochthonous models (Kobayashi et al., 2015). The latter include cancer cell line xenografts (hereby referred to as xenografts) and patient-derived xenografts (PDXs), both of which can be either ectopic or orthotopic (Jung, 2014). Notably, the use of ectopic xenograft or PDX models is much more extended due to their relative ease, rapidity of generation and follow-up. They are frequently used by our team as well as other groups to explore anti-FGFR3 drug response or to validate FGFR3-mediated signaling cascades. By contrast, the generation of orthotopic models presents a higher complexity, especially in the case of a hollow organ such as bladder. Accordingly, an advanced ultrasound-guided intramural inoculation method has been introduced to facilitate orthotopic xenograft of BCa (Jäger et al., 2013). Recently, the first orthotopic PDX model of BCa has been established by Pan et al. (Pan et al., 2015).

Autochthonous models are composed of carcinogen-based and genetically engineered models (mainly transgenic mice). Current carcinogen-based mouse models are mostly generated through the exposure to one of the following chemicals: N-butyl-N-(4-hydroxybutyl)-nitrosamine (BBN), N-[4-(5-nitro-2-furyl)-2-thiazolyl]-formamide (FANFT) or N-methyl-N-nitrosourea (MNU) (Fantini et al., 2018; John and Said, 2017; Kates et al., 2017). BBN has been identified as a bladder carcinogen in rodents since 1980s and is detected in tobacco smoke, as well as in environmental and infectious metabolites. It mainly induces genome instability, particularly massive mutations, to promote formation of high-grade tumors showing *TP53* mutations and high mRNA expression of *EGFR* (John and Said, 2017). However, activating *FGFR3* mutations have not been found in BBN model. Regarding MNU, it is a genotoxic compound that causing persistent DNA methylation. Currently, FANFT is not commonly manipulated as it is hazardous for human health.

Majority of the BCa transgenic mice models have used the mouse Uroplakin II (*UPK2*, UII) promoter, a marker specifically expressed in the urothelium. Accordingly, Zhang et al.

developed a transgenic mouse model, causing urothelial hyperplasia and superficial papillary NMIBC tumor, by targeting the mutated form of UII-HRas Q61L (Zhang et al., 2001). This model is not suitable for investigating FGFR3 signaling, because *RAS* mutations are exclusive to *FGFR3* mutations despite *RAS* mutations are also very common in BCa. Although a transgenic mouse model engineered with Ad5-CMV-Cre inducible expression of FGFR3-TACC3 fusion has been recently produced and shown to develop lung cancer (Best et al., 2018), there is not yet transgenic mice generated with oncogenic 'driver' mutations of *FGFR3* that is alone able to induce tumor transformation in bladder or any other cancer types. Previous works did not observe spontaneous formation of BCa in transgenic mice generated by UII-*FGFR3* S249C or K644E mutation up to a follow-up of 12 and 18 months, respectively (Ahmad et al., 2011; Foth et al., 2018). However, Foth et al. indeed observed an accelerated formation of invasive BCa by combining *FGFR3* S249C expressing transgenic mouse model and BBN exposure. In brief, *in vivo* BCa models suitable for the investigation of FGFR3 signaling remain underdeveloped.

Transcriptomic strategy

Like many other signaling pathways, activated FGFRs transmit the signal from the plasma membrane to the cell nucleus through a series of downstream processes that eventually lead to rapid changes in gene expression, resulting in an appropriate cellular response to a certain stimulus. With the fast development of DNA arrays and high-throughput genomic sequencing technologies during the past two decades, gene expression can be easily quantified in large-scale manner, namely through transcriptomics (mostly messenger RNA). The term Differential expressed gene (DEG) is largely used to determine whether a particular gene is expressed distinctly between two or more conditions. In addition, in recent years, transcriptomics performed on patient tissue of different cancer types led to identification of molecular subtypes that could have different prognosis or reflect targetable driving pathways specific for each subclass (Robertson et al., 2017). Transcriptomics is nowadays one of the most commonly utilized strategy to explore if a biological process is deregulated in a particular condition, both for *in vitro* and *in vivo* systems. In one of our previous works we performed such RNA-based large scale studies to decipher a FGFR3-MAPK11/AKT-MYC regulation axis and identify a FGFR3-MYC positive feedback loop (Mahe et al., 2018).

Proteomic strategy

From transcriptome to proteome. Although transcriptomics can indeed provide a comprehensive view of cancer signaling by itself, the proteome is the final product of the genome and is one of the key functional elements in the cell. The proteome is a highly dynamic entity that presents a great diversification of functionalities not only through changes in expression levels, but by a wide range of posttranslational modifications (PTMs), several cellular localizations and molecular interactions / complexes, all of which remain completely hidden in the static DNA code (Figure 1-14; page 36). Gene expression does not always faithfully translate into protein expression (Forget et al., 2018), probably as a result of numerous factors, such as splicing events, proteins assembly, subcellular localization, protein stability, efficacy of RNA translation and importantly, PTMs (Harper and Bennett, 2016). Particularly, when considering RTK/FGFRs signaling pathways, which are largely dependent on protein-protein interactions (PPIs) and PTM cascades, transcriptomic data provides rather limited insight. Nowadays, proteomics tackels different questions, from the detection of all protein isoforms and modifications, to the PPIs, protein structures and their higher-order complexes.

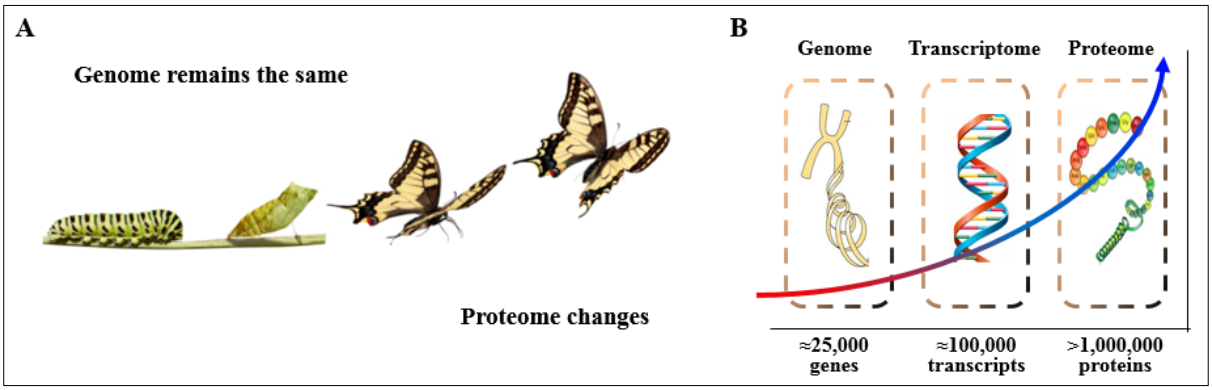


Figure 1-14: *Multiplicity and complexity of proteome.*
A) *Butterfly development reflects the inherent dynamic nature of the proteome with respect to the genome. Picture downloaded online: http://www.iamashcash.com/wp-content/uploads/2011/03/caterpillar-to-butterfly1.jpg.* ***B)*** *As illustration of its complexity, human proteome consists of over one million proteins that originated from ~25000 genes. (Jensen, 2004)*

From targeted strategy to unbiased approach. The analysis of encoded proteins have traditionally been carried out on single molecules or a discrete set of candidates of interest. Latter, protein microarrays were developed as a high-throughput tool to facilitate protein function analyses. There are three major categories of protein microarray: analytical antibody

arrays, functional protein microarrays and Reverse Phase Protein Arrays (RPPA) (Sutandy et al., 2013). Analytical antibody arrays take advantage of highly specific antigen-antibody recognition to build a protein detection system. Functional protein microarrays are developed *via* the immobilization of purified proteins and used for investigating protein functions such as PPIs, enzyme-substrate reactions and immune responses. RPPA, a more recently developed technology compared to the formers, allows for the analysis of many proteins and various tissue/cell lysates obtained at different states. Although the number of analysed proteins remains restrained and restricted to the availability of specific antibodies, these targeted approaches were successfully applied in cancer research. For example, proteomic characterization of BCa and investigation of FGFR3-driven signaling have been widely carried out using RPPA (Mahe et al., 2018; Robertson et al., 2017). Since the early 2000s, a new approach relying on mass spectrometry (MS) analysis has been greatly developed, being able to sensitively identify proteins in a high-throughput, unbiased manner and becoming a valuable tool for the elucidation of signaling pathways initiated by a given stimulus (Aebersold and Mann, 2003; Olsen et al., 2006). Although MS-based approaches present a high plasticity and expanding applications, the most widely used strategy consists in the digestion of proteins into peptides prior to the MS analysis. Specific enzymes, namely trypsin, are used for this purpose, thus ensuring the traceable localization of cleavage sites. Peptide sequences would thereafter be identified through MS and the results would be aligned to known amino acid sequences stored in protein databases (Aebersold and Mann, 2003). The quality of protein identification would hence depend on the quality and number of corresponding sequenced peptides.

Protein quantification based on MS. MS allows not only for protein identification, but also for relative and absolute protein quantification (Ong and Mann, 2005). Comparing the signals from the same peptide under different conditions yields an estimation of relative protein quantification between two proteomes. Technically, three categories of quantitative proteomics could be distinguished: label-free quantification, chemical labeling and metabolic labeling, have been introduced (Bantscheff et al., 2007). Label-free method, even though less accurate than the introduction of particular labels, is widely used due to being less time-consuming, reduced in cost and unlimited in multiplexing, as well as its applicability to samples from virtually all origin (Li et al., 2012b). However, label free accounts for higher technical variability and an increased number of replicates is desirable.

Chemical labeling methods incorporate different isotopic tags to peptides after cell lysis and protein digestion. The most commonly used methods are isotope-coded affinity tags (ICAT)

(Gygi et al., 1999), isotope tags for relative and absolute quantification (iTRAQ) (Ross et al., 2004) and tandem mass tags (TMT) (Thompson et al., 2003). Different conditions in study would be labelled with different tags, allowing the mixture of the distinct peptide extracts in a single sample for its further analysis. At the mass spectrometer, tags would be cleaved and their individual intensities compared to construct relative ratios among conditions.

Stable isotope labeling with amino acids in cell culture (SILAC) is a metabolic labelling strategy that based in the incorporation of stable heavy essential amino acids into living cells while they proliferate (Ong and Mann, 2006). The heavy amino acids would be incorporated in the newly synthesized proteins of the cell, hence allowing the labelling of virtually the whole cellular proteome. Therefore, samples from different conditions can be mixed right after the cell lysis, greatly reducing technical errors. Through MS analysis, peptides from different conditions would present a shift in mass, which permits the calculation of ratios among the doublets/triplets corresponding to the same peptide in distinct labelling states. SILAC represents a very robust approach, although it is so far limited to three conditions and to its theoretical use only in cell lines. Nevertheless, recent strategies have been implemented for the utilization of SILAC principle in mouse fed with heavy-labeled amino acids (Sirvent et al., 2012; Zanivan et al., 2012) and in human-derived tissues (the so-called super-SILAC) (Shenoy and Geiger, 2014). Due to the robustness of the method and its applicability for the purpose of our studies, we decided to implement the use of SILAC for the investigation of altered FGFR3 cascade.

Interactome. PPIs are crucial for all biological processes. Different approaches have been developed to characterize protein complexes and networks, such as the yeast two-hybrid system (Y2H) and high-throughput MS, both of which can contribute in a complementary manner (Mering et al., 2002; Stelzl et al., 2005). In Y2H system, one protein (‘bait’) is fused to a DNA-binding domain, the other (‘prey’) to a transcriptional activator domain. Any direct interaction between them is detected by the formation of a functional transcription factor or activation of reporter genes.

In MS screening, protein extracts are obtained in mild conditions and complexes are further enriched through different methods, namely antibody-based immunoprecipitation of ‘bait’ and corresponding partners. The thereby enriched proteome is then analysed through mass spectrometry for the identification of particular proteins and their quantification according to the chosen strategy. Being the enrichment of an endogenous protein a challenge in many technical aspects, different strategies have been implemented to aid in the recovery of an

interactome of interest. Notably, the introduction of an epitope-tag to the 'bait' protein is widely utilized, allowing its enrichment through tag-directed antibodies or affinity columns.

The combination of the aforementioned techniques with MS-based analysis has been largely applied to the investigation of PPIs in cancer research, such as for the elucidation of EGFR signalling cascade (Blagoev et al., 2003). However, technical challenges in the field are yet to be overcome, mostly due to limited methods to isolate integral membrane proteins while conserving associated complexes intact and to the lack of specific antibodies suitable for enriching endogenous 'bait' proteins in many cases. In addition, the complexity gets even larger when spatial and temporal aspects of interactions are considered.

Phosphoproteomics. PTMs largely enhance the diversity of protein functions by adding moieties of different nature to the canonical sequence of proteins. The MS-based analysis of these modifications requires previous enrichment techniques due to their relatively low abundance in the cell. Reversible phosphorylation of proteins is one of most important PTMs modulating signal transduction pathways, namely downstream RTKs. For example, SILAC coupled with MS has been extensively utilized to reveal EGFR-dependent phosphorylation cascades (Olsen et al., 2006). Two major enrichment methods based on the use of positively charged metal ions have been developed for phosphopeptide enrichment: immobilized metal affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC). Some of these techniques include Fe^{3+}-IMAC, Ti^{4+}-IMAC or TiO2-MOAC (Zhou et al., 2013). These very well implemented protocols allow a major enrichment of phosphoserine and phosphothreonine (pS, pT) peptides but are less efficient for the isolation of phosphotyrosines (pY), due to the relative low abundance of the latter. Currently, the enrichment of pY sites is mostly performed through the use of anti-pY antibodies.

Omics integrative analysis

We are now in the coming era to investigate health and cancer at the omics scale. Notably, a comprehensive understanding of cancer biology will require quantitative information at all levels, from DNA variants to their protein products, and thus integrative analysis of omic data is becoming increasingly used. With the rapid trend of generating both transcriptomics and proteomics data in parallel, it may be a challenge to accurately combine these large-scale approaches. Hopefully, the omics data-centred study of FGFR3 pathway should provide a more complete picture of its signaling cascade and hopefully discover alternative therapeutic targets for a subset of BCa dependent on FGFR3.

Chapter 2 Objectives

OBJECTIVES

BCa ranks the third in terms of bearing tumor mutation burden, just following melanoma and lung cancer. These somatic mutations are cumulative consequence of the endogenous mutation process and / or exogenous mutagenic exposures in addition to functional selection in a Darwinian fashion. APOBEC enzymes are one of endogenous mutagenic sources that leave fingerprints on cancer genome and form characteristic mutation pattern, termed as APOBEC mutational signature. APOBEC mutational signature is the predominant source of overall mutations in BCa.

FGFR3 is one of the most frequently mutated genes in BCa with ~65% of non-muscle-invasive BCa (NMIBCs) and 15% of muscle-invasive BCa (MIBCs) carrying an *FGFR3* mutation. Within 14 recurrent residues (frequency ≥ 2), one hotspot mutation – *FGFR3* S249C – was strongly over-represented compared to other recurrent *FGFR3* mutations (62% of all recurrent mutations). Most of these recurrent mutations have been shown to be oncogenic drivers *in vitro*. However, *in vivo*, genetically engineered mice with oncogenic 'driver' mutations of *FGFR3* alone is insufficient to induce tumor transformation in bladder or any other cancer types. Thus, *in vivo* models suitable to understand the oncogenic nature of FGFR3 and its associated signaling are very limited. *FGFR3* mutations trigger constitutive activation by increased dimerization and trans-phosphorylation within the dimer, resulting in downstream phosphorylation, prolonged signaling cascades and subsequent uncontrolled tumor proliferation. Targeted therapy based on anti-FGFR3 shows increasing interest in BCa treatment. Recently, the first pan-FGFR inhibitor – Erdafitinib/Balversa – has been approved by the FDA for patients with locally advanced or metastatic BCa. However, optimal treatment regimens need a fully understanding of exact signaling pathways to achieve a personalized medical care and to overcome potential acquired drug resistance. Compared to well-studied RTKs, e.g. EGFR or FGFR1, FGFR3-driven signaling pathways are yet poorly understood. Few proteomic studies have been done to uncover the entity of FGFR3 signaling systematically. Therefore, the three main objectives of my thesis are as follows:

1.1) understand why FGFR3 S249C is strongly over-represented and following our finding that the over-representation was due to the APOBEC mutagenesis process, 1.2) look for a possible association between frequent mutations and APOBEC mutational signature;

2) study the *in vivo* oncogenic properties of mutated-*FGFR3* in BCa by establishing a transgenic mouse model expressing *FGFR3* 249C in the urothelium;

3) decipher FGFR3-driven signaling network using a proteomic strategy to identify FGFR3 binding partners and downstream phosphoproteome, and transcriptomic strategy to identify downstream transcription factors as well as regulated genes. A graphical model illustrates this last objective (Figure 2-1; page 44).

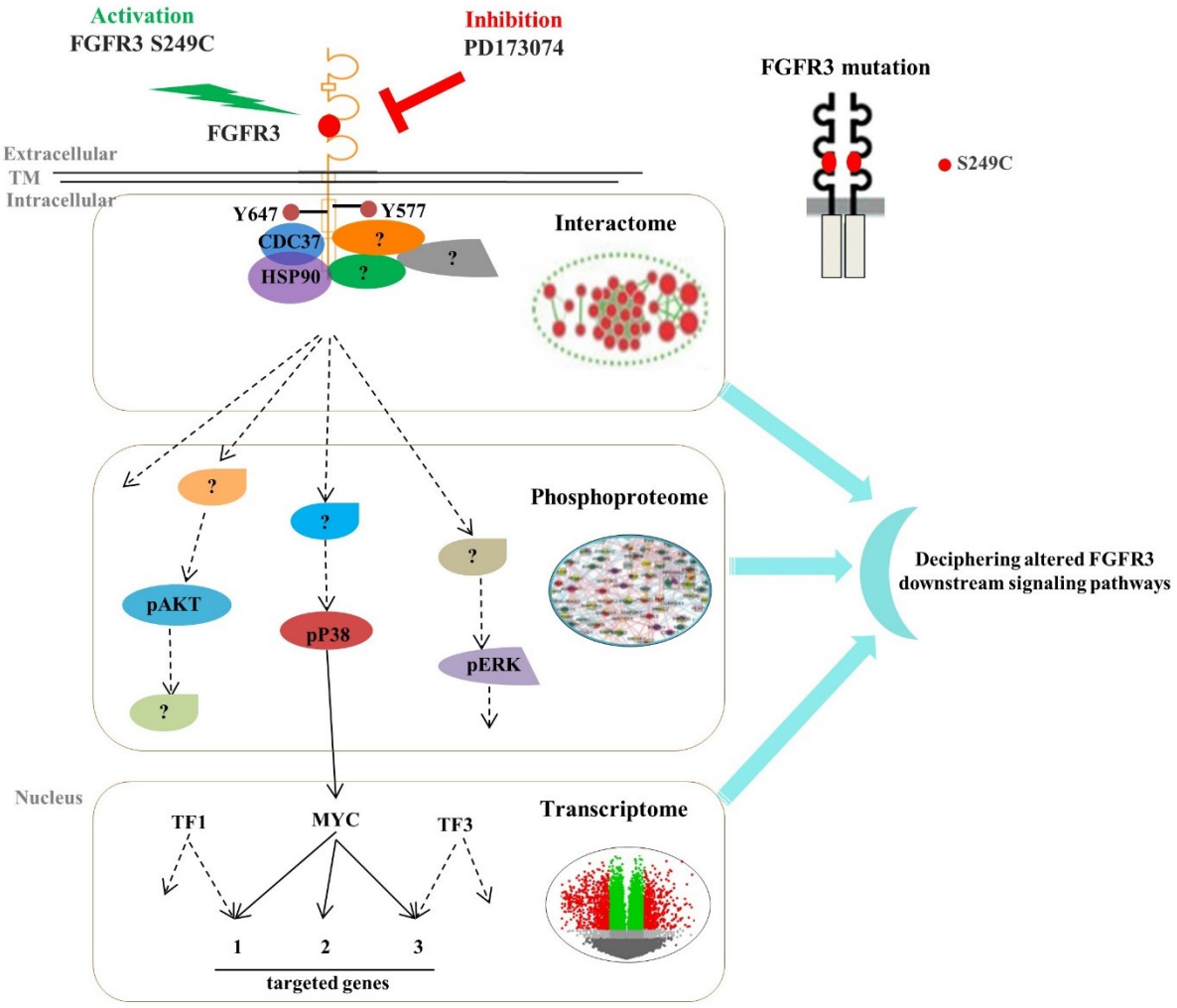


Figure 2-1: *Graphical illustration of 3rd objective.*

The aim is to decipher FGFR3-driven signaling through -omic data generation and integration, using UMUC-14 as working model. Representative known proteins involved in FGFR3 signaling are given.

Chapter 3 Results

3.1 APOBEC mutational signature in bladder cancer

3.1.1 Introduction

Identify prevalent mutations has been accepted as an edge tool to uncover important cancer genes (Greenman et al., 2007), since onco-drivers undergo functional selection and are frequently mutated at the same amino acid positions. However, the endo/exo-genous mutation process may be another factor favouring frequent mutations. Each mutation process leaves a characteristic fingerprint on the cancer genome in a DNA sequence context depending manner, termed as a mutational signature. APOBEC mutational signatures (COSMIC signature 2 and signature 13), both characterized by high prevalence of single-nucleotide C → T transition or C → G transversion at stringent 5'-TCW motif (W = A or T) and present ubiquitously in many cancer types, were proposed to be due to over activity of APOBEC deaminases.

BCa ranks the third in terms of bearing tumor mutation burden (TMB), just following melanoma and lung cancer. In BCa, APOBEC mutational signature (S2 & S13) is the predominant source of overall mutations. Thus, frequent mutations in BCa may also be a result of APOBEC mutagenesis in addition to functional advantage. Indeed, several 'driver' mutations have been identified previously to be associated with APOBEC in BCa, like *PIK3CA* E545K and E542K, *ERBB2* S310F and *TP53* R280T (Cannataro et al., 2019; Henderson et al., 2014; Poulos et al., 2018; Temko et al., 2018). Still, study focusing on a systematic identification of APOBEC-associated mutations has not been performed yet.

FGFR3 is one of the most frequently mutated genes in BCa. For two decades, the strong over-representation of *FGFR3* S249C mutation has been noted in BCa and supposed to be due a greater functional advantage of this mutation compared to the other *FGFR3* recurrent mutations. However, there is minor difference of malignant transforming potential between S249C and the second frequent *FGFR3* mutation – Y375C despite of their biased mutation rates in BCa. We wondered whether APOBEC mutagenesis can account for this biased mutation frequency. Results regarding this hypothesis will be presented in the first part of this chapter as 'Paper 1', which has been recently published. In this paper, we reported that only *FGFR3* S249C (TCC -> TGC, with mutated nucleotide underlined) mutation represents an APOBEC-type motif and is probably caused by theAPOBEC-mediated mutagenic process, accounting for its over-representation. We observed significant enrichment of the APOBEC mutational signature and

overexpression of AID/APOBEC gene family members in bladder tumors with S249C compared to tumors with other recurrent *FGFR3* mutations (frequency ≥ 2). Analysis of replication fork directionality suggests that the coding strand of FGFR3 is predominantly replicated as a lagging strand template that could favor the formation of hairpin structures, facilitating mutagenic activity of APOBEC enzymes. *In vitro* APOBEC deamination assays confirmed S249 as an APOBEC target. We also found that the *FGFR3* S249C mutation was common in three other cancer types with an APOBEC mutational signature (cervical, head and neck and lung cancer), but rare in urothelial tumors without APOBEC mutagenesis (Lynch syndrome associated UTUC) and in two diseases probably related to aging (bone disorders and benign skin tumors).

In the second part of this chapter, referred as 'Paper 2', I and another PhD student (Xiangyu) and we tried to identify all APOBEC-associated frequent mutations in BCa. A manuscript is being preparing for submission. In this paper, I found 44 frequent mutations (frequency ≥ 4) as 'bona fide' candidates that were associated with APOBEC mutagenesis and compatible with characteristics favouring the accessibility to APOBEC enzymes, such as preferable occurrence at lagging-strand template during DNA replication or a loop within hairpin structure. These mutations not only represented nearly half of all frequent mutations identified in BCa, but were also over-represented compared to other mutations within the same gene. Many of these mutations were known 'onco-drivers'. We pinpointed novel potential 'drivers' but also 'frequent passengers'. Interestingly, we revealed frequent loss-of-function mutations can occur on tumor suppressors solely due to APOBEC-related selection. The associations between some of these mutations and APOBEC mutagenesis were also recovered in other cancer types presenting an APOBEC mutation pattern. Further, by applying only these 44 mutations, a group of patients with high APOBEC mutagenesis can be selected, who may benefit from immunotherapies or future anti-APOBEC treatment.

3.1.2 Results

Article 1

APOBEC-mediated Mutagenesis as a Likely Cause of FGFR3 S249C Mutation Over-representation in Bladder Cancer

Ming-Jun Shi*, *Xiang-Yu Meng**, *Philippe Lamy**, *A. Rouf Banday**, *Jie Yang, Aura Moreno-Vega, Chun-Long Chen, Lars Dyrskjøt, Isabelle Bernard-Pierrot, Ludmila Prokunina-Olsson, François Radvanyi*[†]*.*

Published in European Urology as issue cover, 2019 Jul;76(1):9-13.
doi: 10.1016/j.eururo.2019.03.032.

Letter to Editor from Yang and colleagues regarding 'Article 1'

Re: Ming-Jun Shi, Xiang-Yu Meng, Philippe Lamy, et al. APOBEC-mediated Mutagenesis as a Likely Cause of FGFR3 S249C Mutation Overrepresentation in Bladder Cancer

Alexander Yang[a]*, Vincent L. Cannataro*[a]*, Jeffrey P. Townsend*[a,b,c*]*.*

Revised in European Urology, EURUROL-D-19-01062.

Response to Letter from Yang and colleagues

Letter to the Editor: Reply to Yang et al. (corresponding author Dr. Townsend) EURUROL-D-19-01062 Re: APOBEC-mediated mutagenesis as a likely cause of FGFR3-S249C mutation over-representation in bladder cancer

Ming-Jun Shi*, *Xiang-Yu Meng**, *Chun-Long Cheng, Lars Dyrskjøt, Francois Radvanyi, Ludmila Prokunina-Olsson, Isabelle Bernard-Pierrot*[†]*.*

Submitted in European Urology, EURUROL-D-19-01160.

Manuscript 2

APOBEC induces over-represented driver but also passenger hotspot mutations within its target genes in bladder cancer

Ming-Jun Shi*, *Xiang-Yu Meng**, *Chun-Long Chen, François Radvanyi*[†]*, Isabelle Bernard-Pierrot*[†]*.*

Under submission

Article 1

EUROPEAN UROLOGY 76 (2019) 9–13

available at www.sciencedirect.com
journal homepage: www.europeanurology.com

European Association of Urology

Brief Correspondence

APOBEC-mediated mutagenesis as a likely cause of *FGFR3*-S249C mutation over-representation in bladder cancer

Ming-Jun Shi[a,b,i,1], Xiang-Yu Meng[a,b,c,1], Philippe Lamy[d,1], A. Rouf Banday[e,1], Jie Yang[f], Aura Moreno-Vega[a,b], Chun-Long Chen[g,h], Lars Dyrskjøt[d,†], Isabelle Bernard-Pierrot[a,†], Ludmila Prokunina-Olsson[e,†], François Radvanyi[a,†]

a. Institut Curie, CNRS, UMR144, Molecular Oncology team, PSL Research University, Paris, France

b. Paris-Sud University, Paris-Saclay University, Paris, France

c. Department of Urology, Zhongnan Hospital, Wuhan University, Wuhan, China

d. Department of Molecular Medicine, Aarhus University Hospital, Aarhus, Denmark

e. Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA

f. Inserm UMR-1162, Génomique fonctionnelle des Tumeurs solides, Université Paris Diderot, Paris, France

g. Institut Curie, CNRS, UMR3244, PSL Research University, Paris, France

h. Sorbonne Université, Paris, France

i. Department of Urology, Beijing Friendship Hospital, Capital Medical University, Beijing, China

1 These authors contributed equally to this work.

† Joint senior authors.

Correspondence:

François Radvanyi

Institut Curie, CNRS, UMR144, Molecular Oncology team

26 Rue d'Ulm, 75005 Paris, France

TEL: +33 1 42 34 63 40

FAX: +33 1 42 34 63 49

Email: francois.radvanyi@curie.fr

Abstract

FGFR3 is one of the most frequently mutated genes in bladder cancer (BLCA) and a driver of an oncogenic dependency. Here, we report that only the most common recurrent *FGFR3* mutation, S249C (TCC → TGC), represents an APOBEC-type motif and is likely caused by the APOBEC-mediated mutagenic process, accounting for its over-representation. We observed significant enrichment of APOBEC mutational signature and over-expression of *AID*/*APOBEC* gene family members in bladder tumors with S249C compared to tumors with other recurrent *FGFR3* mutations. Analysis of replication fork directionality suggests that the coding strand of *FGFR3* is predominantly replicated as lagging strand template that could favour formation of hairpin structures facilitating mutagenic activity of APOBEC enzymes. *In vitro* APOBEC deamination assays confirmed S249 as an APOBEC target. We also found *FGFR3*-S249C mutation to be common in three other cancer types with APOBEC mutational signature, but rare in urothelial tumors without APOBEC mutagenesis and in two diseases likely related to aging.

***Patient summary**:* We propose that APOBEC-mediated mutagenesis can generate clinically relevant driver mutations even within suboptimal motifs, such as in the case of *FGFR3*-S249C, one of the most common mutations in bladder cancer. Knowledge about etiology of this mutation will improve our understanding of molecular mechanisms of bladder cancer.

Keywords: Bladder cancer, upper urinary tract cancer, *FGFR3* mutation, APOBEC, Lynch syndrome

Text

FGFR3 (fibroblast growth factor receptor 3) is one of the most frequently mutated genes in bladder cancer (BLCA). Over 65% of non-muscle-invasive bladder cancer (NMIBC) and 15% of muscle-invasive bladder cancer (MIBC) carry an *FGFR3* mutation driving an oncogenic dependency [1,2]. We reviewed publicly available data for 10,032 bladder tumors (Fig.S1 and Table S1) and identified 56 different *FGFR3* mutations, including 14 recurrent mutations (detected in ≥ 2 samples, Table S2, Fig.1A). The most common was S249C mutation (TCC → TGC), representing 62% of all recurrent *FGFR3* mutations. We wondered whether this over-representation of *FGFR3*-S249C was associated with some specific mutational processes. Considering all mutational signatures [3], the S249C (TCC → TGC) is most similar to an APOBEC-type mutation (TCN → T[G/T]N, where N = any nucleotide, but most frequently A or T). Of all recurrent *FGFR3* mutations, only S249C presents an APOBEC-type motif (Fig.1A). APOBEC (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like) mutational signature accounts for ~30% and 65% of all mutations in NMIBC and MIBC, respectively [2,4]. Thus, we hypothesized that *FGFR3*-S249C mutation might be caused by the activity of APOBEC enzymes.

We analyzed mutational signatures in NMIBC based on RNA-seq data and observed that only the APOBEC-type signature (S3 scores, represent APOBEC signature fraction score and mutation calling from RNA-seq data, Supplementary method) was significantly higher in tumors with S249C mutation compared to tumors with other recurrent *FGFR3* mutations (Fig.1B), while other RNA-seq derived mutational signatures did not differ between these groups (Fig.S2).

We also analyzed APOBEC mutation load in MIBC in The Cancer Genome Atlas (TCGA). Even though only 13% of MIBC had recurrent *FGFR3* mutations (compared to 67% in NMIBC), S249C was found in similar proportions (60%) in MIBC and NMIBC. We were unable to demonstrate a significant association between overrepresentation of S249C mutation and APOBEC mutation load in the much smaller MIBC subset of tumors with recurrent *FGFR3* mutations (n = 52, Fig.S3) compared to NMIBC (n = 227). To consider higher heterogeneity of MIBC than NMIBC, we took advantage of the previous stratification of MIBC tumors as APOBEC-high, APOBEC-low and APOBEC-no [2]. We observed a significantly higher proportion of S249C mutation in tumors with any APOBEC activity (APOBEC-high and low) compared to APOBEC-no tumors (Fig.1C). In addition, considering the two groups of tumors with APOBEC activity, APOBEC mutation load was overall significantly higher in tumors

with S249C mutation compared to tumors bearing other recurrent *FGFR3* mutations (Fig.1D). Thus, it appears that *FGFR3*-S249C mutation is favored in tumors with APOBEC activity;

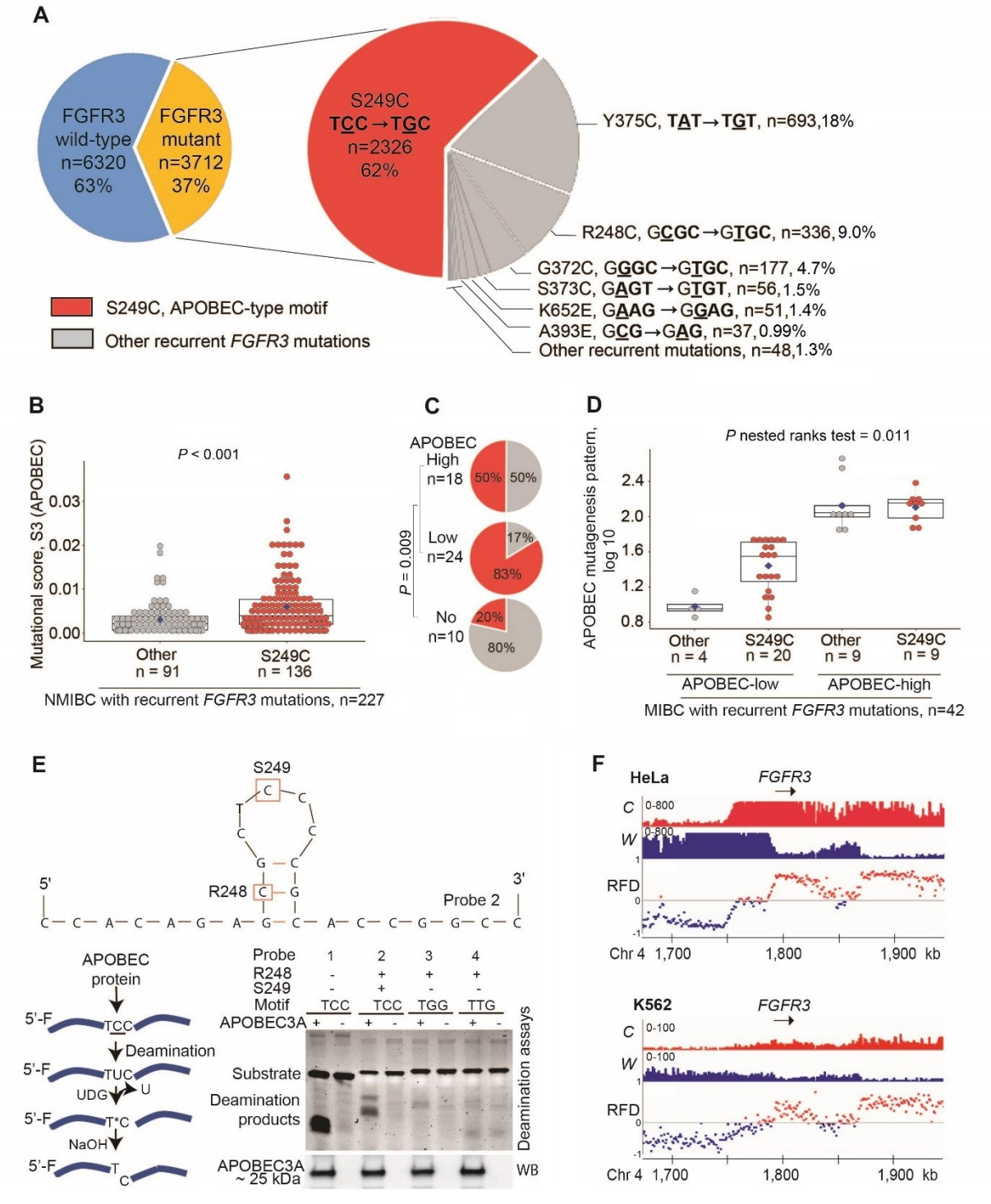


Figure 1. *FGFR3-S249C mutation as a possible outcome of APOBEC-mediated mutagenesis. (A) The rates and distribution of FGFR3 mutations in 10,032 BLCA patients. Shown are recurrent FGFR3 mutations observed in at least two BLCA patients, with some patients carrying several FGFR3 mutations. The mutation numbering corresponds to FGFR3 IIIb as*

the main isoform in cells of epithelial origin. The FGFR3 IIIb isoform contains two more amino acids than the FGFR3 IIIc isoform. The full list of recurrent FGFR3 mutations is provided in Table S2. The most common recurrent FGFR3 mutation hotspot, S249 (TCC) is the only motif possibly targeted by APOBEC-mediated mutagenesis. (B) RNA-seq derived APOBEC mutation score (S3) in 227 NMIBC tumors in relation to recurrent FGFR3 mutations. P-value is for Mann-Whitney U tests between two groups. (C) Distribution of recurrent FGFR3 mutations in 52 TCGA MIBC tumors classified as APOBEC-high, APOBEC-low and APOBEC-no; P-value is for Fisher's exact test comparing APOBEC-no group versus APOBEC presenting groups (high and low). (D) APOBEC mutagenesis pattern (log10) in 42 TCGA MIBC tumors in relation to recurrent FGFR3 mutations in APOBEC-high and APOBEC-low groups. Box-plots show group medians and 50% of all the values, dots represent individual values and group means. P-value is for nested ranks test between all groups of samples. (E) Predicted secondary structure (Mfold) for FGFR3 sequence, with R248 and S249 mutation hotspots marked. APOBEC deamination assays show successful generation of DNA breaks at the cysteine positions only within probe 1 (positive control) and probe 2 in which intact S249 site is located within the single-stranded 5-nucleotide loop, but not within probes 3 and 4, which lack S249 site (negative control). Additional information and secondary structures of all the probes are provided in Fig. S6. (F) Replication fork directionality (RFD) profiles around FGFR3 gene in HeLa and K562 cell lines determined based on mapping of Okazaki fragments to C (Crick) and W (Watson) DNA strands. Red (blue) RFD profiles mark regions where the Watson (Crick) strands are replicated majority as lagging strand templates. Arrows indicate the position of FGFR3 gene (GRCh37_Chr 4: 1,795-1,811 kb) which is predominantly replicated from lagging strand template in both cell types.

BLCA, bladder cancer; NMIBC, non-muscle-invasive bladder cancer; MIBC, muscle-invasive bladder cancer.

APOBEC-low MIBC and NMIBC may have lower background noise than APOBEC-high tumors, making the S249C enrichment more noticeable than in APOBEC-high tumors.

To identify a possible APOBEC mutagen for the *FGFR3*-S249C mutation, we analyzed expression levels of all 11 genes from the AID/APOBEC gene family (Fig.S4). Comparing tumors with *FGFR3*-S249C vs. other recurrent *FGFR3* mutations, only expression of

APOBEC3A and APOBEC3H was significantly different in NMIBC and only expression of APOBEC3A and APOBEC3B in APOBEC-low MIBC (Fig.S5).

APOBEC-mediated mutagenesis preferentially targets lagging DNA strand templates [5], which is consistent with transient excess of single-stranded DNA (ssDNA) during replication process. The efficiency of APOBEC mutagenesis has also been associated with the propensity of ssDNA to form hairpins, with some APOBEC3 enzymes, such as APOBEC3A, preferentially targeting loops in the stem-loop structures [6]. Notably, residue S249 is located in the center of a 5-nucleotide ssDNA loop (Fig.1E). Accordingly, in vitro deamination assays confirmed S249 as a target of the APOBEC deamination activity (Fig.1E). We also performed in silico analysis of genome-wide replication fork directionality (RFD) data in two cancer cell lines [7] (Fig.1F). We conclude that the coding strand of *FGFR3* is replicated predominantly as the lagging strand template, thereby creating an opportunity for ssDNA to form a hairpin and expose S249 to mutagenic activity of APOBEC enzymes.

Interestingly, dominance of *FGFR3*-S249C mutation was reported in sporadic, low-grade upper-tract urothelial carcinomas (UTUC), also enriched in APOBEC-signature mutations. In contrast, when associated with Lynch syndrome (LS), an inherited disorder caused by germline mutations in DNA mismatch repair genes, UTUC lack APOBEC-signature mutations and *FGFR3*-S249C but have high frequency of *FGFR3*-R248C further supporting the link between APOBEC and over-representation of *FGFR3*-S249C [8] (Fig.2A).

We also tested whether the link between the APOBEC-mediated mutagenesis and *FGFR3*-S249C mutation exists in other cancers. We reviewed publicly available data (Table S1) and catalogued *FGFR3* mutations in some other cancer types, including head and neck cancer (HNSCC), cervical cancer (CESC) and non-small cell lung cancer (NSCLC) (Fig.2B-D), in which enrichment of APOBEC-signature mutations has been reported [3]. *FGFR3*-S249C mutation was enriched in all these conditions (Fig. 2B-D). Because APOBEC3s are interferon stimulated genes [9], it is possible that in virally-induced cancers, such as HNSCC and CESC, and in BLCA that may also have infectious etiology, *FGFR3*-S249C mutation is generated as a result of APOBEC3 induction in the course of immune response.

FGFR3 mutations are also detected in benign skin tumors (nevus and seborrheic keratosis) and germline bone disorders (thanatophoric dysplasia). However, in these conditions that have no infectious etiology and have been linked with other causes such as aging [10], R248C (GCG → GTG) is the predominant *FGFR3* mutation (Fig.2E-F).

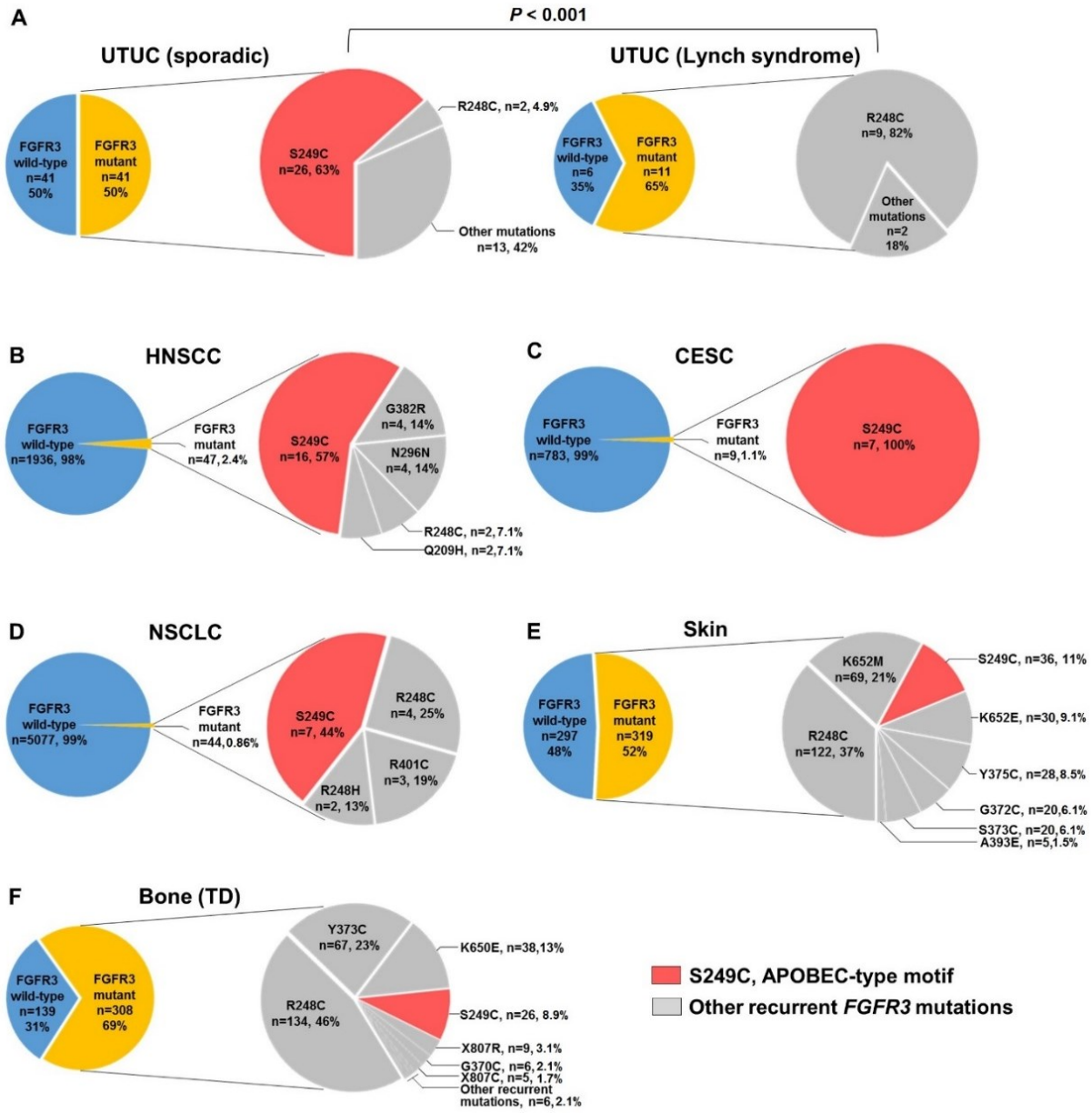


Figure 2. *FGFR3 mutation spectrum across several cancer types, benign skin tumors and bone disorders. (A-F) The rates and distribution of FGFR3 mutations in patients with sporadic (n = 82) and Lynch syndrome-associated (n = 17) UTUC, HNSCC (n = 1983), CESC (n = 792), NSCLC (n = 5121), benign skin tumors (n = 616) and bone disorders (thanatophoric dysplasia) (n = 447). Among all recurrent FGFR3 mutations only FGFR3-S249C mutation motif (TCC) is the possible target of APOBEC-mediated mutagenesis. (A-E) The mutation numbering corresponds to FGFR3 IIIb as the main isoform in cells of epithelial origin. (F) The mutation numbering corresponds to FGFR3 IIIc as the main isoform in chondrocytes. The full list of recurrent FGFR3 mutations with numbering corresponding to both FGFR3 IIIb and IIIc isoforms is provided in Table S2.*

HNSCC, head and neck squamous cell carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; NSCLC, non-small cell lung cancer; benign skin tumors include seborrhoeic keratosis and epidermal nevus; bone disorders include thanatophoric dysplasia-I (TD-I) and II (TD-II). FGFR3 mutations found in cancers and benign skin tumors are somatic, those found in bone disorders are germline but identical to somatic mutations in tumors.

We found that FGFR3-S249C protein has similar potential to transform NIH-3T3 cells compared to FGFR3 with a recurrent non-APOBEC-type mutation Y375C (TAT → TGT, 18% of BLCA, Fig.1A), and FGFR3 with either mutation activates the same transcriptional regulators in bladder cancer cell lines suggesting their comparable functions (Fig.S7). Thus, the over-representation of S249C in APOBEC-related cancers is likely due to increased mutation rate caused by APOBEC3 activity rather than increased tumorigenicity of the S249C mutation.

In conclusion, we demonstrate that *FGFR3*-S249C mutation, despite being a less frequent APOBEC-motif, is likely caused by the APOBEC-mediated mutagenic activity in BLCA and other conditions. Further investigations should explore whether the APOBEC mutagenesis alone generates *FGFR3*-S249C mutation or it requires other factors. Our results also pave the way for further studies to explore other APOBEC-induced driver mutations considering broader definition of motifs targeted by the APOBECs.

Author contributions: François Radvanyi had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Shi, Meng, Dyrskjøt, Prokunina-Olsson, Bernard-Pierrot, Radvanyi.

Acquisition of data: Shi, Chen, Banday, Lamy, Moreno-Vega, Bernard-Pierrot.

Analysis and interpretation of data: Shi, Meng, Banday, Lamy, Yang, Moreno-Vega, Chen, Dyrskjøt, Bernard-Pierrot ,Prokunina-Olsson, Radvanyi.

Drafting of the manuscript: Shi, Meng, Prokunina-Olsson, Bernard-Pierrot, Radvanyi.

Critical revision of the manuscript for important intellectual content: Lamy, Chen, Dyrskjøt.

Statistical analysis: Meng, Yang, Banday, Moreno-Vega.

Obtaining funding: Bernard-Pierrot, Radvanyi.

Administrative, technical, or material support: None.

Supervision: Dyrskjøt, Prokunina-Olsson, Bernard-Pierrot, Radvanyi.

Other (specify): None.

Financial disclosures: François Radvanyi certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

Funding/Support and role of the sponsor: This work was supported by a grant from Ligue Nationale Contre le Cancer (IBP, FR, MJS, XYM, AMV) as an associated team (Equipe labellisée). MJS was supported by a scholarship from China Scholarship Council. XYM was supported by a fellowship from AVIESAN ITMO-cancer INSERM. ARB and LPO were supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, USA. PL and LD were supported by Aarhus University, Denmark.

Acknowledgment statement: We would like to acknowledge several colleagues: Jing Liu for improving R script; Yanish Soorojebally, Florent Dufour and Elodie Chapeaublanc for data and figure preparation; Virginia Sanchez-Quiles for manuscript revision. We have used data partially generated by the TCGA Research Network (http://cancergenome.nih.gov/). We acknowledge all members of TCGA working group.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j. eururo.2019.03.032.

References

[1] Hedegaard J, Lamy P, Nordentoft I, Algaba F, Høyer S, Ulhøi BP, et al. Comprehensive Transcriptional Analysis of Early-Stage Urothelial Carcinoma. Cancer Cell 2016;30:27–42. doi:10.1016/j.ccell.2016.05.004.

[2] Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell 2017;171:540–556.e25. doi:10.1016/j.cell.2017.09.007.

[3] Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Ng AWT, Boot A, et al. The Repertoire of Mutational Signatures in Human Cancer. BioRxiv 2018. doi:https://doi.org/10.1101/322859.

[4] Lamy P, Nordentoft I, Birkenkamp-Demtröder K, Houlberg Thomsen MB, Villesen P, Vang S, et al. Paired exome analysis reveals clonal evolution and potential therapeutic targets in urothelial carcinoma. Cancer Res 2016;76:5894–906. doi:10.1158/0008-5472.CAN-16-0436.

[5] Haradhvala NJ, Polak P, Stojanov P, Covington KR, Shinbrot E, Hess JM, et al. Mutational Strand Asymmetries in Cancer Genomes Reveal Mechanisms of DNA Damage and Repair. Cell 2016;164:538–49. doi:10.1016/j.cell.2015.12.050.

[6] Sharma S, Baysal BE. Stem-loop structure preference for site-specific RNA editing by APOBEC3A and APOBEC3G. PeerJ 2017;5:e4136. doi:10.7717/peerj.4136.

[7] Wu X, Kabalane H, Kahli M, Petryk N, Laperrousaz B, Jaszczyszyn Y, et al. Developmental and cancer-associated plasticity of DNA replication preferentially targets GC-poor, lowly expressed and late-replicating regions. Nucleic Acids Res 2018:1–16. doi:10.1093/nar/gky797.

[8] Donahue TF, Bagrodia A, Audenet F, Donoghue MTA, Cha EK, Sfakianos JP, et al. Genomic Characterization of Upper Tract Urothelial Carcinoma in Patients With Lynch Syndrome. JCO Precis Oncol 2018. doi:10.1200/PO.17.00143.

[9] Middlebrooks CD, Banday AR, Matsuda K, Udquim KI, Onabajo OO, Paquin A, et al. Association of germline variants in the APOBEC3 region with cancer risk and enrichment with APOBEC-signature mutations in tumors. Nat Genet 2016;48:1330–8. doi:10.1038/ng.3670.

[10] Goriely A, Wilkie AOM. Paternal age effect mutations and selfish spermatogonial selection: Causes and consequences for human disease. Am J Hum Genet 2012;90:175–200. doi:10.1016/j.ajhg.2011.12.017.

Supplementary materials

1. Supplementary methods

1.1 Data collection

1.1-1) Databank for FGFR3 mutation spectrum was compiled from three sources: 1) COSMIC portal (https://cancer.sanger.ac.uk/cosmic) [1], 2) cBioPortal for Cancer Genomics (http://www.cbioportal.org/) [2,3], and 3) manual search. We extracted well-documented FGFR3 mutation data in tumors from COSMIC portal and selected 4 cancers and several skin diseases (seborrhoeic keratosis and epidermal nevus) with a significant number of recurrent FGFR3 mutations. As cBioPortal is another important public source of mutation data, we double-checked the records for the selected cancers in cBioPortal. We included the latest data or, if there was an overlap (ie. TCGA-BLCA), we combined the data between these two major sources, otherwise we manually added non-redundant data from cBioPortal. As neither COSMIC nor cBioPortal included data from the two large cohorts of non-muscle-invasive bladder cancer (NMIBC) [4,5], we added them manually. Lastly, we noted that although FGFR3 mutations were common in bone disorders (thanatophoric dysplasias), no systematically pooled data were publicly available. Therefore, we manually reviewed literature of thanatophoric dysplasia and catalogued a comprehensive FGFR3 mutational spectrum for this disease, with all FGFR3 mutations being germline. One article [6] was excluded, because the frequent mutation (G697C) reported in this study was debatable [7] and not observed elsewhere. A graphical workflow of data collection and detailed mutation spectrum are presented in Fig. S1 and Table S1 (separate Excel file).

Recent publications reported enrichment of FGFR3-R248C mutation in upper urinary tract urothelial cancer (UTUC) with Lynch syndrome where APOBEC signature was very low; in contrast, S249C mutation was much more common in the subgroup of UTUC without Lynch syndrome that exhibited APOBEC signature [8,9]. We presented the reported data [8] in Fig. 2A.

1.1-2) NMIBC cohort

The largest NMIBC cohort to date with a total of 476 tumors was published by Hedegaard et al [4]. For these tumors we used RNA-Seq derived scores for six mutational signatures, including APOBEC-like, S3 scores (227 tumors with mutation load adequate for signature extraction, including 136 tumors with FGFR3-S249C mutation and 91 tumors with other recurrent FGFR3 mutations) and RNA expression measured as FPKM (270 tumors subjected to RNA sequencing, including 161 tumors with FGFR3-S249C mutation and 109 tumors with other recurrent FGFR3 mutations).

1.1-3) TCGA-MIBC cohort

Data for the FGFR3 mutation status, log10-transformed APOBEC mutagenesis pattern (represented by APOBEC_MutLoad_MinEstimate) and APOBEC mutagenesis category (no, low, and high) were available in Table S1 of the TCGA bladder cancer paper [10]. RNA-seq data (RSEM) were downloaded from cBioPortal and log2-transformed. There were 52 tumors with recurrent FGFR3 mutations: 31 with S249C versus 21 with other mutations; of those - 10 tumors were classified as APOBEC-no, 24 as APOBEC-low and 18 as APOBEC-high; one tumor lacked RNA-seq data and was not used in expression analysis.

1.2 Deamination assays

Custom-designed 5'-fluorescein-labeled oligonucleotides (probes 1-4) were purchased from Thermo Fisher Scientific. The positive control (probe 1) carrying a TCC motif was previously described [11]. The probe 2 included a 25-nucleotide fragment of FGFR3 centered on S249 (TCC); in negative control probes 3 and 4 the S249 (TCC) sequence was altered to TGG and TTG. Probes 2-4 also carry the R248 site (GCG, underlined italics). Deamination is expected to affect cytosines within the underlined motifs; additional identical sequences not targeted by deamination (small fonts) were added to probes 2-4 to increase their size.

Probe 1:

5′-fluorescein - ATTATTATTATTATTCCCAATTATTTATTTATTTATTTATTT

Probe 2:

5′fluorescein – attattattaCCACAGAGCGCTCCCCGCACCGGCCattattattat - 3'

Probe 3:

5′-fluorescein - attattattaCCACAGAGCGCTGGGCGCACCGGCCattattattat - 3'

Probe 4:

5′-fluorescein – attattattaCCACAGAGCGCTTGGCGCACCGGCCattattattat - 3'

The C-terminally Myc-DDK tagged APOBEC3A expression construct (NM_145699) in the pCMV6 vector was purchased from OriGene (Rockville, MD). The construct was transiently transfected with Lipofectamine 3000 (Thermo Fisher Scientific) into human embryonal kidney HEK293-T cells, seeded in 175 cm2 flasks (Corning) at a density of 4×106 cells/20 mls. Cells were harvested and lysed in CelLytic M buffer (Sigma) 24 hrs post-transfection. To increase concentration of the recombinant APOBEC3A protein, whole-cell lysates were passed through purification step using c-Myc tagged Protein Mild Purification Kit (MBL, Japan) and treated with RNAase A at 37℃ for 30 minutes.

Deamination reactions were performed using a previously described protocol [12]. Briefly, each 10 µl reaction mix contained 1 µl of a probe (5-10 picomoles), 4 µl of semi-purified APOBEC3A recombinant protein (~ 0.25 ug) and 1 µl of 10x deamination buffer (100 µl of 100 mM Tris/HCl, pH 7.5; 100 µl of 500 mM NaCl; 10 µl of 10 mM DTT and 790 µl of water) and 4 µl H2O. Reactions were incubated in water bath at 37°C for 2 hrs, treated with Uracil DNA Glycosylase (UDG) for 40 min at 37°C, followed by addition of 0.6 N NaOH for 20 min at 37°C. After adding 20 µl of 2x RNA loading dye (Thermo Fisher Scientific), the reactions were heated at 95°C for 2-3 min. Of the total reaction volume, 15 µl aliquot was resolved on 15% TBE-urea polyacrylamide gel (Life Technologies) at 150 V for 1 hr and 30 min at room temperature in 1x TBE buffer. Gels were imaged with Gel Doc (Bio-Rad) using Fluorescein/UV settings. Another set of 15 µl aliquots from the same reactions was separately resolved on 4-12% Tris-glycine SDS polyacrylamide gel (Life Technologies) for detection of APOBEC3A with an anti-DDK antibody (F7425; Lot # 086M4803V; Sigma) using the ECL Plus Western blotting detection system (GE Healthcare Life Sciences).

1.3 Analysis of secondary structure of single-stranded DNA (ssDNA)

Mfold tool with default parameters for DNA folding (http://unafold.rna.albany.edu/?q=mfold) [13] was used to evaluate secondary structure of all 4 probes used for deamination assays, focusing on 25 nucleotide sequences centered on FGFR3-S249C as input.

1.4 Functional comparison of FGFR3 with S249C versus Y375C mutations

NIH-3T3 cells (murine fibroblasts) transiently transfected with expression constructs for the human FGFR3 with S249C or Y375C mutations, positive control with high transforming potential (HRAS-Q61R) or the mock control pcDNAI-Neo plasmid (Neo) were established as

previously described [14]. Pools of transfected cells were established by two weeks of selection on 800 µg/ml G418, followed by culturing in DMEM with 10% fetal bovine serum (FBS), 2 mM glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, and 400 µg/ml G418. The ability of the expressed proteins to transform NIH-3T3 cells was evaluated by the anchorage independent growth of the cells in soft agar. Cells (3x104) were seeded in 12-well plates containing DMEM with 10% FBS and 1% agar, in triplicates. The plates were incubated for two weeks and colonies larger than 50 µm in diameter, as measured with a phase-contrast microscope equipped with a measuring grid, were counted.

To identify genes regulated by FGFR3 with different mutations, MGH-U3 and UMUC-14 bladder cancer cells endogenously expressing FGFR3-Y375C and FGFR3-S249C, respectively, were transfected for 72 hrs with three FGFR3 siRNAs (described in Mahe et al [15]). mRNA was extracted and purified with the RNeasy Mini kit (Qiagen). Total RNA (200 ng) from control and siRNA-treated MGH-U3 and UMUC-14 cells was analyzed with the Affymetrix human exon 1.0 ST array and the Affymetrix U133 plus 2 array, respectively, as previously described [15]. Experiments using MGH-U3 cells have been described by Mahe et al [15] and the microarray data were available from GEO (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE84733. Data for the UMUC-14 cells were generated in the current work. The LIMMA algorithm was used to identify genes differentially expressed between FGFR3 siRNA-treated (3 different siRNAs) and Lipofectamine-treated cells (3 replicates) [16]. The p-values were adjusted for multiple testing by Benjamini–Hochberg FDR method. Genes with a log2 fold-change ≥ 0.58, in a positive or negative direction and an FDR p-value below 5%, were considered to be differentially expressed. An analysis of the two lists of FGFR3-regulated genes using the upstream regulator function of the Ingenuity Pathway Analysis (IPA) software identified upstream regulators activated and inhibited by FGFR3-S249C and FGFR3-Y375C.

1.5 Replication fork directionality (RFD) profiling

We used data for RFD profiling in two human cancer cell lines - HeLa and K562 cells [17,18]. In these reports, the authors isolated and sequenced Okazaki fragments (OK-Seq) to determine the whole-genome RFD profiles of a given cell model. RFD was computed as the difference between the proportions of Crick (C) and Watson (W) okazaki fragments in 1 kb windows as: $RFD = (C - W)/(C + W)$. A region majority replicated by right-ward replication forks (Watson strand as lagging strand template) was considered as "+" RFD, and a left-ward replication forks (Crick strand as lagging strand template) was considered as "-" RFD. This directionality determined which strand would be favored as lagging strand template. Analysis of RFD

profiles showed that FGFR3-S249C mutation was located in the lagging strand template, known to be preferentially targeted by APOBEC mutagenesis [19,20].

1.6 Statistical analysis

Non-parametric Mann-Whitney U test was performed to compare APOBEC signature/mutagenesis values and expression of APOBEC genes between groups of tumors with FGFR3-S249C and other recurrent FGFR3 mutations. The mixed model extension of Mann-Whitney U test, i.e. nested ranks test, was used for similar analysis with multiple groups. Dunnett's test was performed to compare the number of soft agar colonies after overexpression of FGFR3-S249C, FGFR3-Y375C, and negative and positive controls in NIH-3T3 cells. Analyses were performed using R version 3.5.2, package 'nestedRanksTest', version 0.2. Fisher's exact test was used to compare differences in distribution of categorical variables. Plots were generated with Microsoft Excel 2016 (pie charts) or R version 3.5.2 using package 'easyGgplot2', version 1.0.0.9000. Figures were assembled in Adobe Illustrator.

2. Supplementary Figures

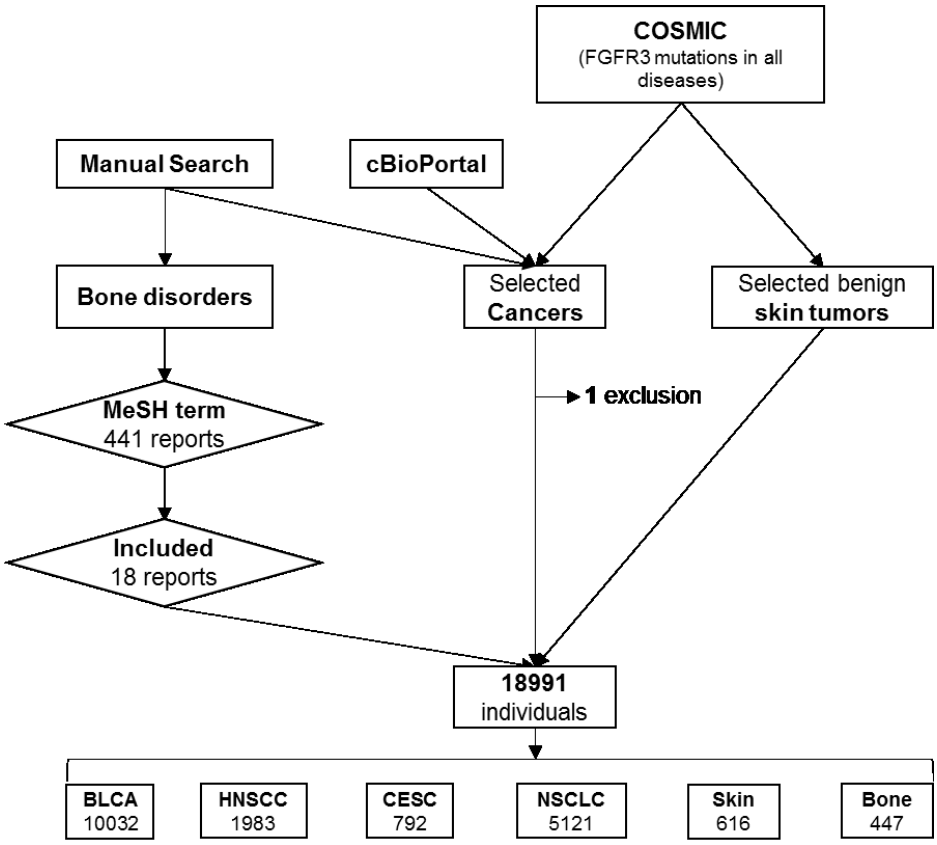


Fig. S1 Workflow of data collection for *FGFR3* mutation spectrum. MeSH terms can be found in **Table S1**. BLCA, bladder cancer; HNSCC, head and neck squamous cell carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; NSCLC, non-small cell lung cancer; Benign skin tumors include seborrhoeic keratosis and epidermal nevus; Bone disorders include thanatophoric dysplasia-I (TD-I) and II (TD-II).

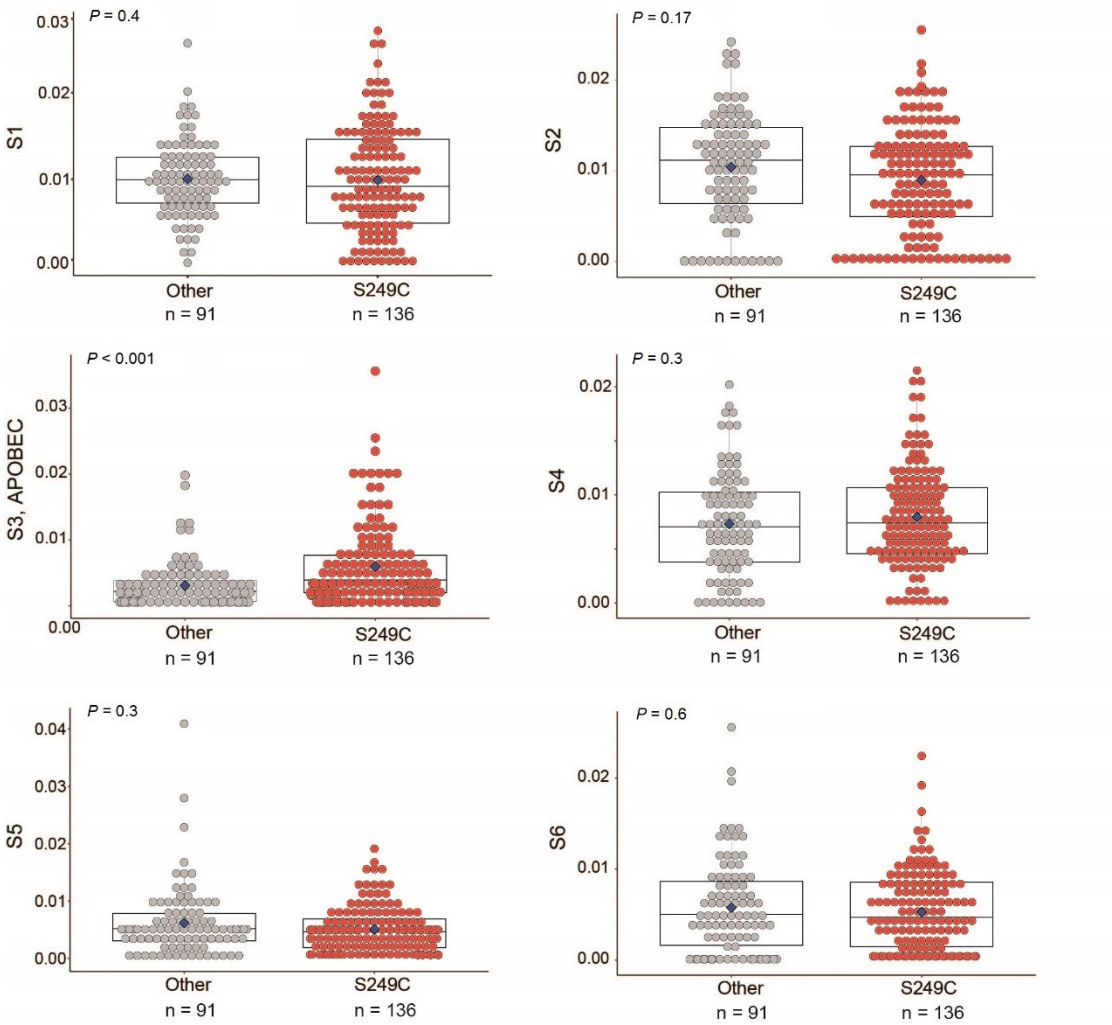


Fig. S2 Distribution of mutational signature scores in 227 tumors from patients with non-muscle-invasive bladder cancer (NMIBC [4]) with recurrent *FGFR3* mutations - S249C vs. other mutations. Recurrent mutations were defined as those found in at least 2 patients in analysis presented in **Fig.1A** and listed in **Table S2**. *P*-values are for Mann-Whitney U test; the result for S3 (APOBEC) signature scores is also plotted in **Fig. 1B.** Box-plots show group medians and 50% of all the values, dots represent individual values and group means.

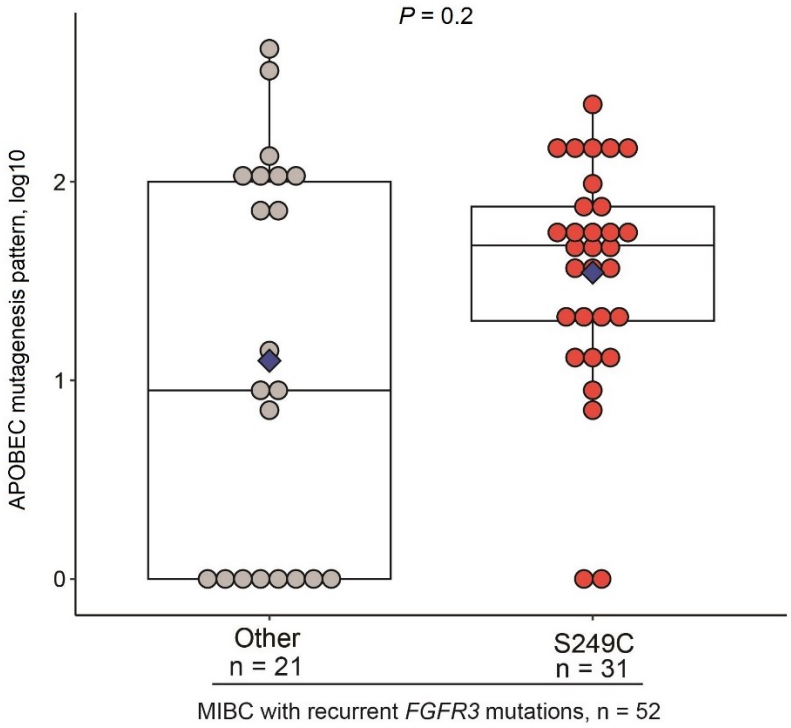


Fig. S3 Distribution of APOBEC mutational pattern in 52 MIBC [10] patients with recurrent *FGFR3* mutations - S249C vs. other mutations. Recurrent mutations were defined as those found in at least 2 patients in analysis presented in **Fig. 1A** and listed in **Table S2**. *P*-values are for Mann-Whitney U test. Box-plots show group medians and 50% of all the values, dots represent individual values and group means.

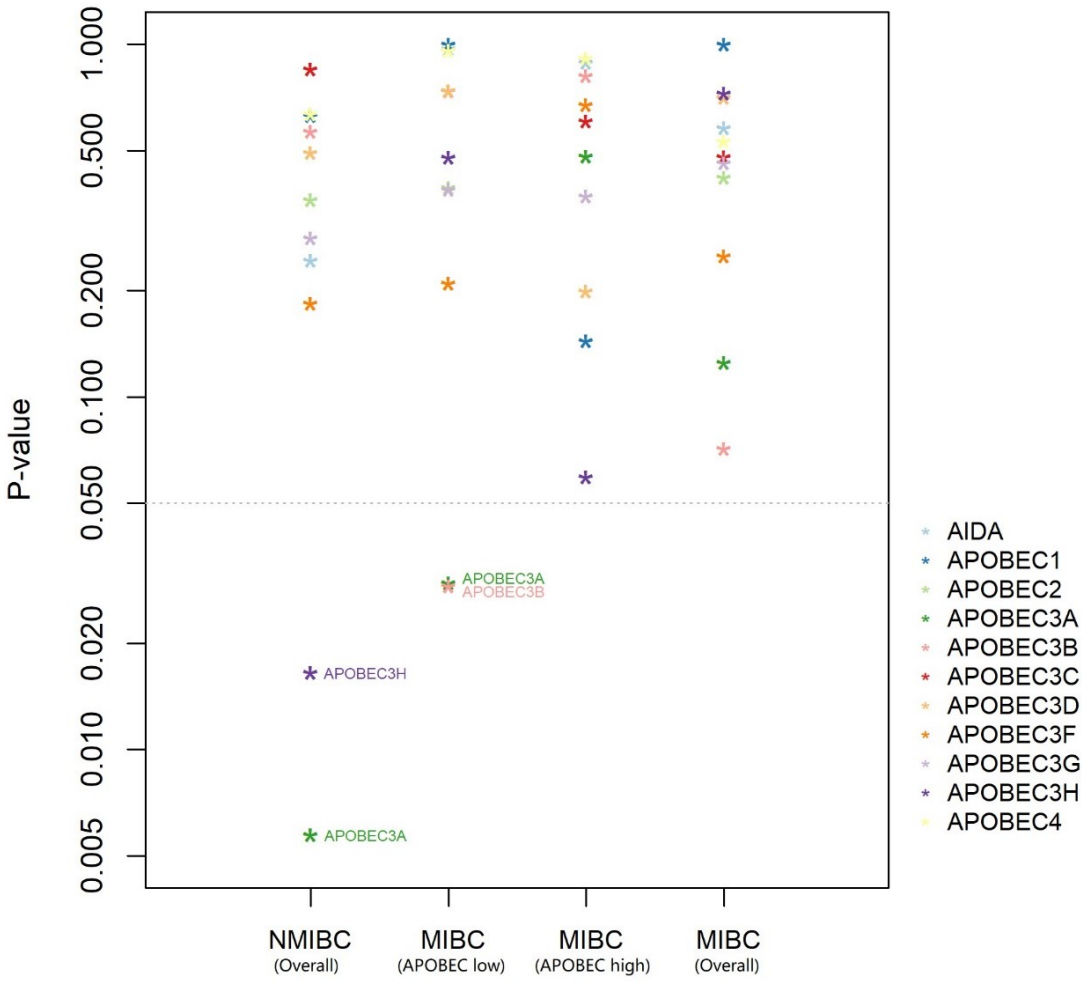


Fig. S4 Statistical significance for the association between *AID/APOBEC* gene expression (FPKM, log2) and recurrent *FGFR3* mutations - S249C vs. other recurrent *FGFR3* mutations in 270 NMIBC [4] and 41 MIBC (RSEM, log2) [10] patients (24 with APOBEC-low and 17 with APOBEC-high tumors). Low and high groups correspond to APOBEC-signature mutation load, as has been previously defined [10]. Recurrent mutations were defined as those present in at least 2 patients in analysis presented in **Fig. 1A** and listed in **Table S2.** There are 11 APOBEC genes that could potentially contribute to APOBEC mutagenesis - *AICDA (AID), APOBEC1 (A1), APOBEC2 (A2), APOBEC3* (*A3A, A3B, A3C, A3D, A3F, A3G and A3H*) and *APOBEC4 (A4)* [21]. *P*-values are for Mann-Whitney U tests between two groups (overall NMIBC, APOBEC-low MIBC and APOBEC-high MIBC) or for nested ranks test between all groups of MIBC samples. The asterisks with different colors represent *P*-values for association between *FGFR3*-S249C and indicated genes in different groups.

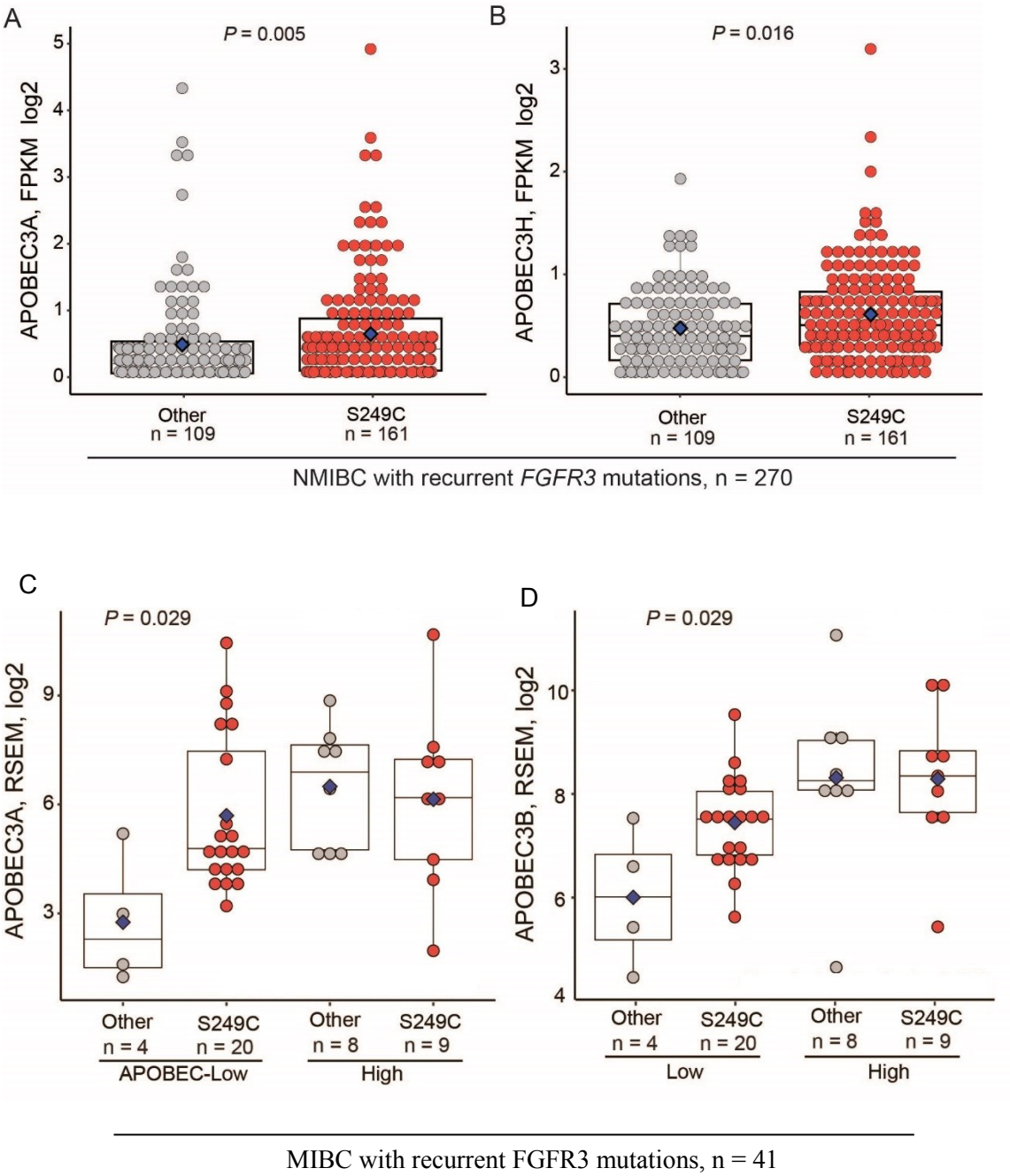


Fig. S5 Distribution of expression values for the *AID/APOBEC* genes significantly associated with recurrent *FGFR3* mutations - S249C vs. other mutations in 270 NMIBC [4] and 41 MIBC [10] patients in analysis presented in Fig. S4. **(A)** *APOBEC3A* in NMIBC tumors. **(B)** *APOBEC3H* in NMIBC tumors. **(C)** *APOBEC3A* in MIBC tumors. **(D)** *APOBEC3B* in MIBC tumors. APOBEC-low and high groups correspond to APOBEC-signature mutation load, as has been previously defined [10]. Recurrent mutations were defined as those present in at least 2 patients in analysis presented in **Fig. 1A** and listed in **Table S2**. Box-plots show group medians and 50% of all the values, dots represent individual values and group means. *P*-values are for Mann-Whitney U tests between two (comparison was conducted only within APOBEC-low groups in Fig.S5C-D).

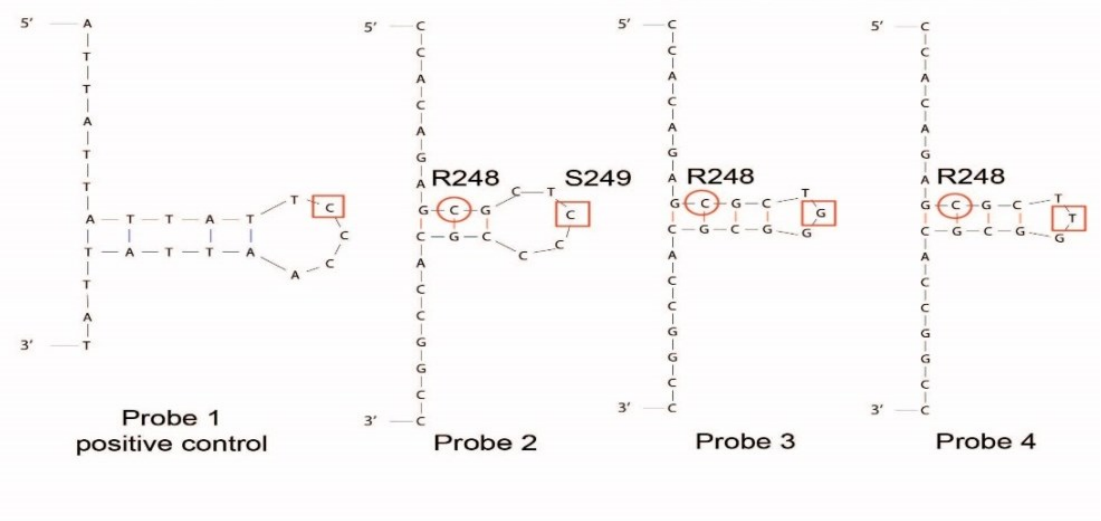
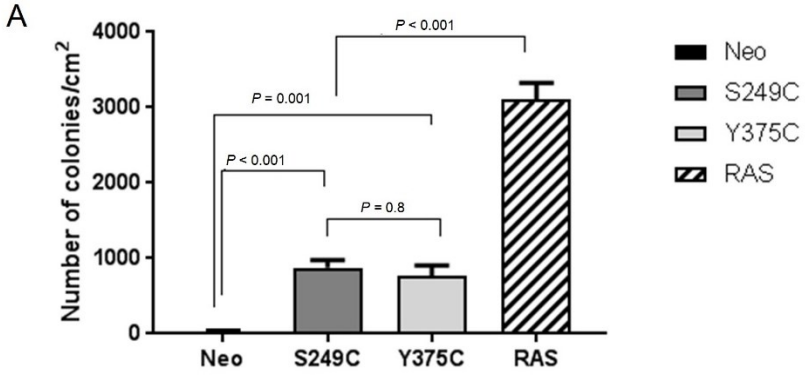
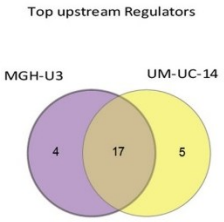


Fig. S6 Mfold analysis of secondary structures of all probes used for deamination assays - S249 is located within a single-stranded 5-nucleotide loop, while R248 is located within the double-stranded hairpin stem; shown are central 25 bp nucleotides of each probe. APOBEC-mediated mutagenesis is accumulated in ssDNA, preferentially targeting hairpin loops [13,22,23]. Loops of more than 3-nt have been shown to aid APOBEC enzyme binding [13,23], with the APOBEC3A binding site requiring bent ssDNA [22].

A



B



	MGH-U3			UM-UC-14		
Upstream Regulator	Predicted state	Activation z-score	Adj pval	Predicted state	Activation z-score	Adj pval
MYC	Inhibited	-4,57	9,43E-11	Inhibited	-5,194	6,37E-15
TRIM24	Inhibited	-2,795	4,23E-03	Inhibited	-4,31	5,10E-10
TBX2	Inhibited	-3,9	2,67E-12	Inhibited	-3,898	9,18E-06
ESR1	Inhibited	-3,511	4,94E-17	Inhibited	-3,814	3,65E-33
TAL1	Inhibited	-2,846	1,50E-08	Inhibited	-3,685	2,48E-03
CCND1	Inhibited	-3,929	1,00E-16	Inhibited	-3,555	6,58E-12
RARA	Inhibited	-3,305	3,73E-09	Inhibited	-3,198	4,25E-12
FOXM1	Inhibited	-3,174	8,78E-06	Inhibited	-3,149	2,07E-03
HSF2	Inhibited	-2,646	1,91E-04	Inhibited	-2,813	7,80E-04
E2F3	Inhibited	-4,096	6,10E-13	Inhibited	-2,213	1,19E-04
RBL1	Activated	2,626	2,02E-10	Activated	2,155	1,92E-02
NFATC2	Activated	2,423	9,33E-03	Activated	3,156	8,04E-05
MXI1	Activated	2,57	2,19E-08	Activated	2,607	2,59E-05
GATA1	Activated	3,477	5,09E-07	Activated	3,448	2,43E-02
FOXO3	Activated	2,9	7,92E-13	Activated	3,694	9,99E-16
TP53	Activated	3,362	1,20E-28	Activated	5,462	3,44E-23
STAT1	Activated	2,817	3,22E-02	Activated	5,525	7,05E-06
E2F1	Inhibited	-3,531	6,50E-17	-	-1,379	3,30E-08
NUPR1	Activated	3,474	1,94E-16	-	-0,108	6,54E-11
HDAC1	Activated	2,534	3,16E-07	-	1,316	1,26E-02
VHL	Activated	2,607	9,31E-03	-	1,246	1,97E-03
NKX2-3	-	-1,414	2,80E-04	Inhibited	-5,011	4,90E-16
IRF3	-	-	-	Activated	3,55	4,15E-06
IRF1	-	1,309	3,72E-02	Activated	3,949	8,46E-07
TCF7L2	-	0,791	2,14E-02	Activated	4,122	4,93E-03
IRF7	-	1,987	4,27E-02	Activated	5,284	6,30E-12

Fig. S7 Similar tumorigenic potential of FGFR3 with S249C and Y375C mutations. **(A)** Overexpression of FGFR3-S249C and FGFR3-Y375C in NIH-3T3 cells shows similar transformation potential and significantly lower compared to HRAS-Q61R (positive control), based on the number of soft agar colonies; Dunnett's multiple comparisons test. Shown are mean+/- SD of one representative experiment conducted in triplicate. Three experiments were performed with three different pools of transiently transfected cells. **(B)** Genes affected by FGFR3 depletion in human MGHU-U3 and UMUC-14 bladder cancer cells endogenously expressing FGFR3-Y375C and FGFR3-S249C, respectively, were identified using expression analysis with Affymetrix arrays. Upstream regulators possibly controlling the expression of these genes were identified using IPA software. Top 10 most activated vs. most inhibited master regulators were compared in both cell lines. The Venn diagram shows a strong overlap of the main master regulators modulated by FGFR3 in the same way with either mutation suggesting their comparable ability to activate the same main signaling pathways. FGFR3 exists as two isoforms, FGFR3 IIIb (main isoform in cells of epithelial origin) and FGFR3 IIIc (expressed in chondrocytes). Due to difference in the size of an alternatively spliced exon in FGFR3 IIIb compared to FGFR3 IIIc, the numbering after S249 shifts by +2. Full list of FGFR3 mutations and their numbering is provided in **Table S2**.

3. Supplementary tables

Table S1. FGFR3 mutation spectrum in human pan-cancer and benign skin tumors and bone disorders among 18,991 individuals. Related to Figure 1A, Figure 2B-F and Supplementary Figure 1. This large matrix is available at https://doi.org/10.1016/j.eururo.2019.03.032.

Table S2. Frequency of recurrent FGFR3 mutations. Related to Figure 1A and Supplementary Figure 2B-2F.

Recurrent *FGFR3* mutations		Pan-cancer and other diseases						Nucleotide context	
MUTATION_CDS	MUTATION_AA*	BLCA	HNSCC	CESC	NSCLC	Skin	Bone (TD)	Code_WT	Code_Mut
c.746C>G	p.S249C	62%(2326)	57%(16)	100%(7)	44%(7)	11%(36)	9%(26)	**TCC**	**TGC**
c.1124A>G	p.Y375C	18%(693)	/	/	/	9%(28)	23%(67)	**TAT**	**TGT**
c.742C>T	p.R248C	9%(336)	7%(2)	/	25%(4)	37%(122)	46%(134)	G**C**GC	G**T**GC
c.1114G>T	p.G372C	5%(177)	/	/	/	6%(20)	2%(6)	G**G**GC	G**T**GC
c.1117A>T	p.S373C	1%(56)	/	/	/	6%(20)	1%(2)	G**A**GT	G**T**GT
c.1954A>G	p.K652E	1%(51)	/	/	/	9%(30)	13%(38)	G**A**AG	G**G**AG
c.1178C>A	p.A393E	1%(37)	/	/	/	2%(5)	/	**G**CG	**G**AG
c.1144G>C or c.1144G>A	p.G382R	1%(20)	14%(4)	/	/	/	/	C**G**GG	C**C/A**GG
c.1955A>T	p.K652M	0.3%(12)	/	/	/	21%(69)	1%(2)	**A**AG	**A**TG
c.1954A>C	p.K652Q	0.1%(5)	/	/	/	/	/	G**A**AG	G**C**AG
c.1955A>C	p.K652T	0.1%(5)	/	/	/	/	/	**A**AG	**A**CG
c.1156T>C	p.F386L	0.1%(2)	/	/	/	/	/	C**T**TC	C**C**TC
c.1178C>T	p.A393V	0.1%(2)	/	/	/	/	/	**G**CG	**G**TG
c.1927G>A	p.D643N	0.1%(2)	/	/	/	/	/	G**G**AC	G**A**AC

**Mutation positions correspond to FGFR3 IIIb, the numbering of FGFR3 IIIc see in the sheet of <ReadMeFirst>. Recurrent mutations (n=14) were defined as present in at least 2 of 3712 patients with bladder cancer (count see in Suppl 2b); One mutations - p.A371A was excluded as it was a silent mutation. BLCA, Bladder cancer; HNSCC, Head and neck squamous cell carcinoma; CESC,Cervical squamous cell carcinoma and endocervical adenocarcinoma; NSCLC, Non-small cell lung cancer; Benign skin tumors, composed of Seborrhoeic keratosis and Epidermal nevus; Bone disorders, composed of Thanatophoric dysplasia-I (TD-I) and TD-II. Codon was shown in bold and mutated nucleotide underlined.*

4. Supplementary references

[1] Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, et al. COSMIC: Somatic cancer genetics at high-resolution. Nucleic Acids Res 2017;45:D777–83. doi:10.1093/nar/gkw1121.

[2] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2:401–4. doi:10.1158/2159-8290.CD-12-0095.

[3] Jianjiong Gao, Bülent Arman Aksoy, Ugur Dogrusoz, Gideon Dresdner, Benjamin Gross, S. Onur Sumer, Yichao Sun, Anders Jacobsen, Rileen Sinha EL, Ethan Cerami, Chris Sander and NS. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. Sci Signal 2013;6:1–34. doi:10.1126/scisignal.2004088.Integrative.

[4] Hedegaard J, Lamy P, Nordentoft I, Algaba F, Høyer S, Ulhøi BP, et al. Comprehensive Transcriptional Analysis of Early-Stage Urothelial Carcinoma. Cancer Cell 2016;30:27–42. doi:10.1016/j.ccell.2016.05.004.

[5] Hurst CD, Alder O, Platt FM, Droop A, Stead LF, Burns JE, et al. Genomic Subtypes of Non-invasive Bladder Cancer with Distinct Metabolic Profile and Female Gender Bias in KDM6A Mutation Frequency. Cancer Cell 2017;32:701–715.e7. doi:10.1016/j.ccell.2017.08.005.

[6] Yan Z, Hiraishi Y, Wang H, Sumi KS, Hayashido Y, Toratani S, et al. Constitutive activating mutation of the FGFR3b in oral squamous cell carcinomas. Int J Cancer 2005;117:166–8. doi:10.1002/ijc.21145.

[7] Aubertin J, Tourpin S, Janot F, Ahomadegbe J-C, Radvanyi F. Analysis of fibroblast growth factor receptor 3 G697C mutation in oral squamous cell carcinomas. Int J Cancer 2007;120:2058–2059; author reply 2060. doi:10.1002/ijc.22285.

[8] Donahue TF, Bagrodia A, Audenet F, Donoghue MTA, Cha EK, Sfakianos JP, et al. Genomic Characterization of Upper Tract Urothelial Carcinoma in Patients With Lynch Syndrome. JCO Precis Oncol 2018. doi:10.1200/PO.17.00143.

[9] Audenet F, Isharwal S, Cha EK, Donoghue MTA, Drill E, Ostrovnaya I, et al. Clonal relatedness and mutational differences between upper tract and bladder urothelial carcinoma. Clin Cancer Res 2018:clincanres.2039.2018. doi:10.1158/1078-0432.CCR-18-2039.

[10] Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell 2017;171:540–556.e25. doi:10.1016/j.cell.2017.09.007.

[11] Kanu N, Cerone MA, Goh G, Zalmas L, Bartkova J, Dietzen M, et al. DNA replication stress mediates APOBEC3 family mutagenesis in breast cancer. Genome Biol 2016:1–15. doi:10.1186/s13059-016-1042-9.

[12] Nair S, Rein A. In vitro Assay for Cytidine Deaminase Activity of APOBEC3 Protein. Bio Protoc 2016;4. doi:10.21769/BioProtoc.1266.

[13] Holtz CM, Sadler HA, Mansky LM. APOBEC3G cytosine deamination hotspots are defined by both sequence context and single-stranded DNA secondary structure. Nucleic Acids Res 2013;41:6139–48. doi:10.1093/nar/gkt246.

[14] Bernard-Pierrot I, Brams A, Dunois-Lardé C, Caillault A, Diez de Medina SG, Cappellen D, et al. Oncogenic properties of the mutated forms of fibroblast growth factor receptor 3b. Carcinogenesis 2006;27:740–7. doi:10.1093/carcin/bgi290.

[15] Mahe M, Dufour F, Neyret-Kahn H, Moreno-Vega A, Beraud C, Shi M, et al. An FGFR3/MYC positive feedback loop provides new opportunities for targeted therapies in bladder cancers. EMBO Mol Med 2018:1–18. doi:10.15252/emmm.201708163.

[16] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47. doi:10.1093/nar/gkv007.

[17] Petryk N, Kahli M, D'Aubenton-Carafa Y, Jaszczyszyn Y, Shen Y, Silvain M, et al. Replication landscape of the human genome. Nat Commun 2016;7:1–13. doi:10.1038/ncomms10208.

[18] Wu X, Kabalane H, Kahli M, Petryk N, Laperrousaz B, Jaszczyszyn Y, et al. Developmental and cancer-associated plasticity of DNA replication preferentially targets GC-poor, lowly expressed and late-replicating regions. Nucleic Acids Res 2018:1–16. doi:10.1093/nar/gky797.

[19] Haradhvala NJ, Polak P, Stojanov P, Covington KR, Shinbrot E, Hess JM, et al. Mutational Strand Asymmetries in Cancer Genomes Reveal Mechanisms of DNA Damage and Repair. Cell 2016;164:538–49. doi:10.1016/j.cell.2015.12.050.

[20] Hoopes JI, Cortez LM, Mertz TM, Malc EP, Mieczkowski PA, Roberts SA. APOBEC3A and APOBEC3B Preferentially Deaminate the Lagging Strand Template during DNA Replication. Cell Rep 2016;14:1273–82. doi:10.1016/j.celrep.2016.01.021.

[21] Swanton C, McGranahan N, Starrett GJ, Harris RS. APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity. Cancer Discov 2015;5:704–12. doi:10.1158/2159-8290.CD-15-0344.

[22] Silvas T V., Hou S, Myint W, Nalivaika E, Somasundaran M, Kelch BA, et al. Substrate sequence selectivity of APOBEC3A implicates intra-DNA interactions. Sci Rep 2018;8:1–11. doi:10.1038/s41598-018-25881-z.

[23] Poulos RC, Wong YT, Ryan R, Pang H, Wong JWH. Analysis of 7,815 cancer exomes reveals associations between mutational processes and somatic driver mutations. PLoS Genet 2018;14:e1007779. doi:10.1371/journal.pgen.1007779.

Letter to Editor

Re: Ming-Jun Shi, Xiang-Yu Meng, Philippe Lamy, et al. APOBEC-mediated Mutagenesis as a Likely Cause of FGFR3 S249C Mutation Overrepresentation in Bladder Cancer

Alexander Yang[a], Vincent L. Cannataro[a], Jeffrey P. Townsend[a,b,c]*

a. Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA

b. Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT, USA

c. Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

*Corresponding author.

Department of Biostatistics, Yale School of Public Health,

135 College Street, New Haven, CT 06510, USA

Tel. +1 203 737 7042

E-mail address: jeffrey.townsend@yale.edu (J.P. Townsend).

Letter

Shi et al [1] cogently demonstrate that S249C is the most frequent among *FGFR3* mutations in bladder cancer, that tumors with the S249C variant tend toward higher APOBEC activity, and that - *in vitro* - the tumorigenicity of this mutation is comparable to others in *FGFR3*. However, there is a more direct test of the question their title poses: whether the over-representation of S249C compared to other recurrent mutations in *FGFR3* is a consequence of a higher benefit of S249C to the proliferation and survival of bladder cancer cell lineages, or is solely a consequence of a higher APOBEC-driven mutation rate. The two relevant forces are mutation (the higher the mutation rate, the higher the representation) and natural selection for the mutated cancer cell lineage (the more the variant increases proliferation and survival, the higher the representation observed [2]). Thus, their question can be restated as whether the cancer effect of S249C is greater than that of other recurrent *FGFR3* mutations? Cancer effect sizes can be calculated by estimating the *FGFR3* mutation rate using synonymous mutations and known covariates of the mutation rate [3], estimating the tumor-specific rate of each trinucleotide change [4], and comparing the expected recurrence based on mutation and neutral

drift to the recurrence observed [2]. S249C in The Cancer Genome Atlas BLCA data set has a mean mutation rate of 1.2×10^{-5} per cancer-competent somatic cell per development to tumor resection, the highest rate of the seven recurrent *FGFR* mutation sites (ranging from a lowest mutation rate of 8.0×10^{-7}). The cancer effect size of S249C is 6.9×10^{3}. The six other recurrent *FGFR3* mutation sites yield effect sizes ranging from 1×10^{3} to 7×10^{3}; the largest cancer effect size is that of S373C, with an effect higher than that of S249C.

These estimates critically depend on the accuracy of the mutation rate estimate. Shi et al [1] showed that the S249C mutation occurs within a 5-nt DNA hairpin, and could be subject to elevated APOBEC3A mutation rates unaccounted for by trinucleotide context alone. Is the cancer effect size of S249C even lower than 6.9×10^{3}? Apparently not; the relative mutability of sites within the loop of a DNA hairpin was recently quantified by Buisson et al [5] and the greatest APOBEC3A mutability is conferred to the 3'-most site within 4-nt loops. Shi et al [1] show that the S249C mutation occurs in the third position of a 5-nt loop. Buisson et al [5] quantified its substrate optimality as slightly less than one 1. Thus, if anything, its mutation rate is slightly lower than would be expected for a typical APOBEC3A site. Accordingly, the S249C mutation is strongly selected within bladder cancer lineages, consistent with our estimate. Nevertheless, the strength of selection is not beyond the range of other *FGFR3* mutations, despite the remarkably high recurrence of S249C. Quantitative analysis of cancer effect sizes allow rigorous testing of the importance of recurrent mutations within cancer driver genes.

References

[1] Shi M-J, Meng X-Y, Lamy P, et al. APOBEC-mediated mutagenesis as a likely cause of FGFR3 S249C mutation over-representation in bladder cancer. Eur Urol 2019;76:9–13.

[2] Cannataro VL, Gaffney SG, Townsend JP. Effect sizes of somatic mutations in cancer. J Natl Cancer Inst 2018;110:1171–7.

[3] Martincorena I, Raine KM, Gerstung M, et al. Universal patterns of selection in cancer and somatic tissues. Cell 2017;173:1823.

[4] Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome Biol 2016;17:31.

[5] Buisson R, Langenbucher A, Bowen D, Kwan EE, Benes CH, Zou L, et al. Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. Science 2019;364:eaaw2872.

Letter to Editor

Letter to the Editor: Reply to Yang et al. (corresponding author Dr. Townsend) EUROL-D-19-01062 Re: APOBEC-mediated mutagenesis as a likely cause of FGFR3-S249C mutation over-representation in bladder cancer

Ming-Jun Shi[a,b,i,1], Xiang-Yu Meng[a,b,c,1], Chun-Long Chen[d,e], Lars Dyrskjøt[f], François Radvanyi[a, †], Ludmila Prokunina-Olsson[g, †], Isabelle Bernard-Pierrot[a,†]

a. Institut Curie, CNRS, UMR144, Molecular Oncology team, PSL Research University, Paris, France

b. Paris-Sud University, Paris-Saclay University, Paris, France

c. Department of Urology, Zhongnan Hospital, Wuhan University, Wuhan, China

d. Institut Curie, CNRS, UMR3244, PSL Research University, Paris, France

e. Sorbonne Université, Paris, France

f. Department of Molecular Medicine, Aarhus University Hospital, Aarhus, Denmark

g. Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA

i. Department of Urology, Beijing Friendship Hospital, Capital Medical University, Beijing, China

1 These authors contributed equally to this work.

† Corresponding author.

Correspondence:

Isabelle Bernard-Pierrot

Institut Curie, CNRS, UMR144, Molecular Oncology team

26 Rue d'Ulm, 75005 Paris, France

TEL: +33 1 42 34 63 40

FAX: +33 1 42 34 63 49

Email: isabelle.bernard-pierrot@curie.fr

Letter

We thank Yang and colleagues for their comments that highlight an additional angle of the multifaceted question addressed by our report [1]. The authors rightly note that the frequency of somatic mutations is the result of two factors: the mutation rate at a specific position and selection intensity on the mutation leading to clonal cell expansion. The former partially depends on endo-/exogenous mutagens, such as APOBEC activity. To quantify the latter, the authors introduced the cancer effect size, which is a measure of a mutation's contribution to

tumour fitness, including its effect on tumorigenic potential, cell proliferation, and survival [2]. Since we found that *FGFR3*-S249C mutation has similar *in vitro* transformation ability compared to a less common, non-APOBEC related mutation *FGFR3*-Y375C, we proposed that, compared to other recurrent *FGFR3* mutations, *FGFR3*-S249C is over-represented in bladder tumors due to its increased mutation rate induced by APOBEC activity [1], although being a suboptimal APOBEC target [3].

Yang and colleagues quantified the cancer effect sizes of seven recurrent *FGFR3* mutations in The Cancer Genome Atlas bladder cancer dataset (TCGA-BLCA). Interestingly, they found that, despite its very high frequency, the cancer effect size of S249C is within the range of other recurrent *FGFR3* mutations. These results are consistent with our experimental observation on the similar transformation ability of S249C and Y375C [1], and extend to other *FGFR3* mutations including S373C which has a very low mutation rate of 1.5%, as compared to 62% for S249C. So, this example implies that rare mutations can have cancer effect sizes comparable to more prevalent mutations. This also suggests a more general question - why the other recurrent *FGFR3* mutations, such as Y375C and R248C, are more common than S373C, the mutation with the strongest calculated cancer effect size? Could it be due to increased mutation rates linked with specific mutagenic processes? To answer this question, we used a recently developed approach by Letouze et al. [4]. This approach is not restricted to the validation of association between a given mutation and signatures showing distinct trinucleotide motifs such as APOBEC signature as we applied initially [1] and enables estimating the probability of each mutation being due to each mutational process considering the mutation category (substitution type and trinucleotide context), and the number of mutations attributed to each process in the corresponding tumour. Applying this algorithm to the TCGA-BLCA dataset revealed a predominant association between S249C and APOBEC signature (29 samples of 31), in line with our report [1]. Interestingly, Y375C was strongly linked with ERCC2-associated signature (7/8) characterized by Kim et al. [5], while R248C - to age-associated signature (3/3). Of note, the paternal age effect has been linked to thanatophoric dysplasia in which R248C is the main recurrent *FGFR3* mutation. Our initial strategy has its own advantage regarding analysis of mutational signatures with distinct trinucleotide motifs and comparison with Letouze and colleagues' method could be an interesting future topic.

Integrating different methods should allow to evaluate the functional importance and a better understanding of etiology of cancer driver mutations through the identification of causal mutational processes.

References

[1] Shi MJ, Meng XY, Lamy P, Banday AR, Yang J, Moreno-Vega A, et al. APOBEC-mediated Mutagenesis as a Likely Cause of FGFR3 S249C Mutation Over-representation in Bladder Cancer. Eur Urol 2019;76:9–13. doi:10.1016/j.eururo.2019.03.032.

[2] Cannataro VL, Gaffney SG, Townsend JP. Effect sizes of somatic mutations in cancer. J Natl Cancer Inst 2018;110:1171–7. doi:10.1093/jnci/djy168.

[3] Buisson R, Langenbucher A, Bowen D, Kwan EE, Benes CH, Zou L, et al. Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. Science (80-) 2019;364:eaaw2872. doi:10.1126/science.aaw2872.

[4] Letouzé E, Shinde J, Renault V, Couchy G, Blanc JF, Tubacher E, et al. Mutational signatures reveal the dynamic interplay of risk factors and cellular processes during liver tumorigenesis. Nat Commun 2017;8:1315. doi:10.1038/s41467-017-01358-x.

[5] Kim J, Mouw KW, Polak P, Braunstein LZ, Kamburov A, Tiao G, et al. Somatic ERCC2 mutations are associated with a distinct genomic signature in urothelial tumors. Nat Genet 2016;48:600–6. doi:10.1038/ng.3557.

Manuscript 2

APOBEC induces over-represented driver but also passenger hotspot mutations within its target genes in bladder cancer

Ming-Jun Shi[1,2,3], Xiang-Yu Meng[1,2,4*], Chun-Long Chen[5,6], François Radvanyi[1,#] and Isabelle Bernard-Pierrot[1,#]*

1. Institut Curie, CNRS, UMR144, Molecular Oncology team, PSL Research University, Paris, France

2. Paris-Sud University, Paris-Saclay University, Paris, France

3. Department of Urology, Beijing Friendship Hospital, Capital Medical University, Beijing, China

4. Department of Urology, Zhongnan hospital of Wuhan University, Wuhan, China

5. Institut Curie, CNRS, UMR3244, PSL Research University, Paris, France

6. Sorbonne Université, Paris, France

* These authors contributed equally.

# These authors contributed equally.

Correspondence:

Xiang-Yu Meng

Email: mengxy_whu@163.com

Isabelle Bernard-Pierrot
Institut Curie, CNRS, UMR144, Molecular Oncology team
26 Rue d'Ulm, 75005 Paris, France
TEL: +33 1 42 34 63 40
FAX: +33 1 42 34 63 49

Email: isabelle.bernard-pierrot@curie.fr

Abstract

Background: APOBEC is the main contributor to mutagenesis in bladder cancer (BCa) accounting for more than half of the mutations. Mutagenesis driven by APOBEC and functional positive selection of cancer cells may synergically drive over-representation of hotspot driver mutations as we reported for *FGFR3* S249C. However, yet only few APOBEC-related driver hotspot mutations have been identified.

Objective: To systematically identify APOBEC-associated hotspot mutations and study the nature of these mutations in BCa.

Design, setting, and participants: We analysed 602 exome-sequenced BCa samples for part of which gene expression data were also available.

Outcome measurements and statistical analysis: APOBEC-related hotspot mutations were identified by motif-mapping, mutation-signature fitting and APOBEC-mediated mutagenesis comparison. Similarity-based joint analysis of DNA-hairpin stability and gene expression was performed to predict driver/passenger hotspot mutations. Difference in APOBEC-mediated mutagenesis between samples with and without APOBEC-related hotspot mutations were investigated.

Results and limitations: APOBEC-related hotspot mutations mostly occurred in the loop of DNA-hairpin structures and included known driver but also likely passenger mutations. They were systematically over-represented compared to other mutations within APOBEC-target genes, independently of their functional impact, in BCa and also in other cancer types with high APOBEC activity. A combination analysis of loop stability and gene expression allowed to distinguish known passenger from driver mutations including loss-of-function mutations affecting tumour suppressor genes, and to predict new driver and passenger mutations. The panel of APOBEC-related hotspot mutations enabled to select BCa patients with high APOBEC-mediated mutagenesis that could benefit from immunotherapies as well as from future anti-APOBEC strategies.

Conclusions: Our study provides a list of APOBEC-associated driver hotspot mutations in BCa, but also challenges the dogma that recurrent mutations are driver and mostly gain-of-function mutations affecting oncogenes. It sheds new lights on the discovery of new driver mutations from analysis of high-throughput sequencing data.

Patient summary: APOBEC is the dominant mutagenic process in BCa. We identified hotspot mutations induced by APOBEC and distinguished the non-functional ones from those

favouring cancer cell expansion. This work could lead to a better understanding of bladder cancer biology and aetiology and provide new therapeutic targets. The set of APOBEC-associated hotspot mutations could allow selecting patients for emerging treatments.

Keywords: Bladder cancer, APOBEC, mutagenesis, stem-loop, driver mutation, passenger mutation, immunotherapy

Introduction

Bladder cancer (BCa) has a very high overall mutation load and the number of driver mutations per BCa sample is also ranked in the top 4 across various cancer types [1,2]. Identifying such driver mutations helps a better understanding of cancer biology and can provide new therapeutic targets. Driver hotspot mutations do not accumulate randomly but occur at specific sites as a result of two factors: the mutation rate at this position and the functional advantage given by this mutation leading to clonal cell expansion. The mutation rate is partially impacted by endo-/exogenous mutagenic processes leaving characteristic fingerprints on the cancer genome in a DNA sequence context-dependent manner, such as APOBEC (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like)-mediated mutagenesis that is related to APOBEC deaminase activity [3,4]. Pan-cancer analysis has revealed that APOBEC-mediated mutagenesis makes a significant contribution to the overall mutations in several cancer types, particularly in BCa being the cancer with the second highest abundance of APOBEC-induced mutations after cervical cancer [3]. We could therefore assume that APOBEC contributed to the selection of driver hotspot mutations in BCa. In line with this hypothesis, we recently showed that the frequent mutation targeting the fibroblast growth factor receptor gene, *FGFR3*-S249C, is over-represented in BCa as compared to other *FGFR3* recurrent activating mutations due to increased mutation rate induced by APOBEC rather than increased tumorigenicity of this mutation [5]. However, only few APOBEC-related driver hotspot mutations have been identified in BCa [5–9].

Here, we systematically looked for APOBEC-related hotspot mutations in BCa. We found that the over-representation of APOBEC-related mutations was common in APOBEC-target genes and could also be associated not only with driver mutations but also with passenger mutations. We investigated the DNA replication fork directionality (RFD) and the DNA-hairpin structures of these hotspot mutations, given their link with APOBEC enzyme accessibility as previously reported [10–15]. We proposed a model to predict new driver and

passenger within APOBEC-related hotspot mutations. Finally, we explored the potential clinical implication based on this targeted sequencing panel with APOBEC-related mutations.

Materials and methods

Whole exome sequencing (WES) data of 602 BCa and 3,751 other APOBEC-related cancer (cervical, head and neck, breast and lung cancer) samples were downloaded from cBioPortal database. Only single-nucleotide variants (SNVs) were considered. Mutations corresponding to 5'-TCN motifs (TCN → T[G/T]N, N = any nucleotide) were considered as APOBEC-type mutations. Count distribution of SNVs (overall and APOBEC-type mutations) was analysed and an optimal threshold was determined to define hotspot mutations. For each patient, the arithmetic sum of signature fraction scores for COSMIC (Catalogue Of Somatic Mutations In Cancer) signature 2 and 13 was defined as the parameter to evaluate APOBEC-mediated mutagenesis. To determine APOBEC-associated hotspot mutations, comparison using Wilcoxon rank sum test was made between subjects with a given candidate mutation (APOBEC-type mutation with count ≥ 4) and with none candidate mutation.

In BCa, probability of clonal events were compared between APOBEC-related hotspot mutations and all other mutations of APOBEC-target genes, using Fisher's test and generalised linear mixed model (GLMM). For APOBEC-related hotspot mutations, the probability of locating in lagging-strand template and loop of DNA-hairpin was calculated (Supplementary Methods) and compared against random (i.e. 0.5), using Wilcoxon signed rank test and z test (rate logit-transformed), respectively. Normalised gene expression rank was calculated and compared between APOBEC-target genes of known and unknown functional importance. Normalised stem-loop stability score was calculated and compared between known drivers and passengers. Similarity-based joint analysis of stem-loop stability and gene expression was performed for driver/passenger prediction. The capacity of these mutations for selecting patients with high APOBEC-mediated mutagenesis was investigated in the primary BCa dataset (n = 602), and in additional 27 exome-sequenced metastatic BCa tumours and 39 BCa-derived cell lines for validation.

A $P < 0.05$ was considered statistically significant. R version 3.5.2, the sigfit version 1.3.1 and the ggpubr version 0.2 package were used for relevant analyses and visualisation.

A detailed version of materials and methods is provided in the Supplementary Data.

Results

Identification of APOBEC-associated hotspot mutations in BCa

More than half of the SNVs in 602 BCa patients were APOBEC-type mutations and their frequency followed a long-tail, or more precisely, a power-law distribution (Fig. 1A upper panel, Fig. S1 and Method). We determined that a frequency ≥ 4 was an optimal threshold to distinguish the 'head' (as hotspot mutations) from the 'tail' within the distribution and to define candidate APOBEC-associated hotspot mutations (n = 59) (Fig. 1A lower panel, Supplementary Table 1, Method). Assuming that tumours with a genuine APOBEC-associated mutation should present significantly higher APOBEC-mediated mutagenesis than tumours without any of the 59 candidate mutations, we classified 44 mutations (mapping to 33 genes) as APOBEC-associated ones (Fig. 1B, Method). These mutations included all the five previously identified APOBEC-associated mutations in BCa (*PIK3CA* E545K and E542K, *FGFR3* S249C, *ERBB2* S310F and *TP53* R280T) [5–9]. Consistently, tumours with larger numbers of APOBEC-related mutations presented higher APOBEC-mediated mutagenesis (Fig. 1C).

APOBEC-related hotspot mutations are located in DNA lagging-strand template and in loop of DNA-hairpin structures

In addition to motif specificity, APOBEC enzymes also preferably target single-strand DNA (ssDNA) that confers spatial accessibility and APOBEC-related mutations are dominated by replicative but not transcriptional mutational asymmetries [10–15]. Accordingly, we did not observe coding strand bias within 44 mutations classified (Supplementary Table 1), and their likelihoods being induced by APOBEC were supported by the observation that most (42/44, except for the *ERBB4* E317K and *SF3B1* E902K mutations) more likely occurred in lagging-strand templates during DNA replication (Fig. 1D, across 9 cell lines [16,17], median probability = 0.73, P = 0.003, Method) and/or were located within the loop of DNA-hairpin structures (Fig. 1E, 39/44, probability = 0.89, P = 7.6 × 10-6, Method) against random, i.e. 0.5. Representative examples of genome-wide RFD and DNA-hairpin are shown for the *ERBB2* S310F mutation (Fig. 1F-G) and details for other mutations are summarised in Fig. S2 and Supplementary Table 1.

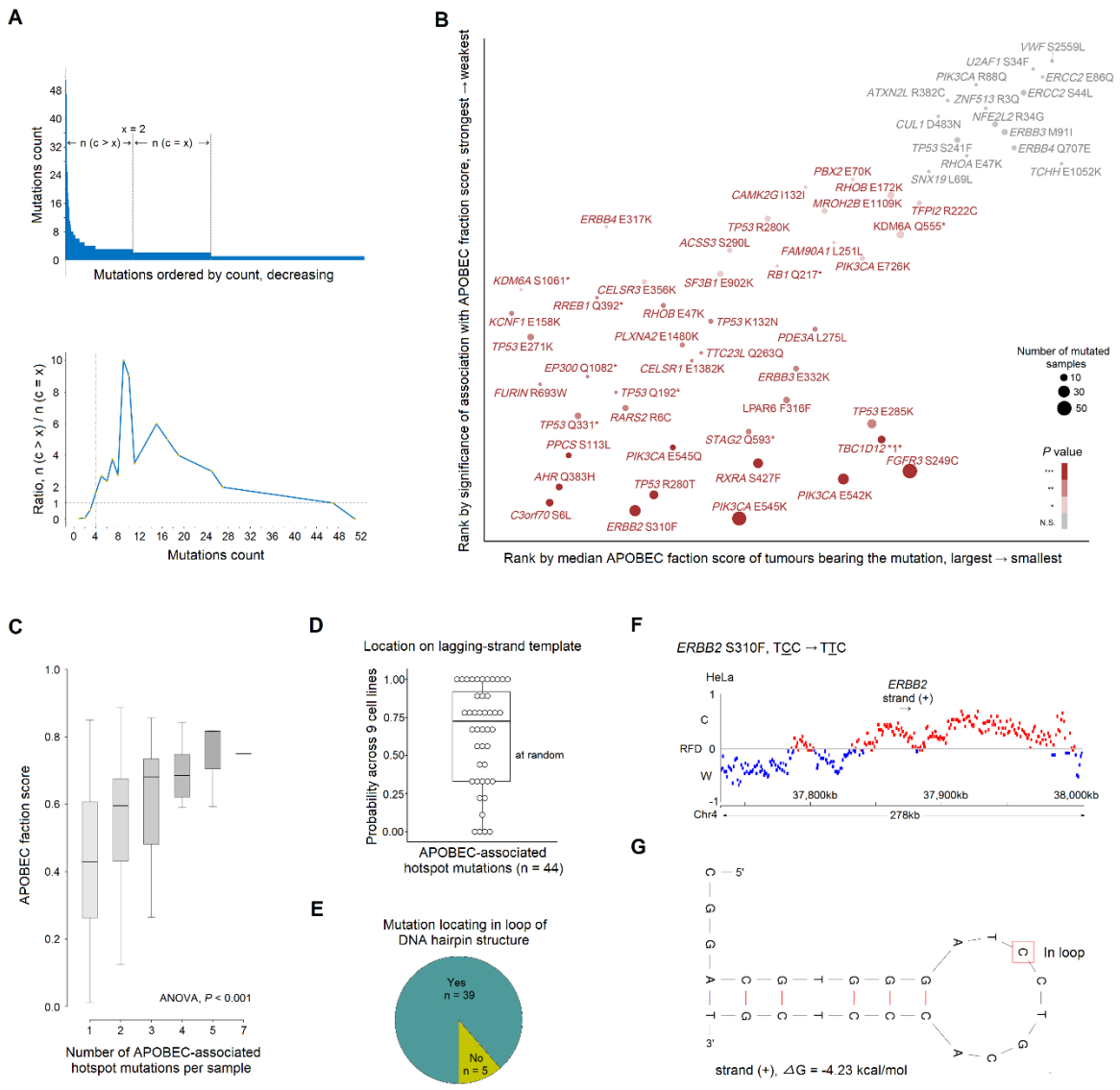


Figure 1. Identification of APOBEC-associated hotspot mutations in bladder cancer and characterisation of their spatial features for the accessibility of APOBEC enzymes. A) *The frequency of APOBEC-type mutations (TCN → T[G/T]N) in 602 bladder cancers follows a long-tail distribution (upper panel). Ratios of the number of mutations with a count larger than this integer (n(c>x)) to the number of mutations with a count equal to this integer (n(c=x)) (lower panel).* ***B)*** *Association between 59 candidate APOBEC-related hotspot mutations and APOBEC-mediated mutagenesis of tumour samples. APOBEC fraction score, the sum of signature fraction scores of COSMIC signature 2 and 13 [4]. P value: Wilcoxon's test; scale in red for significant associations (n = 44): *, < 0.05; **, < 0.01, ***, < 0.001; grey for non-significant (N.S.) associations (n = 15). Circle size for mutation frequency.* ***C)*** *Correlation between the number of APOBEC-associated hotspot mutations per sample and APOBEC-*

mediated mutagenesis. P value: ANOVA test. ***D)*** *Probability of locating on lagging-strand-template for 44 APOBEC-associated hotspot mutations across nine cancer/normal cell lines [16,17].* ***E)*** *Status of locating in loop of DNA hairpin structure (25nt ssDNA centred on mutated site) for 44 APOBEC-associated hotspot mutations.* ***F)*** *Representative replication-fork-directionality (RFD) around the ERBB2 gene in HeLa cells, as determined by mapping Okazaki fragments to C (Crick) and W (Watson) DNA strands. Red (blue) RFD profiles mark indicate the regions in which Watson (Crick) strands are replicated mostly as lagging-strand-templates. Black arrow under gene symbol for transcriptional direction.* ***G)*** *Representative predicted stem-loop structure for the ERBB2 gene. Red rectangle marks mutation. Free energy parameter – ΔG (kcal/mol) for loop stability. Strand (+) indicates cytosine (C) mutation, whereas strand (-) indicates guanine (G) mutation.*

APOBEC-mediated mutagenesis shapes landscape of hotspot mutations in BCa and mutation spectra in its target genes

Given the considerable proportion of mutations induced by APOBEC-mediated mutagenesis in BCa [18,19], we considered the possibility of this process acting as a major source of hotspot mutations in BCa. Similar to the pan-cancer genome-wide mutation profile [20], the frequency of all SNVs from aforementioned WES data in BCa followed a long-tail/power-law distribution (Fig. S3, Method). The optimal frequency threshold was again ≥ 4 (Fig. S3, Method) and we identified so 130 hotspot mutations mapping to 75 genes (Supplementary Tables 1&2). APOBEC-related mutation events accounted for almost half the total number of hotspot mutation events (Fig. 2A), and their frequency was significantly higher than that of all other hotspot mutations (Fig. 2B), suggesting that APOBEC is the leading source of hotspot mutations in BCa. Strikingly, as previously observed for *FGFR3* S249C, almost all the APOBEC-associated hotspot mutations (30/33, except for three mutations in *TP53*, *ERBB3* and *ERBB4*) were also significantly over-represented relative to other mutations within the same gene (Fig. 2C), suggesting that APOBEC-mediated mutagenesis shapes the landscape of mutations in its target genes in BCa. Representative examples of mutation spectra are shown for the *ERBB2* and *KDM6A* genes (Fig. 2D) and others are summarised in Fig. S2. In addition, within all mutations mapped to the 33 genes targeted by APOBEC, APOBEC-associated hotspot mutations showed higher probability of clonal event than the other ones (Fig.2E, Method), indicating APOBEC-mediated mutagenesis is an early event in BCa tumorigenesis as previously reported [21].

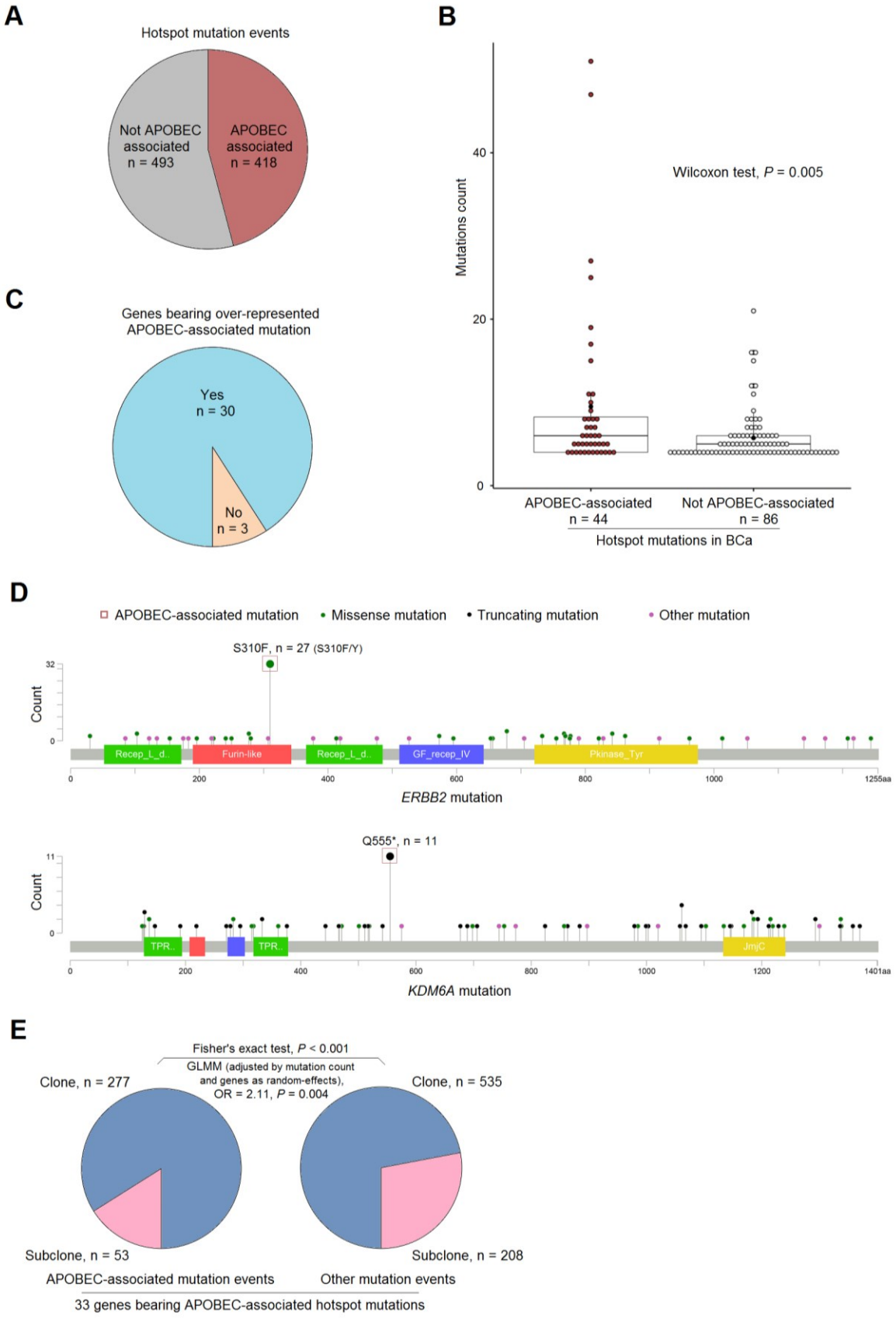


Figure 2. APOBEC-mediated mutagenesis shapes the landscape of hotspot mutations in bladder cancer (BCa) and the mutation spectra for its target genes. ***A)*** *Proportion of*

APOBEC-associated hotspot mutation events in tumours bearing 130 hotspot mutations (counts ≥ 4) in BCa. ***B)*** *Comparison of mutation frequencies between APOBEC-associated hotspot mutations and other hotspot mutations. P value: Wilcoxon test.* ***C)*** *Proportion of APOBEC-associated hotspot mutations over-represented within the genes (n = 33) targeted by APOBEC.* ***D)*** *Representative mutation spectra for the ERBB2 and KDM6A gene in 602 BCa. Red rectangles indicate APOBEC-associated hotspot mutations; green dots mark missense mutations; black dots mark truncating mutations; pink dots mark other mutations. ERBB2 S310F and KDM6A Q555* are APOBEC-related hotspot mutations and strongly over-represented relative to other mutations within their gene size.* ***E)*** *Distribution of clone and subclone events for APOBEC-related hotspot and other mutations of genes targeted by APOBEC (n = 33). P value: Fisher's exact test. GLMM, generalized linear mixed model, with mutation count as covariate and genes as random effects; OR, odds ratio.*

Identification of APOBEC-associated hotspot mutations in other APOBEC-related cancer types

We investigated whether these features of APOBEC-associated hotspot mutations could be generalised to other cancer types presenting relatively high APOBEC-mediated mutagenesis (cervical, head and neck, breast and lung cancer) [3]. Pooling together all the APOBEC-type mutations from these four cancer types (from 3,751 patients), we observed that the frequency of these SNVs also followed a long-tail/power-law distribution, and a frequency count ≥ 4 was again identified as the optimal threshold for the identification of candidate APOBEC-associated hotspot mutations (n = 112) (Fig. S4, Supplementary Table 3 and Method). Using this threshold, we considered the 78 mutations (mapping to 55 genes) that were significantly associated with higher APOBEC-mediated mutagenesis as APOBEC-associated hotspot mutations (Fig. 3A, Supplementary Table 3 and Method). As in BCa, these mutations were more likely to occur in lagging-strand templates (median probability = 0.78, $P = 2.0 \times 10^{-6}$, Method) and/or within loop structures (55/78, probability = 0.71, P = 0.0002, Method) respectively against random, increasing the likelihood of being induced by APOBEC enzymes (Supplementary Table 3, Method). Although APOBEC-mediated mutagenesis also contributed significantly to hotspot mutations in these cancer types, it was less common than in BCa, highlighting the particular importance of APOBEC in BCa (Fig. 3B). However, as in BCa, in 93% cases, APOBEC-associated hotspot mutations were significantly over-represented relative to other mutations within the same gene targeted by APOBEC (Fig. 3C and Fig. S5).

Finally, even though some tumour specificity was observed, 32% of the APOBEC-related hotspot mutations identified in BCa were also found in other cancer types (Fig. 3D).

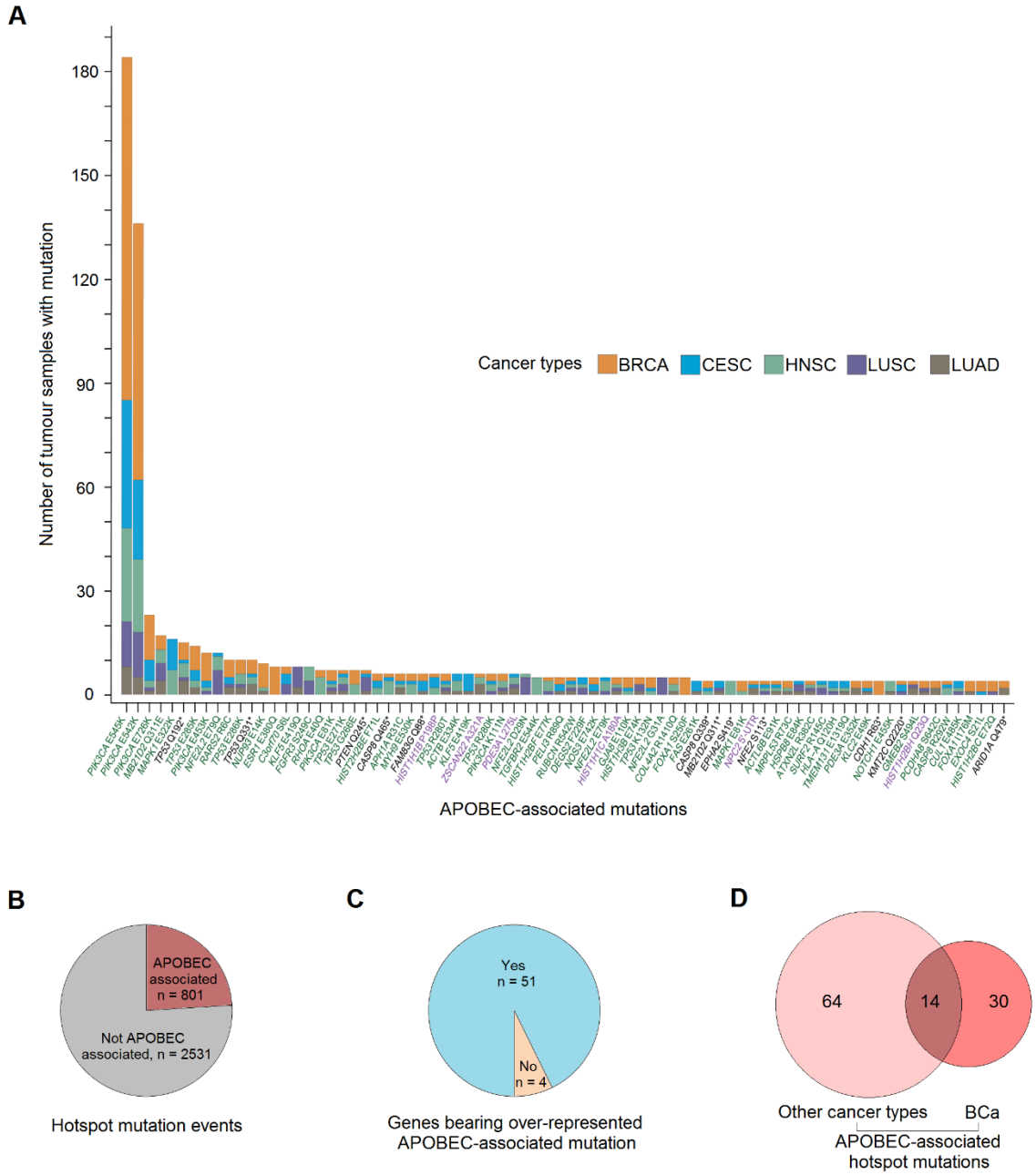


Figure 3. Identification of APOBEC-associated hotspot mutations in other cancer types presenting relatively high APOBEC-mediated mutagenesis, and comparison with those observed in bladder cancer (BCa). A) *Distribution of the frequencies of 78 significantly APOBEC-associated hotspot mutations in 3,744 tumours from cervical, head and neck, breast and lung cancer.* ***B)*** *Proportion of APOBEC-associated hotspot mutation events in tumours bearing 344 hotspot mutations (counts ≥ 4) identified in four cancer types.* ***C)*** *Proportion of APOBEC-associated hotspot mutations over-represented within the genes (n = 55) targeted by*

APOBEC. ***D)*** *Intersection between APOBEC-associated hotspot mutations identified in BCa and those identified in other cancer types. BRCA = breast cancer; CESC = cervical squamous cell carcinoma; HNSC = head and neck squamous cell carcinoma; LUAD = lung adenocarcinoma; LUSC = lung squamous cell carcinoma and BCa = bladder cancer.*

Functional characteristics of APOBEC-associated hotspot mutations

It is generally assumed that hotspot mutations are most likely to be gain-of-function mutations affecting oncogenes and that loss-of-function mutations affecting Tumour Suppressor Genes (TSGs) are non-recurrent, with the exception of dominant-negative mutations [22,23]. We investigated whether this was the case for APOBEC-related hotspot mutations. We focused on the 44 APOBEC-associated hotspot mutations in BCa, taking advantage of the findings of a recent comprehensive study [24] in which 299 cancer genes and 579 driver mutations were functionally annotated, for characterisation of the mutations (driver or passenger) and/or the genes affected (TSG or proto-oncogene) (Supplementary Table 1). As expected, we observed gain-of-function driver mutations (n = 4) affecting proto-oncogenes, such as *FGFR3* S249C and *ERBB2* S310F, and mutations with undetermined function (n = 7) but mapping to known oncogenes, such as *RXRA* S427F and *RHOB* E172K. Interestingly, we observed nonsense hotspot mutations mapping to five TSGs (*TP53*, *KDM6A*, *STAG2*, *EP300* and *RB1*) that are not likely to display dominant-negative activity since the mRNA levels of these TSGs were significantly lower in tumours bearing APOBEC-associated nonsense mutations than in non-mutated tumours (Fig. S6A) and two TSGs were even located in X chromosome. We also identified seven hotspot mutations that were very likely to be passengers, including five silent mutations, one mutation affecting the transit peptide (*RARS2* R6C), and another mutation of the *MROH2B* gene showing an absence of mRNA in BCa (Fig. 4A, Supplementary Table 1 and Method).

Prediction for new APOBEC-related driver and passenger hotspot mutations

To distinguish APOBEC-associated driver from passenger hotspot mutations in BCa, we evaluated the stability of DNA-hairpin structures (estimated with free energy parameter – ΔG [25]) for these APOBEC-associated hotspot mutations and mRNA expression level in cancer samples of the genes bearing these mutations. We showed that known oncogenes/TSGs were quite systematically highly expressed in tumours whereas expression levels were very heterogeneous for genes of unknown function (Fig. 4B and Method). We also showed that

known passenger mutations were more significantly located in more stable loops than known

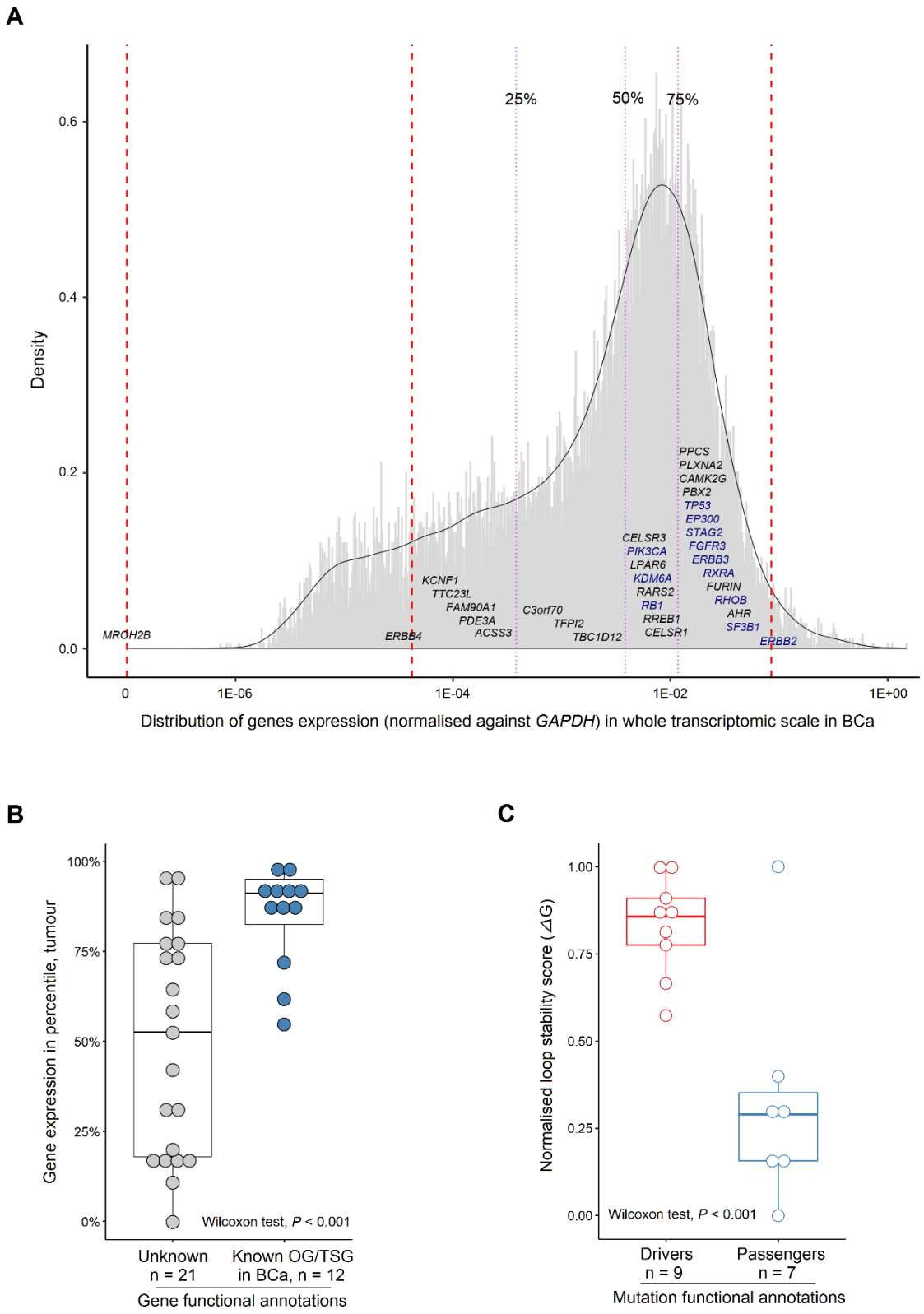


Figure 4. mRNA expression level of APOBEC-target genes and DNA loop stability of APOBEC-associated known driver and passenger mutations in bladder cancer (BCa). A)

Distribution of mRNA levels (normalised against GAPDH) for APOBEC-target genes at the whole-transcriptome scale in BCa from The Cancer Genome Atlas (TCGA) [18]. The percentile ranks of genes were converted into quartiles. Blue and black font indicate known OG/TSGs and other genes, respectively. ***B)*** *Distribution of the expression ranks in percentile in BCa between APOBEC-target known OG/TSGs and genes with unknown function.* ***C)*** *Distribution of the loop stability between APOBEC-associated known driver and passenger mutations. Higher normalised Δ G (kcal/mol) scores reflect lower loop stability. Passengers include 5 silent mutations, 1 mutation within transit peptide and 1 missense mutation on gene with an absence of mRNA expression. The functional annotation for mutations and genes are curated from a recent publication [24].* ***B-C)*** *P value: Wilcoxon test. OG, oncogene; TSG, tumour suppressor gene.*

driver mutations (Fig. 4C and Method). We then combined these two parameters to predict the nature of unknown mutations. We used an iterative approach based on similarity (measured with Euclidean distance) with known driver and passenger mutations (Fig. 5, Supplementary Table 1 and Method). As a result, we were able to classify 23 of 28 unknown mutations: 17 as drivers and 6 as passengers (Fig. 5). Most of the mutations with an unknown functional impact based on a large-scale functional annotation study of mutations [24], but affecting known oncogenes such as *RXRA* S427F, or TSGs such as *STAG2* Q593*, were predicted to be driver mutations. Supporting our prediction model, *RXRA* S427F mutation was recently demonstrated to induce ligand-independent activation of PPARG/RXR pathway and to display pro-tumorigenic activity in bladder cancer [26]. We also predicted some driver mutations in genes that were not reported to display oncogenic or tumour suppressive properties in BCa such as *AHR* Q383H and *RREB1* Q392*, and some passenger mutations such as *C3orf70* S6L. We suspected that *RREB1* Q392* might be a hotspot loss-of-function mutation and the gene itself a TSG in BCa, given that samples with the *RREB1* Q392* mutation displayed significantly lower levels of expression for this gene than non-mutated tumours (Fig. S6B).

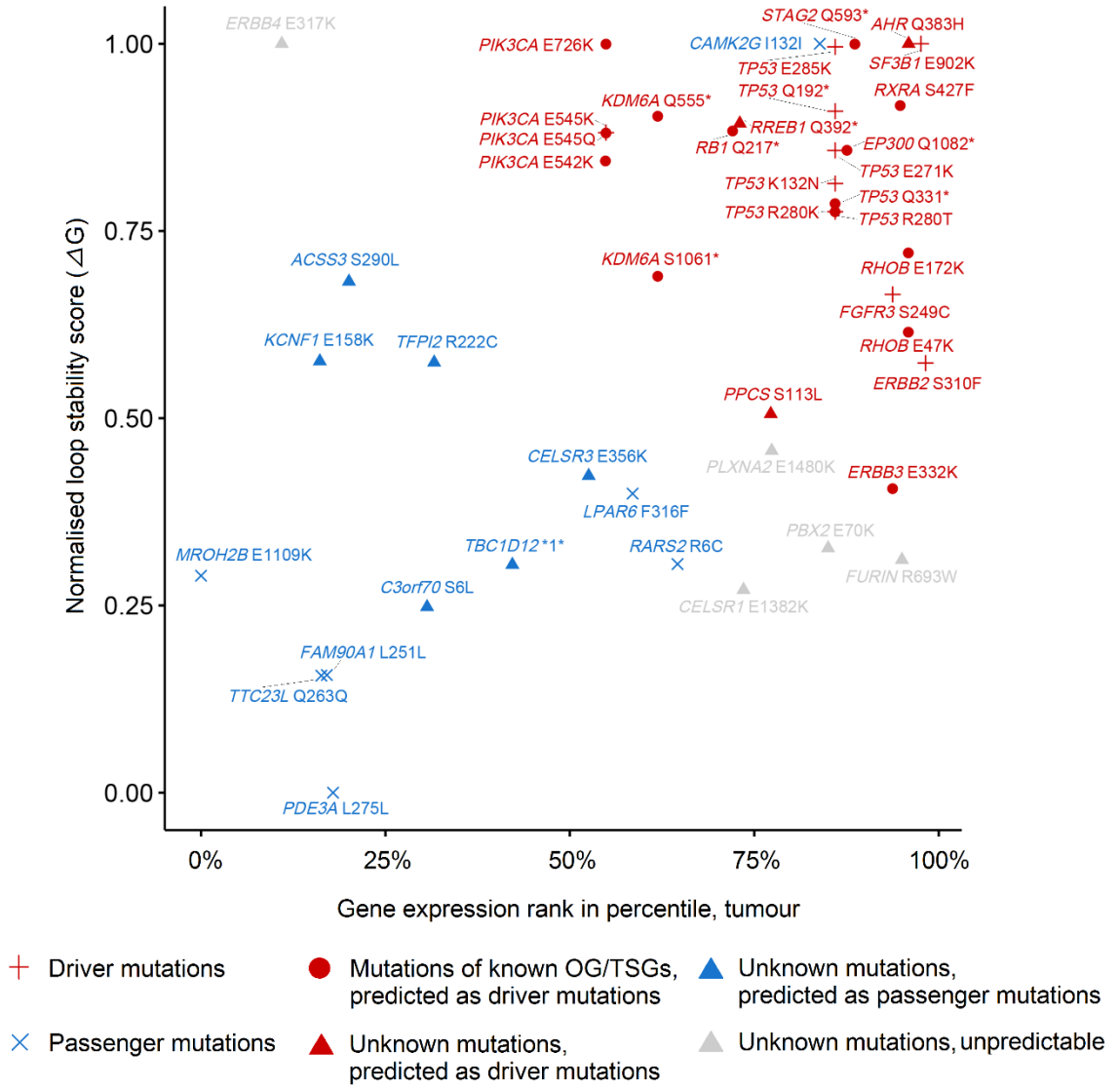


Figure 5. Prediction for new 'drivers' and 'hotspot passengers' with combination analysis of DNA loop stability and mRNA expression levels for APOBEC-target genes in bladder cancer (BCa). *Higher normalised Δ G (kcal/mol) scores reflect lower loop stability. Gene expression level (normalised against GAPDH gene) was presented by their rank at the whole-transcriptome scale in tumours samples. OG, oncogene; TSG, tumour suppressor gene. Passengers include 5 silent mutations, 1 mutation within transit peptide and 1 missense mutation on gene with an absence of mRNA expression. The functional annotation for mutations and genes are curated from a recent publication [24].*

Clinical implication potency of the set of APOBEC-related hotspot mutations

Tumours with a high APOBEC-mediated mutagenesis were recently shown to present better responses to immunotherapy than other tumours in bladder and lung cancer [27,28], highlighting the therapeutic value of identifying such tumours. We showed that patients bearing

any of the 44 APOBEC-associated hotspot mutations in BCa presented a significantly higher APOBEC-mediated mutagenesis than the other patients in two independent cohorts of 602 and 27 BCa (Fig. 6A-B) and in 39 BCa-derived cell lines (Fig. 6C and Methods). A targeted sequencing approach of 44 APOBEC-related hotspot mutations should be faster and cheaper than exome-sequencing for the selection of patients with high levels of APOBEC-mediated mutagenesis, who may be more likely to benefit from immunotherapy and, in the future, from anti-APOBEC treatment [29].

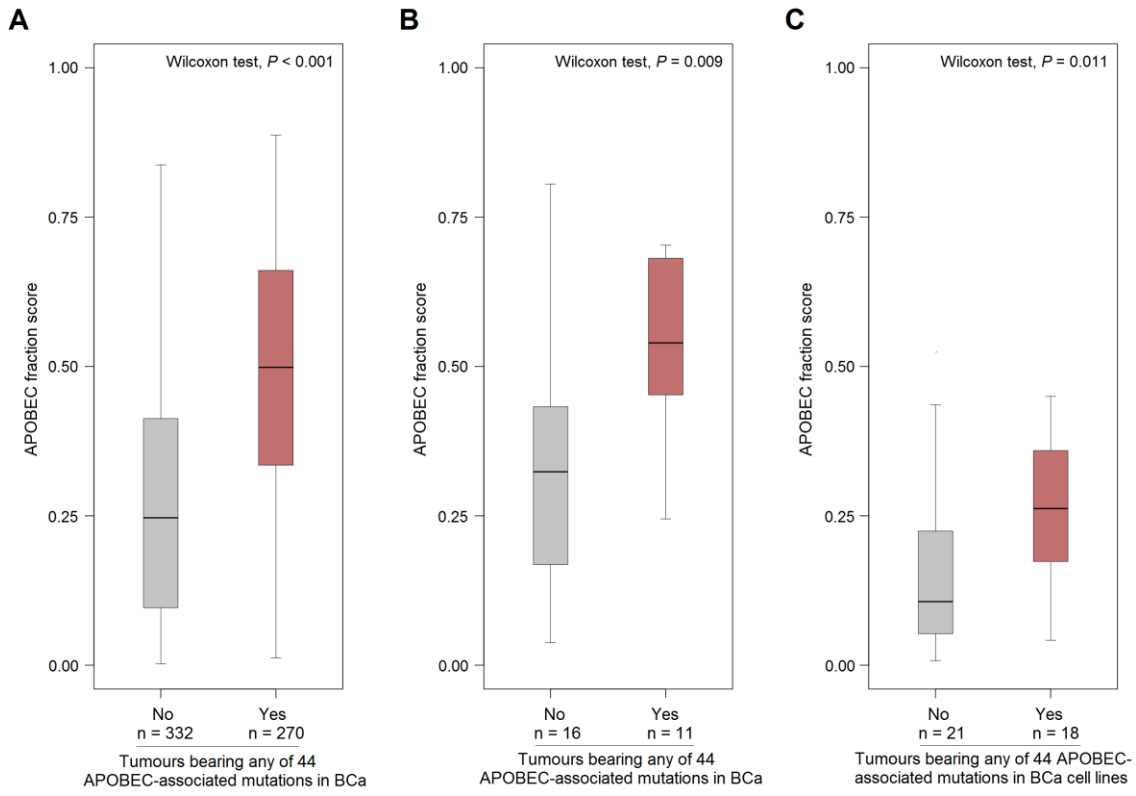


Figure 6. Ability of 44 APOBEC-associated hotspot mutations to predict enrichment of APOBEC-mediated mutagenesis A) *in 602 bladder cancer (BCa) cohort of investigation,* ***B)*** *in an additional 27 BCa cohort with available whole-exome sequencing (WES) data* ***C)*** *in 39 BCa-derived cell lines with available WES data. APOBEC fraction score, the sum of signature fraction scores of Catalogue Of Somatic Mutations In Cancer (COSMIC) signature 2 and 13 which have been demonstrated as associated with APOBEC activity [3,4]. P value: Wilcoxon test.*

Discussion

Here, we identified 44 APOBEC-associated hotspot mutations in BCa, a much larger number than in previous studies aiming to identify APOBEC-associated mutations [5–9]. Our

strategy for association validation was original in that we compared APOBEC-mediated mutagenesis between tumours bearing a given candidate hotspot mutation corresponding to APOBEC motifs and tumours without any of such candidate mutations. This comparative strategy is more adapted to systematically identify APOBEC-related mutations than previous studies where comparing tumours harbouring a candidate APOBEC-related mutation with tumours bearing other recurrent mutations [5] within the same target gene or wild-type samples [6–9] regarding APOBEC-mediated mutagenesis.

The identified APOBEC-induced mutations were almost systematically over-represented within APOBEC-target genes in not only BCa but also other APOBEC-related cancer types studied. Despite being hotspot mutations, not all of them were gain-of-function affecting oncogenes. We also found hotspot loss-of-function nonsense mutations affecting TSGs without obvious dominant-negative properties but also passenger mutations. It implies that synergistic selection between functional advantage and mutagenic process may account for APOBEC-related driver mutations whereas mutagenic process alone drives APOBEC-related passenger hotspot mutations. The fact that APOBEC-mediated mutagenesis seems to be an early event during BCa tumorigenesis could additionally contribute to this phenomenon. Altogether, our findings allow a better understanding of BCa biology and aetiology and challenge the dogma that recurrent mutations are much more likely to be driver and mostly gain-of-function mutations and should lead to caution in the way of identifying driver mutations from the analysis of high-throughput sequencing data.

Considering an optimal ssDNA sequence (25nt length) centred on the mutated nucleotide for APOBEC-related hotspot mutations, we noticed that only a few APOBEC-related mutations were not located within a loop of DNA-hairpins. However, the stability of the loop, measured with a simple and easily accessible parameter (free energy – ΔG, kcal/mol) [25] calculated by Mfold tool, differed for likely passengers and known drivers mutations. The former being mostly located in a very stable loop whereas the latter in less stable loops. By a combination analysis of this loop stability parameter and mRNA expression level of genes bearing these mutations, we were able to better distinguish APOBEC-related driver from passenger hotspot mutations and highlighted 17 new potential driver mutations that will worth to be further validated by functional studies and could be new therapeutic targets in BCa. During the submission process of our manuscript, Lawrence and colleagues also proposed that APOBEC3A-associated mesoscale genomic features could lead to hotspot passenger mutations

[30]. Notably, both our studies based on different approaches highlighted common new driver mutations and identified passenger mutations associated with APOBEC.

We proposed that the targeted panel of APOBEC-related hotspot mutations allows identifying patient with high APOBEC activity and that it may be a way to select eligible patients for treatment with immune-checkpoint inhibitors. However, a limitation of this study is the lack of drug response data for BCa samples treated with immunotherapies. Consequently, we were not able to detect the power of this targeted sequencing panel to predict immunotherapy efficacy and it should be investigated in future studies.

Conclusion

Our study, in addition to proposing new APOBEC-induced driver mutations in BCa and contributing so to a better understanding of BCa biology and etiology, highlights a general feature of APOBEC-mediated mutagenesis, namely the over-representation of APOBEC-related hotspot mutations within APOBEC-target genes. Our work shows that APOBEC can also favor passenger hotspot mutations locating in optimal DNA loops. It therefore challenges the dogma that all recurrent mutations are likely drivers.

Author contributions: Isabelle Bernard-Pierrot had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Shi, Meng, Bernard-Pierrot, Radvanyi.
Acquisition of data: Shi, Meng.
Analysis and interpretation of data: Shi, Meng, Chen, Bernard-Pierrot, Radvanyi.
Drafting of the manuscript: Shi, Meng, Bernard-Pierrot.
Critical revision of the manuscript for important intellectual content: Bernard-Pierrot, Chen, Radvanyi.
Statistical analysis: Meng, Shi.
Obtaining funding: Bernard-Pierrot, Radvanyi.
Administrative, technical, or material support: None.
Supervision: Bernard-Pierrot, Radvanyi.
Other (specify): None.

Financial disclosures: Isabelle Bernard-Pierrot certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

Funding/Support and role of the sponsor: This work was supported by a grant from *Ligue Nationale Contre le Cancer* (IBP, FR, MJS, XYM) as an associated team (*Equipe labellisée*). MJS was supported by a scholarship from *China Scholarship Council*. XYM was supported by a fellowship from ITMO Cancer AVIESAN within the framework of Cancer Plan.

Acknowledgment statement: We thank The Cancer Genome Atlas (TCGA) Network for generating the RNA-seq and somatic mutation data and for providing open access, and we thank the COSMIC Signatures of Mutational Processes in Human Cancer working group for the collection of established mutation signature catalogue. We would like to thank Clarice Groeneveld and Jacqueline Fontugne for their help for manuscript improvement.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at xxxxx.

References

[1] Lawrence MS, Stojanov P, Polak P, Kryukov G V, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 2013;499:214–8. doi:10.1038/nature12213.

[2] Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P, et al. Universal Patterns of Selection in Cancer and Somatic Tissues. Cell 2017;171:1029-1041.e21. doi:10.1016/j.cell.2017.09.042.

[3] Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat Genet 2013;45:970–6. doi:10.1038/ng.2702.

[4] Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin A V., et al. Signatures of mutational processes in human cancer. Nature 2013;500:415–21. doi:10.1038/nature12477.

[5] Shi MJ, Meng XY, Lamy P, Banday AR, Yang J, Moreno-Vega A, et al. APOBEC-mediated Mutagenesis as a Likely Cause of FGFR3 S249C Mutation Over-representation in Bladder Cancer. Eur Urol 2019;76:9–13. doi:10.1016/j.eururo.2019.03.032.

[6] Henderson S, Chakravarthy A, Su X, Boshoff C, Fenton TR. APOBEC-Mediated Cytosine Deamination Links PIK3CA Helical Domain Mutations to Human Papillomavirus-Driven Tumor Development. Cell Rep 2014;7:1833–41. doi:10.1016/j.celrep.2014.05.012.

[7] Temko D, Tomlinson IPM, Severini S, Schuster-Böckler B, Graham TA. The effects of mutational processes and selection on driver mutations across cancer types. Nat Commun 2018;9:1857. doi:10.1038/s41467-018-04208-6.

[8] Poulos RC, Wong YT, Ryan R, Pang H, Wong JWH. Analysis of 7,815 cancer exomes reveals associations between mutational processes and somatic driver mutations. PLoS Genet 2018;14:1–20. doi:10.1371/journal.pgen.1007779.

[9] Cannataro VL, Gaffney SG, Sasaki T, Issaeva N, Grewal NKS, Grandis JR, et al. APOBEC-induced mutations and their cancer effect size in head and neck squamous cell carcinoma. Oncogene 2019;38:3475–87. doi:10.1038/s41388-018-0657-6.

[10] Haradhvala NJ, Polak P, Stojanov P, Covington KR, Shinbrot E, Hess JM, et al. Mutational Strand Asymmetries in Cancer Genomes Reveal Mechanisms of DNA Damage and Repair. Cell 2016;164:538–49. doi:10.1016/j.cell.2015.12.050.

[11] Sharma S, Baysal BE. Stem-loop structure preference for site-specific RNA editing by APOBEC3A and APOBEC3G. PeerJ 2017;5:e4136. doi:10.7717/peerj.4136.

[12] Holtz CM, Sadler HA, Mansky LM. APOBEC3G cytosine deamination hotspots are defined by both sequence context and single-stranded DNA secondary structure. Nucleic Acids Res 2013;41:6139–48. doi:10.1093/nar/gkt246.

[13] Adolph MB, Love RP, Feng Y, Chelico L. Enzyme cycling contributes to efficient induction of genome mutagenesis by the cytidine deaminase APOBEC3B. Nucleic Acids Res 2017;45:11925–40. doi:10.1093/nar/gkx832.

[14] Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016;534:47–54. doi:10.1038/nature17676.

[15] Shi K, Carpenter MA, Banerjee S, Shaban NM, Kurahashi K, Salamango DJ, et al. Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. Nat Struct Mol Biol 2017;24:131–9. doi:10.1038/nsmb.3344.

[16] Wu X, Kabalane H, Kahli M, Petryk N, Laperrousaz B, Jaszczyszyn Y, et al. Developmental and cancer-associated plasticity of DNA replication preferentially targets GC-poor, lowly expressed and late-replicating regions. Nucleic Acids Res 2018;46:10157–72. doi:10.1093/nar/gky797.

[17] Petryk N, Kahli M, D'Aubenton-Carafa Y, Jaszczyszyn Y, Shen Y, Silvain M, et al. Replication landscape of the human genome. Nat Commun 2016;7:10208. doi:10.1038/ncomms10208.

[18] Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell 2017;171:540-556.e25. doi:10.1016/j.cell.2017.09.007.

[19] Lamy P, Nordentoft I, Birkenkamp-Demtroder K, Thomsen MBH, Villesen P, Vang S, et al. Paired Exome Analysis Reveals Clonal Evolution and Potential Therapeutic Targets in Urothelial Carcinoma. Cancer Res 2016;76:5894–906. doi:10.1158/0008-5472.can-16-0436.

[20] Garraway LA, Lander ES. Lessons from the cancer genome. Cell 2013;153:17–37. doi:10.1016/j.cell.2013.03.002.

[21] Nordentoft I, Lamy P, Birkenkamp-Demtröder K, Shumansky K, Vang S, Hornshøj H, et al. Mutational context and diverse clonal development in early and late bladder cancer. Cell Rep 2014;7:1649–63. doi:10.1016/j.celrep.2014.04.038.

[22] Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Jr. LAD, Kinzler KW. Cancer Genome Landscapes. Science (80-) 2013;339:1546–58. doi:10.1126/science.1235122.

[23] Buljan M, Blattmann P, Aebersold R, Boutros M. Systematic characterization of pan-cancer mutation clusters. Mol Syst Biol 2018;14:e7974. doi:10.15252/msb.20177974.

[24] Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell 2018;173:371-385.e18. doi:10.1016/j.cell.2018.02.060.

[25] Mathews DH, Sabina J, Zuker M, Turner DH. Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure. J Mol Biol 1999;288:911–40. doi:10.1006/jmbi.1999.2700.

[26] Halstead AM, Kapadia CD, Bolzenius J, Chu CE, Schriefer A, Wartman LD, et al. Bladder-cancer-associated mutations in RXRA activate peroxisome proliferator-activated receptors to drive urothelial proliferation. Elife 2017;6:pii:e30862. doi:10.7554/eLife.30862.

[27] Miao D, Margolis CA, Vokes NI, Liu D, Taylor-Weiner A, Wankowicz SM, et al. Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. Nat Genet 2018;50:1271–81. doi:10.1038/s41588-018-0200-2.

[28] Wang S, Jia M, He Z, Liu XS. APOBEC3B and APOBEC mutational signature as potential predictive markers for immunotherapy response in non-small cell lung cancer. Oncogene 2018;37:3924–36. doi:10.1038/s41388-018-0245-9.

[29] Olson ME, Harris RS, Harki DA. APOBEC Enzymes as Targets for Virus and Cancer Therapy. Cell Chem Biol 2018;25:36–49. doi:10.1016/j.chembiol.2017.10.007.

[30] Buisson R, Langenbucher A, Bowen D, Kwan EE, Benes CH, Zou L, et al. Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. Science (80-) 2019;364:eaaw2872. doi:10.1126/science.aaw2872.

Supplementary data

1. Supplementary materials & methods

SNV data. All available tumour datasets for SNVs from whole-exome sequencing were downloaded from cBioPortal for Cancer Genomics [1,2] for 602 BCa [3–7], 281 cervical cancer [8], 648 head and neck cancer [8–11], 1,425 breast cancer [8,12–14], 150 The Metastatic Breast Cancer Project and 1,247 lung cancer [8,15,16] samples. Duplicated samples from time series or multiple position sampling from the same subject were removed. A validation cohort of 27 bladder cancer patients from a recent publication was also included [17]. WES-derived SNVs for 39 bladder cancer cell lines were downloaded from the DepMap: The Cancer Dependency Map Project data repository.

Long-tail/power-law distribution. We ranked the frequency of mutations for all SNVs and for APOBEC-type SNVs (corresponding to TCN → T[G/T]N mutations) in the BCa tumours (n = 602). Both these frequencies followed a long-tail distribution. We visualised, on a log-log scale, the relationship between the rank of a given frequency and the frequency itself, and tested the log-log linearity in-between by least-squares linear regression, to check for a potential power-law relationship [18,19]. An empirical threshold was determined for distinguishing the 'head' from the 'tail' within the distribution. This threshold was defined as the smallest integer for which the ratio of the number of mutations with a frequency larger than this integer to the number of mutations with a frequency equal to this integer was > 1. We visualised the ratios across mutation frequencies and for the determined threshold. Mutations with a frequency that was not less than this empirical threshold were considered 'genuine' hotspot mutations that had undergone specific selection. We repeated this analysis for the

dataset combining other cancer types presenting relatively high APOBEC-mediated mutagenesis (cervical, head and neck, breast and lung cancer) (n = 3,751) [20–25].

Mutational signature fitting. We conducted non-negative matrix factorisation (NMF) by fitting the 30 established Catalogue Of Somatic Mutations In Cancer (COSMIC) signatures to the SNVs obtained for the primary BCa dataset (n = 602), the other APOBEC-related cancer types [20–25] (n = 3,744, samples with only 1 mutation removed), the BCa validation dataset (n = 27), and the BCa cell line dataset (n = 39). For each patient, the arithmetic sum of signature fraction scores for COSMIC signature 2 and 13, which have been demonstrated to be associated with APOBEC activity [21], was defined as the parameter to evaluate APOBEC-mediated mutagenesis (termed as APOBEC fraction score in all figures). R version 3.5.2 and Bioconductor package *sigfit* version 1.3.1 were used for these analyses and for the associated visualisation.

Association between the APOBEC signature and mutations. For the identification of APOBEC-associated hotspot mutations in BCa (n = 603) and other APOBEC-related cancer types [20–25] (n = 3,744), we compared the APOBEC-mediated mutagenesis of tumours bearing one given candidate APOBEC-associated hotspot mutation (defined as an APOBEC-type mutation with a frequency $\geq$ 4) with tumours free of candidate APOBEC-associated hotspot mutations. Mutations displaying no association with the APOBEC-mediated mutagenesis were labelled ‘undetermined’.

Clonality of APOBEC-associated hotspot mutations. The data on mutation event clonality was extracted from the TCGA BCa WES dataset [3]. We compared the probability of clonal event between the 44 APOBEC-related hotspot mutations and all other mutations of these APOBEC-target genes, using Fisher’s exact test. Considering that mutation frequency may also be associated with clonality, we further conducted multivariate analysis using generalized linear mixed model (GLMM) taking mutation frequency as covariate and genes as random-effects.

Replication fork directionality (RFD) profiling and stem-loop structures for ssDNA. We performed analysis of published genome-wide replication fork directionality (RFD) data in nine human cancer/normal cell lines —— HeLa, IMR90, TLSE19, K562, TF1, GM06990, BL79, IARC385 and Raji cells [26,27] —— to identify the strand that would be favoured as the lagging-strand template. Considering APOBEC enzymes specifically deaminated ‘C’ to ‘U’, we expected the complementary strand to be the lagging-strand template

if mutations were of the NGA → N[C/A]A type. Data availability and interpretation have been described elsewhere [26–28]. RFD profiles were determined by mapping Okazaki fragments to C (Crick) and W (Watson) DNA strands. Red (blue) RFD profiles mark indicate the regions in which Watson (Crick) strands are replicated mostly as lagging-strand templates. We simply assigned a value of '1' (or '-1') to mutations occurring on Watson (or Crick) strands replicated mostly as lagging-strand templates for each cell line (Supplementary Table 1&3). For each APOBEC-associated hotspot mutation, we then calculated the probability of locating in lagging-strand template across all the nine cell lines, and we compared the probabilities of all these mutations against random, i.e. 0.5, using one-side Wilcoxon signed rank test ('greater' hypothesis). Figures were visualised with Integrative Genomics Viewer (IGV) software.

The Mfold tool with the default parameters for DNA folding [29] was used for evaluating secondary structures of single strand DNA (ssDNA) for all APOBEC-associated hotspot mutations, with 25-nucleotide sequences centred on the mutation sites as input. A thermodynamic parameter [30] —— free energy (ΔG) —— widely used to evaluate the stabilities of stem-loop structures, was calculated, as summarised in Supplementary Table 1&3. Our preliminary investigation has determined 25 nt as the appropriate sequence length for stem-loop structure prediction as well as ΔG calculation, as we found that with the sequence length increasing (starting from 13nt, 4nt increase per escalation, centred around the mutation site), the completion of the primary stem-loop structure harbouring the mutation site always occurred before or at 25nt length and the formation of neighbouring/secondary stem-loop structure not associated with the mutation site always after 25nt, considering all the 6 passenger mutations (except for *CAMK2G* I132I mutation which was not located in loops of any sequence length we tested) probably locating on a stem-loop structure when ssDNA formed (Fig. S7A-B). Given that APOBEC enzymes specifically deaminate 'C' to 'U', we expected the complementary strand to form loop structures if mutations were of the NGA → N[C/A]A type. We then predicted the stem-loop structure formation as well as the ΔG value for each APOBEC-associated mutation, calculated the probability that these mutations were located in loop, and tested whether this probability was significantly larger than random, i.e. 0.5, using logit transformation and z test.

Gene expression analysis. Gene expression data were available only in TCGA datasets (BCa, n = 406) [3]. The corresponding RNA-Seq data (RSEM) were downloaded from cBioPortal. Comparisons were made between tumours bearing a given APOBEC-associated

hotspot mutation of a known/suspected TSG and tumours devoid of any mutation (wild-type) of this gene, in terms of its expression level, using Wilcoxon rank sum test.

Association between gene functional importance (known oncogenes/TSGs *vs.* genes of unknown function, which were annotated with aforementioned method) and their expression level were analysed in BCa tumours (n = 406). In brief, the expression value of each gene (a total of > 20,000 genes) in a given sample was first divided by that of the housekeeping gene *GAPDH*, and the genes were then ranked on a percentile scale according to the median relative expression (normalised against *GAPDH*) level across BCa tumours (n = 406). We then compared this parameter between known oncogenes/TSGs and genes of unknown function, using Wilcoxon rank sum test. Functional annotations for APOBEC-associated hotspot mutations and APOBEC-target genes were curated from a recent comprehensive study [31] in which 299 cancer genes and 579 driver mutations were functionally annotated.

Stem-loop stability. For the standardised evaluation of mutation sites in terms of ssDNA structure-related APOBEC mutagen accessibility, we calculated the normalised loop stability score for the 44 APOBEC-associated hotspot mutations in BCa, as shown below:

$$Normalised\ loop\ stability\ score = \begin{cases} \dfrac{\Delta G - \min(\boldsymbol{\Delta G})}{\max(\boldsymbol{\Delta G}) - \min(\boldsymbol{\Delta G})}, & if\ mutation\ site\ in\ loop \\ 1, & if\ mutation\ site\ not\ in\ loop \end{cases}$$

in which ΔG denotes the exact ΔG value of the loop structure in which a given mutation is located, and $\boldsymbol{\Delta G}$ denotes the vector containing all ΔG values. Lower values are associated with easier formation of more stable loops, with greater accessibility to APOBEC mutagens. We compared normalised loop stability score (ΔG) between APOBEC-associated known driver and passenger hotspot mutations.

Similarity-based driver/passenger prediction by joint analysis of stem-loop stability and gene expression. Given the findings that the normalised loop stability score and gene expression rank can respectively distinguish functional importance of mutations and genes, we combined these two parameters to predict the driverness for the remaining mutations, using an iterative similarity-based approach. In brief, initially for each mutation, we calculated the mean difference between the Euclidean distance with known drivers and passengers, and determined the statistical significance using two-sided Student's t-test with heteroscedastic variances, given the known drivers were close to each other but the known passengers more dispersed in the two-dimensional space. For mutations showing statistically significant difference between distances with drivers and passengers, we predict them as drivers if closer

to known drivers otherwise as passengers. We then repeated this process iteratively for the mutations not determined in previous iterations, by taking into consideration also the driver/passenger labels predicted in the previous iterations. The iteration was stopped once any of the following criteria reached: i) all mutations were predicted as driver or passenger; ii) remaining mutations had no significant difference between distances with drivers and passengers both known and predicted.

Translational potential of APOBEC-associated mutations. We investigated the capacity of the 44 APOBEC-associated hotspot mutations to select patients with high APOBEC-mediated mutagenesis who may be likely to benefit from immunotherapy, as assessed by comparing APOBEC fraction score between samples with and without these mutations. Further verifications were performed with two external datasets, an additional cohort [17] of BCa tumours ($n = 27$) and sequenced BCa cell lines (data from DepMap Project, $n = 39$).

Statistical & bioinformatics analysis. Wilcoxon's rank sum test was used for the comparisons. A value of $P < 0.05$ in two-tailed tests was considered statistically significant. R version 3.5.2 and the *ggpubr* version 0.2 package were used for all analyses and for the associated visualisation.

2. Related URLs

URLs. cBioPortal database, http://www.cbioportal.org/; The Metastatic Breast Cancer Project, https://www.mbcproject.org/; COSMIC Signatures of Mutational Processes in Human Cancer website, http://cancer.sanger.ac.uk/cosmic/signatures; DepMap, The Cancer Dependency Map Project data repository, https://depmap.org/portal/download/; Integrative Genomics Viewer (IGV) software, https://software.broadinstitute.org/software/igv/; Mfold tool, http://unafold.rna.albany.edu/?q=mfold; R software, http://www.r-project.org/.

3. Supplementary figures and tables

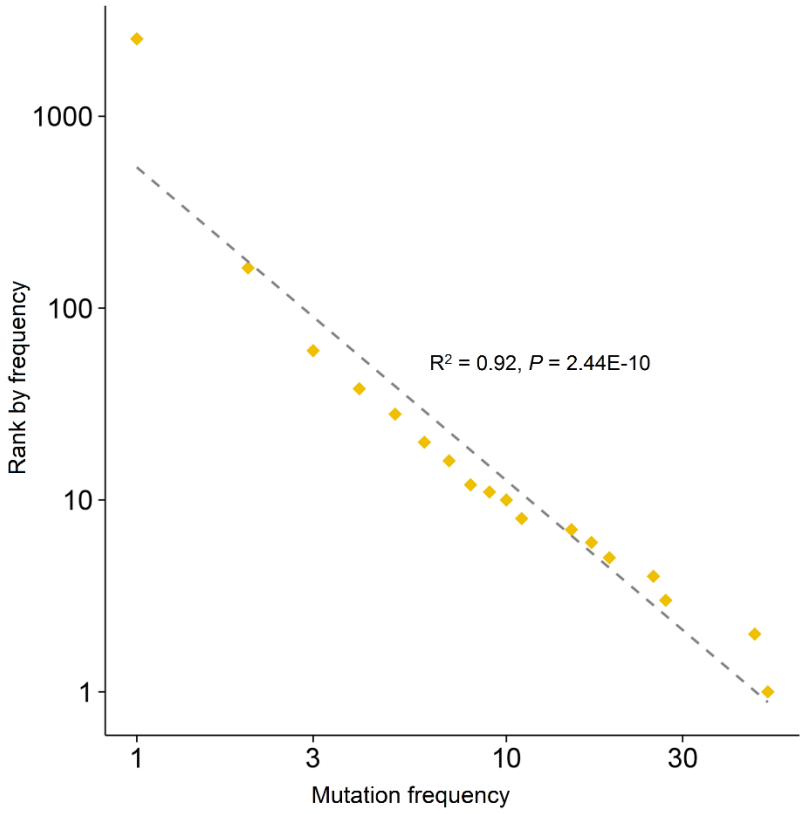


Fig. S1. The frequency of APOBEC-type mutations follows a power-law distribution. Mutations within 5'-TCN motifs (TCN → T[G/T]N, N = any nucleotide) are considered to be APOBEC-type mutations. A strong log-log linear correlation was found between the frequency of APOBEC-type mutations and their frequency rank was found. The R^2 and *P* values of the least-squares linear regression are shown. Related to Figure 1A.

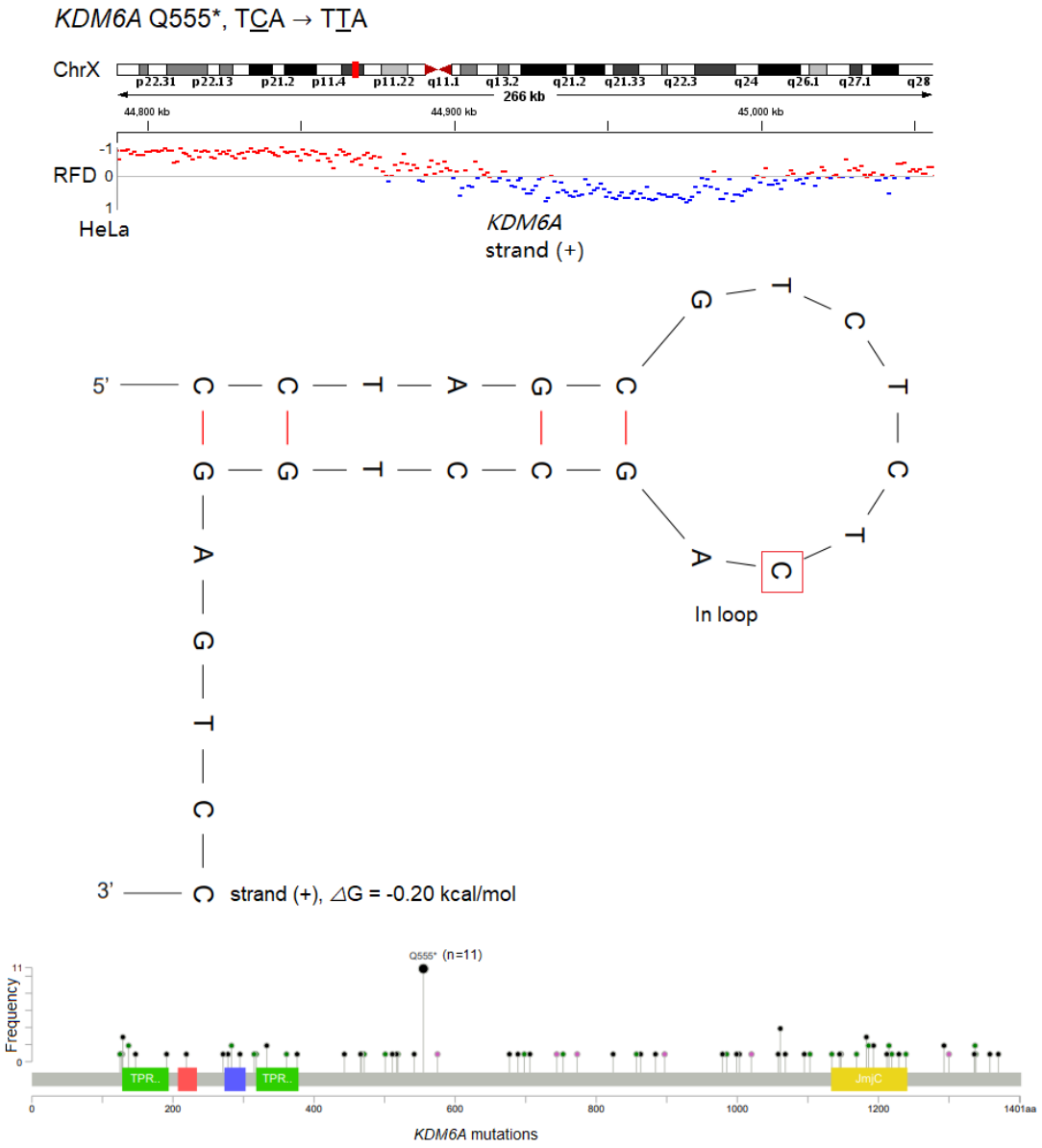


Fig. S2. APOBEC-mediated mutagenesis confers special characteristics to its target mutations. Plots related to Figure 1D-E and Figure 2C. For each of the included 43 APOBEC-associated hotspot mutations (mapping to 32 genes) identified in bladder cancer, we show replication fork directionality (RFD) during DNA replication in the HeLa cell line (upper panel), predicted stem-loop structure (middle panel) and the mutation spectrum of the corresponding gene (lower panel). DNA stem-loop structures were predicted with 25 nt length ssDNA centred on mutated site. The interpretation of the results is similar to that of the example of *ERBB2* S310F shown in the main figures (Figure 1F-G and Figure 2D). The RFDs of eight other cell lines are summarised in Supplementary Table 1.

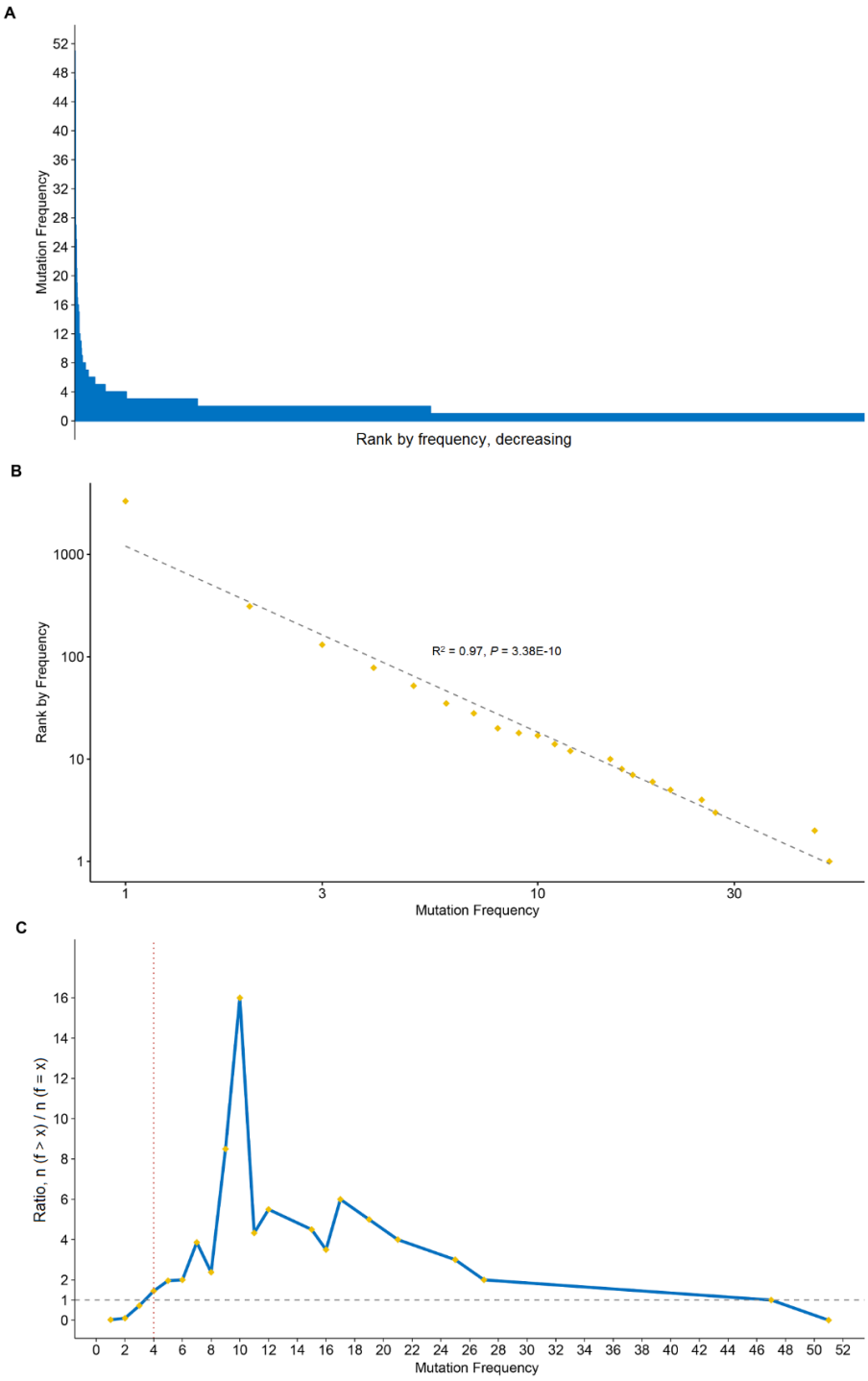


Fig. S3. The frequency of all mutations in 602 bladder cancers follows a long-tail distribution (upper panel), or, more precisely, a power-law distribution (middle panel). Identification of the optimal frequency threshold (counts $\geqslant$ 4) for 130 hotspot mutations (lower panel). Related to Figure 2A-B. Details of these hotspot mutations are summarised in Supplementary Tables 1 & 2. The method of analysis and interpretation of the results are similar to that of aforementioned identification of candidate APOBEC-associated hotspot mutations in Figure 1A and Fig. S1.

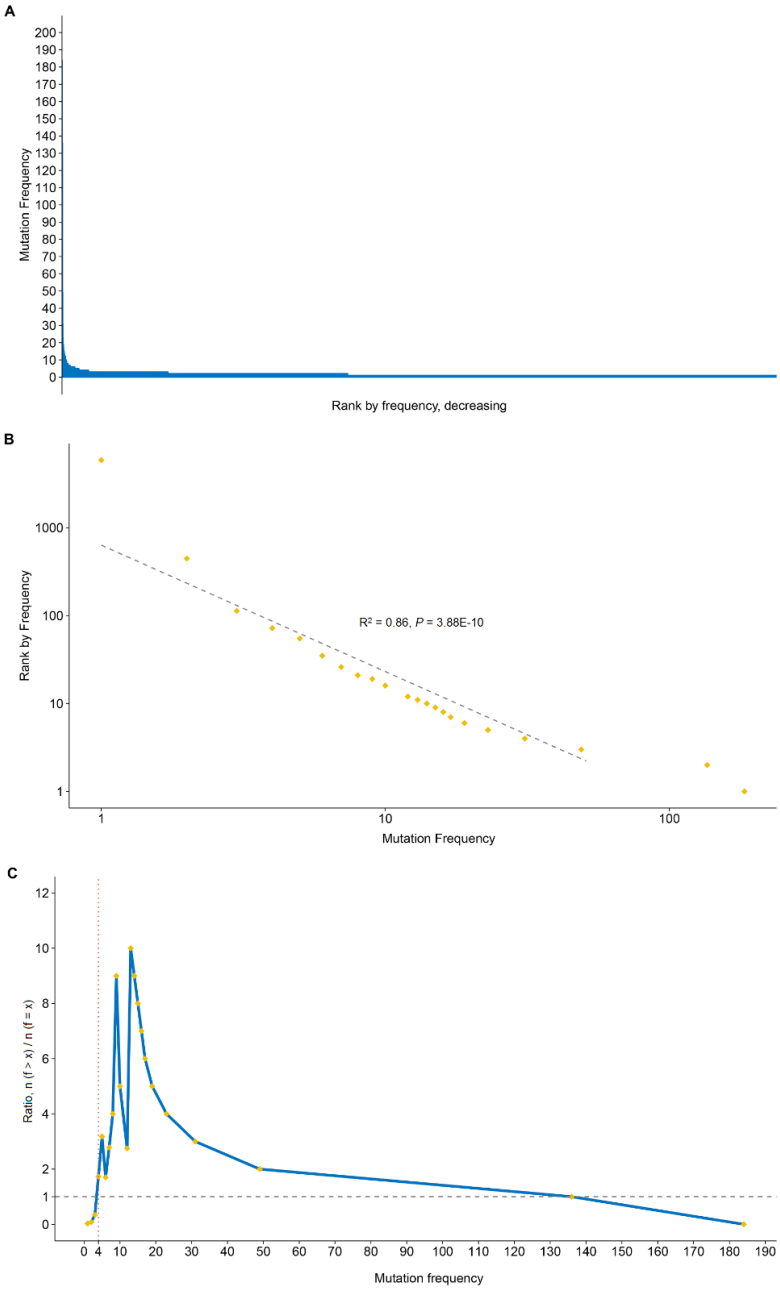


Fig. S4. The frequency of the APOBEC-type mutations identified in other cancer types presenting high APOBEC-mediated mutagenesis follows a long-tail distribution (upper panel), or, more precisely a power-law distribution (middle panel). Identification of the optimal frequency threshold (counts $\geqslant$ 4) for 112 candidate APOBEC-associated hotspot mutations (lower panel). Related to Figure 3A. Details of these hotspot mutations are summarised in Supplementary Table 3. The other cancer types correspond to 3751 tumours from patients with cervical, head and neck, breast and lung cancers. The method of analysis and interpretation of the results are similar to that of aforementioned identification of candidate APOBEC-associated hotspot mutations in bladder cancer in Figure 1A and Fig. S1.

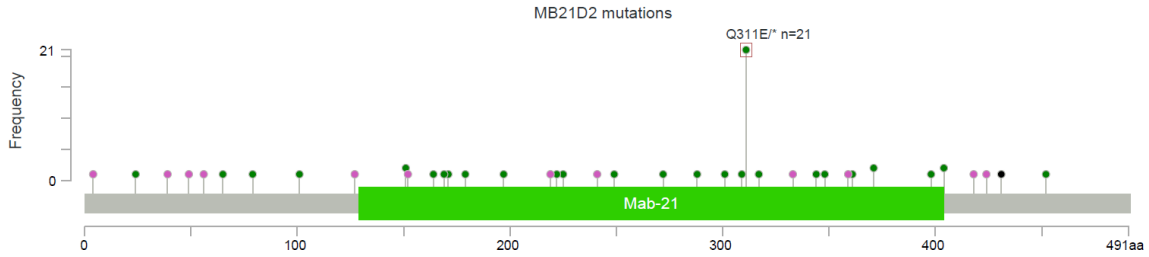


Fig. S5. Mutation spectra for the 55 APOBEC-target genes identified from other cancer types. Related to Figure 3C. One of representative example, *MB21D2*. Red rectangles indicate APOBEC-associated hotspot mutations. The other cancer types corresponded to 3,751 tumours from patients with cervical, head and neck, breast and lung cancer. Most of the over-represented mutations are associated with APOBEC-mediated mutagenesis within its target genes.

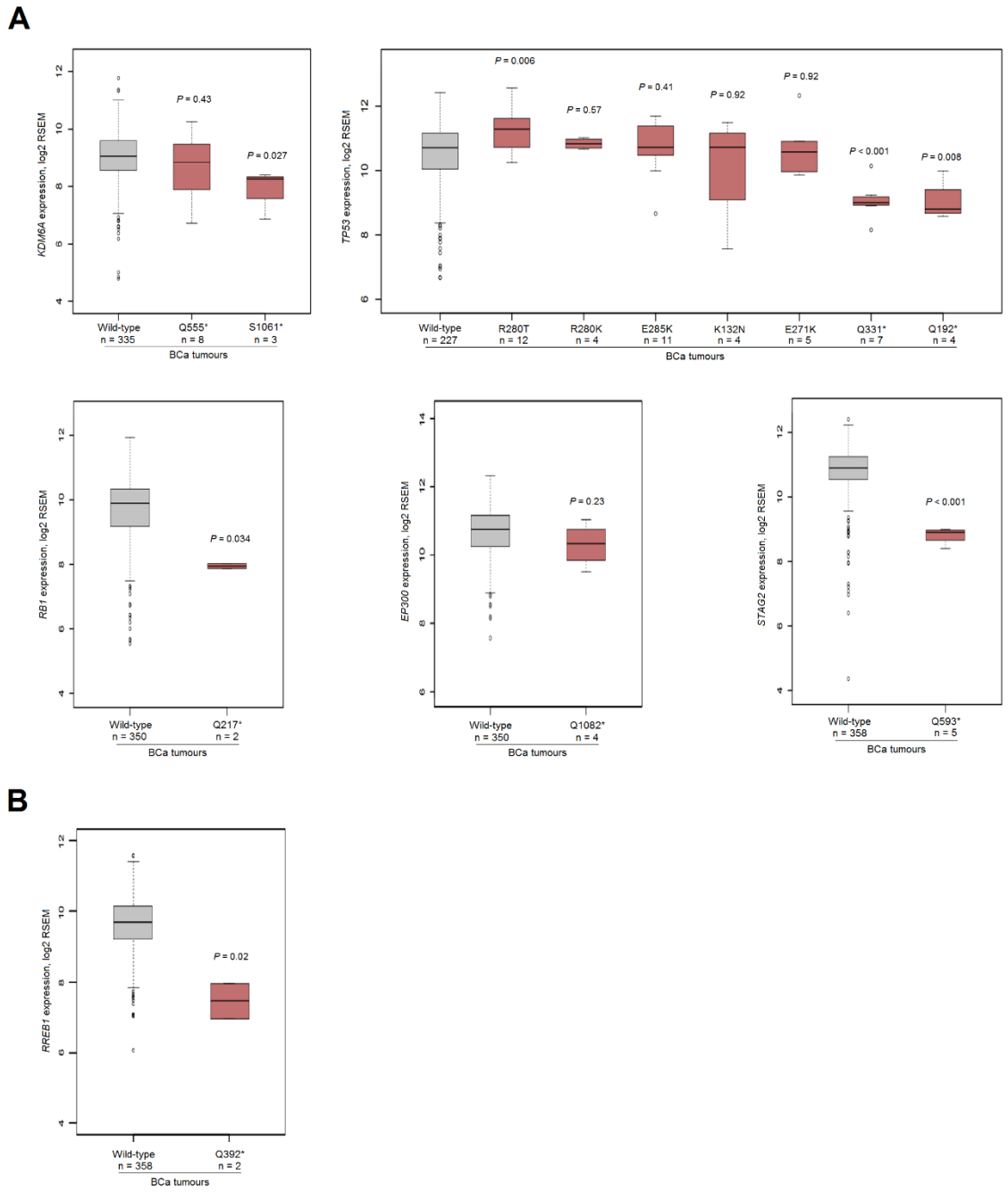


Fig. S6. mRNA levels (RSEM, log2) for **(A)** known and **(B)** suspected tumour suppressor genes (TSGs) according to the APOBEC-associated mutations of these TSGs – tumours with a given APOBEC-associated hotspot mutation *vs.* tumours with wild-type alleles for each known/suspected TSG in bladder cancer. *RREB1* as a suspected TSG. RNA-Seq data are available in 406 bladder cancers from The Cancer Genome Atlas (TCGA). *P* value: Wilcoxon test between two groups.

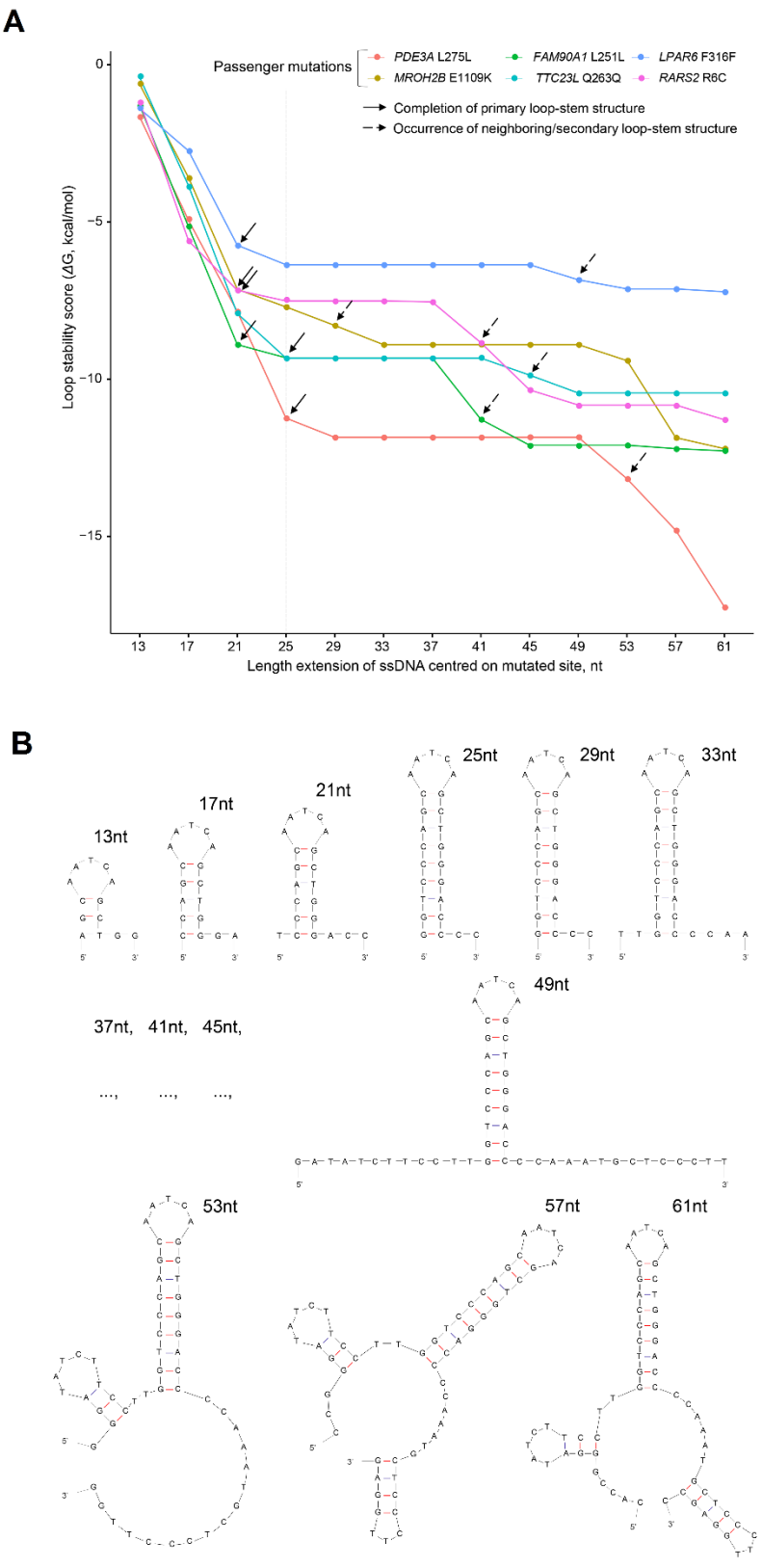


Fig. S7. Optimal ssDNA sequence length – 25nt – for stem-loop structure prediction. **A)** Loop stability scores (*Δ*G, kcal/mol) with gradually increased sequence length of ssDNA (starting from 13nt, 4nt increase per escalation, centred on the mutation site) considering 6 passenger

mutations. Passengers include 4 silent mutations, 1 mutation within transit peptide and 1 missense mutation on gene with an absence of mRNA expression in bladder cancer. Colours represent different passenger mutations. Solid and dashed arrows represent the completion of primary and occurrence of neighbouring/secondary stem-loop structures, respectively. **B)** Examples of predicted stem-loop structures for *PDE3A* L275L mutation across series of ssDNA lengths.

Supplementary Table 1. List of 59 candidate APOBEC-associated hotspot mutations identified in 602 bladder cancers. This table also includes statistical significances for APOBEC-mediated mutagenesis (Related to Figure 1B), gene transcriptional direction, predicted DNA stem-loop structures, replication fork directionality (RFD) data for nine cancer/normal cell lines, detailed data generating Figure 4&5 and functional annotations for these mutations.

Code_Short	Freq	Cancer	APOBEC -type	Variation_Class	DNA_C	Protein_C
FGFR3 S249C	51	BLCA	S13	Missense	c.746C>G	p.S249C
PIK3CA E545K	47	BLCA	S2	Missense	c.1633G>A	p.E545K
ERBB2 S310F	27	BLCA	S2	Missense	c.839C>T	p.S310F
PIK3CA E542K	25	BLCA	S2	Missense	c.1624G>A	p.E542K
RXRA S427F	19	BLCA	S2	Missense	c.989C>T	p.S427F
TP53 E285K	17	BLCA	S2	Missense	c.853G>A	p.E285K
TP53 R280T	15	BLCA	S13	Missense	c.839G>C	p.R280T
KDM6A Q555*	11	BLCA	S2	Nonsense	c.1663C>T	p.Q555*
TBC1D12 *1*	11	BLCA	S2	Targeted_Region	NA	*1*
C3orf70 S6L	10	BLCA	S2	Missense	c.17C>T	p.S6L
RHOB E172K	9	BLCA	S2	Missense	c.514G>A	p.E172K
AHR Q383H	8	BLCA	S13	Missense	c.1149G>C	p.Q383H
LPAR6 F316F	8	BLCA	S2	Silent	c.948C>T	p.F316=
TP53 Q331*	8	BLCA	S2	Nonsense	c.991C>T	p.Q331*
TP53 E271K	8	BLCA	S2	Missense	c.811G>A	p.E271K
RARS2 R6C	7	BLCA	S2	Missense	c.16C>T	p.R6C
SF3B1 E902K	7	BLCA	S2	Missense	c.2704G>A	p.E902K
TP53 R280K	7	BLCA	S2	Missense	c.839G>A	p.R280K
ERBB3 E332K	6	BLCA	S2	Missense	c.994G>A	p.E332K
MROH2B E1109K	6	BLCA	S2	Missense	c.3325G>A	p.E1109K
PIK3CA E545Q	6	BLCA	S13	Missense	c.1633G>C	p.E545Q
PPCS S113L	6	BLCA	S2	Missense	c.338C>T	p.S113L
STAG2 Q593*	6	BLCA	S2	Nonsense	c.1777C>T	p.Q593*
ACSS3 S290L	5	BLCA	S2	Missense	c.869C>T	p.S290L
CELSR3 E356K	5	BLCA	S2	Missense	c.1066G>A	p.E356K
KCNF1 E158K	5	BLCA	S2	Missense	c.472G>A	p.E158K
PDE3A L275L	5	BLCA	S2	Silent	c.825G>A	p.L275=

PIK3CA E726K	5	BLCA	S2	Missense	c.2176G>A	p.E726K
PLXNA2 E1480K	5	BLCA	S2	Missense	c.4438G>A	p.E1480K
RHOB E47K	5	BLCA	S2	Missense	c.139G>A	p.E47K
TFPI2 R222C	5	BLCA	S2	Missense	c.664C>T	p.R222C
TP53 K132N	5	BLCA	S13	Missense	c.396G>C	p.K132N
CAMK2G I132I	4	BLCA	S2	Silent	c.396C>T	p.I132=
CELSR1 E1382K	4	BLCA	S2	Missense	c.4144G>A	p.E1382K
EP300 Q1082*	4	BLCA	S2	Nonsense	c.3244C>T	p.Q1082*
ERBB4 E317K	4	BLCA	S2	Missense	c.949G>A	p.E317K
FAM90A1 L251L	4	BLCA	S13	Silent	c.753C>G	p.L251=
FURIN R693W	4	BLCA	S2	Missense	c.2077C>T	p.R693W
KDM6A S1061*	4	BLCA	S13	Nonsense	c.3182C>G	p.S1061*
PBX2 E70K	4	BLCA	S2	Missense	c.208G>A	p.E70K
RB1 Q217*	4	BLCA	S2	Nonsense	c.649C>T	p.Q217*
RREB1 Q392*	4	BLCA	S2	Nonsense	c.1174C>T	p.Q392*
TP53 Q192*	4	BLCA	S2	Nonsense	c.574C>T	p.Q192*
TTC23L Q263Q	4	BLCA	S2	Silent	c.789G>A	p.Q263=
ERBB3 M91I	7	BLCA	S2	Missense	c.273G>A	p.M91I
ERCC2 S44L	6	BLCA	S2	Missense	c.131C>T	p.S44L
NFE2L2 R34G	6	BLCA	S13	Missense	c.100C>G	p.R34G
TP53 S241F	6	BLCA	S2	Missense	c.722C>T	p.S241F
ERBB4 Q707E	5	BLCA	S13	Missense	c.2119C>G	p.Q707E
ATXN2L R382C	4	BLCA	S2	Missense	c.1144C>T	p.R382C
CUL1 D483N	4	BLCA	S2	Missense	c.1447G>A	p.D483N
ERCC2 E86Q	4	BLCA	S13	Missense	c.256G>C	p.E86Q
PIK3CA R88Q	4	BLCA	S2	Missense	c.263G>A	p.R88Q
RHOA E47K	4	BLCA	S2	Missense	c.139G>A	p.E47K
SNX19 L69L	4	BLCA	S2	Silent	c.207C>T	p.L69=
TCHH E1052K	4	BLCA	S2	Missense	c.3154G>A	p.E1052K
U2AF1 S34F	4	BLCA	S2	Missense	c.101C>T	p.S34F
VWF S2559L	4	BLCA	S2	Missense	c.7676C>T	p.S2559L
ZNF513 R3Q	4	BLCA	S2	Missense	c.8G>A	p.R3Q

Supplementary Table 2. List of 71 hotspot mutations (counts ⩾ 4, excluding 59 candidate APOBEC-associated hotspot mutations) identified in 602 bladder cancers and their functional annotations.

Code_Short	Freq	Cancer	Variation_Class	DNA_C	Protein_C
TP53 R248Q	21	BLCA	Missense	c.743G>A	p.R248Q
MDC1 G1558C	6	BLCA	Missense	c.4672G>T	p.G1558C
KRAS G12D	6	BLCA	Missense	c.35G>A	p.G12D
ZAN P717L	6	BLCA	Missense	c.2150C>T	p.P717L
AC008132.13 V285A	5	BLCA	Missense	c.854T>C	p.V285A
LINC00273 *232*	5	BLCA	RNA	n.694A>G	*232*
LINC00273 *231*	5	BLCA	RNA	n.692G>C	*231*
TCHH E1052K	4	BLCA	Missense	c.3154G>A	p.E1052K

VWF S2559L	4	BLCA	Missense	c.7676C>T	p.S2559L
CDKN2A R80*	4	BLCA	Nonsense	c.238C>T	p.R80*
PAQR9 Y323Y	4	BLCA	Silent	c.969C>T	p.Y323=
KRT16P2 *240*	4	BLCA	RNA	n.720G>A	*240*
TYRO3 R824G	4	BLCA	Missense	c.2470A>G	p.R824G
ZNF721 T565A	4	BLCA	Missense	c.1693A>G	p.T565A
ZNF83 E293E	16	BLCA	Silent	c.879G>A	p.E293=
ZNF83 E293V	16	BLCA	Missense	c.878A>T	p.E293V
FGFR3 Y373C	15	BLCA	Missense	c.1118A>G	p.Y373C
ZNF83 G267G	12	BLCA	Silent	c.801A>C	p.G267=
ERCC2 N238S	11	BLCA	Missense	c.713A>G	p.N238S
HRAS Q61R	8	BLCA	Missense	c.182A>G	p.Q61R
FGFR3 G370C	7	BLCA	Missense	c.1108G>T	p.G370C
FGFR3 R248C	7	BLCA	Missense	c.742C>T	p.R248C
ERCC2 S44L	6	BLCA	Missense	c.131C>T	p.S44L
KRTAP5-10 G105G	5	BLCA	Silent	c.315T>C	p.G105=
ERCC2 E86Q	4	BLCA	Missense	c.256G>C	p.E86Q
U2AF1 S34F	4	BLCA	Missense	c.101C>T	p.S34F
ZNF513 R3Q	4	BLCA	Missense	c.8G>A	p.R3Q
TP53 R282W	4	BLCA	Missense	c.844C>T	p.R282W
GNA13 R200G	4	BLCA	Missense	c.598A>G	p.R200G
ZNF91 H305R	4	BLCA	Missense	c.914A>G	p.H305R
TP53 R175H	6	BLCA	Missense	c.524G>A	p.R175H
ATXN2L R382C	4	BLCA	Missense	c.1144C>T	p.R382C
ERBB4 S1289A	12	BLCA	Missense	c.3865T>G	p.S1289A
TBC1D12 *1*	9	BLCA	5'UTR	c.-3C>T	*1*
ACTB G158R	8	BLCA	Missense	c.472G>A	p.G158R
FBXW7 R505G	8	BLCA	Missense	c.1513C>G	p.R505G
PIK3CA H1047R	8	BLCA	Missense	c.3140A>G	p.H1047R
ERBB3 M91I	7	BLCA	Missense	c.273G>A	p.M91I
ZNF83 K290R	7	BLCA	Missense	c.869A>G	p.K290R
NFE2L2 R34G	6	BLCA	Missense	c.100C>G	p.R34G
TP53 S241F	6	BLCA	Missense	c.722C>T	p.S241F
HRAS G13R	6	BLCA	Missense	c.37G>C	p.G13R
AHNAK V1940A	6	BLCA	Missense	c.5819T>C	p.V1940A
AC018720.10	6	BLCA	RNA	n.906C>T	NA
CD209 T197T	6	BLCA	Silent	c.591C>T	p.T197=
ZNF83 H311H	6	BLCA	Silent	c.933T>C	p.H311=
ERBB4 Q707E	5	BLCA	Missense	c.2119C>G	p.Q707E
TP53 R273C	5	BLCA	Missense	c.817C>T	p.R273C
HRAS Q61L	5	BLCA	Missense	c.182A>T	p.Q61L
TP53 E285*	5	BLCA	Nonsense	c.853G>T	p.E285*
PIK3CA H1047L	5	BLCA	Missense	c.3140A>T	p.H1047L
ERBB2 S310Y	5	BLCA	Missense	c.929C>A	p.S310Y
KRAS G12C	5	BLCA	Missense	c.34G>T	p.G12C
ERBB3 V104L	5	BLCA	Missense	c.310G>T	p.V104L
ERBB3 H228Q	5	BLCA	Missense	c.684T>G	p.H228Q
ANKRD30A P319A	5	BLCA	Missense	c.955C>G	p.P319A

ERBB4 Q707Q	5	BLCA	Silent	c.2121A>G	p.Q707=
ZNF721 A537T	5	BLCA	Missense	c.1609G>A	p.A537T
ZAN P913L	5	BLCA	Missense	c.2738C>T	p.P913L
CUL1 D483N	4	BLCA	Missense	c.1447G>A	p.D483N
PIK3CA R88Q	4	BLCA	Missense	c.263G>A	p.R88Q
RHOA E47K	4	BLCA	Missense	c.139G>A	p.E47K
SNX19 L69L	4	BLCA	Silent	c.207C>T	p.L69=
NRAS Q61R	4	BLCA	Missense	c.182A>G	p.Q61R
RB1 X405_splice	4	BLCA	Splice_Site	c.1215+1G>A	p.X405_splice
TP53 E336*	4	BLCA	Nonsense	c.1006G>T	p.E336*
FGFR3 K650E	4	BLCA	Missense	c.1948A>G	p.K650E
ERBB2 R678Q	4	BLCA	Missense	c.2033G>A	p.R678Q
FGFR3 S371C	4	BLCA	Missense	c.1111A>T	p.S371C
TP53 R248W	4	BLCA	Missense	c.742C>T	p.R248W
HRAS G12D	4	BLCA	Missense	c.35G>A	p.G12D
NBPF3 F424C	4	BLCA	Missense	c.1271T>G	p.F424C
HRAS Q61K	4	BLCA	Missense	c.181C>A	p.Q61K
HRAS G12S	4	BLCA	Missense	c.34G>A	p.G12S
ZNF721 E561V	4	BLCA	Missense	c.1682A>T	p.E561V
ZNF479 T466K	4	BLCA	Missense	c.1397C>A	p.T466K
NR3C2 P697P	4	BLCA	Silent	c.2091A>C	p.P697=
FAM47C Q225E	4	BLCA	Missense	c.673C>G	p.Q225E
STK39 Y182C	4	BLCA	Missense	c.545A>G	p.Y182C
ANKRD30A A353P	4	BLCA	Missense	c.1057G>C	p.A353P
ZNF443 P615P	4	BLCA	Silent	c.1845G>A	p.P615=
MAP4 A391A	4	BLCA	Silent	c.1173T>C	p.A391=
ZNF721 T341A	4	BLCA	Missense	c.1021A>G	p.T341A
FMN2 P1108P	4	BLCA	Silent	c.3324T>C	p.P1108=
NOTCH2 A21T	4	BLCA	Missense	c.61G>A	p.A21T
ESRRA A378A	4	BLCA	Silent	c.1134G>A	p.A378=

Supplementary Table 3. List of 112 candidate APOBEC-associated hotspot mutations identified in other cancer types presenting relatively high APOBEC mutagenesis (a total of 3751 cervical, head and neck, breast and lung cancer tumours). This table also includes statistical significances for APOBEC-mediated mutagenesis, gene transcriptional direction, predicted DNA stem-loop structures, and replication fork directionality (RFD) data for nine cancer/normal cell lines, and functional annotations for these mutations.

Code_Short	Freq	Cancers	APOBEC-type	Variation	DNA_C	Protein_C
PIK3CA E545K	184	4 types	S2	Missense	c.1633G>A	p.E545K
PIK3CA E542K	136	4 types	S2	Missense	c.1624G>A	p.E542K
PIK3CA E726K	23	4 types	S2	Missense	c.2176G>A	p.E726K
MB21D2 Q311E	17	4 types	S13	Missense	c.931C>G	p.Q311E
MAPK1 E322K	16	4 types	S2	Missense	c.964G>A	p.E322K
TP53 Q192*	15	4 types	S2	Nonsense	c.574C>T	p.Q192*

TP53 E285K	14	4 types	S2	Missense	c.853G>A	p.E285K
NFE2L2 E79Q	12	4 types	S13	Missense	c.235G>C	p.E79Q
PIK3CA E453K	12	4 types	S2	Missense	c.1357G>A	p.E453K
RARS2 R6C	10	4 types	S2	Missense	c.16C>T	p.R6C
TP53 Q331*	10	4 types	S2	Nonsense	c.991C>T	p.Q331*
TP53 E286K	10	4 types	S2	Missense	c.856G>A	p.E286K
NUP93 E14K	9	4 types	S2	Missense	c.40G>A	p.E14K
C3orf70 S6L	8	4 types	S2	Missense	c.17C>T	p.S6L
ESR1 E380Q	8	4 types	S13	Missense	c.1138G>C	p.E380Q
FGFR3 S249C	8	4 types	S13	Missense	c.746C>G	p.S249C
KLF5 E419Q	8	4 types	S13	Missense	c.1255G>C	p.E419Q
PIK3CA E81K	7	4 types	S2	Missense	c.241G>A	p.E81K
PTEN Q245*	7	4 types	S2	Nonsense	c.733C>T	p.Q245*
RHOA E40Q	7	4 types	S13	Missense	c.118G>C	p.E40Q
TP53 E271K	7	4 types	S2	Missense	c.811G>A	p.E271K
TP53 G266E	7	4 types	S2	Missense	c.797G>A	p.G266E
ACTB E334K	6	4 types	S2	Missense	c.1000G>A	p.E334K
APH1A R31C	6	4 types	S2	Missense	c.91C>T	p.R31C
CASP8 Q465*	6	4 types	S2	Nonsense	c.1393C>T	p.Q465*
FAM83G Q88*	6	4 types	S2	Nonsense	c.262C>T	p.Q88*
HIST1H1B P198P	6	4 types	S2	Silent	c.594G>A	p.P198=
HIST2H2BE F71L	6	4 types	S13	Missense	c.213C>G	p.F71L
KLF5 E419K	6	4 types	S2	Missense	c.1255G>A	p.E419K
MYH9 E530K	6	4 types	S2	Missense	c.1588G>A	p.E530K
NFE2L2 D29N	6	4 types	S2	Missense	c.85G>A	p.D29N
PDE3A L275L	6	4 types	S2	Silent	c.825G>A	p.L275L
PIK3CA K111N	6	4 types	S13	Missense	c.333G>C	p.K111N
TP53 R280T	6	4 types	S13	Missense	c.839G>C	p.R280T
TP53 R280K	6	4 types	S2	Missense	c.839G>A	p.R280K
ZSCAN22 A321A	6	4 types	S2	Silent	c.963G>A	p.A321=
COL4A2 R1410Q	5	4 types	S2	Missense	c.4229G>A	p.R1410Q
DEGS2 F229F	5	4 types	S2	Silent	c.687C>T	p.F229F
FOXA1 S250F	5	4 types	S2	Missense	c.749C>T	p.S250F
GJA8 E110K	5	4 types	S2	Missense	c.328G>A	p.E110K
HIST1H1C A180A	5	4 types	S2	Silent	c.540G>A	p.A180=
HIST1H2BF E77K	5	4 types	S2	Missense	c.229G>A	p.E77K
HIST1H3B E74K	5	4 types	S2	Missense	c.220G>A	p.E74K
RUBCN R542W	5	4 types	S2	Missense	c.1624C>T	p.R542W
NFE2L2 E79K	5	4 types	S2	Missense	c.235G>A	p.E79K
NFE2L2 G31A	5	4 types	S13	Missense	c.92G>C	p.G31A
NOS3 E742K	5	4 types	S2	Missense	c.2224G>A	p.E742K
PELI3 R89Q	5	4 types	S2	Missense	c.266G>A	p.R89Q
TGFBR2 E544K	5	4 types	S2	Missense	c.1630G>A	p.E544K
TP53 K132N	5	4 types	S13	Missense	c.396G>C	p.K132N
ACTL6B E91K	4	4 types	S2	Missense	c.271G>A	p.E91K
ARID1A Q479*	4	4 types	S2	Nonsense	c.1435C>T	p.Q479*
ATXN2L R382C	4	4 types	S2	Missense	c.1144C>T	p.R382C
CASP8 R292W	4	4 types	S2	Missense	c.874C>T	p.R292W

CASP8 Q339*	4	4 types	S2	Nonsense	c.1015C>T	p.Q339*
CDH1 R63*	4	4 types	S2	Nonsense	c.187C>T	p.R63*
CUL1 E485K	4	4 types	S2	Missense	c.1453G>A	p.E485K
EPHA2 S419*	4	4 types	S13	Nonsense	c.1256C>G	p.S419*
EXOC4 S21L	4	4 types	S2	Missense	c.62C>T	p.S21L
FAS E261K	4	4 types	S2	Missense	c.781G>A	p.E261K
FOXA1 I176M	4	4 types	S13	Missense	c.528C>G	p.I176M
GMEB2 S494L	4	4 types	S2	Missense	c.1481C>T	p.S494L
HIST1H2BC E72Q	4	4 types	S13	Missense	c.214G>C	p.E72Q
*HIST1H2B*H Q23Q	4	4 types	S2	Silent	c.69G>A	p.Q23=
HLA-A Q120H	4	4 types	S13	Missense	c.360G>C	p.Q120H
HSPB8 E84K	4	4 types	S2	Missense	c.250G>A	p.E84K
KLC2 E49K	4	4 types	S2	Missense	c.145G>A	p.E49K
KMT2C Q2220*	4	4 types	S2	Nonsense	c.6658C>T	p.Q2220*
MAPK1 E81K	4	4 types	S2	Missense	c.241G>A	p.E81K
MB21D2 Q311*	4	4 types	S2	Nonsense	c.931C>T	p.Q311*
*MRPL5*3 R73C	4	4 types	S2	Missense	c.217C>T	p.R73C
NFE2 S113*	4	4 types	S13	Nonsense	c.338C>G	p.S113*
NOTCH1 E455K	4	4 types	S2	Missense	c.1363G>A	p.E455K
NPC2 *11*	4	4 types	S2	5'UTR	c.-11C>T	*11*
PCDHA8 S642C	4	4 types	S13	Missense	c.1925C>G	p.S642C
PDE3A E352K	4	4 types	S2	Missense	c.1054G>A	p.E352K
SURF2 R145C	4	4 types	S2	Missense	c.433C>T	p.R145C
TMEM131 E1319Q	4	4 types	S13	Missense	c.3955G>C	p.E1319Q
AKT1 E17K	49	4 types	S2	Missense	c.49G>A	p.E17K
TP53 R213*	31	4 types	S2	Nonsense	c.637C>T	p.R213*
U2AF1 S34F	19	4 types	S2	Missense	c.101C>T	p.S34F
EP300 D1399N	13	4 types	S2	Missense	c.4195G>A	p.D1399N
NFE2L2 D29H	12	4 types	S13	Missense	c.85G>C	p.D29H
TP53 P278S	12	4 types	S2	Missense	c.832C>T	p.P278S
ERBB2 S310F	9	4 types	S2	Missense	c.929C>T	p.S310F
PTEN R130Q	8	4 types	S2	Missense	c.389G>A	p.R130Q
AP3B1 S31L	7	4 types	S2	Missense	c.92C>T	p.S31L
CDKN2A X153_splice	7	4 types	S2	Splice_Site	c.*151-1G>A	p.X153_splice
CDKN2A D84N	7	4 types	S2	Missense	c.250G>A	p.D84N
NFE2L2 R34G	7	4 types	S13	Missense	c.100C>G	p.R34G
CTNNB1 S37F	6	4 types	S2	Missense	c.110C>T	p.S37F
NFE2L2 R34P	6	4 types	S13	Missense	c.101G>C	p.R34P
NFE2L2 R34Q	6	4 types	S2	Missense	c.101G>A	p.R34Q
PIK3CA R88Q	6	4 types	S2	Missense	c.263G>A	p.R88Q
TP53 P278A	6	4 types	S13	Missense	c.832C>G	p.P278A
TP53 S183*	6	4 types	S13	Nonsense	c.548C>G	p.S183*
NFE2L2 L30F	5	4 types	S2	Missense	c.88C>T	p.L30F
PTEN G129E	5	4 types	S2	Missense	c.386G>A	p.G129E
TBC1D12 *1*	5	4 types	S2	5'UTR	c.-1G>A	*1*
BRAF D594N	4	4 types	S2	Missense	c.1780G>A	p.D594N
FBXW7 R479Q	4	4 types	S2	Missense	c.1436G>A	p.R479Q
HIST1H2BH E77K	4	4 types	S2	Missense	c.229G>A	p.E77K

KEAP1 R470C	4	4 types	S2	Missense	c.1408C>T	p.R470C
MAP2K4 S184L	4	4 types	S2	Missense	c.551C>T	p.S184L
MSN E334E	4	4 types	S2	Silent	c.1002G>A	p.E334=
MYC S161L	4	4 types	S2	Missense	c.482C>T	p.S161L
NLRP4 R886W	4	4 types	S2	Missense	c.2656C>T	p.R886W
PTPRF E1721K	4	4 types	S2	Missense	c.5161G>A	p.E1721K
RYR3 T1693T	4	4 types	S2	Silent	c.5079G>A	p.T1693T
SF1 R380W	4	4 types	S2	Missense	c.1138C>T	p.R380W
TP53 E286Q	4	4 types	S13	Missense	c.856G>C	p.E286Q
TP53 S127F	4	4 types	S2	Missense	c.380C>T	p.S127F

4. Supplementary references

[1] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2:401–4. doi:10.1158/2159-8290.CD-12-0095.

[2] Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. Sci Signal 2013;6:pl1. doi:10.1126/scisignal.2004088.

[3] Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell 2017;171:540-556.e25. doi:10.1016/j.cell.2017.09.007.

[4] Van Allen EM, Mouw KW, Kim P, Iyer G, Wagle N, Al-Ahmadie H, et al. Somatic ERCC2 mutations correlate with cisplatin sensitivity in muscle-invasive urothelial carcinoma. Cancer Discov 2014;4:1140–53. doi:10.1158/2159-8290.CD-14-0623.

[5] Guo G, Sun X, Chen C, Wu S, Huang P, Li Z, et al. Whole-genome and whole-exome sequencing of bladder cancer identifies frequent alterations in genes involved in sister chromatid cohesion and segregation. Nat Genet 2013;45:1459–63. doi:10.1038/ng.2798.

[6] Faltas BM, Prandi D, Tagawa ST, Molina AM, Nanus DM, Sternberg C, et al. Clonal evolution of chemotherapy-resistant urothelial carcinoma. Nat Genet 2016;48:1490–9. doi:10.1038/ng.3692.

[7] Hurst CD, Alder O, Platt FM, Droop A, Stead LF, Burns JE, et al. Genomic Subtypes of Non-invasive Bladder Cancer with Distinct Metabolic Profile and Female Gender Bias in KDM6A Mutation Frequency. Cancer Cell 2017;32:701-715.e7. doi:10.1016/j.ccell.2017.08.005.

[8] Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas. Cell 2018;173:321-337.e10. doi:10.1016/j.cell.2018.03.035.

[9] Pickering CR, Zhang J, Yoo SY, Bengtsson L, Moorthy S, Neskey DM, et al. Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers. Cancer Discov 2013;3:770–81. doi:10.1158/2159-8290.CD-12-0537.

[10] Agrawal N, Westra WH, Koch WM, Califano JA, Gibbs RA, Wheeler DA, et al. Exome Sequencing of Head and Neck Squamous Cell Carcinoma Reveals Inactivating Mutations in NOTCH1. Science (80-) 2011;333:1154–7. doi:10.1126/science.1206923.

[11] Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, et al. The Mutational Landscape of Head and Neck Squamous Cell Carcinoma. Science (80-) 2011;333:1157–60. doi:10.1126/science.1208130.

[12] Lefebvre C, Bachelot T, Filleron T, Pedrero M, Campone M, Soria JC, et al. Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. PLoS Med 2016;13:1–18. doi:10.1371/journal.pmed.1002201.

[13] Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature 2012;486:400–4. doi:10.1038/nature11017.

[14] Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL, Frederick AM, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. Nature 2012;486:405–9. doi:10.1038/nature11154.

[15] Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell 2012;150:1107–20. doi:10.1016/j.cell.2012.08.029.

[16] Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Mutational landscape determines sensitivity to PD-1 blockade in non–small cell lung cancer. Science (80-) 2016;348:124–9. doi:10.1126/science.aaa1348.

[17] Miao D, Margolis CA, Vokes NI, Liu D, Taylor-Weiner A, Wankowicz SM, et al. Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. Nat Genet 2018;50:1271–81. doi:10.1038/s41588-018-0200-2.

[18] Barabasi A-L, Albert R. Emergence of Scaling in Random Networks 1999;286:509–12.

[19] Newman MEJ. Power laws, Pareto distributions and Zipf's law. Contemp Phys 2005;46:323–51. doi:10.1080/00107510500052444.

[20] Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat Genet 2013;45:970–6. doi:10.1038/ng.2702.

[21] Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin A V., et al. Signatures of mutational processes in human cancer. Nature 2013;500:415–21. doi:10.1038/nature12477.

[22] Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016;534:47–54. doi:10.1038/nature17676.

[23] Burk RD, Chen Z, Saller C, Tarvin K, Carvalho AL, Scapulatempo-Neto C, et al. Integrated genomic and molecular characterization of cervical cancer. Nature 2017;543:378–84. doi:10.1038/nature21386.

[24] Faden DL, Thomas S, Cantalupo PG, Agrawal N, Myers J, DeRisi J. Multi-modality analysis supports APOBEC as a major source of mutations in head and neck squamous cell carcinoma. Oral Oncol 2017;74:8–14. doi:10.1016/j.oraloncology.2017.09.002.

[25] De Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science (80-) 2014;346:251–6. doi:10.1126/science.1253462.

[26] Wu X, Kabalane H, Kahli M, Petryk N, Laperrousaz B, Jaszczyszyn Y, et al. Developmental and cancer-associated plasticity of DNA replication preferentially targets GC-poor, lowly expressed and late-replicating regions. Nucleic Acids Res 2018;46:10157–72. doi:10.1093/nar/gky797.

[27] Petryk N, Kahli M, D'Aubenton-Carafa Y, Jaszczyszyn Y, Shen Y, Silvain M, et al. Replication landscape of the human genome. Nat Commun 2016;7:10208. doi:10.1038/ncomms10208.

[28] Shi MJ, Meng XY, Lamy P, Banday AR, Yang J, Moreno-Vega A, et al. APOBEC-mediated Mutagenesis as a Likely Cause of FGFR3 S249C Mutation Over-representation in Bladder Cancer. Eur Urol 2019;76:9–13. doi:10.1016/j.eururo.2019.03.032.

[29] Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003;31:3406–15. doi:10.1093/nar/gkg595.

[30] Mathews DH, Sabina J, Zuker M, Turner DH. Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure. J Mol Biol 1999;288:911–40. doi:10.1006/jmbi.1999.2700.

[31] Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell 2018;173:371-385.e18. doi:10.1016/j.cell.2018.02.060.

3.1.3 Discussion

Even though APOBEC-type mutations were much more prevalent within stringent TCW motif (W = A or T) (Alexandrov et al., 2013b; Roberts et al., 2013), multiple lines of evidence (Cannataro et al., 2019; Seplyarskiy et al., 2016; Silvas et al., 2018) as well as our results implied that many mutations within non-canonical TCS motif (S = C or G) were indeed associated with APOBEC mutagenesis and should not be overlooked. Thus, identification of APOBEC-associated mutations in cancer genome should better use 5'-TCN motif. However, the extension to include TCS motifs in the APOBEC signature should be treated with caution, as it is cancer context dependent and interfered by certain conditions: 1) TCC → TTC mutation is the most remarkable characteristic of UV-associated signature (COSMIC S7) that is predominantly found in skin cancer; 2) TCC → TTC mutation rate is increased in Europeans and reflected a population-specific evolution (Harris and Pritchard, 2017); 3) TCG → TTG mutations are frequent in POLE-associated signature (COSMIC S10) or DNA mismatch repair deficiency associated signature (COSMIC S6) that are predominantly found in colorectal and endometrial cancers (Bonneville et al., 2017; Le et al., 2017). Of note, C → T mutation at spontaneously 5-methylated CpG site, which was proposed to be linked to age (COSMIC S1), would not be a major interference for APOBEC-associated TCG → TTG mutations since increasing evidence demonstrated that APOBEC enzymes favor non-methylated TCG sites (Nabel et al., 2012; Seplyarskiy et al., 2016). Consistently, according to our results, all the APOBEC-associated mutations occurring at TCG sites were inversely correlated to age-associated signature and vice versa.

Our strategy for analyzing the association between an given mutation and a mutational signature was original in that we compared APOBEC-mediated mutagenesis between tumours bearing a given candidate hotspot mutation corresponding to APOBEC motifs and tumours without any of such candidate mutations. This comparative strategy is more adapted to systematically identify APOBEC-related mutations than previous studies where comparing tumours harbouring a candidate APOBEC-related mutation with tumours bearing other recurrent mutations (Shi et al., 2019) within the same target gene or wild-type samples (Cannataro et al., 2019; Henderson et al., 2014; Poulos et al., 2018; Temko et al., 2018) regarding APOBEC-mediated mutagenesis. Very recently, Letouzé and colleagues developed a method that enables estimating the probability of each mutation being due to each mutational process considering the mutation category (substitution type and trinucleotide context), and the

number of mutations attributed to each process in the corresponding tumor (Letouzé et al., 2017). This approach is more straightforward and not restricted to the analysis of association between a given mutation and signatures showing distinct trinucleotide motifs such as APOBEC signature as we demonstrated previously. The comparison between both methods could be an interesting future direction.

It is of importance to identify mutagens accounting for APOBEC mutagenesis. It has been shown that the mRNA levels of *APOBEC3A* (A3A) and *APOBEC3B* (A3B) are positively correlated with the enrichment of APOBEC mutagenesis in various cancer types (Lim et al., 2017; Roberts et al., 2013). Nevertheless, this correlation coefficient is relatively low (R^2 are less than 0.15). It may be due to a lack of correlation between APOBEC enzyme activity and mRNA expression levels of *APOBEC3A* and *APOBEC3B* (Buisson et al., 2017). Another possibility is that mutation events may indeed occur when the level of expression of the APOBEC enzymes could have been different, but this time window had been much earlier than the static time you extracted samples and detected mutations when no difference for mRNA expression of APOBEC enzymes could be observed. In addition, we cannot overlook the existence of a potential cooperation network, among APOBEC enzymes and other enzymes, like *REV1* and *UNG* genes, participating in the stepwise reactions of deamination (Helleday et al., 2014). Of note, A3A and A3B may play distinct roles. A3A enzyme is over tenfold more proficient than A3B to generate APOBEC mutations (Chan et al., 2015). A3A and A3B exhibit different preferences for targeted sequence context: A3A favors YTCA sites (Y = C or T), whereas A3B favors RTCA sites (R = A or G) (Chan et al., 2015). Further, A3B is dominantly associated with cancer cell proliferation and induced by DNA-damaging drugs whereas A3A displays specificity for adaptive immunity and induced by antiviral interferon-stimulated response (Middlebrooks et al., 2016; Ng et al., 2019). According to our results, we found the mRNA expression levels of *APOBEC3A* and *APOBEC3H* were significantly associated with APOBEC-associated mutations (including *FGFR3* S249C). In future, I think it will be interesting to explore the factors at the origin of the overexpression of *APOBEC3A* in BCa. Even though virus infection is rare in BCa, as we mentioned previously, under immunosuppression state, BK polyomavirus can be reactivated and thereby induces *APOBEC3A* expression to promote cancer formation. On the other hand, frequent urinary bacterial infection may also interact and induce the expression of APOBEC genes to initiate tumorigenesis.

Another interesting finding is the association between frequent mutations at certain sites on tumor suppressor genes (TSGs) and APOBEC mutagenesis. Frequent mutations on TSGs are uncommon because many sites can provoke loss-of-function except if they act as a dominant-negative mutation as well (Vogelstein et al., 2013). We confirmed that nearly all APOBEC-associated mutations identified on TSGs were loss-of-function in BCa, thus supporting the idea that frequent inactivating mutations could occur on TSGs merely due to a positive selection by the mutagenic process rather than functional advantage. Not coincidently, the identification of APOBEC-associated 'frequent passenger' mutations also reflected a positive selection by APOBEC mutagenesis in BCa.

Mutation signatures associated with homologous recombination deficiency in breast cancer, either by decomposing from WGS or by evaluating a small set of targeted gene panels, have shown potency to predict drug susceptibility to anti-PARP treatment (Davies et al., 2017; Gulhan et al., 2019). Interestingly, increasing *in vitro* evidence has shown that multiple cancer cell lines with high A3A activity are selectively lethal to ATR but not to ATM inhibitors (Buisson et al., 2017; Green et al., 2017). More recently, it has been reported that APOBEC-associated tumors exhibit greater likelihood of complete or partial response for immunotherapy in bladder and head and neck cancer (Miao et al., 2018). Therefore, evaluating the enrichment of APOBEC signature can serve as a preliminary guide to select patient for optimal clinical treatment. However, to calculate the enrichment of APOBEC signature for a given patient, you need to have whole-exome sequencing data first and subsequently perform sophisticated bioinformatic analysis, which is time- and money-consuming as well as technically challengeable. By applying our list of 44 APOBEC-associated mutations, we were able to distinguish not only patients with higher likelihood of enrichment of APOBEC signature but also patients with higher mutation burden from their corresponding counterparts. This small mutation panel is pretty cheap, simple and fast to detect, hereby has a huge potential and advantage for predicting the enrichment of APOBEC signature for a given patient. More interestingly, we found that the prediction for ATRi susceptibility based on current 44 APOBEC-associated mutations was even better than applying the median of either APOBEC fraction score or expression of APOBEC3 family mutagens (data not shown in current thesis). Actually, the median of either APOBEC fraction score or expression of APOBEC3 family mutagens will largely varied depending on the sample size, thus a clear cut-off for clinical translation is far from determined. From this point of view, our mutation panel, represented by dichotomous variables, has distinct advantage for translational significance.

3.2 Transgenic mouse model based on UII-*FGFR3* S249C

3.2.1 Introduction

FGFR3 is one of the most frequently mutated genes in BCa with ~65% of NMIBCs and 15% of MIBCs carry an *FGFR3* mutation (Hedegaard et al., 2016; Robertson et al., 2017). FGFR3 translocations leading to the production of FGFR3-TACC3 and FGFR3-BAIAP2L1 fusion proteins are identified in 3% of MIBCs (Nakanishi et al., 2015a; Williams et al., 2013; Wu et al., 2013). Most of these genetic alterations are gain-of-function and lead to a constitutively active FGFR3, whose oncogenic properties have been well demonstrated *in vitro* (Bernard-Pierrot et al., 2006; Nakanishi et al., 2015a; Williams et al., 2013).

Cancer cell lines based xenografts and patient-derived xenografts (PDXs) are commonly used *in vivo* models in cancer biology. Many of such non-autochthonous models, harbouring gain-of-function genetic alternations of *FGFR3,* have been successfully applied to reveal onco-driver dependency of FGFR3 and its driven signaling pathways (Mahe et al., 2018). Alternatively, autochthonous models such as genetically engineered mice (GEM) have distinct advantages to explore the oncogenic property of specific cancer driver genes. In BCa, several available GEM models exist but many of which are a combined generation with more than one target (John and Said, 2017). Majority of these BCa GEM models have used the mouse Uroplakin II (*UPK2*, UII) promoter, a marker specifically expressed in the urothelium. In the early 2000s, Zhang et al. developed a transgenic mouse model, causing urothelial hyperplasia and superficial papillary NMIBC tumor, by targeting the mutated form of UII-HRas Q61L alone (Zhang et al., 2001). *HRAS* mutations are also frequent in BCa but exclusive to *FGFR3* mutations. Recently, a GEM model engineered with Ad5-CMV-Cre inducible expression of FGFR3-TACC3 fusion has been produced and shown to develop lung cancer (Best et al., 2018). Previous studies failed to observe spontaneous formation of BCa in transgenic mice generated by UII-*FGFR3* S249C or K644E mutation up to a follow-up of 12 and 18 months, respectively (Ahmad et al., 2011; Foth et al., 2018). In fact, to date, there is not yet transgenic mouse generated with oncogenic 'driver' mutations of *FGFR3* that is alone able to induce tumor transformation in bladder or any other cancer types.

In current study, I collaborated and co-contributed with other two PhD students, Aura and Jacqueline. We found mice expressing human UII-*FGFR3* S249C at the urothelium developed

hyperplastic lesions at 6-8 months of age and low-grade papillary tumors from 18 months of age. Tumor frequency was dependent on hFGFR3-S249C zygosity. Immunohistochemical and transcriptomic analyses confirmed that the murine tumors were close to their human counterparts. Thus, we proposed the first GEM model of spontaneous low-grade papillary tumor formation and demonstrated that a mutated *FGFR3* alone can alter urothelium differentiation leading to hyperplasia and act as an oncogene leading to neoplastic transformation *in vivo*. The transformation process was associated with increased genome instability, MYC activation and enhanced angiogenesis probably mediated by hypoxia-inducing factor (*HIF1A*).

3.2.2 Results

Manuscript 3

A mutated-FGFR3 induced mouse model of luminal bladder tumors reveals importance of FGFR3 expression for in vivo oncogenic activity and regulation of HIF1A by FGFR3

Aura Moreno-Vega, **Ming-Jun Shi***, Jacqueline Fontugne*, Florent Dufour, Xiang-Yu Meng, Claire Dunois-Larde, Aurelie Kamoun, Philipe Lamy, Audrey Rapinat, Claire Beraud, Myriam Lasalle, Elodie Chapeaublanc, Anna Almeida, Aurelien De Reynies, Philipe Lluel, Lars Dyrskjøt, Yves Allory, François Radvanyi, and Isabelle Bernard-Pierrot†.*

Prepared for submission

Manuscript 3

A mutated-FGFR3 induced mouse model of luminal bladder tumors reveals importance of FGFR3 expression for in vivo oncogenic activity and regulation of HIF1A by FGFR3

Aura Moreno-Vega[1,2,], Ming-Jun Shi[1,2,3,*], Jacqueline Fontugne[1,2,*], Florent Dufour[1,2], Xiang-Yu Meng[1,2], Claire Dunois-Larde[1,2], Aurelie Kamoun[4], Philipe Lamy[5], Audrey Rapinat[6], Claire Beraud[7], Myriam Lasalle[7], Elodie Chapeaublanc[1,2], Anna Almeid[6], Aurelien De Reynies[4], Philipe LLuel[7], Lars Dyrskjøt[5], Yves Allory[1,8], François Radvanyi[1,2] and Isabelle Bernard-Pierrot[1,2,#]*

[1] Institut Curie, PSL Research University, CNRS, UMR144, Equipe Labellisée Ligue contre le Cancer, Paris, France

[2] Sorbonne Universités, UPMC Université Paris 06, CNRS, UMR144, Paris, France

[3] Department of Urology, Beijing Friendship Hospital, Capital Medical University, Beijing, China

[4] La ligue contre le cancer, Paris, France

[5] Department of Molecular Medicine, Aarhus University Hospital, Aarhus, Denmark

[6] Institut Curie, Department of translational research, Paris, France

[7] Urosphere, Toulouse, France

[8] Institut Curie, Department of Pathology, Saint-Cloud, France

* These authors contributed equally to the work

Correspondence:

Dr Isabelle Bernard-Pierrot
Institut Curie
12 rue Lhomond
75005 Paris
E-mail: isabelle.bernard-pierrot@curie.fr
Tel: +33 1 42 34 63 40, Fax: +33 1 42 34 63 49

Running title: mutated-FGFR3 induce bladder tumors in mice

Abstract

Somatic mutations of the fibroblast growth factor receptor 3 (FGFR3) are one of the most frequent genetic alterations in bladder carcinomas (~70% of cases). They are mostly associated with low-grade papillary non invasive carcinomas and with the luminal papillary subtype of muscle invasive bladder tumors. The oncogenic dependency induced by FGFR3 is an Achile's heel targeted by FGFR3 inhibitors in the clinic, but the tumorigenicity of a mutated FGFR3 has never been demonstrated *in vivo*. We report here that the mutated FGFR3 expression in urothelial cells of transgenic mice induces urothelial hyperplasia and spontaneous genomically unstable low-grade papillary tumor formation. Transcriptomic analyses of these tumors shows they resemble the human counterparts and highlights the activation of a HIF1A/VEGFA pathway in FGFR3-mutated mice and human tumors. Inhibition of FGFR3 in human cell line xenografts and patient derived xenografts expressing a mutated-FGFR3 confirmed the regulation of HIF1A by FGFR3, which could contribute to FGFR3-induced tumorigenesis by favoring angiogenesis. Strategies to inhibit HIF1A in association with anti-FGFR3 therapies may be of interest for FGFR3-mutated bladder tumors.

Significance: Our study represents the first murine model of FGFR3-induced spontaneous bladder carcinomas, demonstrating the tumorigenicity of FGFR3 mutations *in vivo*. This model of FGFR3-mutated tumors resembles the human counterparts and should allow a better understanding of FGFR3 oncogenic properties. This model shed light on a HIF1A activation by FGFR3, which could contribute to tumor formation through the regulation of angiogenesis and possibly cell metabolism.

Introduction

Bladder cancer is the sixth most common cancer in men worldwide, with an even higher incidence in European countries and North America (4th most common cancer in men) (1). At first diagnosis, the majority of tumors are non-muscle-invasive urothelial carcinomas (NMIBC) (70%). In spite of their favorable prognosis, NMIBCs have a high recurrence rate (70%) and are able to progress (10-15%) to the more aggressive form of disease, muscle-invasive bladder carcinoma (MIBC). Different molecular classifications have been established in both NMIBC and MIBC in order to identify different biological processes to support patient stratification and more adapted therapies (2–6).

Fibroblast growth factor receptor 3 (FGFR3) is a tyrosine kinase receptor with frequent genetic alterations in bladder cancer (3,5,7). Point mutations (observed in ~70 % of NMIBC and 15% of MIBC) or chromosomal translocations (affecting ~5% of MIBC) resulting in protein fusions, lead to a constitutively active FGFR3. The oncogenic properties of the altered FGFR3 have been shown in vitro andan FGFR3 oncogenic dependency for tumor growth as demonstrated in vitro as well as in vivo(cell lines or patient derived xenografts) (8–12). Several clinical trials have shown a clinical benefit of FGFR3 inhibition in terms of patient survival (NCT02365597; NCT03473743 and NCT03390504), which has recently led to the FDA approval of the first FGFR inhibitor Erdafinitib, as a treatment of patients with locally advanced or metastatic bladder cancer with an FGFRs alteration.

To determine the functional role of mutated-FGFR3 in bladder cancer in vivo, several teams have developed FGFR3-altered genetically engineered mice (GEM). So far, results suggest that although FGFR3 activation alone is not sufficient to induce tumorigenesis (13–16), it can promote tumor formation when associated with other molecular alterations (p53/pRB deficiency (16); PTEN loss (15)) or carcinogen treatment (14).

In this study, we report for the first time a GEM model overexpressing the human FGFR3b-S249C mutant specifically in the urothelium, in which mice developed both hyperplasic lesions and low-grade papillary bladder carcinomas, with genomic instability, resembling human luminal papillary tumors at the histological and transcriptomic level. This model allowed us to identify an FGFR3-mediated regulation of HIF1A that could contribute to FGFR3-induced tumorigenesis by favoring angiogenesis.

Results

FGFR3-S249C expression in Uroplakin2-expressing cells induces urothelial hyperplasia and non-muscle-invasive low-grade urothelial carcinoma.

To determine the role of a constitutively activated mutated FGFR3 in bladder tumorigenesis, we generated transgenic mice expressing a mutated receptor in the urothelium. We focused on the FGFR3-S249C mutation, the most common FGFR3 mutation in both NMIBC and MIBC (7), and used the uroplakin II gene promoter to target its expression in urothelial cells (Fig.1A). We selected two founders, numbers 569 and 538 that expressed the highest level of the human FGFR3 transgene in the urothelium as evidenced by RT-qPCR (Supplementary Fig 1A). In situ hybridization using a human FGFR3-specific probe showed expression of hFGFR3 mRNA in the supra-basal and intermediate cell layers and in very few basal cells of the urothelium (Supplementary Fig.1B). Moreover, human FGFR3 mRNA expression levels in the urothelium were respectively 4 and 1.5-fold higher than the level of endogenous mouse Fgfr3 in founders 569 and 538, respectively, as assessed by radioactive PCR (Supplementary Fig 1C). These two founders were viable and fertile and transmitted the transgene to their offspring in a Mendelian fashion. Following propagation of founder lines, we examined the bladder of transgenic mice aged 1 to 24 months old. Histological analysis showed hyperplastic lesions defined by a thickened urothelium, with an increase in cell layers, lacking cytologic atypia. The penetrance of the phenotype was complete from 6 months of age in both lines. UII-hFGFR3-S249C mice urothelium exhibited seven to ten cell layers and focally more (ten to twenty) at 18 months (Fig.1B). In contrast, normal mouse urothelium presented only three to four cell layers (Fig.1B). Macroscopically, focal papillary lesions were observed after 15 months with a low penetrance in both lines (~10% and 4% for L569 and L538, respectively). Histological analysis of these lesions revealed they were superficial low-grade papillary urothelium carcinomas. Carcinomas displayed a papillary tumor architecture, characterized by either exophytic or mixed (exophytic and inverted) growth patterns, and low-grade tumor cell cytology, with homogeneous nuclei size (Fig.1B). We focused then on the L569 line presenting a higher penetrance of the phenotype and further characterized these lesions. Hyperplastic lesions were similar to normal urothelium in terms of proliferation, as measured by Mki67 expression levels (Fig.1C), and in terms of transcriptomic profile, determined using Affymetrix mouse exon array (Fig.1D). In contrast, tumors presented a significantly higher proliferation rate (Fig.1C) and principal component analysis highlighted a different transcriptomic profile compared to normal and hyperplastic urothelium (Fig. 1D). In good agreement with low-grade tumors, although higher than in normal urothelium or

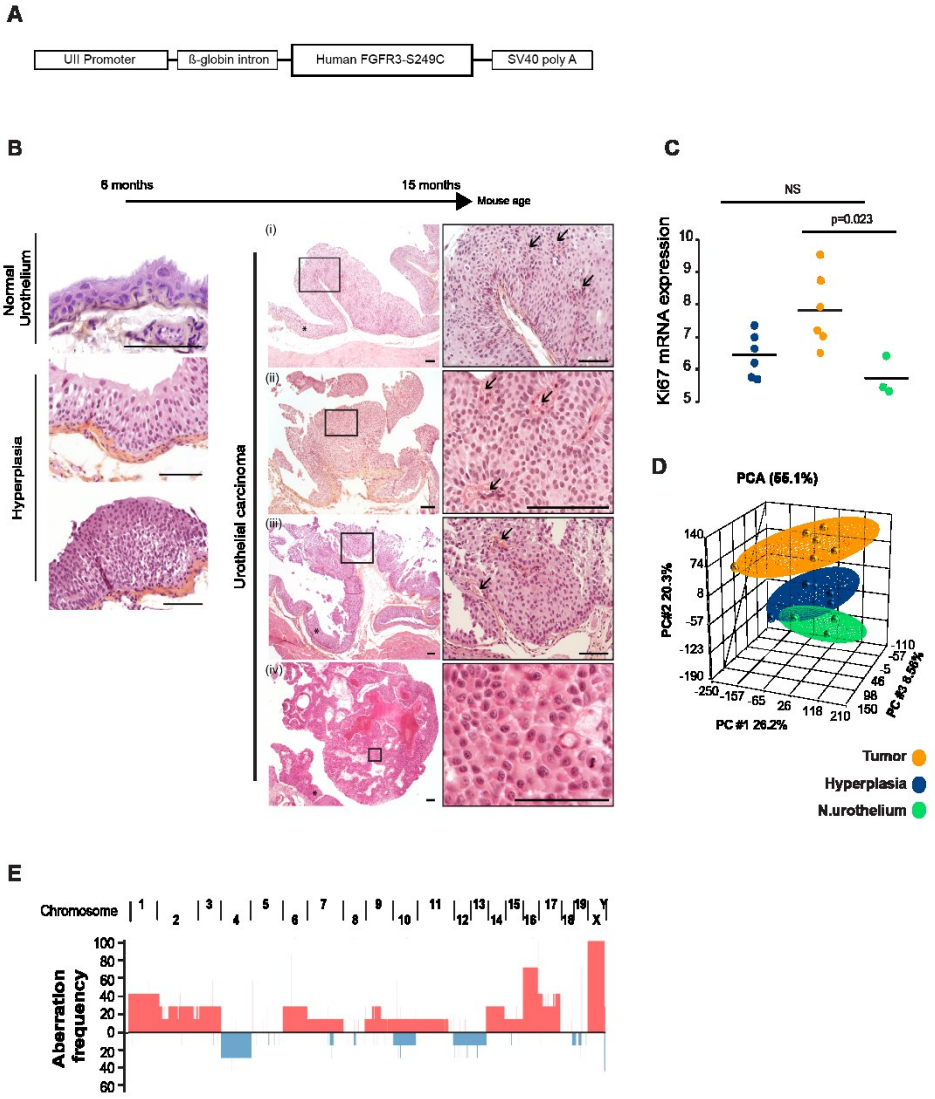


Figure 1. *UII-FGFR3-S249C transgenic mice develop urothelial hyperplasia and non-muscle-invasive low-grade urothelial carcinoma.* ***A)*** *Chimeric construct used to generate transgenic mice, consisting of a 3.6-kb mouse UII gene promoter and a 2.1-kb human FGFR3b cDNA carrying the mutation S249C.* ***B)*** *Representative H&E histology of urothelial lesions in hFGFR3-S249C mice. Hyperplastic lesions (left panel) or low grade papillary urothelial carcinomas (right panel) developed in hFGFR3-S249C mice from 6 months and 15 months of age, respectively. Stars show tumor-adjacent urothelial hyperplasia. Arrows point to papillae fibrovascular cores. Scale bar: 100µm.* ***C)*** *mKi67 mRNA expression levels (Affymetrix Mouse Exon 1.0 ST. Array signal) in tumor and hyperplastic urothelium from UII-hFGFR3-S249C mice and in normal urothelium from control littermates.* ***D)*** *Principal component analysis of all genes expressed on the Affymetrix Mouse Exon 1.0 ST. Array from tumor and hyperplastic urothelium from UII- hFGFR3-S249C mice and from normal urothelial samples from control littermate mice (n= 6 tumors, 6 hyperplastic lesions, 3 normal urothelium).* ***E)*** *Frequency of*

chromosomal copy number alterations in tumors from UII-hFGFR3 S249C. (red = gain; blue= loss).

hyperplasia, proliferation rate in tumors was low, with <10% of Ki67-labelled cells by immunohistochemistry (Supplementary Fig.2A). Whole Exome sequencing analysis for 7 tumors did not reveal any recurrent mutations induced by hFGFR3-249C expression but showed recurrent copy number alterations, the most common being chromosome 16 amplification in 5 out of 7 tumors (Fig.1E). We selected 3 genes (Trat1, Erbb4, Fkbp5) located in 3 amplified regions (chr16, chr1, chr17, respectively) and verified their recurrent amplification by qPCR on genomic DNA, in the tumors previously analyzed by whole exome sequencing and in 4 additional tumors (Supplementary Fig.2B). Taken together, our results showed that hFGFR3-S249C is oncogenic in vivo, inducing genomic instability leading to tumor formation in bladder urothelium.

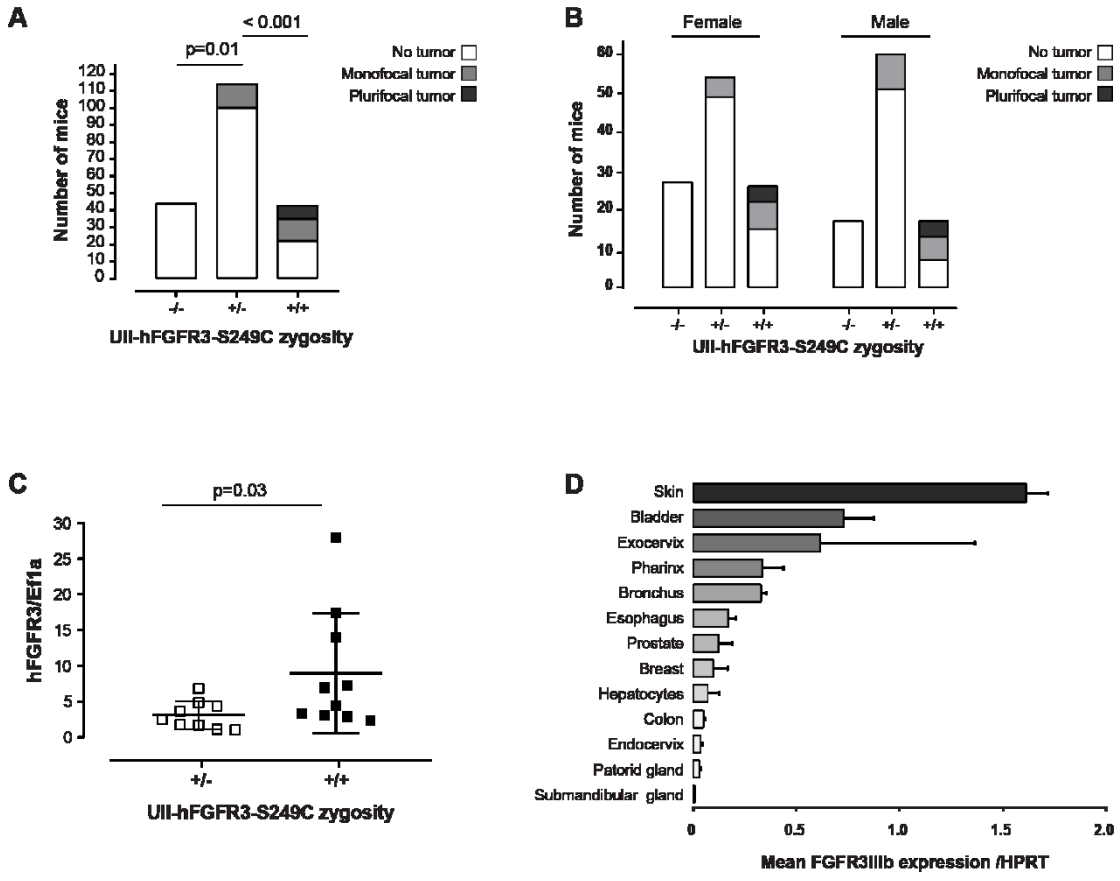


Figure 2. *FGFR3-induced tumor development is dependent on FGFR3 expression levels.* ***A)*** *Frequency of unifocal or plurifocal bladder tumor development in hFGFR3-S249C homozygous (+/+) or heterozygous (+/-) mice versus control littermates (-/-).* ***B)*** *Frequency of*

bladder tumors in male and female UII-hFGFR3-S249C mice. A-B-Proportions were compared using Fisher exact test. ***C)*** *hFGFR3 mRNA expression evaluated by RT-qPCR in hFGFR3 S249C homozygous (+/+) or heterozygous (+/-) for mice . Results were normalized using EF1a expression levels. The statistical significance of differences was assessed using Wilcoxon test.* ***D)*** *FGFR3 expression levels as assessed by RT-qPCR in different epithelia obtained after laser-microdissection. Results were normalized using TBP expression levels.*

FGFR3 expression level impacts tumor formation in UII-hFGFR3-S249C mice and could account for tissue specificity of FGFR3-activated induced tumors.

We then studied an important series of mice (N= 402) and compared the frequency of tumors in 18-month-old UII-hFGFR3-S249C heterozygous and homozygous mice. The frequency of tumors was significantly higher in homozygous compared to heterozygous mice (~ 40% and 10%, respectively) (Fig. 2A) and we did not observe any significant difference between males and females (Fig.2B). Strikingly, plurifocal tumors were specifically identified in homozygous mice, whereas heterozygous mice only developed unifocal tumors, re-enforcing the fact that the urothelium of homozygous mice was more sensitive to spontaneous tumor development than heterozygous mice. We hypothesized that the increased sensitivity to tumor development could be linked to the significantly higher expression level of hFGFR3 in homozygous compared to heterozygous mice, as assessed by RT-qPCR (Fig.2C). Following this hypothesis, we measured FGFR3 expression levels in different normal epithelia, including urothelium, obtained after microdissection. Interestingly, epithelia presenting high expression levels of FGFR3 were those in which FGFR3-mutated tumors are described (bladder, skin, exocervix) (Fig.2D) (3, 17–19). Our data suggest that FGFR3 mutations require an epithelium with a high expression of FGFR3 to induce tumor formation. Nevertheless, although FGFR3 gene dosage in mice influenced tumor frequency, it did not reduce tumor development latency or induce progression towards muscle invasive bladder tumors. No histopathological difference was observed between hFGFR3-S249C- induced tumors from heterozygous or homozygous mice.

UII-hFGFR3-S249C model is a luminal papillary model of human bladder cancer.

Given the papillary nature of hFGFR3-S249C-induced tumors, we hypothesized that they recapitulate a luminal-like human bladder cancer molecular phenotype. We and others previously showed that N-Butyl-N (4-hysdroxybutyl) (BBN)-induced bladder tumors represent

a model of basal-like bladder tumors (20,21). To classify hFGFR3-S249C tumors, we first applied a molecular classifier allowing to distinguish between 3 classes of NMIBC (5). The 6 hFGFR3-S249C-induced tumors had the highest correlations to the class 1 centroid of gene expression (Fig.3A). Consistent with our histological analysis, this class of NMIBC is enriched in low-grade, low-risk, papillary, differentiated tumors (5). We also applied the BASE47 classifier to distinguish between luminal and basal bladder tumor subtypes (22). FGFR3-induced bladder tumors classified as luminal, whereas our previously obtained BBN-induced tumors (20) classified as basal subtype (Fig.3B). To further validate our results, we performed a cross-species comparison study by co-clustering the hFGFR3-S249C and BBN mice tumors with human tumors from our CIT cohort (n=96 MIBC and 99 NMIBC) (23), using genes from a recently developed consensus classifier for basal and luminal-papillary human bladder tumors (2) with the corresponding orthologues across the species. We found that hFGFR3-S249C and BBN tumors co-clustered with human luminal papillary and basal-like tumors, respectively (Fig.3C). These results are in good agreement with luminal papillary tumors being enriched in FGFR3 mutations. Class 1 of NIMBC are characterized by a lower immune response and infiltrating immune cells activity than class 2 tumors (5) and luminal-papillary MIBC also display lower immune infiltration signals than basal tumors (2). Using Micro-environment Cell Population counter methods and transcriptomic data, we estimated the immune cell infiltration in hFGFR3-S249C and BBN tumors (24). Consistent with human tumors, we estimated a weak infiltration of hFGFR3-S249C low-grade luminal papillary tumors by any type of immune cells, whereas BBN basal tumors presented a higher infiltration by monocytes, neutrophils, T and cytotoxic lymphocytes (Fig.3D). The low infiltration of hFGFR3-S249C tumors by immune cells is consistent with FGFR3 mutations synergizing with BBN by suppressing acute inflammation (14). We confirmed that hFGFR3-S249C expression promotes BBN-induced tumor formation and showed that BBN-hFGFR3-S249C tumors retained features of the basal molecular subtype (supplementary Fig.3). Taken together, our data suggest that hFGFR3-S249C mouse tumors recapitulate the human luminal papillary subtype of bladder tumors and could be a useful model to decipher the role of FGFR3 in bladder tumor formation.

VEGFA expression and hypoxia/HIF1A activation are increased in hFGFR3-S249C-induced mouse tumors and in human FGFR3-mutated tumors.

To unravel the molecular mechanisms that could trigger the progression from hyperplasia to low-grade papillary carcinoma, we compared transcriptomic data from 6 hyperplasias and 6 tumors using the LIMMA algorithm. We observed 989 differentially

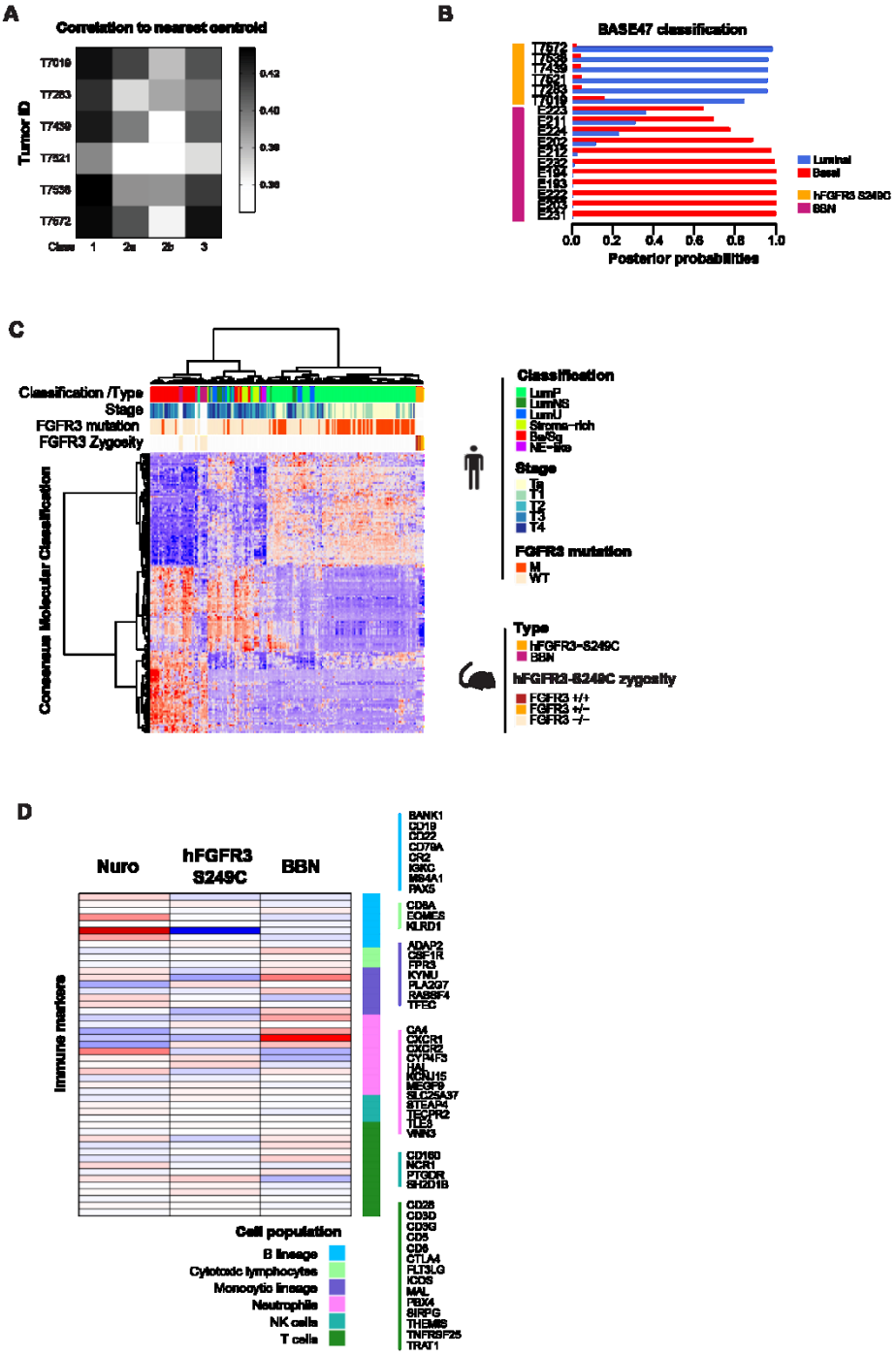


Figure 3. *Mouse hFGFR3-S249C recapitulates the human luminal papillary molecular subtypes of bladder cancer.* ***A)*** *Correlation to the 3 Non-Muscle Invasive Carcinoma classes centroids and* ***B)*** *to the basal centroid (BASE47 classifier) for BBN and/or hFGFR3-S249C.* ***C)*** *Cross-species, unsupervised hierarchical clustering of mice hFGFR3 S249C tumors (n=6) and human bladder tumors across genes in a consensus classifier for basal and luminal tumors.* ***D)*** *Heatmap of MCP counter signature for estimation of infiltration of different immune populations based on transcriptomic data from hFGFR3 S249C tumors (n=6). For panels C*

and D, Red indicates high and blue indicates low mRNA expression respectively (normalized mRNA expression levels).

expressed genes (Supplementary Fig. 4A) that were enriched in pathways or biological processes related to cell adhesion and migration but also to angiogenesis (Fig. 4A and supplementary Fig.4B). As previous studies have reported a potential link between FGFR3 and angiogenesis in multiple myeloma (mediated by VEGFA, vascular endothelial growth factor), bladder cancer and liver cancer (25,26) (Liu et al., 2016, Bertz et al., 2014), we further explored this pathway. Among angiogenesis-related genes, Vegfa was over-expressed in tumors compared to hyperplasia (Log2FC = 1.1; p-value = 8.09E-3). VEGFA is often induced by HIF1A under hypoxic conditions. Using a hypoxia activity score based on the analysis of the expression levels of 120 hypoxia-related genes (see methods), we confirmed that hFGFR3-S249C tumors were more hypoxic than hyperplastic urothelia (Fig.4B). Using Ingenuity Pathway software, Hif1a transcription factor was predicted as an upstream regulator in hFGFR3-S249C tumors compared to hyperplasia (p-value = 2.25E-6). Western-blot analysis confirmed the overexpression of Vegfa and Hif1a at the protein level in hFGFR3-S249C tumors compared to control littermate urothelium and to hFGFR3-S249C hyperplasic urothelium (Fig.4C). Western-bot analysis also supported the activation of c-Myc by FGFR3 in mouse tumors as we recently described in human FGFR3-mutated models in vitro and in vivo (12). Our data suggest therefore that FGFR3-S249C-induced tumor formation relies on c-Myc activation, inducing cell hyper-proliferation, and at least partially on a Hif1a/Vegfa-induced angiogenesis. Interestingly, we also identified a potential relevance of our hypothesis in human tumors. Hypoxia activity score (Fig.4D), VEGFA and carbonic anhydrase 9 (CA9, another HIF1A target gene and surrogate marker of angiogenesis) expression levels were significantly higher in FGFR3-mutated tumors compared to wild-type FGFR3 tumors in publicly available data for 417 NMIBCs (5) (Fig.4E, left panel) and/or in 128 luminal papillary MIBC from TCGA (3) (Fig.4E, right panel).

Mutated-FGFR3 regulates HIF1A and VEGFA protein expression levels.

Due to the low penetrance and high latency of the phenotype in our UII-hFGFR3-S249C model, we took advantage of cancer bladder cell line-derived (MGH-U3) and patient-derived xenograft (F659) models, harboring the most frequent FGFR3 activating mutations (Y375C and S249C, respectively), to modulate FGFR3 activity. Each model was treated with a pan-FGFR inhibitor that repressed tumor growth (PD173074 or BGJ398, respectively) (12).

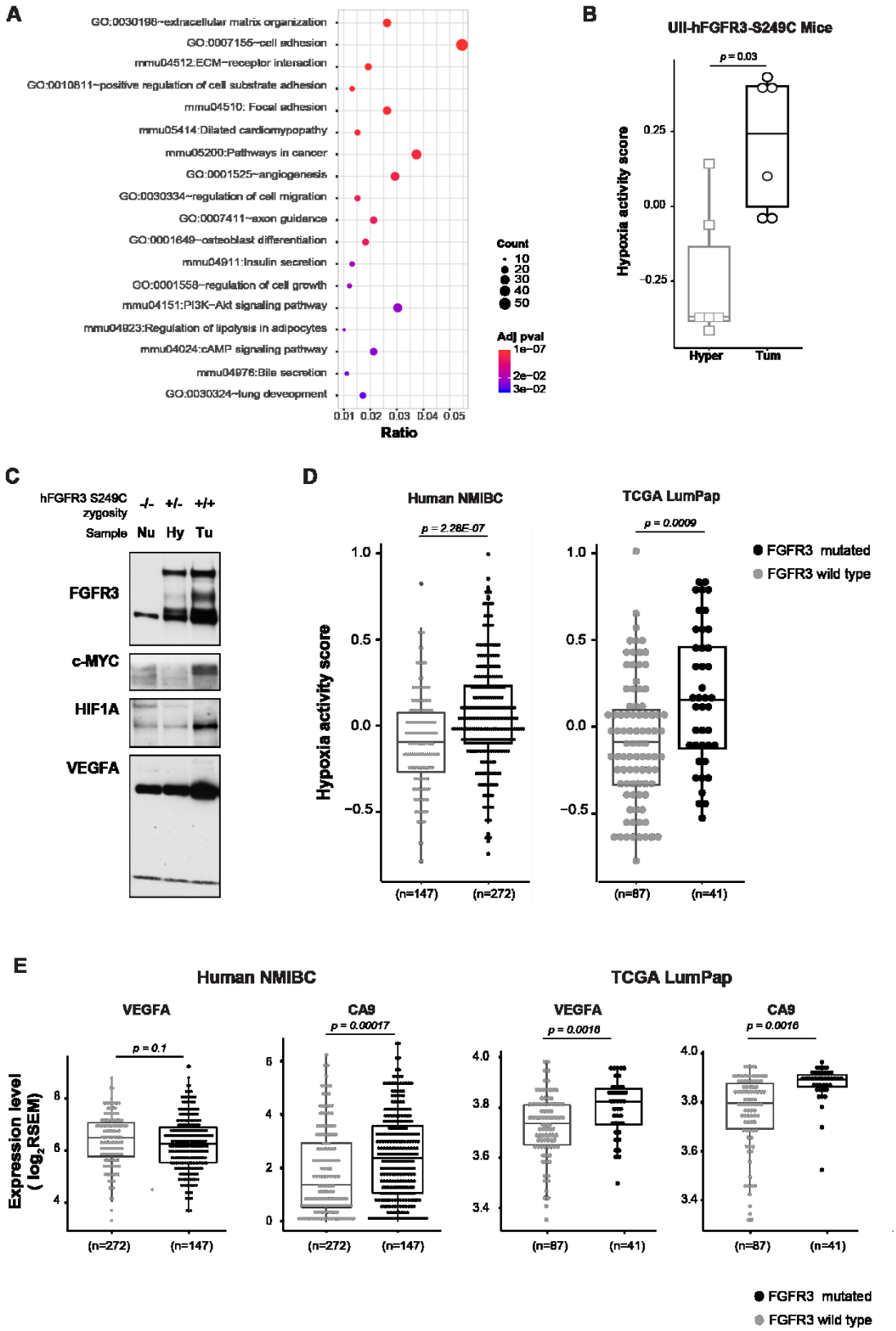


Figure 4. *High HIF1A activity and VEGFA expression are associated with mutated-FGFR3 is mouse and human tumors.* ***A)*** *Plot of top 18 Gene Ontology Biological Processes (GO) and murine KEGG (KE) deregulated pathways usinga set of 989 differentially expressed genes*

obtained by comparison of hFGFR3 S249C tumors and normal mouse urothelium (|log2FC|>0.58; adj.p-val<0.05) The adjusted p-value of each enriched term, as well as the number of genes assigned to each term (count) and the ratio of assigned genes to total number of genes belonging to a term are displayed). ***B)*** *Hypoxia activity prediction derived from Gene Set Variation Analysis (GSVA) enrichment scores calculated using transcriptomic data of tumor and hyperplastic urothelia from UII-hFGFR3-S249C mice (n = 6 tumors, 6 hyperplasic lesions, hypoxia gene signature = 120 genes).* ***C)*** *Western blot comparing FGFR3, c-MYC, VEGFA and HIF1 expression levels in tumor and hyperplastic urothelium in UII-hFGFR3-S249C mice and in normal urothelium from control littermates.* ***D-E)*** *Hypoxia activity score derived from GSVA enrichment scores (D) and expression levels of two hypoxia inducible genes involved in angiogenesis (VEGFA and CA9) (E) determined using RNA-seq data from a published cohort of 416 NMIBCs (5) and of XX luminal papillary MIBCs from TCGA (3). Hypoxia activity scores and expression levels between wild type and FGFR3 mutated tumors were compared using Wilcoxon rank sum test.*

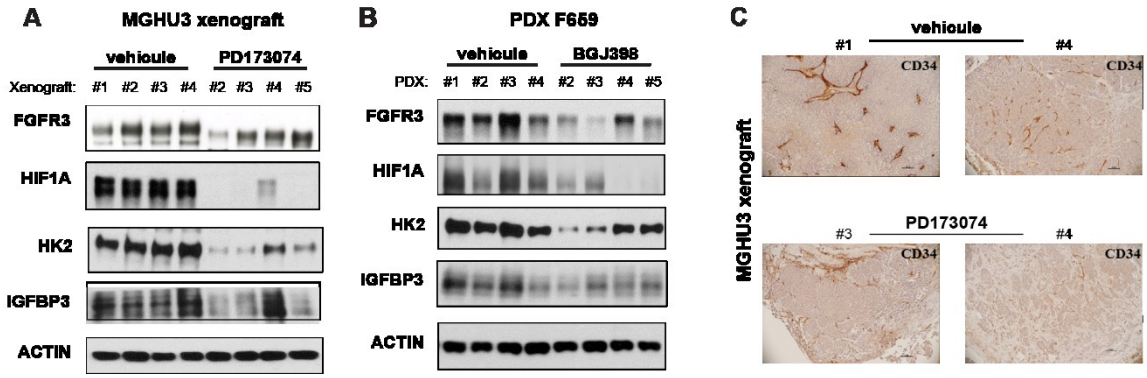


Figure 5. *Inhibition of tumor growth induced by FGFR3 inhibitors is associated with a decrease of HIF1A level and reduced angiogenesis.* ***A)*** *Western-blot evaluation of FGFR3, HIF1A, IGFBP3 and HK2 expression levels upon FGFR3 inhibition in MGH-U3 xenografts expressing FGFR3-Y375C obtained from mice treated for 9 days with a vehicle or an FGFR3 inhibitor (PD173074; 25mg/kg/day) and* ***B)*** *in Mice bearing PDX (F659-model) expressing an FGFR3-S249C mutation, treated for 4 days with a vehicle or an FGFR3 inhibitor (BGJ398; 30mg/kg/day).* ***C)*** *CD34 labelling of endothelial cells in sections from MGH-U3 xenografts with or without pan-FGFR inhibition (PD173074).*

To determine whether HIF1A activation and subsequent VEGFA expression/ regulation of angiogenesis was directly linked to FGFR3 activation, we couldn't use easily this

to modulate FGFR3 activity due. In both models, western-blot analysis showed that FGFR3-inbition induced a loss of HIF1A and IGFBP3, also involved in angiogenesis regulation, (Fig.5 A and 5B). Endothelial cell labelling with CD34 antibody by immunohistochemistry showed a decreased number/size of blood vessels in MGH-U3 xenografts treated with a pan-FGFR inhibitor, supporting the role of an FGFR3/HIF1A-mediated angiogenesis for tumor growth (Fig. 5C). Notwithstanding, we cannot exclude that HIF1A down-regulation could also impact cell metabolism, possibly leading to an impair of tumor growth. Indeed, we also observed an inhibition of HK2 expression, another key target gene of HIF1A involved in metabolism regulation upon FGFR inhibition (Fig.5A and 5B).

Discussion

We described here the first transgenic mouse model demonstrating a tumorigenic activity in vivo of a mutated-FGFR3. Expression of an FGFR3-S249C in uroplakin-2 expressing cells induces spontaneous low-grade papillary tumor formation and favors BBN carcinogen-induced tumor development. We observed FGFR3-S249C-induced tumor formation in two different transgenic lines suggesting that the observed effect was likely induced by the transgene expression itself rather than to an alteration of an endogenous key gene due to the insertion of the transgene. Surprisingly, mutated-FGFR3 has already been targeted to urothelial cells using the same promoter without any spontaneous tumor formation being observed (13–16). However, expression of the mutated receptor did promote bladder tumor formation when induced by exposure to carcinogen (BBN) (14) or in collaboration with Pten loss (15) or P53/pRB deficiency (16). This discrepancy between the previously developed GEM models and our GEM model could be linked to the FGFR3 mutation considered (S249C here, K644E in two previous studies (13,15) or to the use of an inducible model for the expression of FGFR3-S249C in the study by Zhou et al. and Foth et al. (14,16). We have additionally shown that FGFR3-S249C expression levels impact the frequency of tumor formation, suggesting that a lower expression of the transgene in the previously developed GEM could also account for the absence of tumor formation in these models. We used here the most frequent mutation of FGFR3 in bladder cancer but we have recently shown that this overrepresentation of the mutation was likely due to APOBEC mutagenesis rather than an increased tumorigenicity of such mutation as compared to other FGFR3 mutations (7). We can therefore suppose that other FGFR3-mutants would induce bladder tumor formation as well.

Recently, the first pan-FGFR inhibitor – Erdafitinib/Balversa – has been approved by the FDA for patients with locally advanced or metastatic BCa. Considering the increasing interest of targeting FGFR3 for BCa treatment, having a model; such as ours, that resembles human counterparts at histological and transcriptomic levels may have clinical translational value to evaluate drug response and to understand acquired drug resistance mechanisms. In particular, the model we present here is the first immunocompetent model for FGFR3-mutated carcinomas. FGFR3 mutated tumors are non-T cell inflamed and have been associated to a poor immune-infiltrated immune-environment being therefore less prone to respond to immunotherapy (3,27,28). To confirm this hypothesis, a phase 1b/2 clinical trial (NCT03123055) comparing the efficacy of an anti-FGFR3 therapy (B-701, specific monoclonal antibody targeting FGFR3) coupled with immunotherapy (pembrolizumab) in advanced BCa patients presenting an altered FGFR3 is ongoing. In line with the literature, our GEM model showed poor infiltration of different immune cell populations. Hence, this model, or allografts obtained from this model, should help a better understanding of immune-escape or immune-suppression mechanisms driven by a mutated/active FGFR3 and allow evaluation of combined therapies using FGFR and check-point inhibitors.

This model of FGFR3-induced tumors should also allow for a better understanding of the signaling pathways activated by FGFR3 during tumor progression. Targeting simultaneously different proteins forming part of the same signaling pathway could increase treatment efficacy and limit resistance as observed with the combination of B-Raf and MEK inhibitors for the treatment of melanoma (29). We recently demonstrated that MYC activation was crucial for FGFR3 oncogenic activities, pointing to a positive feedback loop of potential therapeutic value in bladder cancer (12). Our GEM model confirmed the stabilization of the c-MYC protein by FGFR3, which could contribute to an FGFR3-induced tumorigenesis through the promotion of cell hyperproliferation. Our model also highlighted a key activation of a HIF1A/VEGFA axis by FGFR3 that could contribute to an FGFR3-induced tumorigenicity via an increased angiogenesis. Such discoveries are in good agreement with the immunohistochemistry staining of vascularization markers in 61 human BCa showing that increased angiogenesis is more common in non-muscle-invasive tumors compared to muscle-invasive cancer, and is also positively correlated with FGFR3 expression and mutation status (30). In hepatocellular carcinoma, FGFR3 has also been shown to promote angiogenesis-dependent metastasis via the facilitation of MCP1 mediated vascular formation (31). It remains thus questionable whether HIF1A alone drives FGFR3-regulated angiogenesis and whether

HIF1A solely participates in the regulation of angiogenesis without being involved in the regulation of other biological processes such as metabolism. Nonetheless, therapeutic strategies targeting HIF1A and FGFR3 are worth being evaluated for FGFR3-mutated bladder tumors.

Methods

Mouse models

All animals were housed and cared for in accordance with the institutional guidelines of the French National Ethics Committee (Ministère de l'Agriculture et de la Forêt, Direction de la Santé et de la Protection Animale, Paris, France). All experiments were reviewed and approved by the institute curie Animal Care and Use Committee.

Generation of UII-hFGFR3-S249C transgenic mice

The expression of a human FGFR3IIIb carrying the S249C mutation was targeted to the urothelium of mice by using the 5' regulatory region of the mouse uroplakin II promoter. The UII-FGFR3b-S249C construct was obtained by inserting the 3.6 kb murine uroplakin II promoter (UII) (Lin et al., 1995) excised with SalI and BamHI into the same restriction sites of the vector containing the β-globin intron 2 and the 3' polyadenylation sequences of SV40 (Ramirez et al., 1994) followed by the insertion of a human S249C mutated FGFR3 cDNA excised with XbaI and HindIII into the SmaI site of this vector. All PCR-generated segments were verified by sequencing both strands. The UII-hFGFR3b-S249C constructs excised with KpnI were purified and microinjected into fertilized B6D2 oocytes. Genomic DNA was extracted from mouse tails and screened by PCR for integration of the transgene. Two lines were selected, L569 and L538, and mice were back-crossed five times to a C57BL/6J mice. Mice were of a mixed background and littermates were used as control. Bladder from mice aged 1 to 24 months were examined for macroscopic lesions followed by a histopathological analysis when required. Mice were then intercrossed to obtain hetero-and homozygous mice for the transgene.

Carcinogen treatment

BBN (N-butyl-N-(4-hydroxybutyl)-nitrosamine) was purchased from Tokyo Kasei Kogyo (Tokyo, Japan). Animals were housed in plastic cages in a controlled-environment room maintained at 22°C ±1°C with 12h light-12h dark cycles. All animals received food ad libitum. The UPII-hFGFR3-S249C mice and control mice

were aged 8-10 weeks old at the time of first carcinogen administration. The BBN was diluted at 0.05% in drinking water (ad libitum) for 8 weeks (the BBN solution was freshly prepared every 2-3 days). After withdrawal of BBN administration, drinking water without added chemicals was available ad libitum. Tumor formation and progression was followed weekly by echography. Mice were sacrificed when tumors reached 80% of bladder volume or when weight moss was greater than 20% of body weight.

Xenograft and PDX models

We used an MGH-U3 bladder cancer cell line-derived xenograft and a PDX model that were previously established in our laboratory (12).

RT-qPCR analyses

Total RNA from mouse urothelium was obtained using the Rneasy mini kit (Qiagen, Courtaboeuf, France) according to the manufacturer's instructions. One µg of total RNA was reverse transcribed using random hexamers (20 pmol) and 200 units of MMLV reverse transcriptase. The expression levels of the human FGFR3 transgene in urothelium and other tissues of transgenic mice were determined by real time PCR analysis. The mouse Ef1a gene was used as a control gene. Quantitative real time PCR was performed using a SYBR green PCR master Mix according to the manufacturer's instructions (Applied Biosystems, Foster city, CA, USA), on an ABI prism 7900 sequence detection system (Applied Biosystems). FGFR3 expression levels were calculated using the comparative Ct method normalized to Ef1a mRNA expression levels. The sequence of these primers used were as follow:

FGFR3, forward 5'-AGTCCTGGATCAGTGAGAG-3',
reverse 5'-CTGCTCGGGGCCCGTGAACG-3';
Eef1a1 forward 5'-CTGGAGCCAAGTGCTAATATGCC-3',
reverse 5'-GCCAGGCTTGAGAACACCAGTC-3'

Radioactive PCR

To compare the relative expression of the FGFR3 transgene to that of the endogenous murine Fgfr3 in the transgenic urothelium, transgenic urothelium cDNA was amplified in presence of 32P dCTP using the primers forward 5'-GCAGGCATCCTCAGCTAC-3' and reverse 5'-TGGACTCGACCGGAGCGTA-3' which recognized both human and mouse FGFR3. The 107 bp amplified products were then digested with RsAI and HinP1I. The human amplified product possesses a RsAI restriction site and the mouse amplified product a HinP1I restriction site. After digestion, two fragments of 88 bp and 19 bp were obtained from the amplified human FGFR3 cDNA and two fragments of 59 bp and 48 bp were obtained from the amplified mouse Fgfr3 cDNA. The digested products were subjected to polyacrylamide gel

electrophoresis and the intensities of the bands were quantified with a Molecular Dynamics Storm PhosphorImager (Molecular Dynamics/Amersham, Sunnyvale, CA, USA).

Histological and immunohistochemical analyses

UPII-hFGFR3-S249C mutant and control mice bladders were fixed in 10 % formalin, embedded in paraffin and cut at 4-µm thick slides for histological and immunohistochemical analyses. Histological hematoxylin and eosin-stained (H&E) slides were reviewed by two genito-urinary pathologists. Immunohistochemical evaluation of the proliferation rate was performed using a polyclonal rabbit anti-Ki67 antibody (1:2500, Abcam ab15580) using a BOND RX automated stainer (Leica Biosystems). Vascularization was evaluated by staining endothelial cells of blood vessels with anti-CD34 antibody (1:200, Abcam, cat.no. ab81289).

Whole exome sequencing and identification of copy number alterations

DNA from UII-hFGFR3-S249C mouse normal urothelium, hyperplastic urothelia and urothelial carcinomas was extracted using phenol-chlororoform. Whole exome libraries were prepared by Integragen (Evry, France)

Raw sequence alignment and variant calling were carried out using Illumina CASAVA 1.8 software (mm10 mouse reference genome). Each variant was annotated according to its presence in the 1000Genome, Exome Variant Server (EVS) or Integragen database, and according to its functional category (synonymous, missense, nonsense, splice variant, frameshift or in-frame indels). Reliable somatic variants were identified as those having a sequencing depth in ≥10 reads in tumor and normal urothelium samples, with ≥3 variant calls representing ≥15% total reads in the tumor, ≤1 variant calls representing <5% total reads in the normal urothelium, and a QPHRED score ≥20 for both SNP detection and genotype calling (≥30 for indels).

Copy number alterations (CNAs) were identified using coverage data to calculate the log ratio of the coverage in each tumor sample as compared to a normal urothelium sample. Log-ratio profiles were then smoothed using the circular binary segmentation algorithm as implemented in the Bioconductor package DNAcopy. The most frequent smoothed value was considered to be the zero level of each sample. Segments with a smoothed log ratio above zero + 0.15 or below zero − 0.15 were considered to have gains and deletions, respectively. High-level amplification and homozygous deletion thresholds were defined as the mean +7 s.d. of smoothed log ratios in regions with gains and deletions, respectively.

The identified frequent chromosomal gains or deletions were further validated by qPCR using genomic DNA. Primers targeting exonic regions from different genes found in the most frequently altered chromosomes were designed. A Taqman qPCR (Applied Biosystems) was carried out on gDNA to compare expression levels between normal urothelium and tumors from UII-hFGFR3-S249C mice. Normalization was performed using genes present on chromosomes without genomic alteration. The designed primers were the following:

Most frequently altered chromosomes			
		Primer sequence 5′ - 3′	
Gene	*Exon*	*Forward*	*Reverse*
Erbb4	26	TGCAACGGCTGAGATGTTT	GTGCCACTGGCTTTCGTAG
Trat1	6	GGCCCAGGAAACAGAATACTAA	GAGAAACGTTGGCATCCATT
Fkbp5	9	AGGCCGTGATTCAGTACAGG	TCTGACAGGCCGTATTCCAT
Control chromosomes w/o genomic alteration			
		Primer sequence 5′ - 3′	
Gene	*Exon*	*Forward*	*Reverse*
Tgfbr3	13	TTGTGTTCAAGTCCGTGTTCA	TTCCTAGAGCACAGCGTCAG
Inpp4b	15	GCTACAACCTCTCATAGCAACTCA	TCAGGCTGTCTGGAGAACG

DNA array

Total RNA (200ng) from UPII-hFGFR3-S249C mouse normal urothelium, hyperplastic lesions and urothelial carcinomas was analyzed with the Affymetrix Mouse Exon 1.0 ST. Array gene expression was RMA normalized and annotated to the GRCm38 genome version. The LIMMA algorithm was applied to calculate the genes having a significant change of expression between urothelial carcinomas and normal urothelium or hyperplastic lesions. Genes were considered to be differentially expressed when they presented an absolute log2FC >0.58 and an adjusted p-value <0.05. P-values were adjusted for multiple comparisons using the Benjamini-Hochberg correction.

Pathway and Gene Ontology Biological Processes Enrichment

Genes with an absolute log2FC of at least 0.58 and an adjusted p-value inferior to 0.05 were used to carry out an enrichment analysis of KEGG Pathways and Gene Ontology Biological Processed. The enrichment analysis was done using David 6.8, mus musculus Affy Exon 1.0 ST background. Significantly enriched pathways were considered when they had an adjusted p-value (Benjamini and Hochberg) inferior to <0.05.

Cross-species hierarchical clustering

DNA array data from UII-hFGFR3-S249C and BBN mice was combined with DNA array data from human bladder tumors (CIT; Affymetrix Exon 1.0 ST; 96 MIBC and 99 NMIBC). Batch effects due to data combination were corrected using the surrogate variable analysis R package. The protocol used for co-clustering of the two species was that of previously described by Saito et al 2018. Hierarchical clustering was done using a gene signature derived for the consensus molecular classification of MIBC (2).

Transcriptome classifier

Subtype calls were done on murine hFGFR3-S249C and previously established murine BBN induced tumor transcriptomes. Samples were classified using the BASE47 classification algorithm and the median centered expression of the murine orthologues found in the BASE47 signature, as previously described (2,21).

Hypoxia activity score

We only reviewed literature reporting experimentally validated hypoxia-regulated genes and collected 11 datasets including 290 genes in total (27–35). A consensus hypoxia metagene set was obtained by including only genes reported in at least two of the datasets described above, among which 35 genes were classified as down-regulated and 85 as up-regulated in hypoxic conditions. We applied the gene set variation analysis (GSVA) (36) to calculate sample specific enrichment score (ssES) for the down- and up-regulated hypoxia gene sets, and then calculated a hypoxia activity score defined as the arithmetic difference between the two enrichment scores (up versus down), as a bioinformatics measurement of biological hypoxia activity. This was done for each sample in three datasets: our mouse hFGFR3-S249C model (n=6 tumors and 6 hyperplastic lesions), the TCGA luminal papillary MIBC (n = 128, classification based on Kamoun et al. (2)), and the UROMOL NMIBC dataset (n = 419) (3,5). We compared the predicted hypoxia activity score between different sample

groups using a two-sided Wilcoxon rank sum test. A value of $P < 0.05$ in two-tailed tests was considered statistically significant. R version 3.5.2 and the ggpubr version 0.2 package were used for all statistical analyses and for the associated visualization

Cell culture and transfection

The human bladder cancer derived cell lines RT112 and UM-UC-14 were obtained from DSMZ (Heidelberg, Germany). MGH-U3 cells were kindly provided by Dr. Francisco Real (CNIO, Madrid). All cells were cultured at 37°C in an atmosphere of 5%CO2 in DMEM medium except for RT112 cells which were cultured in RPMI medium. All cell media were supplemented with 10% fetal bovine serum (FBS).

The model of transfected murine NIH-3T3 fibroblasts expressing the mutated human FGFR3IIIb receptor was established in a previous study (Bernard-Pierrot et al 2006). Cells were cultured at 37°c in an atmosphere of 5% CO2 in DMEM medium. Cellular media was supplemented with 10% newborn calf serum, 2 mM glutamine, 100 U/ml penicillin and 100 mg/ml streptomycin. Control cells were transfected with the empty pcDNAI-Neo plasmid (clones Neo1.5, Neo 2.1).

The expression of FGFR3 was knocked down following siRNA transfection of UM-UC-14, MGH-U3 and RT112 cells. Cells were seeded in six-well plates at the following densities: 250,000 cells/well for UM-UC-14 and 200,000 cells/well for MGH-U3 and RT112 cells. Cells were transfected with 10nM of siFGFR3 (siRNAs #3 and #4) in the presence of Lipofectamine RNAiMax reagent (Invitrogen) following the manufacturer's protocol. siRNAs targeting FGFR3 and luciferase (control siRNA; Qiagen SI03650353) were purchased from Ambion and Qiagen.

The siRNA sequences used were:

	Sequence 5'-3'	
siRNA	Sense	Antisense
FGFR3 #3	CCGUAGCCGUGAAGAUGC	AGCAUCUUCACGGCUACGG
FGFR3 #4	CCUGCGUCGUGGAGAACA	UUGUUCUCCACGACGCAGG

FGFR3 small molecule inhibitors

The PD173074 inhibitor was purchased from Calbiochem (Merck Eurolab, Fontenay Sous Bois, France) and the BGJ398 inhibitor was purchased from LC Laboratories (USA).

FGFR3 activity inhibition

UM-UC-14, MGH-U3 and RT112 cells were treated with the pan-FGFR inhibitor PD173074 [500nM] for 40hrs. DMSO was used as control.

Immunoblotting

Urothelia were separated mechanically from the underlying muscle layer, frozen in liquid nitrogen and stored at –80°C until used for protein extraction. Tissue was homogenized and resuspended in Laemmli lysis buffer (50 mM Tris-HCl pH 6.8, 2 mM DTT, 2.5 mM EDTA, 2.5 mM EGTA, 2% SDS, 5% glycerol with protease inhibitors and phosphatase inhibitors) (Roche). Obtained lysates were clarified by centrifugation. Protein concentration was determined by the BCA method (Thermo Scientific) and 5µg of protein per sample were resolved by SDS-PAGE in 7.5 or 10% polyacrylamide gels. Following electrotransfer of proteins to a nitrocellulose membrane, membranes were probed with antibodies against the extracellular domain of FGFR3 (Abcam, ref: ab133644, diluted 1/2000), MYC (Cell Signaling Technology ref: 9402, diluted 1/500), HIF1A (Cell Signaling Technology ref: 14179, diluted 1/500), VEGFA(Abcam ref: 46154, diluted 1/1000), and loading control protein β-ACTIN (Sigma Aldrich ref: ref A2228, diluted 1/20 000). The secondary antibodies used were anti-mouse IgG, HRP-linked, and anti-rabbit IgG, HRP-linked antibody (Cell Signaling Technology ref:7076 and ref 7074, diluted 1/3000- respectively). Signal was detected using SuperSignal West Femto (ThermoFisher) or Clarity Western ECL (BioRad) substrates followed by exposure on X-ray film (ThermoFischer).

Proteins of interest were evaluated in bladder cancer cell lines, xenograft and PDX models following the previously described immunoblotting protocol for 20, 5 and 9µg of protein lysate for respectively. Proteins of interest were analyzed with the following antibodies: anti-FGFR3 (previously cited), anti-HIF1A (previously cited), anti-DUSP6 (Abcam, ref: ab76310, diluted 1/500) and anti α-TUBULIN (Sigma-Aldrich T9026, diluted 1/20,000).

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. Kamoun A, Reyniès A de, Allory Y, Sjödahl G, Robertson AG, Seiler R, et al. A Consensus Molecular Classification of Muscle-Invasive Bladder Cancer. bioRxiv. 2018;

3. Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell. 2017;171:540-556.e25.

4. Sjödahl G, Eriksson P, Liedberg F, Höglund M. Molecular classification of urothelial carcinoma: global mRNA classification versus tumour-cell phenotype classification. J Pathol. 2017;242:113–25.

5. Hedegaard J, Lamy P, Nordentoft I, Algaba F, Høyer S, Ulhøi BP, et al. Comprehensive Transcriptional Analysis of Early-Stage Urothelial Carcinoma. Cancer Cell. 2016;30:27–42.

6. Kim W-J, Kim E-J, Kim S-K, Kim Y-J, Ha Y-S, Jeong P, et al. Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. Mol Cancer. 2010;9:1–9.

7. Shi MJ, Meng XY, Lamy P, Banday AR, Yang J, Moreno-Vega A, et al. APOBEC-mediated Mutagenesis as a Likely Cause of FGFR3 S249C Mutation Over-representation in Bladder Cancer. Eur Urol. 2019;76:9–13.

8. Porębska N, Latko M, Kucińska M, Zakrzewska M, Otlewski J, Opaliński Ł. Targeting Cellular Trafficking of Fibroblast Growth Factor Receptors as a Strategy for Selective Cancer Treatment. J Clin Med. 2018;8:7.

9. Nakanishi Y, Akiyama N, Tsukaguchi T, Fujii T, Satoh Y, Mizuno H, et al. Mechanism of oncogenic signal activation by the novel fusion kinase FGFR3-BAIAP2L1. Cancer Res. 2015;75:123–123.

10. Williams S V., Hurst CD, Knowles MA. Oncogenic FGFR3 gene fusions in bladder cancer. Hum Mol Genet. 2013;22:795–803.

11. Bernard-Pierrot I, Brams A, Dunois-Lardé C, Caillault A, Diez de Medina SG, Cappellen D, et al. Oncogenic properties of the mutated forms of fibroblast growth factor receptor 3b. Carcinogenesis. 2006;27:740–7.

12. Mahe M, Dufour F, Neyret-Kahn H, Moreno-Vega A, Beraud C, Shi M, et al. An FGFR3/MYC positive feedback loop provides new opportunities for targeted therapies in bladder cancers. EMBO Mol Med. 2018;10:pii: e8163.

13. Ahmad I, Singh LB, Foth M, Morris C-A, Taketo MM, Wu X-R, et al. K-Ras and b-catenin mutations cooperate with Fgfr3 mutations in mice to promote tumorigenesis in the skin and lung, but not in the bladder. Dis Model Mech. 2011;4:548–55.

14. Foth M, Ismail NFB, Kung JSC, Tomlinson D, Knowles MA, Eriksson P, et al. FGFR3 mutation increases bladder tumourigenesis by suppressing acute inflammation. J Pathol. 2018;246:331–43.

15. Foth M, Ahmad I, Van Rhijn BWG, Van Der Kwast T, Bergman AM, King L, et al. Fibroblast growth factor receptor 3 activation plays a causative role in urothelial cancer pathogenesis in cooperation with Pten loss in mice. J Pathol. 2014;233:148–58.

16. Zhou H, He F, Mendelsohn CL, Tang MS, Huang C, Wu XR. FGFR3b extracellular loop mutation lacks tumorigenicity in vivo but collaborates with p53/pRB deficiency to induce high-grade papillary urothelial carcinoma. Sci Rep. 2016;6:1–11.

17. Logié A, Dunois-Lardé C, Rosty C, Levrel O, Blanche M, Ribeiro A, et al. Activating mutations of the tyrosine kinase receptor FGFR3 are associated with benign skin tumors in mice and humans. Hum Mol Genet. 2005;14:1153–60.

18. Hafner C, Van Oers JMM, Vogt T, Landthaler M, Stoehr R, Blaszyk H, et al. Mosaicism of activating FGFR3 mutations in human skin causes epidermal nevi. J Clin Invest. 2006;116:2201–7.

19. Rosty C, Aubriot MH, Cappellen D, Bourdin J, Cartier I, Thiery JP, et al. Clinical and biological characteristics of cervical neoplasias with FGFR3 mutation. Mol Cancer. 2005;4:2–9.

20. Rebouissou S, Bernard-Pierrot I, Reyniès A de, Lepage M-L, Krucker C, Chapeaublanc E, et al. EGFR as a potential therapeutic target for a subset of muscle-invasive bladder cancers presenting a basal-like phenotype. Sci Transl Med. 2014;6:244ra91-244ra91.

21. Saito R, Smith CC, Utsumi T, Bixby LM, Kardos J, Wobker SE, et al. Molecular subtype-specific immunocompetent models of high-grade urothelial carcinoma reveal differential neoantigen expression and response to immunotherapy. Cancer Res. 2018;78:3954–68.

22. Damrauer JS, Hoadley KA, Chism DD, Fan C, Tiganelli CJ, Wobker SE, et al. Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology. Proc Natl Acad Sci. 2014;111:3110–5.

23. Biton A, Bernard-Pierrot I, Lou Y, Krucker C, Chapeaublanc E, Rubio-Pérez C, et al. Independent Component Analysis Uncovers the Landscape of the Bladder Tumor Transcriptome and Reveals Insights into Luminal and Basal Subtypes. Cell Rep. 2014;9:1235–45.

24. Becht E, De Reyniès A, Giraldo NA, Pilati C, Buttard B, Lacroix L, et al. Immune and stromal classification of Colorectal cancer is associated with molecular subtypes and relevant for precision immunotherapy. Clin Cancer Res. 2016;22:4057–66.

25. Zhu L, Somlo G, Zhou B, Shao J, Bedell V, Slovak ML, et al. Fibroblast growth factor receptor 3 inhibition by short hairpin RNAs leads to apoptosis in multiple myeloma. Mol Cancer Ther. 2005;4:787–98.

26. Qian S, Somlo G, Zhou B, Zhu L, Mi S, Mo X, et al. Ribozyme Cleavage Leads to Decreased Expression of Fibroblast Growth Factor Receptor 3 in Human Multiple Myeloma Cells, Which Is Associated with Apoptosis and Downregulation of Vascular Endothelial Growth Factor. Oligonucleotides. 2005;15:1–11.

27. Kardos J, Chai S, Mose LE, Selitsky SR, Krishnan B, Saito R, et al. Claudin-low bladder tumors are immune infiltrated and actively immune suppressed. JCI Insight. 2016;1:e85902.

28. Sweis RF, Spranger S, Bao R, Paner GP, Stadler WM, Steinberg G, et al. Molecular drivers of the non-T cell-inflamed tumor microenvironment in urothelial bladder cancer. Cancer Immunol Res. 2016;4:563–8.

29. Flaherty KT, Infante JR, Daud A, Gonzalez R, Kefford RF, Sosman J, et al. Combined BRAF and MEK Inhibition in Melanoma with BRAF V600 Mutations. N Engl J Med. 2012;367:1694–703.

30. Bertz S, Abeé C, Schwarz-Furlan S, Alfer J, Hofstädter F, Stoehr R, et al. Increased angiogenesis and FGFR protein expression indicate a favourable prognosis in bladder cancer. Virchows Arch. 2014;465:687–95.

31. Liu X, Jing X, Cheng X, Ma D, Jin Z, Yang W, et al. FGFR3 promotes angiogenesis-dependent metastasis of hepatocellular carcinoma via facilitating MCP-1-mediated vascular formation. Med Oncol. Springer US; 2016;33:46.

32. Benita Y, Kikuchi H, Smith AD, Zhang MQ, Chung DC, Xavier RJ. An integrative genomics approach identifies Hypoxia Inducible Factor-1 (HIF-1)-target genes that form the core response to hypoxia. Nucleic Acids Res. 2009;37:4587–602.

33. Chi JT, Wang Z, Nuyten DSA, Rodriguez EH, Schaner ME, Salim A, et al. Gene expression programs in response to hypoxia: Cell type specificity and prognostic significance in human cancers. PLoS Med. 2006;3:395–409.

34. Elvidge GP, Glenny L, Appelhoff RJ, Ratcliffe PJ, Ragoussis J, Gleadle JM. Concordant regulation of gene expression by hypoxia and 2-oxoglutarate-dependent dioxygenase inhibition: The role of HIF-1α, HIF-2α, and other pathways. J Biol Chem. 2006;281:15215–26.

35. Eustace A, Mani N, Span PN, Irlam JJ, Taylor J, Betts GNJ, et al. A 26-gene hypoxia signature predicts benefit from hypoxia-modifying therapy in laryngeal cancer but not bladder cancer. Clin Cancer Res. 2013;19:4879–88.

36. Harris BHL, Barberis A, West CML, Buffa FM. Gene Expression Signatures as Biomarkers of Tumour Hypoxia. Clin Oncol. 2015;27:547–60.

37. Sørensen BS, Toustrup K, Horsman MR, Overgaard J, Alsner J. Identifying pH independent hypoxia induced genes in human squamous cell carcinomas in vitro. Acta Oncol (Madr). 2010;49:895–905.

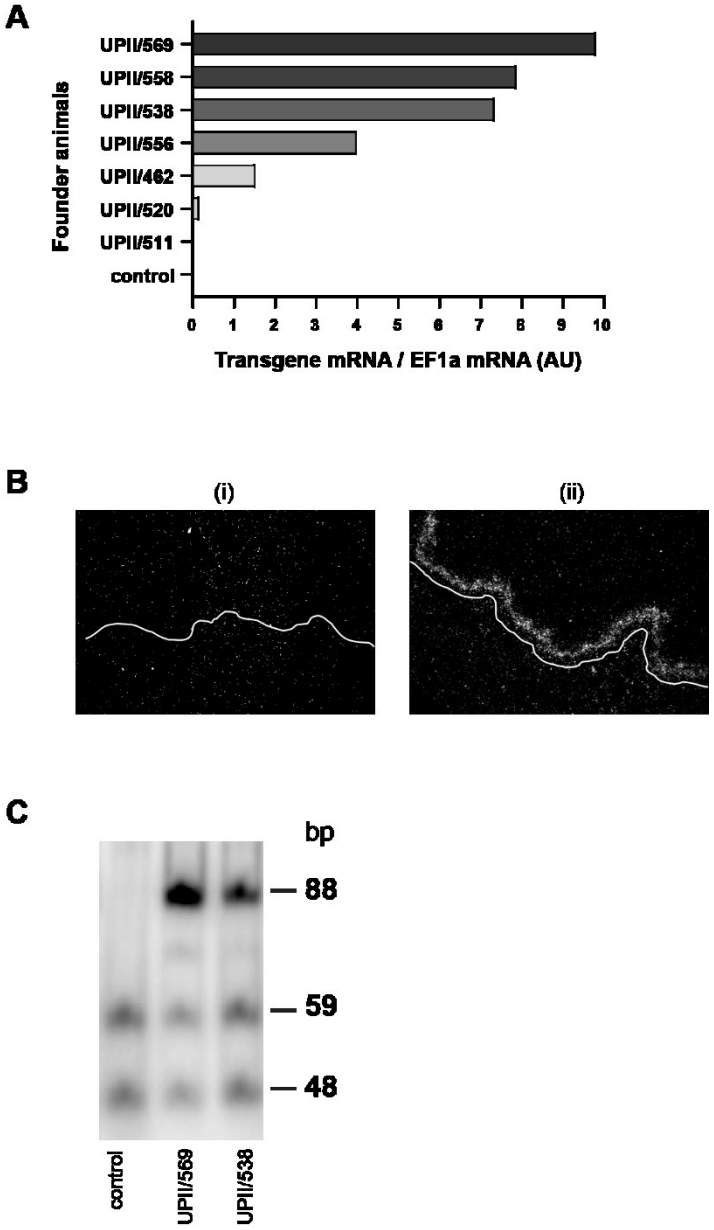
38. Toustrup K, Sørensen BS, Nordsmark M, Busk M, Wiuf C, Alsner J, et al. Development of a hypoxia gene expression classifier with predictive impact for hypoxic modification of radiotherapy in head and neck cancer. Cancer Res. 2011;71:5923–31.

39. Winter SC, Buffa FM, Silva P, Miller C, Valentine HR, Turley H, et al. Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. Cancer Res. 2007;67:3441–9.

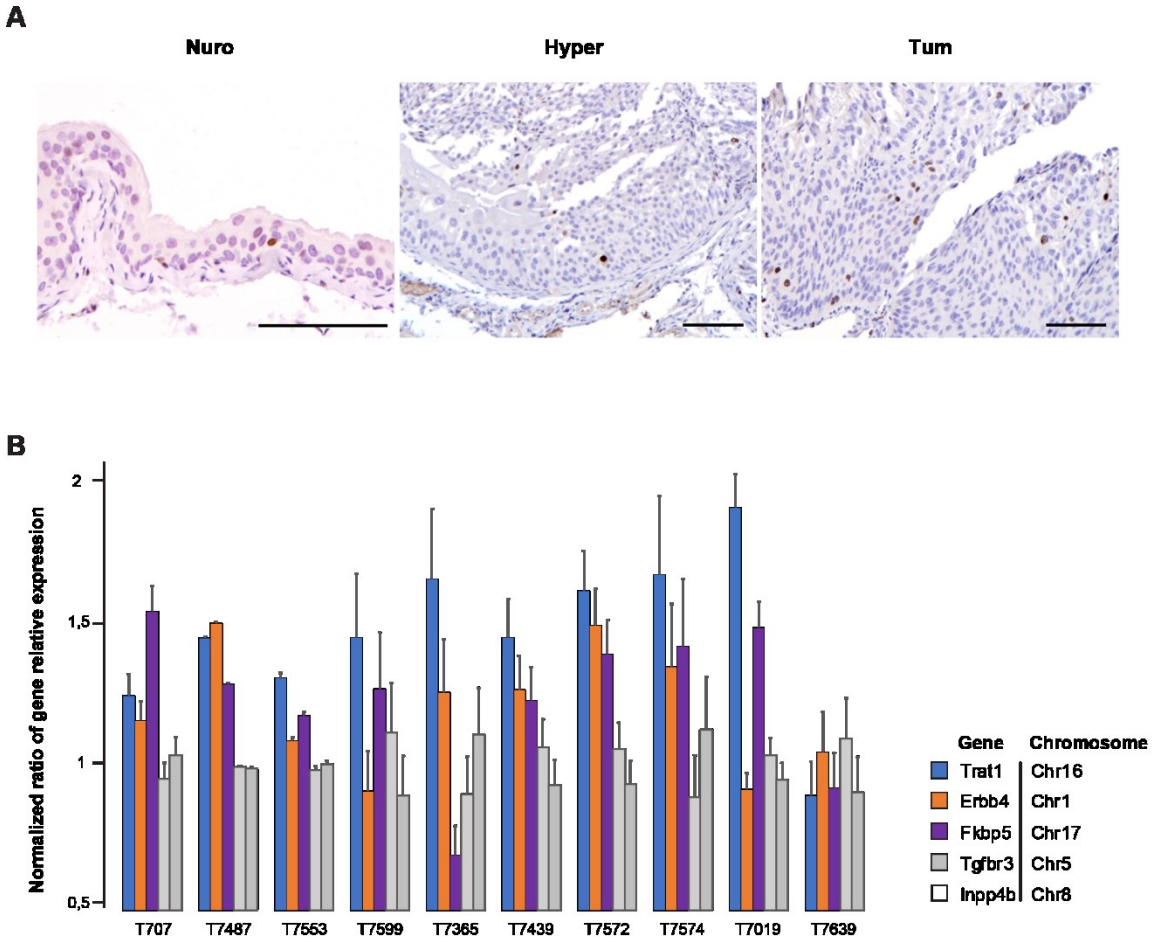
40. Yang L, Roberts D, Takhar M, Erho N, Bibby BAS, Thiruthaneeswaran N, et al. Development and Validation of a 28-gene Hypoxia-related Prognostic Signature for Localized Prostate Cancer. EBioMedicine. The Authors; 2018;31:182–9.

41. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013;14:1–15.

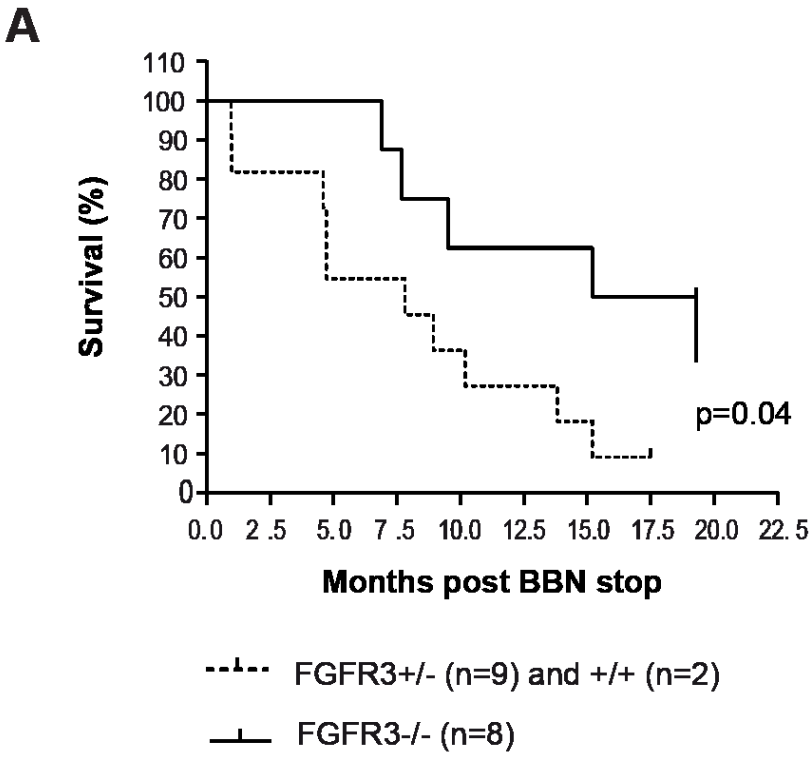
Supplementary Figures



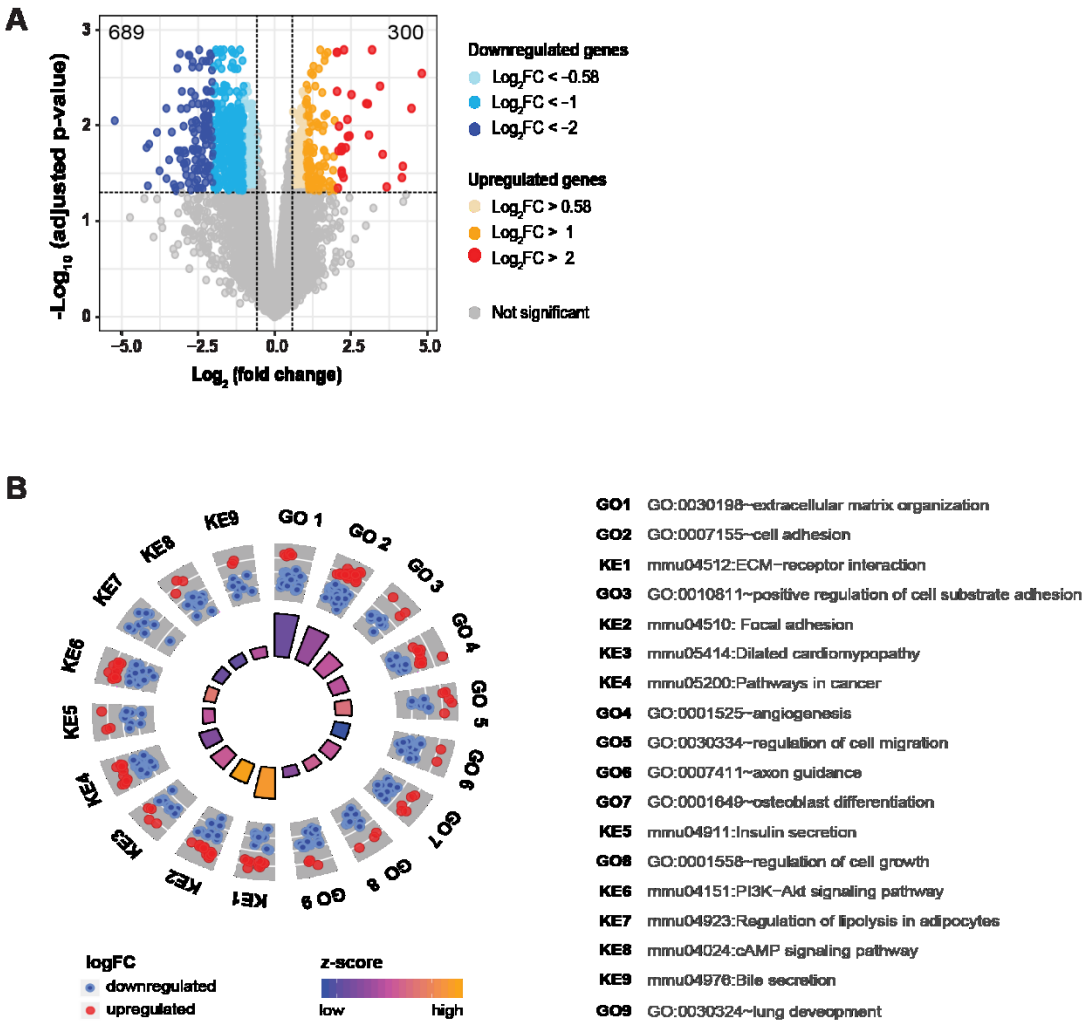
Supplementary Figure 1. A) Validation of the expression of the relative mRNA expression levels of the human FGFR3 transgene in the urothelium of transgenic UII-FGFR3-S249C mice. B) *In situ* hybridization showing expression of the human FGFR3 transgene at the supra-basal and intermediate cell layers of UII-FGFR3-S249C mice urothelium (4 months of age). Magnification x100. C) Radioactive PCR showing the expression of both human and mouse FGFR3 digested amplicons (cDNA) in control and UII-FGFR3-S249C (line 569 and 538) mice. The bands of 59 and 48 bp correspond to mouse endogenous FGFR3, and the band of 88 bp to the human FGFR3 transgene.



Supplementary Figure 2. A) Representative immunohistochemistry showing Ki67 expression in hyperplastc urothelium (middle panel) and baldder tumor (right panel) of UII-hFGFR3-S249C mice and in normal urothelia from control littermates (right panel). B) Genomic DNA qPCR validation of genes found in frequently altered regions (chromosomes 1, 16 and 17) of tumors from UII-hFGFR3-S249C mice. Shown is the ratio of relative expressions of exonic regions of genes found in altered chromosomes (Trat1, Erbb4, Fkbp5) against the genes found in stable chromosomes (Tgfbr3, Inpp4b). Each relative expression value was calculated using the $2^{-\Delta\Delta Ct}$ method and values were normalized to control urothelia for each sample.



Supplementary Figure 3. A) Survival plot of UII-hFGFR3-S249C mice (FGFR3 +/- or +/+) versus control mice from littermates (FGFR3-/-) following treatment with 0.05% BBN in drinking water for 8 weeks.



Supplementary Figure 4. A) Volcano plot of the set of differentially expressed genes (DEGs; 989 genes; |log2FC|>0.58; adjpval<0.05) by comparing gene expression in hFGFR3-S249C mice tumors compared to normal control urothelium. B) Circle plot of the top 18 Gene Ontology Biological Processes (GO) and murine KEGG (KE) pathways inferred from the DEGs. The outer circle represents a scatter plot of the DEGs assigned for each GO and KE term. Red and blue circles represent up- or down-regulated genes, respectively. Bars at the inner circle represent the adjusted p-value of the enriched term (the higher the bar, the more significant the p-value). All p-values are <0.05. The color of the bars represents a predicted activation score (z-score) of the enriched term based on the expression levels of the assigned genes (circles on scatter plot).

3.2.3 Discussion

There are still some experiments are ongoing to finish this manuscript. I will examine whether VEGFA protein will be decreased after FGFR3 inhibition as HIF1A do in MGH-U3 xenografts. We also have some new PDX models harbouring FGFR3 mutations and treated with FGFR3 inhibitor, I will check and compare angiogenesis status by IHC staining in those samples.

For many years, it has been believed that mutated *FGFR3* alone is insufficient to induce tumorigenesis. While *FGFR3* mutations coupling with other genetic modifications or additional carcinogens exposure promote tumor development. Ahmad et al. found that when combining FGFR3 mutations with the overexpression of *KRAS* or β-catenin inactivating mutations, transgenic mice developed skin and lung tumors but not BCa (Ahmad et al., 2011). Foth et al. did not observe spontaneous formation of BCa in transgenic mice generated by UII-*FGFR3* S249C mutation up to a follow-up of 12 months, but observed an accelerated formation of invasive BCa by additional exposure with BBN carcinogen (Foth et al., 2018). However, in our GEM model, we indeed observed papillary tumor of BCa by 18 months and herein demonstrated the first evidence showing *FGFR3* S249C mutation alone is enough to induce BCa tumorigenesis. Recently, the first pan-FGFRs inhibitor – Erdafitinib/Balversa – has been approved by the FDA (April 2019) for patients with locally advanced or metastatic BCa. Considering the increasing interest of targeting FGFR3 for BCa treatment, our model may have clinical translational value to evaluate drug response and to understand acquired drug resistance mechanism.

The molecular signature of BBN induced BCa model is similar to basal-like MIBCs, which are known as T-cell-inflamed and highly immunogenic tumors (Fantini et al., 2018; Saito et al., 2018; Sweis et al., 2016). As a result, BBN models are valuable resource to examine cancer biology and immunotherapy. FGFR3 mutations are much more common in NMIBCs or luminal papillary type tumors of MIBCs. Opposite to basal-like MIBCs, the luminal papillary-like tumors are non-T-cell-inflamed and present poorly infiltrated immune-environment and therefore are more likely less responsive to immunotherapy (Kardos et al., 2016; Rosenberg et al., 2016; Sweis et al., 2016). To confirm this hypothesis, phase 1b/2 clinical trial (NCT03123055), comparing the efficacy of anti-FGFR3 therapy (B-701, specific monoclonal antibody targeting FGFR3) and immunotherapy (pembrolizumab) in patients with *FGFR3* altered advanced BCa, is ongoing. In line with literatures, our GEM model showed poor infiltration of different immune cell populations by MCP counter program as well as decreased

$CD8^+$ T-cell by IHC staining. Hence, this model is useful to understand immune-escape or immune-suppression mechanism driven by mutated/active FGFR3.

Hypoxia is a common feature of tumorigenesis. The hypoxic environment in tumours is the result of uncontrolled proliferation and excessive oxygen/nutrition demand outstripping angiogenesis. One of the best-known transcription factor – hypoxia inducing factor (*HIF1A*) – rapidly activate a subset of angiogenic genes to enable tumor cells adapting hypoxic environment and continue to grow (Harris et al., 2015). There are some evidence showing that FGFRs can modulate HIF1A signaling and angiogenetic process. In lung cancer cells, FGF2 ligand stimulation or FGFR1 dependent activation of PI3K/AKT pathway can induce HIF1A activation (Fumarola et al., 2017). In hepatocellular carcinoma, FGFR3 promotes angiogenesis-dependent metastasis via facilitating *MCP1* mediated vascular formation (Liu et al., 2016). From immunohistochemistry staining of vascularization markers in 61 BCa, increased angiogenesis is more common in non-muscle-invasive tumors compared to muscle-invasive cancer, and is also positively correlated with FGFR3 expression and mutation status (Bertz et al., 2014). Similar results have been achieved in our GEM model, where we proposed a regulation axis *FGFR3*/*HIF1A/VEGFA*/angiogenesis driving this BCa tumorigenesis which was also validated in several other *in vivo* models.

The FGFR3/MYC regulation axis and a positive feedback loop between them have been confirmed in BCa (Mahe et al., 2018). Considering a complex regulatory network among FGFR3, MYC and HIF1A may exist, it remains questionable whether HIF1A alone drives FGFR3-regulated angiogenesis and whether HIF1A solely participates in angiogenesis but is not involved in other biological process such as metabolism. First, it is known that HIF1A can interact with MYC *via* its N-terminal region, suppressing MYC as well as MYC-targeted genes (Gordan et al., 2007; Koshiji et al., 2004). Second, in opposite direction, it is reported that MYC is able to stabilize HIF1A as well as induce HIF1A-targeted genes by *MYC* overexpression (Doe et al., 2012). Third, considering the ability of both HIF1A and MYC to modulate angiogenesis and glycolysis metabolism by sharing many common target genes (Semenza, 2010; Yu et al., 2017), such as *VEGFA* (surrogate marker of vascularization) and *HK2* (enzyme involved in the first step of glucose metabolism), they can cooperate with each other to promote these processes (Kim et al., 2007). Herein, future efforts should be made to figure out the crosstalk between MYC and HIF1A since both of which are regulated by FGFR3 in BCa. In addition, to identify potential downstream kinase or effector of FGFR3 mediating the regulation between FGFR3 and HIF1A is of importance.

3.3 FGFR3 signaling in bladder cancer

3.3.1 Introduction

As we previously described in Chapter 1, the major identified FGFR3-driven signaling pathways so far are RAS-MAPK, PI3K-AKT, PLCG1-PKC and JAK-STATs cascades. RAS-MAPK and PI3K-AKT pathways are commonly documented in BCa harbouring FGFR3 activating mutations or fusions whereas JAK-STATs cascades are often reported in the context of multiple myeloma or lung cancer (Dudka et al., 2010; Lee et al., 2014). Some of them are initially identified based on prior knowledge, for example supposed signaling similarities among FGFRs family or other RTKs. With the advent of gene array or RNA sequencing, a lot of comparative transcriptomic studies have been done in models with either artificial overexpression of FGFR3 or endogenous knockdown *in vitro* and *in vivo*. It largely provides us in-depth understanding of other signaling that may be relevant to FGFR3. For instance, recently we generated comparative transcriptomic analyses from two BCa cell lines (MGH-U3 and RT-112) and reported MYC as a key master regulator of proliferation in the aberrantly activated FGFR3 pathway (Mahe et al., 2018). These cell lines were derived from human bladder tumors, and they endogenously express a mutated activated form of FGFR3 (*FGFR3* Y375C, the second most frequent mutation in bladder tumors) and the FGFR3-TACC3 fusion protein (the most frequent FGFR3 fusion protein in bladder tumors), respectively.

FGFR3 is a receptor tyrosine kinase, thus the propagation of downstream signaling necessarily depends on protein-protein interactions (PPIs) and posttranslational modifications (especially tyrosine phosphorylation). Unfortunately, very little is known about FGFR3 binding partners and its phosphorylated substrates, even though the proteomics coupled with mass spectrometry (MS) strategy has been greatly advanced in the last two decades. To date, there are only three proteomics studies focusing on FGFR3 in different contexts: one sole study aiming to identify FGFR3 interactors (Balek et al., 2018) and two studies aiming to identify phosphoproteomics of FGFR3 (Lombardi et al., 2017; St-Germain et al., 2009). They are summarized in Table 3-1; page 164. There is no study exploring FGFR3 interactome in BCa and the only one conducted by Balek et al. in 293T cells using artificial overexpression of *FGFR3*. Although both phosphoproteomics studies applied endogenously FGFR3 activating models, only experiments focusing on oncogenic fusion of FGFR3-TACC3 in BCa or *FGFR3* mutations plus fusion

alterations in multiple myeloma were investigated previously. In fact, it is still not clear whether the downstream signaling induced by FGFR3 activating mutations are similar to those induced by activating FGFR3-TACC3 fusion, in addition the downstream signalling pathways should be context specific. Thus, prior knowledge of proteomics from oncogenic fusion of FGFR3 in other cancers is not fully representative for the profile driven by activating *FGFR3* mutations in BCa.

Table 3-1: *Literatures of proteomics studies focused on FGFR3.*
Summarized literatures of proteomics studies based on perturbation of FGFR3 protein in different models.(Balek et al., 2018; Lombardi et al., 2017; St-Germain et al., 2009)

Proteomics	Context	Working_model	Comparing_groups	Methods	Number_identified	Reference
Interactome	Not specific	**293T** cell: Overexpress FGFR3-C'_flag (V5);	K650E/M *vs.* WT	IP-V5 and IP-4G10 for pY proteins;	**55***	Balek et al, 2018
Author considered as Interactome'	Not specific	**RCS** (rat chondrosarcoma) cell: endogenously expressed FGFR2 and FGFR3;	FGF2(+) *vs.* FGF2(-)	IP-4G10 for pY proteins;	**36***	Balek et al, 2018
Phosphoproteome	Multiple myeloma	**KMS11** cell: endogenously harbour *FGFR3* Y373C; **OPM2** cell: endogenously harbour *FGFR3* K650E; **LP1** cell: endogenously express WT of FGFR3; All above three cell lines are positive for the t(4;14) translocation.	PD173074(+) *vs.* PD173074(-) in KMS11 cell with 2 replicates; FGF1(+) *vs* . FGF1(-) in 4 repetitions with KMS11 and two each with the LP1 and OPM2 cells;	IP-pY100 for pY peptides;	Downregulated pY by PD173074 (FC>2): **45** proteins (57 pY sites); #Upregulated pY by FGF1: **34** proteins (40 pY sites);	St-Germain et al, 2009
Phosphoproteome	Urothelial cancer	TERT-**NHU** cell: Overexpress FGFR3 (WT); **RT-112** cell: endogenously harbour FGFR3-TACC3 fusion(FUS);	C1: WT *vs.* FUS (NHU cell *vs.* RT-112); C2: WT+FGF1 *vs.* FUS+FGF1; C3: WT *vs.* WT+FGF1;	IP-pY1000 for pY peptides; TiO2 for enrichment of pS/T peptides;	Inclusion criteria: at least >1 replicate and log2FC>0.5 or <-0.5. The numbers as followed: C1: **24**; C2: **26**; C3: **19**.	Lombardi et al, 2017

55*:In total of six V5, ten FLAG and ten 4G10 IP experiments, at least 4 out of 26 experiments except for proteins previously known to interact with FGFRs as inclusion criteria.
36*:In total of six repetitions, at least 2 out of 6 considered except for proteins previously known to interact with FGFRs.
#: All these pY sites are regulated by PD173074 group and fold changes by FGF1 are not necessarily >2. PD173074, selective pan-FGFRs inhibitors; WT, wild-type.

With the rapid development of both transcriptomics and proteomics techniques in parallel, investigating cell signaling pathways with the -omics integrating analysis is a new trend. In BCa, the most common *FGFR3* mutation is S249C, representing 62% of all recurrent *FGFR3* mutations and being a driver of an oncogenic dependency (Bernard-Pierrot et al., 2006; Shi et al., 2019). We also showed the over-representation of S249C was likely due to increased mutation rate caused by APOBEC mutagenesis rather than increased tumorigenicity of the S249C mutation (Shi et al., 2019). In current study, we performed comparative analysis by perturbating FGFR3 activation, generated three -omic datasets (interactome, phosphoproteome and transcriptome) and aimed to provide a comprehensive characterization of altered FGFR3 signalling pathways by data integrating analysis in BCa. Of note, we worked with an endogenously activated FGFR3 *in vitro* BCa model (UMUC-14 cell line) that harbours the most frequent mutation – *FGFR3* S249C. Subsequently, the selected candidates and their

relevance with FGFR3 were examined and/or validated *in vitro* and *in vivo* in several BCa models.

3.3.2 Materials & Methods

Cell culture & SILAC labeling

The human bladder-derived cell lines UMUC-14 and RT-112 were obtained from DSMZ (Heidelberg, Germany). MGH-U3 cells were kindly provided by Dr. Paco Real (CNIO, Madrid). I used UMUC-14 cells for all large-scale experiments. UMUC-14 cells bear a homozygous *FGFR3* S249C mutation, MGH-U3 cells harbour a homozygous *FGFR3* Y375C mutation and RT-112 cells have a FGFR3-TACC3 fusion together with a wild-type allele. A comprehensive genomic characterization of these cells has been reported (Earl et al., 2015). UMUC-14 and MGH-U3 cells were cultured in DMEM, whereas RT-112 cells were cultured in RPMI (Life Technologies). Media were supplemented with 10% fetal calf serum (Lonza) and 1% Pen-Strep (Gibco). Cells were incubated at 37°C, under an atmosphere containing 5% CO2. The identity of the cell lines used was checked by analyzing genomic alterations with comparative genomic hybridization arrays (CGH array), and *FGFR3* and *TP53* mutations, checked with the SNaPshot technique (for *FGFR3*) or by classical sequencing (for *TP53*), the results obtained being compared with the initial description of the cells. We routinely checked for mycoplasma contamination (Lonza, cat.no. LT07-118).

Stable isotope labeling by amino acids in cell culture (SILAC) media (depleted methionine, lysine and arginine, cat.no. 88420) and dialysed fetal bovine serum (FBS, cat.no. 26400-044) were from Thermo-scientific (Ozyme, France); L-Methionine (cat.no. M5308), L-Proline (cat.no. P5607), normal 'light (L)' amino acids labeling: L-lysine (Lys0, cat.no. L8662) and L-arginine (Arg0, cat.no. A8094) hydrochloride ($^{12}C/^{14}N$, K0/R0), 'medium (M)' labeling ($^{2}H4/^{13}C6$, K4/R6; cat.no. 616192 and 643440, respectively) and 'heavy (H)' labeling ($^{13}C6^{15}N2/^{13}C6^{15}N4$, K8/R10; cat.no. 608041 and 608033, respectively) amino acids were provided by Sigma. Stock solutions were prepared by dissolving amino acids in PBS or non-restituted culture medium for 1000x concentration (Arginine, 84 mg/ml; lysine, 146 mg/ml; and methionine, 30 mg/ml) and filtered (0.22um syringe filter) before use. Increased molecular weight of the amino acids in medium or heavy media was taken into account to give equimolar amounts. For instance, L-arginine-$^{13}C6$ was prepared at a concentration of 87.4 mg/ml. UMUC-14 cells were grown in light, medium and heavy SILAC media, with 10% FBS and

without antibiotics, for different experimental purposes. Cell lines were grown for more than six or seven cell doublings in the labeling media to ensure complete incorporation.

Kinase and protein inhibitors and FGFRs ligand

A selective pan-FGFRs inhibitor (PD173074, cat.no. HY-10321), PP2A inhibitor (okadaic acid, OA, cat.no. 459620) and proteasome inhibitor (MG-132, cat.no. 474790) were purchased from Calbiochem (Merck Eurolab, Fontenay Sous Bois, France). FGFRs inhibitor (BGJ398, cat.no. S2183) and CK2 inhibitor (silmitasertib, CX-4945, cat.no. S2248) were obtained from Selleckchem (Euromedex, Souffelweyersheim, France). IC50 for all inhibitors were tested prior to use. Acidic fibroblast growth factor (FGF1) and heparin were purchased from R&D Systems.

Cell viability assay

Cells were seeded in a 96-wells plate at a density of 10000 cells/well in complete medium (without antibotics) and incubated for 24 h. They were treated with increasing concentrations of corresponding inhibitors. After incubating the plates for 72 h, cell viability assay was detected by CellTiter-Glo kit (Promega) according to the manufacturer's directions. Luminescence was measured using the FLUOstar OPTIMA plate reader (BMG LABTECH). The viability index was expressed by the relative value to control DMSO-treated cells.

RNA interference assay

For siRNA transfection, UMUC-14, MGH-U3 and RT-112 cells were used to seed six-well or 24-well plates at a density of 250,000 cells/well for UMUC-14 and MGH-U3 cells and 200,000 cells/well for RT-112 cells. Cells were transfected with optimal concentrations (10 nM for *FGFR3* siRNA #3 and #4; 20 nM for *FGFR3* siRNA #1 and #2) of several siRNAs in the presence of Lipofectamine RNAiMax reagent (Invitrogen), in accordance with the manufacturer's protocol. The cells were collected after 24 h, 40 h or 72 h for further experiments. siRNAs were purchased from Ambion and Qiagen. For the control siRNA, we used a Qiagen control targeting luciferase (SI03650353).

FGFR3 siRNA#1 and siRNA#2 targeted exon 19 of *FGFR3* (NM_001163213). They therefore knocked down the expression of wild-type and mutated *FGFR3*, but not of the *FGFR3*-fusion gene containing the first 18 exons of *FGFR3*. Conversely, siRNA#3 and siRNA#4 targeted exons 12 and 6 of *FGFR3* (NM_001163213), respectively, knocking down both wild-type and FGFR3-TACC3 expression in RT-112 cells.

The detailed information and sequences of the siRNAs were as follows:

Gene_siRNAs	Brand	Reference	Targeted_exon	Sequences	
				Sense strand	Antisense strand
FGFR3 siRNA #1	Ambion	144467	exon 19 of *FGFR3* (NM_001163213)	5'-GCUUUACCUUUUAUGCAA-3'	5'-UUGCAUAAAAGGUAAAGGC-3'
FGFR3 siRNA #2	Ambion	144468	exon 19 of *FGFR3* (NM_001163213)	5'-GGGAAGCCGUGAAUUCAGU-3'	5'-ACUGAAUUCACGGUUCCC-3'
FGFR3 siRNA #3	Ambion	s5167	exon 6 & 12 of *FGFR3* (NM_001163213)	5'-CCGUAGCCGUGAAGAUGC-3'	5'-AGCAUCUUCACGGCUACGG-3'
FGFR3 siRNA #4	Ambion	s5168	exon 6 & 12 of *FGFR3* (NM_001163213)	5'-CCUGCGUCGUGGAGAACA-3'	5'-UUGUUCUCCACGACGCAGG-3'

RNA extraction and reverse transcription

The total mRNA was isolated from transfected cells using the RNAeasy kit (Qiagen). Five-hundreds micrograms of total RNA was reverse-transcribed in 1X RT Buffer using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems) in a final volume of 20 µL containing 1X dNTP Mix (4 mM), 1X RT Random Primers, RNase Inhibitor (1 U/µl) and MultiScribe RT (2.5 U/µl). Reaction was run in a Mastercycler pro PCR System (Eppendorf) as follows: 25 °C (10 sec), 37 °C (120 min), 85 °C (5sec).

Real-time quantitative PCR (RT-qPCR)

Primer design was performed using Primer3 plus online software (https://primer3plus.com/cgi-bin/dev/primer3plus.cgi). RT-qPCR was carried out in a LightCycler 480 Instrument (Roche) in a final volume of 20µL containing forward and reverse primers, 1X Probes Master Mix (Roche) and 10ng of cDNA. Thermal cycling conditions included a pre-incubation step at 95°C (5min), followed by 45 cycles of amplification at 95 °C (10 sec), 60 °C (30 sec) and 72 °C(1sec), and cooling at 40 °C (30 sec). The efficacy of primers were validated prior to use. Analysis was performed with the LightCycler 480 Software. Fold differences were calculated according to the $2^{-\Delta\Delta Ct}$ method and normalized against the endogenous expression of *TBP* gene.

Sequences of primers used are described below:

Gene	Concentration (nM)	Probe_Roche	Primers sequences	
			Forward 5'-3'	Reverse 5'-3'
FGFR1	500	#67	ACAACCTGCCTTATGTCCAGA	ACAACCTGCCTTATGTCCAGA
FGFR2	500	#14	GGACCCAAAATGGGAGTTTC	ACCACTTGCCCAAAGCAA
FGFR3	500	#63	AGAAGGCCTTTTGGCTGAG	GATGCCTGCATACACACTGC
FGFR4	500	#18	CAGAGGCCTACCTTCAAGCA	GAAGGTCAGGCGGAGGTC
MYC	500	#34	CACCAGCAGCGACTCTGA	GATCCAGACTCTGACCTTTTGC
DUSP6	500	#66	CGACTGGAACGAGAATACGG	AATGTACCAAGACACCACAGTTCT
GATA3	900	#71	CTCATTAAGCCCAAGCGAAG	TCTGACAGTTCGCACAGGAC
TIMP2	500	#43	GAAGAGCCTGAACCACAGGT	CGGGGAGGAGATGTAGCAC
TBP	TBP control: 20x single mix (FAM™/MGB probe, non-primer limited)			

Western blot analysis

Total protein was extracted in lysis buffer (50 mM Tris-HCl (pH 7.5), 250 mM NaCl, 1% SDS), with proteases and phosphatase inhibitors (Roche)). Protein concentration of the supernatants was determined with the BCA protein assay (ThermoFisher). Twenty micrograms of proteins or conditioned media were loaded on SDS-polyacrylamide gels using 4-15% Tris-glycine precast gels (BioRad) and transferred to nitrocellulose membranes using a Trans-Blot Turbo transfer system (BioRad). Membranes were stained with 1x Naphthol Blue Black for rapid staining of protein bands (AmidoBlack staining, Sigma) and then blocked for 1h with 5% non-fat milk or bovine serum albumin in PBST or TBST buffer at room temperature. Next, membranes were incubated with the primary antibody overnight at 4°C. Secondary anti-mouse or anti-rabbit antibodies (1/3000 dilution) were incubated for 1h at room temperature. Signal detection was performed using SuperSignal West Femto (ThermoFisher) or Clarity Western ECL (BioRad) substrates followed by exposure on X-ray film (ThermoFischer) or using the BioRad ChemiDoc MP instrument. Image analysis was performed using the Image Lab Software (BioRad).

Primary and secondary antibodies are listed below:

Gene	Protein	Supplier	Cat.no.	Isotype	MW(KDa)	Buffer	Dilution
ACTB	β-Actin	Sigma	A2228	Mouse monoclonal	42	5% BSA PBST	1/20000
AKT1/2/3	p-Akt (S473)	CST	4060	Rabbit polyclonal	60	5% BSA TBST	1/1000
AKT1/2/3	Akt	CST	2920	Mouse monoclonal	60	5% BSA PBST	1/1000
CAMK2A/B /D/G	CaMKII	CST	4436	Rabbit monoclonal	50-60	5% BSA PBST	1/1000
CAMK2A/B /D/G	CaMKII (Pan)	CST	3362	Rabbit polyclonal	50,60	5% BSA PBST	1/1000
CAMK2A /B/G	p-CaMKII (T286)	CST	12716	Rabbit monoclonal	50,60	5% BSA TBST	1/1000
CCNB1	Cyclin B1 (V152)	CST	4135	Mouse monoclonal	60	5% BSA PBST	1/1000
CDC37	Cdc37	SCBT	sc-17758	Mouse monoclonal	≈50	5% BSA PBST	1/500
CDK1	p-CDK1 (Y15)	CST	9111	Rabbit polyclonal	34	5% BSA TBST	1/1000
CDK1	p-CDK1 (T161)	CST	9112	Rabbit polyclonal	34	5% BSA TBST	1/1000
CDK1	CDK1	CST	9116	Mouse monoclonal	34	5% BSA PBST	1/1000
CREB1	p-CREB1 (S133)	CST	9198	Rabbit monoclonal	43	5% BSA TBST	1/2000
CREB1	CREB1	CST	9197	Rabbit monoclonal	43	5% BSA PBST	1/1000
CSNK2A1	CK2a (1AD9)	SCBT	sc-12738	Mouse monoclonal	37-40	5% BSA PBST	1/300

DUSP6	DUSP6	Abcam	ab76310	Rabbit monoclonal	37	5% BSA PBST	1/3000
FGFR3	FGFR3	Sigma	F3922	Rabbit monoclonal	110	5% milk PBST	1/1000
FGFR3	FGFR3	Sigma	F0425	Rabbit monoclonal	110	5% milk PBST	1/1000
FGFR3	FGFR3	Abcam	ab133644	Rabbit monoclonal	110	5% milk PBST	1/2000
FRS2	FRS2a	Thermo-Fisher	PA1-24685	Rabbit polyclonal	92-95	5% milk PBST	1/2000
GSK3B	p-GSK3β (S9)	CST	5558	Rabbit monoclonal	46	5% BSA TBST	1/2000
GSK3B	GSK-3β	CST	9832	Mouse monoclonal	46	5% BSA PBST	1/1000
HBEGF	HBEGF (E-10)	SCBT	sc-74526	Mouse monoclonal	18	5% BSA PBST	1/200
HIF1A	HIFa	CST	14179	Rabbit monoclonal	120	5% BSA PBST	1/500
HIST1H3A	Histone H3	Abcam	ab1791	Rabbit polyclonal	17	5% milk PBST	1/2000
HK2	HK2	CST	2867	Rabbit monoclonal	62	5% BSA PBST	1/2000
HMGB1	HMGB1	R&D	MAB1690	Mouse monoclonal	26	5% milk PBST	1/1000
HSP90AA1	HSP90	Abcam	ab13492	Mouse monoclonal	90	5% milk PBST	1/1000
IGFBP3	IGFBP3 (E-9)	SCBT	sc-374365	Mouse monoclonal	40/44	5% BSA PBST	1/2000
MAP3K7	p-TAK1 (T184/187)	CST	4531	Rabbit polyclonal	60-80	5% BSA TBST	1/1000
MAP3K7	TAK1	CST	5206	Rabbit monoclonal	60-80	5% milk PBST	1/1000
MAPK1/3	p-Erk1/2 (T202/Y204)	CST	9101	Rabbit polyclonal	42/44	5% BSA TBST	1/5000
MAPK1/3	Erk1/2	CST	9102	Rabbit polyclonal	42/44	5% BSA PBST	1/2000
MYC	c-Myc	CST	5605	Rabbit monoclonal	57-65	5% milk PBST	1/1000
PPARG	PPARr	Abcam	ab41928	Mouse monoclonal	50-55	5% milk PBST	1/1500
PPP2CA	p-PP2A (Y307)	Thermo-Fisher	PA5-36874	Rabbit polyclonal	36	5% BSA TBST	1/1000
PPP2CA	PP2A	Sigma	SAB4200266	Mouse monoclonal	36	5% milk PBST	1/1000
PTEN	p-PTEN (S380)	CST	9551	Rabbit polyclonal	54	5% BSA TBST	1/3000
PTEN	PTEN	CST	9559	Rabbit monoclonal	54	5% BSA PBST	1/1000
RPS6KA3	p-RSK2 (S380)	CST	11989	Rabbit monoclonal	90	5% BSA TBST	1/2000
RPS6KA3	RSK2	CST	5528	Rabbit monoclonal	90	5% BSA PBST	1/2000
RPS6KB1/2	p-S6K (T389)	CST	9205	Rabbit polyclonal	70,85	5% BSA TBST	1/500
RPS6KB1/2	S6K	CST	9202	Rabbit polyclonal	70,85	5% BSA PBST	1/2000
SERPINE1	PAI-1 (C-9)	SCBT	sc-5297	Mouse monoclonal	50-55	5% BSA PBST	1/300

STRN	STRN	Sigma	HPA017286	Rabbit polyclonal	100	5% milk PBST	1/2000
TUBA4A	Tubulin alpha	Sigma	T9026	Mouse monoclonal	≈50	5% BSA PBST	1/20000
VEGFA	VEGFA	Abcam	ab46154	Rabbit polyclonal	23-27	5% BSA PBST	1/1000
VHL	VHL	CST	68547	Rabbit polyclonal	18-22	5% BSA PBST	1/2000
YBX1	YB1	CST	4202	Rabbit polyclonal	49	5% BSA PBST	1/5000
/	Phospho-tyrosine (4G10)	Millipore	05-321	Mouse monoclonal	/	/	/
/	Rabbit IgG control	R&D	AB-105-C	/	/	/	/
/	Mouse IgG control	R&D	MAB002	/	/	/	/
Secondary Antibodies	Anti-rabbit IgG, HRP-linked	CST	7074	/	/	/	1/3000
	Anti-mouse IgG, HRP-linked	CST	7076	/	/	/	1/3000
	EasyBlot antiMouse IgG (HRP)	GeneTex	GTX221667-01	/	/	/	1/1000
	EasyBlot antiRabbit IgG (HRP)	GeneTex	GTX221666-01	/	/	/	1/1000
	VeriBlot anti Mouse IgG (HRP)	Abcam	ab131368	/	/	/	1/1000
	Jackson anti Rabbit IgG (HRP)	Jackson Immunoresearch	211-032-171	/	/	/	1/1000

Immunoprecipitation & co-immunoprecipitation (IP/Co-IP)

Different lysis buffers for the immunoprecipitation of FGFR3 were compared to select the optimal method. As listed below:

Lysis buffer 1: 50 mM HEPES pH 7.5, 250 mM NaCl, 1% Triton X-100 + protease and phosphatase inhibitors;

Lysis buffer 2: 25 mM Tris pH 7.5, 50 mM NaCl, 0.1% NP40 + protease and phosphatase inhibitors;

Lysis buffer 3: 50 mM Tris pH 7.5, 150 mM NaCl, 1% NP40 + protease and phosphatase inhibitors;

Lysis buffer 4: 50 mM Tris pH 7.5, 150 mM NaCl, 0.1% NP40 + protease and phosphatase inhibitors.

cOmplete™, EDTA-free Protease inhibitor cocktail (1 tablet for a volume of 50 ml solution) and PhosStop-Phosphatase inhibitor cocktail (1 tablet for a volume of 50 ml solution) were purchased from Sigma (cat.no. 11873580001 and 4906837001, respectively). Cell lysates were clarified by centrifugation and the protein concentration was determined with BCA Protein Assay kit (ThermoFisher). 100 µl of each sample were kept as an input control. After pre-cleaning (4 h at 4 °C) of lysate using protein G/A beads (GE healthcare), FGFR3-bound complexes were immunoprecipitated using two specific antibodies against the receptor and incubated overnight at 4 °C. 20-30ul beads were added to the mixtures next morning and continually rotated the mixtures for 1 h at 4 °C. The mixtures were rinsed five times with lysis buffer. The protein complex was eluted from the beads with 25 ul 4 x Laemmli loading buffer (BioRad) and boiled for 10 min at 95 °C. Tyrosine phosphorylated complexes were similarly enriched with the aforementioned process with specific antibody (anti-pY, 4G10). The immunoprecipitated proteins were separated by SDS-PAGE using 4-15% Tris-glycine precast gel (BioRad) and analyzed by Western blotting as described above.

Protein antibody array

The analysis of phosphorylation profiles of 43 kinases and their protein substrates was performed with the Human Phospho-Kinase Array as described in the manufacturer's instructions (R&D Systems, cat.no. ARY003B). Capture and control antibodies are spotted in duplicate on nitrocellulose membranes. 200-600 ug cell lysates were diluted and incubated overnight with the Human Phospho-Kinase Array. The array was washed to remove unbound proteins followed by incubation with a cocktail of biotinylated detection antibodies. Streptavidin-HRP and chemiluminescent detection reagents were applied and a signal was produced at each capture spot corresponding to the amount of phosphorylated level of a protein. Image J software was used to quantify the intensity of these capture spots.

Kinases included in Human Phospho-Kinase Array kit are as followed:

Simultaneously detect the relative phosphorylation of these proteins in a single sample.		
Akt 1/2/3 (S473)	Hck (Y411)	PLC gamma-1 (Y783)
Akt 1/2/3 (T308)	HSP27 (S78/S82)	PRAS40 (T246)
AMPK alpha1 (T183)	HSP60	Pyk2 (Y402)
AMPK alpha2 (T172)	JNK 1/2/3 (T183/Y185, T221/Y223)	RSK1/2/3 (S380)
beta-Catenin	Lck (Y394)	Src (Y419)
Chk-2 (T68)	Lyn (Y397)	STAT2 (Y689)
c-Jun (S63)	MSK1/2 (S376/S360)	STAT3 (S727)
CREB (S133)	p27 (T198)	STAT3 (Y705)
EGF R (Y1086)	p38 alpha (T180/Y182)	STAT5a (Y699)
eNOS (S1177)	p53 (S15)	STAT5a/b (Y699)
ERK1/2 (T202/Y204, T185/Y187)	p53 (S392)	STAT5b (Y699)
FAK (Y397)	p53 (S46)	STAT6 (Y641)
Fgr (Y412)	P70 S6 Kinase (T389)	TOR (S2448)
Fyn (Y420)	p70 S6 Kinase (T421/S424)	WNK-1 (T60)
GSK-3 alpha/beta (S21/S9)	PDGF R beta (Y751)	Yes (Y426)

The analysis of expression profiles of 55 angiogenesis-related proteins was performed with the membrane-based sandwich immunoassay (Human Angiogenesis Array) as described in the manufacturer's instructions (R&D Systems, cat.no. ARY007). Capture and control antibodies were spotted in duplicate on nitrocellulose membranes. 100-300 ug mixture of samples and a cocktail of biotinylated detection antibodies were diluted and incubated overnight with the Human Angiogenesis Array kit. Streptavidin-HRP and chemiluminescent detection reagents were applied and a signal was produced at each capture spot corresponding to the amount of protein expression level. Image J software was used to quantify the intensity of these capture spots.

Angiogenesis-related proteins included in Human Angiogenesis Array kit are as followed:

Simultaneously detect the relative levels of these angiogenesis-related proteins in a single sample.		
Activin A	FGF-7/KGF	PD-ECGF
ADAMTS-1	GDNF	PDGF-AA
Angiogenin	GM-CSF	PDGF-AB/PDGF-BB
Angiopoietin-1	HB-EGF	Persephin
Angiopoietin-2	HGF	CXCL4/PF4
Angiostatin/Plasminogen	IGFBP-1	PlGF
Amphiregulin	IGFBP-2	Prolactin
Artemin	IGFBP-3	Serpin B5/Maspin
Tissue Factor/Factor III	IL-1 beta	Serpin E1/PAI-1
CXCL16	CXCL8/IL-8	Serpin F1/PEDF
DPPIV/CD26	LAP (TGF-beta 1)	TIMP-1
EGF	Leptin	TIMP-4
EG-VEGF	CCL2/MCP-1	Thrombospondin-1
Endoglin/CD105	CCL3/MIP-1 alpha	Thrombospondin-2
Endostatin/Collagen XVIII	MMP-8	uPA
Endothelin-1	MMP-9	Vasohibin
FGF acidic	NRG1-beta 1	VEGF
FGF basic	Pentraxin 3	VEGF-C
FGF-4		

Immunohistochemistry (IHC)

Tumor xenografts were fixed in alcohol-formaldehyde-acetic acid (AFA) or buffered formol and then paraffin-embedded and processed into 4 µm thick cuts and placed on glass slides. These slides were deparaffinized in xylene at 37 °C and rinsed in medicinal−graded ethanol (100, 100, 95, 70 and 50%). Subsequently, antigen retrieval was achieved by heating in unmasking buffer (Tampon citrate pH 6) for 20 min at 95 °C. Next, 3% H2O2 (Sigma, cat.no. 216763) was used to inactivate endogenous peroxidase and slides were incubated in blocking buffer (QUANTO protein block, MM France) for 1h. Primary and secondary antibody reactions were followed for 2 h and 1 h, respectively. Finally, a freshly prepared DAB colorant (MM France) was used for dyeing, followed by quick staining with haematoxylin solution

(Sigma), and then dehydrated in medicinal−graded ethanol (70, 95, 100 and 100%) and xylene. Photos were taken with an OLYMPUS VANOX micrographic system. The results were assessed by an experienced pathologist specialized in urothelial carcinoma. The following antibodies were used for staining to assess vascularization: CD34 (1/200; Abcam, ab81289) and ERG (1/200; Abcam, ab92513).

DNA array

For the identification of genes displaying changes in expression after the depletion of *FGFR3* in UMUC-14 cells, we transfected the cells for 40 h with *FGFR3* siRNA #3, *FGFR3* siRNA #4 or lipofectamine alone as control. Total RNA (200 ng) from control and siRNA-treated UMUC-14 cells was analyzed with the Affymetrix U133 plus 2 DNA array. The microarray data described here are available from GEO (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE125547. The LIMMA algorithm was used to identify genes differentially expressed between *FGFR3* siRNA-treated and lipofectamine-treated cells (three replicates) (Ritchie et al., 2015). The P-values were adjusted for multiple testing by Benjamini-Hochberg FDR methods. Genes with a log2 fold-change of at least 0.58, in a positive or negative direction, with a FDR below 5%, were considered to be differentially expressed.

Sample preparation for mass spectrometry

Interactome. UMUC-14 cells were lysed in modified RIPA buffer (Lysis buffer 3). Proteins extracted from triple SILAC (L/M/H) labeled conditions were mixed 1:1:1 (8 mg proteins in total) and followed by IP-CoIP of FGFR3. The protein elution from immunoprecipitation was separated by 1D-PAGE electrophoresis gel (7.5%, BioRad) and then the gel was stained by coomassie dye (LabSafe GEL Blue™, cat.no. 786-35) and was reduced to a thickness of 1-2 mm. Next a classical 'In-gel digestion' methods (Granvogl et al., 2007) was applied following several steps below:

1) destaining of coomassie dye with 25 mM ammonium bicarbonate (AmBic) in 50% acetonitrile; 2) reduction of disulfide bridges (e.g. incubate with 5 mM DTT in 100 mM AmBic for 45 min at 56 °C) and alkylation of the SH groups (e.g. incubate in the dark with 10 mM CAA in 100 mM AmBic for 45 min at room temperature); 3) shrink of gel pieces in 50 µl of acetonitrile; 4) trypsinization of proteins with 10 ng/µL Trypsin/LysC solution (prepared in 50 mM AmBic with 10% acetonitrile, Promega) for 15-30 min and recover of gel fragments in 50

mM AmBic overnight at 30 °C; 5) collection of peptides with desalting stepsand dryness using SpeedVac. Sample is now ready for mass spectrometry.

Phosphoproteome. UMUC-14 cells were lysed in urea buffer (8 M urea in 50 mM AmBic, supplemented with protease and phosphatase inhibitors). Proteins extracted from triple SILAC (L/M/H) labeled conditions were mixed 1:1:1 (3 mg proteins in total). Firstly, samples were reduced, alkylated and digested as aforementioned steps. Next, peptides were desalted through Sep-Pak C18 cartridges (Waters) prior to their enrichment or fractionation. Titanium dioxide (TiO2) beads (GL Sciences) were used for the enrichment of phosphorylated peptides, as described before (Larsen et al., 2005). Briefly, desalted peptides and TiO2 beads mixture were packed by centrifugation in equilibrated C18 spin columns (ThermoFisher, cat.no. 164535). Beads were sequentially washed with 300 µL of glycolic acid solution, 50% acetonitrile, AmBic solution (20 mM AmBic pH 6.8 in 50% acetonitrile) and 50% acetonitrile. Afterwards, phosphopeptides were eluted using NH4OH, acidified, dried in a SpeedVac and, after reconstitution, analyzed by mass spectrometry.

Liquid Chromatography coupled to Mass Spectrometry (LC-MS/MS)

Peptides were resuspended in solvent A (2% acetonitrile, 0.085% formic acid), separated and analyzed by nanoLC-MS/MS using an UltiMate 3000 RSLCnano system (ThermoFisher) coupled to an Orbitrap Fusion mass spectrometer (Q-OT-qIT, ThermoFisher). Samples were trapped on a C18 µ-precolumn (75 µm inner diameter × 2 cm; nanoViper Acclaim PepMap^{TM} 100, ThermoFisher, cat.no. 164942) at 2.5 µl/min in solvent A. After a desalting of 4 min, the precolumn was switched on the C18 column (75 µm i.d. × 50 cm, packed with Acclaim PepMap100 C18, 3 µm, 100 Å; ThermoFisher, cat.no. 164535) equilibrated in 99% solvent A and 1% solvent B (100% acetonitrile, 0.085% formic acid) for the phosphorylated peptides and in 5% solvent B for the proteome fractions. Bound enriched modified-peptides were eluted by using three step linear gradients of 211 min (phospho-enriched fractions) or linear gradients of 100 min (interactome fractions) at a 300 nl/min flow rate and an oven temperature of 55°C. Proteome elution: 100 min from 5 to 30% (v/v) of solvent B. Phospho-enriched 211 min elution: 61 min from 1 to 5% (v/v), 120 min from 5 to 20% (v/v) and 30 min from 20 to 40% (v/v) of solvent B. We acquired Survey MS scans at a resolution set to a value of 120,000, with a mass range of m/z 400–1500 and a 4 × 105 ion count target. Each scan was recalibrated in real time by co-injecting an internal standard from ambient air (445.12003 m/z) into the C-trap. Tandem MS was performed by isolation at 1.6 Th or 1.2 h (for the modified enriched samples) with the

quadrupole, HCD fragmentation with normalized collision energy of 28, and rapid scan MS analysis in the ion trap. The MS2 ion count target was set to 1 × 104 and the max injection time was 100 ms and only those precursors with charge state from 2 to 7 were sampled for MS/MS acquisition. The dynamic exclusion duration was set to 30 or 15 s (for the modified enriched peptides) with a 10 ppm tolerance around the selected precursor and its isotopes. The instrument was run in top speed mode with 3 s cycles.

In vivo models

Xenograft models. Six-week-old female Swiss *nu/nu* mice (Charles River Laboratories) were raised in the animal facilities of Institut Curie, in specific pathogen-free conditions. They were housed and cared for in accordance with the institutional guidelines of the French National Ethics Committee (Ministère de l'Agriculture et de la Forêt, Direction de la Santé et de la Protection Animale, Paris, France), under the supervision of authorized investigators. Mice received a subcutaneous injection, into each flank (dorsal region), of $5x10^6$ MGH-U3 or RT-112 BCa cells in 100 µl PBS. For each study, with each of the cell lines, mice were randomly separated into two groups when tumors reached a volume of 100 mm^3 (±20). For FGFR3 inhibition studies, the mice were treated daily for 9 days, by oral gavage with PD173074 (25 mg/kg; n = 4) in one group and with vehicle (0.05 M acetate buffer) in the other (n = 4). The tumors were then removed. Part of the tumor was flash-frozen in liquid nitrogen for protein extraction in Laemmli buffer. For each treatment, the tumor was measured twice weekly with calipers, and its volume in mm^3 was calculated with the formula: $\pi/6$ x (largest diameter) x $(\text{shortest diameter})^2$.

Patient-derived Tumor Xenograft (PDX) model (F659). A patient-derived bladder cancer xenograft model (F659) was established as follow. A fresh specimen was collected from a patient diagnosed with a muscle-invasive bladder carcinoma with two positive perivesical lymph nodes (pT3bN2Mx), in accordance with French regulations concerning patient information and consent and then xenografted subcutaneously in the interscapular space of 5-week-old male Swiss *nu/nu* mice (Charles River Laboratories) and serially passaged into male Swiss *nu/nu* mice (Charles River Laboratories). DNA was isolated from snap-frozen tumor from the patient and from the PDX tumor (at passage 3 in mice), with a classical phenolchloroform-isoamyl alcohol extraction protocol. *FGFR3* mutations were studied by the SNaPshot method, as previously described (Van Oers et al., 2005), and a *FGFR3* S249C heterozygous mutation was detected in both samples. PDX (F659) tumor tissue at passage 4 in

mice was cut into small pieces (5 mm^3) and subcutaneously xenografted into multiple mice in the interscapular region. When tumor sizes reached 100-200 mm^3, mice were randomly divided into two groups and treated by daily oral gavage with BGJ398 (30 mg/kg, LC Laboratories) or vehicle (0.05 M acetate buffer). Tumor growth was measured twice weekly with an electronic caliper, and tumor volume was calculated and expressed relative to the initial size of the tumor. Two experiments were conducted as follows: one for a long-term treatment (29 days; n = 5 animal per group) in which tumors were monitored for two additional weeks after the end of treatment, and one for a short-term treatment over a period of 4 days (n = 4 animal per group). The mice were sacrificed at the end of the experiments. Their tumors were harvested and flash-frozen for further experimental exploration.

Data processing and bioinformatic analysis

Protein-peptide identification and quantitation. MS data were acquired using the Xcalibur software (v 3.0) and the resulting raw data files were analysed with MaxQuant (Cox and Mann, 2008) version 1.5.2.8, which includes the Andromeda search engine. Peak lists were searched against human UniProt database version 2016.01.21, 42045 entries. Variable modifications for the searches included N-terminal protein acetylation, methionine oxidation, deamidation of asparagine/glutamine and phosphorylation on serine/threonine/tyrosine residues. Carbamidomethylation of cysteine was set as fixed modification. The maximum number of modifications accepted per peptide was six and the minimum peptide length was set to seven amino acids. Trypsin was used as protease with a maximum of two missed cleavages. Both peptide and protein maximum false discovery rates (FDR) were set to 0.01 based on the target-decoy approach. Known common contaminants, as specified in MaxQuant, were also included in the search. Proteins were considered identified and quantified confidently with presence of at least two unique peptides and at least 2 MS counts in two out of three replicates. Only phosphosites that were identified in more than two replicates and with a localization probability of at least 0.75 across all replicates were further screened. We then considered a modification site as well-quantified if it presents at least 2 MS counts in two out of three replicates.

Bioinformatic analysis. Perseus (Tyanova and Cox, 2018) and Venny 2.1 web application (http://bioinfogp.cnb.csic.es/tools/venny/) were used for the construction of hierarchical clustering, principal component analysis (PCA) and Venn diagrams, respectively. The Gene Ontology (GO), pathways enrichment and upstream regulator prediction analysis were performed in Enrichr (Kuleshov et al., 2016) and Ingenuity Pathway Analysis (IPA) software.

Protein-protein interaction (PPI) network analysis was conducted in STRING (https://string-db.org/) and Agile Protein Interactomes DataServer (APID) (Alonso-López et al., 2016). Upstream kinase prediction was performed in NetworKIN 3.0 (Horn et al., 2014) and Kinase Enrichment Analysis (KEA) (Lachmann and Ma'ayan, 2009). Enrichr was used for upstream phosphatase prediction. Motif enrichment was assessed by applying motif-X (Chou and Schwartz, 2011) to phosphosites that were regulated by FGFR3. A probability value of 0.05 was considered as significant. Plots were generated with Microsoft Excel 2016, GraphPad Prism 7.02 or R version 3.5.2 using package 'easyGgplot2', version 1.0.0.9000. Figures were assembled in Adobe Illustrator or Microsoft PowerPoint 2016.

3.3.3 Results and Discussion

Working cell model selection for -omic screening

In our available BCa cell lines bank in the laboratory, only several cell lines with *FGFR3* genetic alterations showed a detectable endogenous protein expression of FGFR3 (Figure 3-1A; page 179). RT-112 and RT4 cells harbour FGFR3-TACC3 translocation and showed the highest expression of FGFR3, followed by UMUC-14 and MGH-U3 cells who bear S249C and Y375C mutations, respectively. Although HT1197 and J82 cells contained S249C and K652E mutations, respectively, we did not observe a significant FGFR3 expression. FGFR3 protein was not detectable in other cell lines with wild-type FGFR3, including those EGFR-dependent cell lines. Given that *FGFR3* mutations are much more frequent than translocations in BCa, we were more interested in investigating aberrant signaling induced by mutated forms of FGFR3. Herein, UMUC-14 (*FGFR3* S249C) and MGH-U3 (*FGFR3* Y375C) cells were suitable to decipher FGFR3 regulated signaling driven by mutations at endogenous level. Considering the amount of work and the cost for our large-scale -omic screening, it's more realistic to select only one cell line to generate -omics datasets for further analysis. For this aim, I compared UMUC-14 and MGH-U3 cell lines in terms of FGFR3 activation (tyrosine phosphorylation) levels. However, the challenge is that no publicly available antibodies could be used to detect reliably phosphorylated FGFR3. We thus intended to perform forward and reversed IP-CoIP

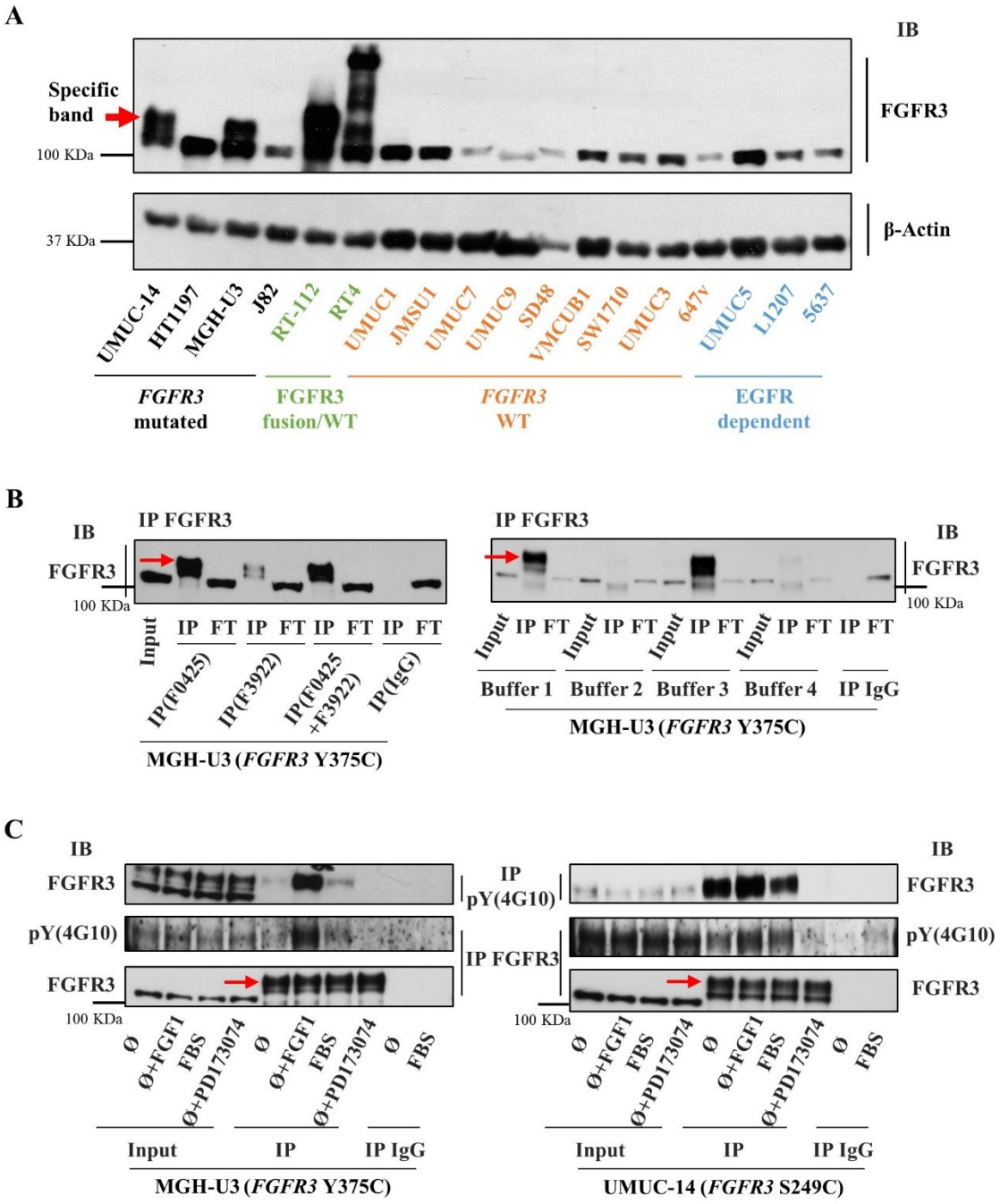


Figure 3-1: *Working cell model selection.*

A) *Detection of FGFR3 protein expression in multiple bladder cancer cell lines. Genetic alternations of each cell line were defined according to previous reports (described in Method section).* ***B)*** *Left, test of antibodies combination for immunoprecipitating endogenous FGFR3, IgG as negative control; Right, test for IP-CoIP efficiency with different lysis buffers. F0425, cytoplasmic antibody anti-FGFR3; F3922, extracellular antibody anti-FGFR3; Detailed lysis buffers' composition see Method section.* ***C)*** *Detection of FGFR3 activation (level of tyrosine phosphorylation) status by forward and reverse IP-CoIPs in different conditions (starving,*

FGF1 stimuli, FGFR3 inhibition) and horizontal comparison between MGH-U3 (left part) and UMUC-14 (right part) cell lines (both are FGFR3 mutated cells). IP, immunoprecipitation; FT, flow through (IP fractions); IB, immunoblotting; Ø, starving cells for 16 h;FGF1, 40 ng/ml for 10 min in presence of Heparin 30ug/ml; PD173074, selective pan-FGFRs inhibitor (used 1 uM for 1 h).

with anti-tyrosine antibody (4G10) to assess phosphorylated tyrosine level of FGFR3. First, we optimized basic conditions for achieving a successful IP-CoIP of FGFR3 endogenously in MGH-U3 cells, for example, the use of antibodies combination or not for IP of FGFR3 and the impact of different lysis buffers (Figure 3-1B; page 179). We found that the F3922 (extracellular antibody anti-FGFR3) was not efficient for IP of FGFR3, instead, the application of 1:1 mixture of extracellular (F3922) and intracellular (F0425) antibodies of anti-FGFR3 worked well (Figure 3-1B, left; page 179). The most successful IP of FGFR3 was obtained from the lysis buffer 3 (50 mM Tris pH 7.5, 150 mM NaCl, 1% NP40 + protease and phosphatase inhibitors) while other buffers were much less efficient (Figure 3-1B, right; page 179). Consequently, throughout our following IP experiments, we utilized combined F3922+F0425 antibodies for IP of FGFR3 and lysis buffer 3 to obtain whole cell extract for IP-CoIPs. We found that FGF1 ligand stimulated full activation of FGFR3 and FGFR3 inhibitor (PD173074) completely inhibit FGFR3 activity in both MGH-U3 and UMUC-14 cells. However, in the absence of FGF1 stimuli (see in starvation condition in our blot results), there was already a basal and constitutive activation of FGFR3 in UMUC-14 but not in MGH-U3 cells (Figure 3-1C; page 179). Notably, the starvation condition could decrease much of stimuli interferences and may provide a more pure environment to explore FGFR3-driven signaling. In brief, UMUC-14 cells were selected as working cell model for -omic screening because of several advantages against MGH-U3.

MS-based FGFR3 interactome

Workflow for MS-based FGFR3 interactome. Triple SILAC strategy was applied for UMUC-14 cells. All cells were starved (0% FBS) for 16 h before treatment, ‘Light (L)’ and ‘Medium (M)’ cells were treated with DMSO and ‘Heavy (H)’ cells were treated with 500 nM of FGFR3 inhibitor (PD173074) for 1 h. ‘L’ condition corresponded to IP of IgG negative control, ‘M’ and ‘H’ conditions corresponded to IPs of FGFR3. Following steps for sample preparation and MS data acquisition are described in the Methods section. A simplified workflow is presented in Figure 3-2A; page 181. Of note, we also optimized the amount of antibodies and proteins

used for IPs. We finally took 1 µg antibodies mixture (F3922/F0425 = 1 : 1) per microgram of proteins and 5 mg proteins for each condition as optimal parameters (Figure 3-2B&C; page 181).

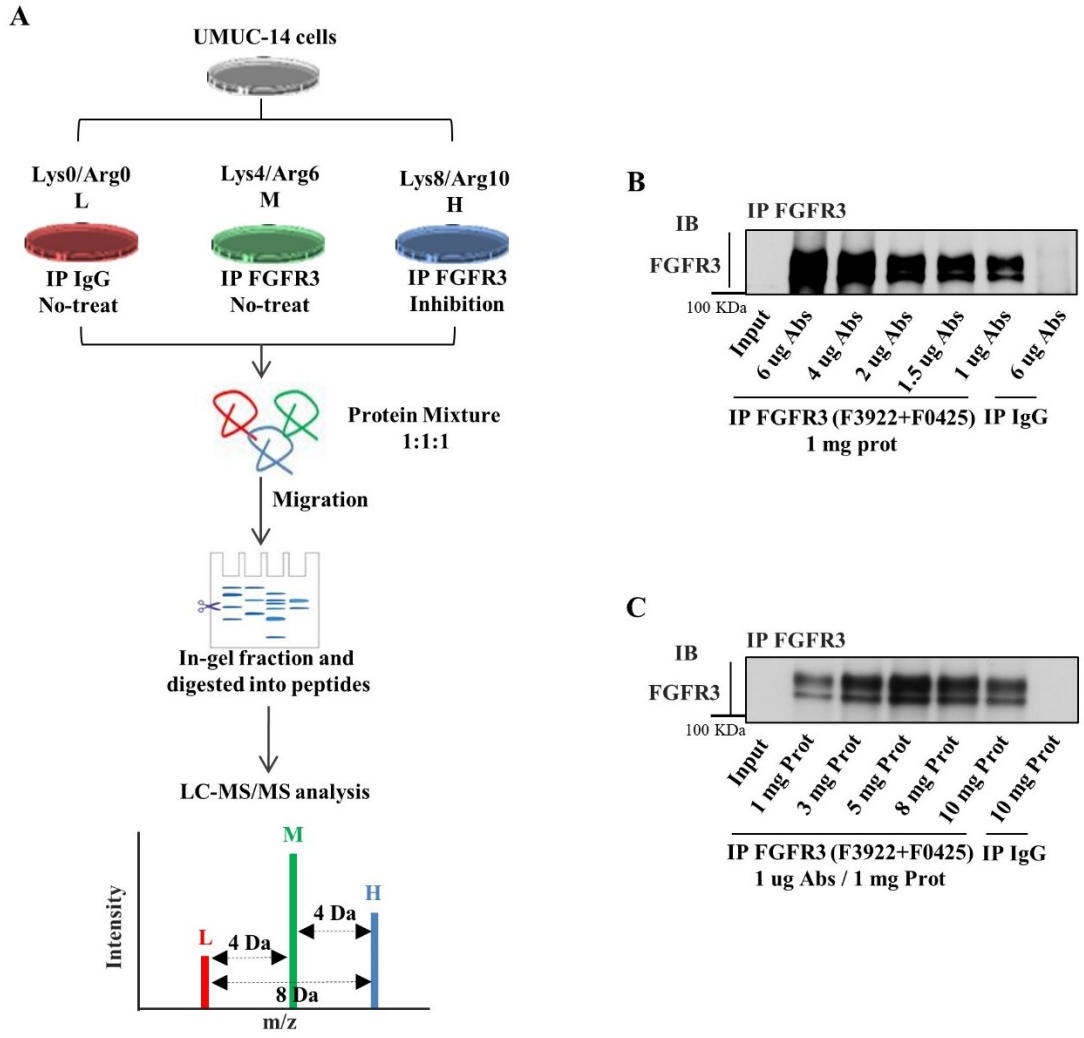


Figure 3-2: *Workflow and optimized conditions for MS-based FGFR3 interactome.*
A) *Graphical workflow for FGFR3 interactome. No-treat, DMSO; FGFR3 inhibition, using PD173074 inhibitor at 500 nM for 1 h.* ***B)*** *Optimization for the amount of antibodies (Abs) for IP with fixed 1 mg protein. The F3922 and F0425 were mixed with equal quantity.* ***C)*** *Optimization for the amount of proteins (Prot) for IP with fixed Abs quantity.*

67 highly confident partners of FGFR3. We can expect the binding partners of FGFR3 will be dynamically changed (detached, constitutive or newly recruited) upon its activating or inhibiting status as suggested in schematic diagram (Figure 3-3A; page 182). According to data

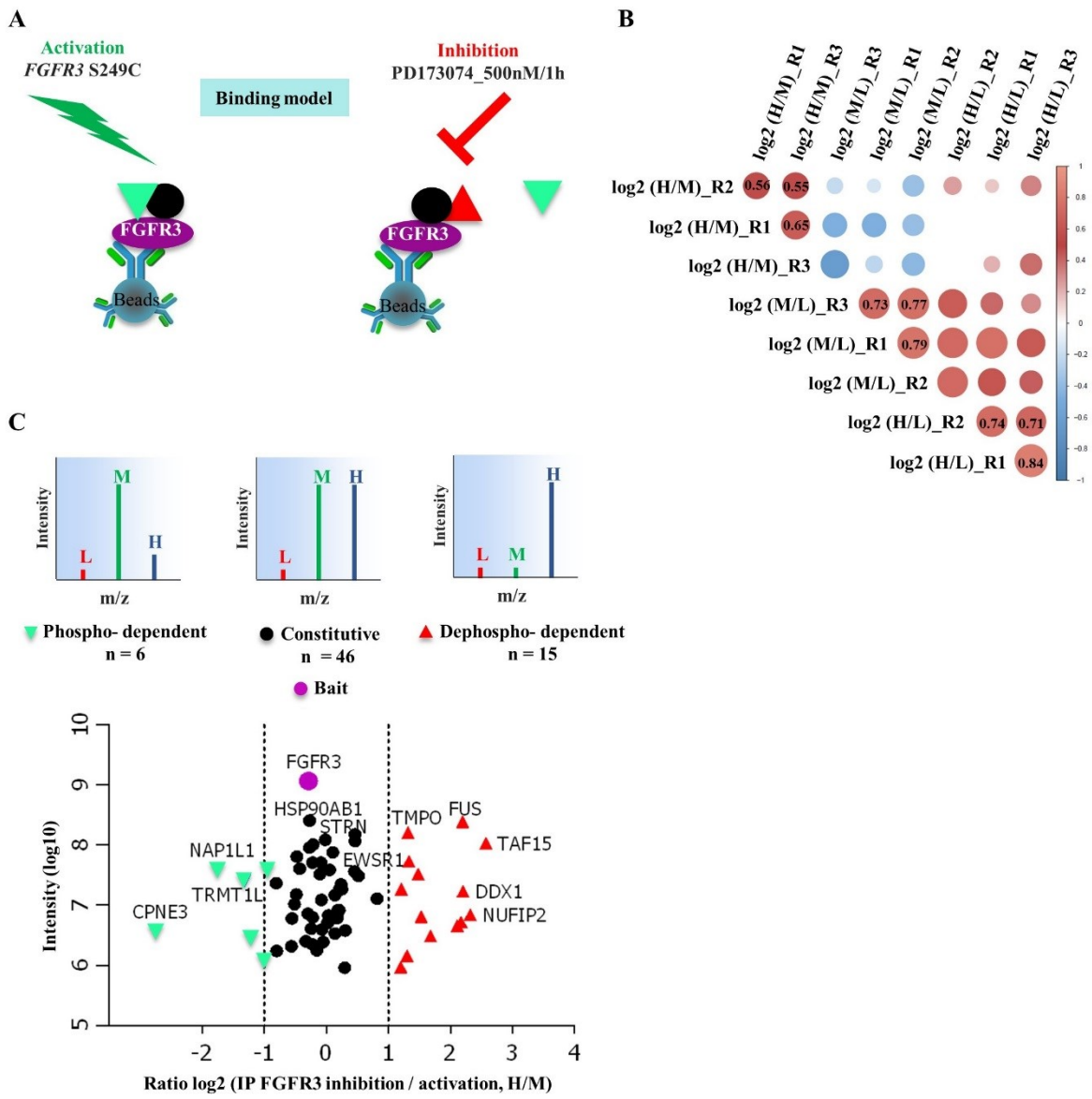


Figure 3-3: *67 highly confident partners' binding profiles.*

A) *Three types of dynamic binding models upon FGFR3 activation or inhibition.* ***B)*** *Correlations among triplicate reflect experimental reproducibility (411 identified and quantified proteins). L, IP-IgG control; M, IP-FGFR3 of activated FGFR3; H, IP-FGFR3 of FGFR3 inhibition; R, replicate.* ***C)*** *67 interactors were considered as highly confident by taking cut-off of either log2(M/L) or log2(H/L) over than '1'. They were divided into three types: 6 phospho-dependent (or detached), 46 constitutive and 15 dephospho-dependent (or new) binders.*

processing criteria described in the method section, we finally obtained a list of 411 proteins that were identified and quantified from MS in overall three replicates. Because there were few variations for those 411 potential candidates when comparing IP-FGFR3 inhibition versus IP-

FGFR3 activation (H/M) condition in all replicates, the strong correlations among replicates would not be expected (Figure 3-3B; page 182). The remaining results from all replicates were significantly correlated with each other, thus indicating a reproducible protocol (Figure 3-3B; page 182). To pinpoint highly confident interactors, we took a stringent cut-off ratio of either log2 (IP-FGFR3 activation / IP-IgG control, M/L) or log2 (IP-FGFR3 inhibition / IP-IgG control, H/L) over than '1' and obtained 68 candidates (including FGFR3 itself) eventually. Further, by comparing FGFR3 inhibition versus activation condition, we divided these 67 partners into three groups by a cut-off of '$|\log2 (H/M)| \geq 1$': 6 phospho-dependent (or detached), 46 constitutive and 15 dephospho-dependent (or new) binders of FGFR3 (Figure 3-3C; page 182). A detailed list of these partners was presented in Table 3-2, page 183.

Table 3-2: *List of 67 highly confident interactors of FGFR3.*
*Identified from UMUC-14 cells in triplicate experiment. *, combined value from triplicate. Evidence level: +++, reached quantification threshold in all three replicates; ++, reached quantification threshold in two replicates and with only one MS count in the third replicate; +, reached quantification threshold in two replicates but without MS count in the third replicate.*

Num	Genes	IDs	Type	Unique peptides *	Ratio MS count*	IP-FGFR3 inhibition / IP-FGFR3 activation	IP-FGFR3 activation / IP-IgG control	IP-FGFR3 inhibition / IP-IgG control	Evidence level
						fold change H/M*	fold change M/L*	fold change H/L*	
1	FGFR3	P22607-2	/	29	101	0.82	51.24	40.06	+++
2	BTBD10	Q9BSF8	Constitutive	7	10	0.94	26.26	27.44	+++
3	STRN	O43815	Constitutive	27	46	0.99	17.46	18.92	+++
4	SEC24C	P53992	Constitutive	19	22	1.04	12.83	12.84	+++
5	CUL7	Q14999	Constitutive	11	11	0.82	12.82	9.44	+++
6	WTAP	Q15007	Constitutive	16	45	1.38	10.77	15.11	+++
7	KIAA1429	Q69YN4-3	Constitutive	30	51	0.74	10.49	12.99	+++
8	CDC37	Q16543	Constitutive	7	11	0.72	10.28	5.68	+++
9	TNFRSF21	O75509	Constitutive	6	11	1.02	9.66	7.86	+++
10	KCTD5	Q9NXV2	Constitutive	8	25	0.83	8.88	5.86	+++
11	CAMK2G	Q13555-4	Constitutive	9	26	1.08	8.00	7.97	+++
12	CAMK2B	Q13554	Constitutive	9	26	1.08	8.00	7.97	+++
13	CBLL1	Q75N03	Constitutive	8	10	1.19	7.55	7.84	+++
14	CAMK2D	Q13557-8	Constitutive	6	8	1.16	7.52	8.87	+++
15	SEC23A	Q15436	Constitutive	11	17	1.18	7.51	8.80	+++
16	ITGB4	P16144-4	Constitutive	17	40	0.57	6.15	3.42	+++
17	TFRC	P02786	Constitutive	25	77	0.86	5.00	4.24	+++

18	YBX3	P16989-2	Constitutive	6	12	0.95	3.73	2.95	+++
19	YBX1	P67809	Constitutive	7	35	0.86	3.53	2.01	+++
20	PPP2CA	P67775	Constitutive	6	10	1.14	3.26	3.63	+++
21	PPP2CB	P62714	Constitutive	6	10	1.14	3.26	3.63	+++
22	PPP2R1A	P30153	Constitutive	11	17	1.10	3.10	3.16	+++
23	RBM15	Q96T37-4	Constitutive	18	39	1.37	2.90	4.96	+++
24	HSP90AA1	P07900	Constitutive	15	47	0.72	2.71	1.99	+++
25	EWSR1	Q01844-6	Constitutive	6	29	1.43	2.36	2.31	+++
26	PABPC1	P11940	Constitutive	17	77	1.38	2.36	4.12	+++
27	HSP90AB1	P08238	Constitutive	16	143	0.83	2.29	2.15	+++
28	FXR1	P51114-2	Constitutive	10	19	1.76	2.24	4.51	+++
29	MOV10	Q9HCE1-2	Constitutive	8	14	1.10	2.17	2.75	+++
30	UPF1	Q92900-2	Constitutive	15	27	1.16	2.02	2.29	+++
31	KCTD20	Q7Z5Y7	Constitutive	8	5	0.93	20.32	12.14	++
32	DHX34	Q14147	Constitutive	15	7	0.70	18.04	14.57	++
33	FMNL3	Q8IVF7	Constitutive	8	6	0.68	14.77	10.85	++
34	TNKS1BP1	Q9C0C2	Constitutive	18	9	1.07	11.47	10.96	++
35	STRIP1	Q5VSL9	Constitutive	5	5	1.24	9.30	16.51	++
36	SEC23B	Q15437	Constitutive	5	6	0.95	6.89	6.21	++
37	MRPS22	P82650	Constitutive	3	5	0.68	4.77	2.21	++
38	STX6	O43752	Constitutive	4	7	0.86	4.74	3.66	++
39	EPHA2	P29317	Constitutive	5	11	0.86	3.93	2.18	++
40	SLTM	Q9NWH9	Constitutive	3	5	0.97	3.55	4.67	++
41	LMAN2	Q12907	Constitutive	5	6	0.90	3.08	2.23	++
42	PGAM5	Q96HS1	Constitutive	8	10	1.13	2.78	3.42	++
43	HSP90AB2P	Q58FF8	Constitutive	2	5	0.85	2.72	2.11	++
44	PTRF	Q6NZI2	Constitutive	3	5	1.01	2.65	2.10	++
45	PLXNB2	O15031	Constitutive	6	7	0.58	4.57	2.41	+
46	EGFR	P00533	Constitutive	3	7	0.79	3.66	2.52	+
47	TUBGCP3	Q96CW5-2	Constitutive	3	4	1.23	2.84	3.71	+
1	TAF15	Q92804-2	New	5	34	5.95	0.42	3.48	+++
2	NUFIP2	Q7Z417	New	7	13	4.99	1.04	4.98	+++
3	DDX1	Q92499	New	13	28	4.60	1.08	5.55	+++
4	FUS	P35637-2	New	7	65	4.59	0.83	3.68	+++
5	ATXN2L	Q8WWM7-6	New	8	17	2.89	1.46	4.47	+++
6	XRCC6	P12956	New	25	52	2.79	0.92	2.39	+++
7	DDX3X	O00571	New	23	58	2.51	1.21	2.68	+++
8	DDX3Y	O15523	New	23	58	2.51	1.21	2.68	+++
9	TMPO	P42166	New	17	94	2.50	0.88	2.58	+++

10	KHDRBS1	Q07666-2	New	4	6	4.32	0.84	3.78	++
11	CAPRIN1	Q14444-2	New	4	6	3.21	1.30	4.32	++
12	RTCB	Q9Y3I0	New	11	12	2.32	1.43	4.00	++
13	C14orf166	Q9Y224	New	6	6	4.51	1.12	4.23	+
14	EDC4	Q6P2E9	New	6	7	2.46	0.83	2.18	+
15	TRIM25	Q14258	New	3	4	2.30	1.32	2.71	+
1	NAP1L1	P55209-2	Detached	7	23	0.30	5.15	1.27	+++
2	TRMT1L	Q7Z2T5	Detached	17	20	0.40	13.76	4.46	+++
3	PLEKHA7	Q6IQ23	Detached	8	10	0.43	3.59	1.77	+++
4	RPL23	P62829	Detached	9	33	0.52	4.25	1.77	+++
5	CPNE3	O75131	Detached	3	5	0.15	5.63	1.09	+
6	HMGB1	P09429	Detached	2	4	0.50	2.08	0.91	+

Intersection of our list with literature reported FGFR3/FGFRs partners. As summarized previously, the only MS-based study exploring FGFR3 partners was conducted by Balek et al. in an overexpressed model and 45 reliable partners were screened (Balek et al., 2018). Balek et al. also summarized all validated partners of FGFR1-4 by IP or yeast two-hybrid (Y2H) methods in literature up to 2015. I updated this list up to 2019 by adding recently validated partners. Overall, I achieved a list of 85 known interactors. I overlapped these two datasets with my list as shown in Figure 3-4; page 186. FGFR3 was well-known as a strong client interactor of Hsp90/Cdc37 complex (CDC37 is known as regulatory subunit of Hsp90) (Taipale et al., 2012). Consistently, we found Hsp90/Cdc37 complex was common among all lists. 13/67 partners of my list can be recovered in the list from Balek et al. CBLL1, which was validated as FGFR2 partners in literature, was also found in my list. However, the most known adaptors like FRS2, GRB2, GAB1 and PLCG1, etc, were neither found in my list nor in Balek et al.'s list. FRS2 is a specific substrate of FGFRs, constitutively binds to common conserved region of FGFRs independent of FGF ligand stimuli or receptor activation / inhibition. Thus, it was very puzzling for the reason why FRS2 was not in my list. I suppose it may be matter of technique issue rather than context or receptor specificity. In brief, we identified many new partners of FGFR3 through endogenous working system in BCa that may enlarge our knowledge of FGFR3 signaling.

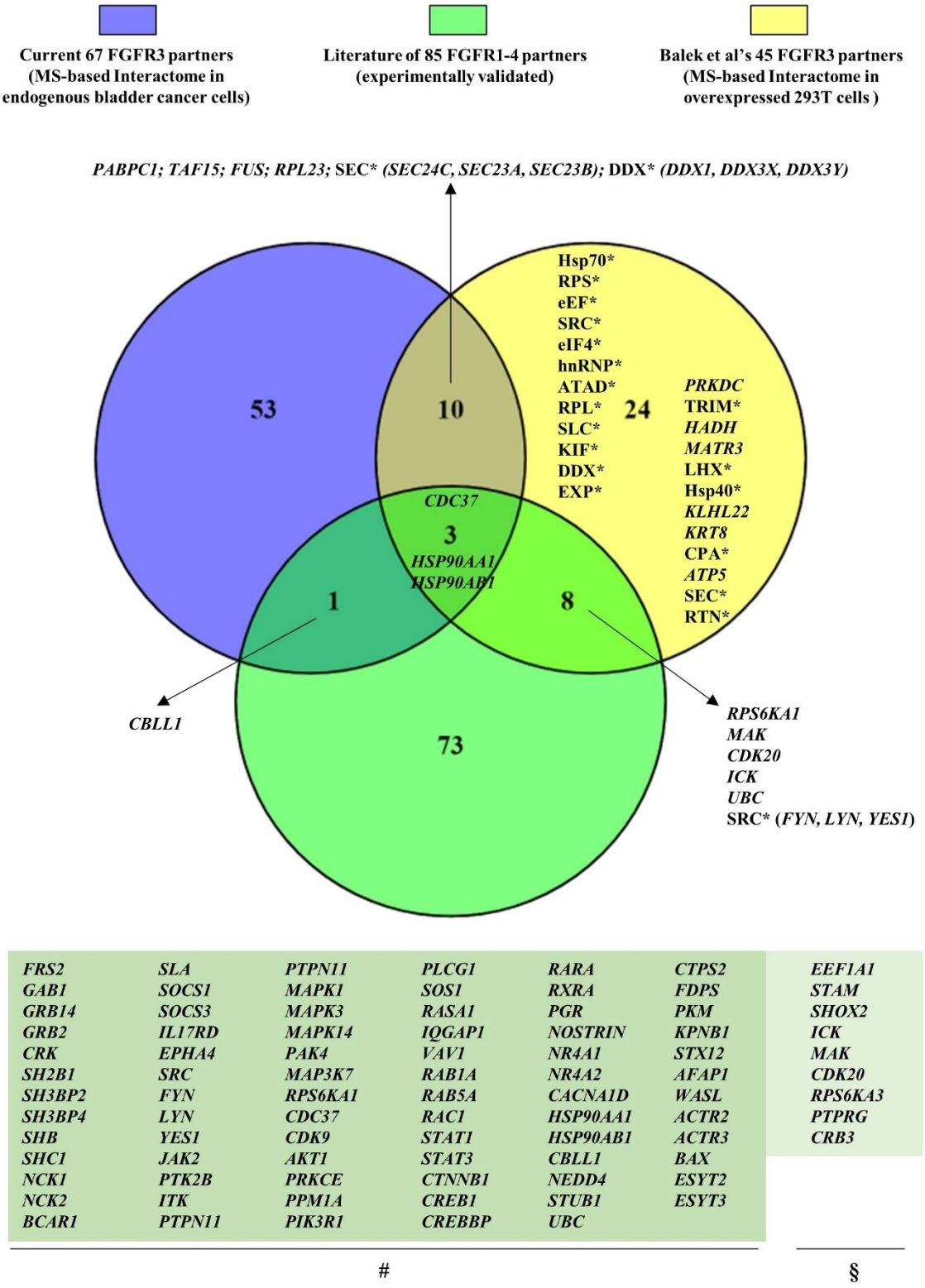


Figure 3-4: *Intersection of FGFR3/FGFRs partners.*

These partners including 67 currently identified FGFR3 partners by MS through endogenous bladder cancer cell line (blue), 45 MS-based FGFR3 partners through overexpression in 293T cell line that was reported by Balek et al (yellow) (Balek et al., 2018) and 85 experimentally

*validated FGFR1-4 partners by IP or yeast two-hybrid (Y2H) methods that were crawled in literature (green). The yellow list from Balek et al was reduced with more stringent criteria: 1) I did not consider the proteins identified from IP-pY(4G10) as partners; 2) a reliable partners should be identified at least in two out of sixteen experiments. *, proteins containing multiple family members; #, literature survey (1992 - 2015) of proteins interacting with FGFR1-4, see supplementary table S2 of Balek et al (2018); §, newly added partners of FGFR1-4 through updated literature (up to 2019) from below publications: (Balek et al., 2018; Iioka et al., 2019; Kostas et al., 2018; Kunova Bosakova et al., 2019; Nadratowska-Wesolowska et al., 2014).*

Protein complexes within my list of FGFR3 partners. We explored PPIs among 67 FGFR3 partners in STRING database to identify potential protein complexes. One third of them showed significant interactions (confidence level: 0.4) and formed distinct protein complexes (Figure 3-5A&B, page 188). The gene ontology (GO) analysis indicated an enrichment of cell-cell adhesion signal and some potential nuclear FGFR3 interactors as well. PI3K/AKT pathway was the most enriched one in KEGG pathway analysis, but interestingly, HIF1 signaling which has been validated in our transgenic mouse model, can also be enriched (Table 3-3; page 187).

Table 3-3: *Pathways enrichment analysis with FGFR3 partners.*

Pathways enrichment analysis in Enrichr web application (input n = 68 proteins, FGFR3 & its partners).

Term	Overlap	P-val (Adj)	Z-score	Genes
PI3K-Akt signaling	**10/341**	**1.89E-05**	-2.07	PPP2CA;PPP2CB;HSP90AA1;HSP90AB1;PPP2R1A;ITGB4;CDC37;FGFR3;EGFR;EPHA2
Oocyte meiosis	**6/123**	**1.45E-04**	-1.86	PPP2CA;CAMK2B;PPP2CB;CAMK2D;PPP2R1A;CAMK2G
Protein processing in endoplasmic reticulum	**6/169**	**3.25E-04**	-1.69	HSP90AA1;SEC23A;HSP90AB1;LMAN2;SEC24C;SEC23B
mRNA surveillance	**5/91**	**2.50E-04**	-1.62	PPP2CA;UPF1;PPP2CB;PPP2R1A;PABPC1
HIF-1 signaling	**5/103**	**3.25E-04**	-1.68	CAMK2B;CAMK2D;TFRC;CAMK2G;EGFR
ErbB signaling	**4/87**	**2.10E-03**	-1.66	CAMK2B;CAMK2D;CAMK2G;EGFR
Hepatitis C	**4/133**	**8.03E-03**	-1.58	PPP2CA;PPP2CB;PPP2R1A;EGFR
Tight junction	**4/139**	**8.53E-03**	-1.41	PPP2CA;PPP2CB;PPP2R1A;YBX3
Long-term depression	**3/60**	**8.03E-03**	-1.53	PPP2CA;PPP2CB;PPP2R1A

Wnt signaling	**3/142**	**3.08E-02**	-1.02	CAMK2B;CAMK2D;CAMK2G

***P* < 0.05**

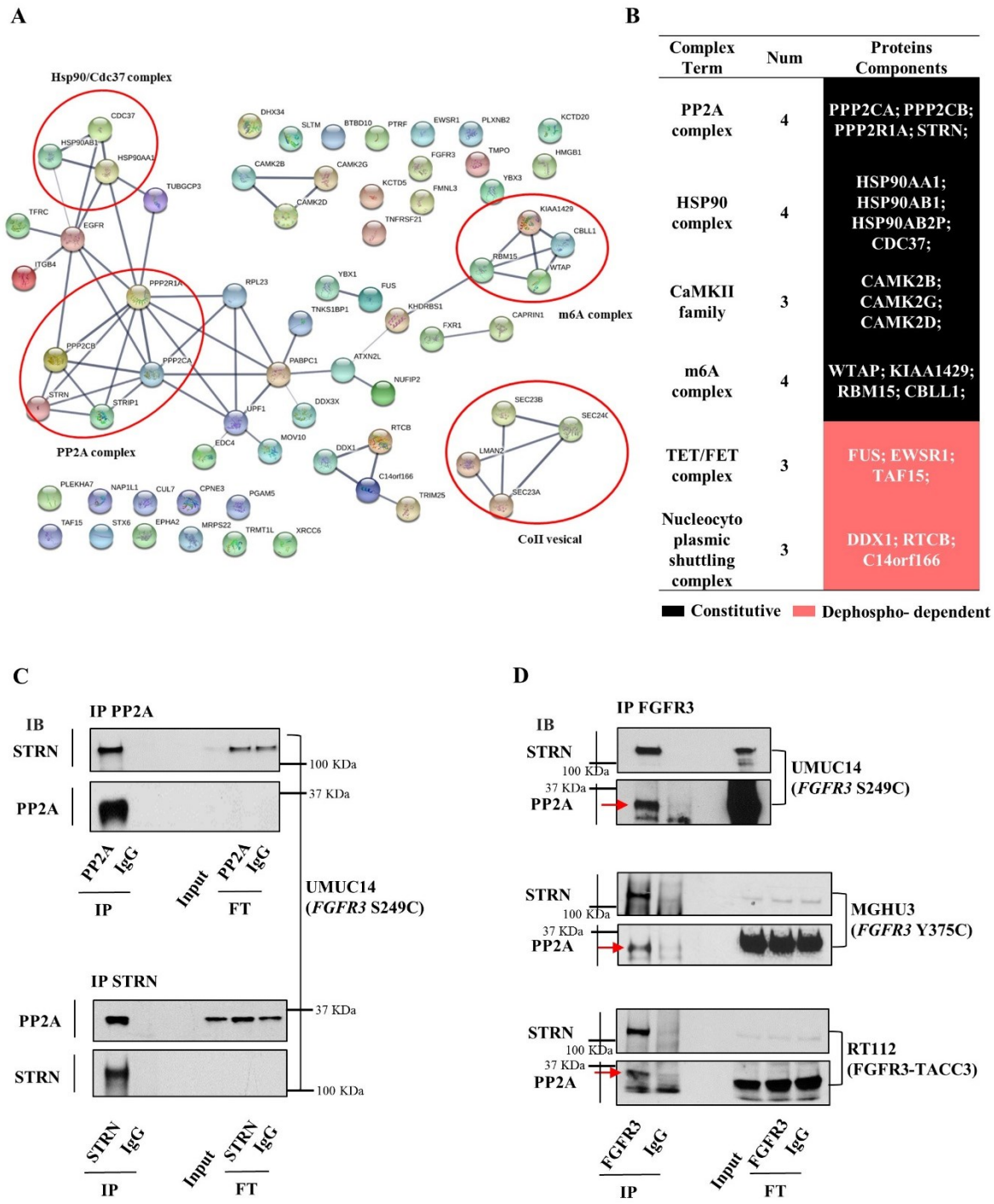


Figure 3-5: *Protein complexes & partners validation.*

A) *PPI networks of 68 FGFR3 partners generated from STRING database. Only experimental or database reported interactions were considered and confidence level was set as 0.4.* ***B)*** *Subunits and their binding types for several representative protein complexes.* ***C)*** *Validation*

of interaction between PP2A and STRN by IP in UMUC-14 cells. ***D)*** *Validation of interaction between FGFR3 and PP2A or STRN in three FGFR3-dependent cell models – UMUC14, MGH-U3 and RT-112 cells. IP-IgG is corresponding negative control.*

Except for the known Hsp90/Cdc37 complex, we found that protein phosphatase 2 (PP2A) complex and RNA N^6-methyladenosine modification (m6A) complex as well as some others were enriched in our list. The m6A is the most abundant reversible modification that is specially enriched in the 3′-UTRs near the stop codons of eukaryotic mRNAs and may affect gene expression and cell fate decisions by modulating multiple RNA-related cellular pathways (Jia et al., 2013; Meyer and Jaffrey, 2014). Methyltransferase-like 3 and 14 (METTL3 and METTL14) and their cofactors, WTAP, KIAA1429 and RBM15, compose the complex to catalyse m6A modification as the m6A 'writer'. Some m6A demethylases, such as FTO and ALKBH5, are known m6A 'erasers' that can remove m6A modification from RNA in a dynamic balance. Members of the YT521-B homology (YTH) domain family of proteins, like YTHDF12/3 and YTHDC1/2 are classical m6A 'readers' (Deng et al., 2018). Very recently, insulin-like growth factor 2 mRNA-binding proteins (IGF2BP1/2/3) were identified as new m6A 'readers' and regulated the stability of oncogenic diver – *MYC* gene – in an m6A-dependent manner (Huang et al., 2018). In our list, we found an enrichment of m6A 'writers' (WTAP, KIAA1429, RBM15 and CBLL1) that constitutively interacted with FGFR3 independently of FGFR3 inhibition. Another strong complex identified in our list is the PP2A complex. PP2A is a serine/threonine phosphatase that makes up 1% of all cellular proteins and along with protein phosphatase 1 (PP1) accounts for over 90% of all Ser/Thr phosphatase activity in the cell (Eichhorn et al., 2009). Several studies have highlighted the role of PP2A as a tumor suppressor by the fact that okadaic acid (OA), a selective but not specific inhibitor of PP2A, promotes tumor growth in mice and endogenous PP2A inhibitors (like SET and CIP2A) are up-regulated in various human cancer types (Fujiki and Suganuma, 1993; Fujiki et al., 2018; Suganuma et al., 1990). Structurally, PP2A is a trimeric holoenzyme that consists of three subunits: a catalytic 'C' (PP2A-C), a scaffolding 'A' (PP2A-A or PR65), and a regulatory 'B' subunit. More than fifteen regulatory B-subunits have been described. These B-subunits contain putative substrate-binding pockets and are crucial to decide substrate specificity regulated by PP2A (Kaur and Westermarck, 2016). Striatin (STRN) is a 780 amino acid protein with four protein–protein interaction domains including a caveolin-binding domain, a coiled-coil domain, a Ca^{2+}/calmodulin (CaM)-binding domain, and a tryptophan–aspartate (WD)-

repeat domain, and is proposed as one of the regulatory subunit of PP2A in 2000s (Moreno et al., 2000). The exact role of STRN on cancer phenotype is still far from clear. Previous studies discussed that STRN might be a marker of neuronal polarity due to strict expression in spine and be necessary scaffold platform for multi-protein complexes establishment due to its multi-domains (Hwang and Pallas, 2014). STRN may also be involved in endocytosis and vesical trafficking (Baillat et al., 2001) and control of cell migration (Gordon et al., 2011; Goudreault et al., 2008). **Of note, we can find all subunits of PP2A within our list: PPP2CA/PPP2CB as catalytic subunits, PPP2R1A as scaffold subunit and STRN as regulatory subunit.**

Experimental validation for partners & mechanism of regulation. There is no evidence yet to demonstrate the interaction between PP2A and STRN endogenously in a BCa context. I thus performed endogenous IP of either catalytic subunit of PP2A or STRN in UMUC-14 cells, and showed a clear interaction between each other (Figure 3-5C, page 188). By immunoprecipitating endogenous FGFR3, I confirmed that FGFR3 indeed interacted with both STRN and PP2A not only in UMUC-14 cells but also in two other FGFR3 dependent cell models (MGH-U3 and RT-112 cells) (Figure 3-5D, page 188, from top to down). Next, we asked two questions: 1) how does FGFR3 regulate PP2A activity considering PP2A/STRN are constitutive binders? 2) what are the impacts of PP2A on downstream signaling of FGFR3? There is no SH2/PTB domain in any of PP2A subunits, thus the validated interactions between FGFR3 and PP2A or STRN by IP-CoIP were possibly indirect. Alternatively, we supposed that PP2A complex could also be activated by FGFR3 through an intermediate FGFR3 adaptor protein, which can present SH2 domain that subsequently enables tyrosine phosphorylation of PP2A. Although such adaptor protein would be difficult to pinpointed, it has been intensively reported that the phosphorylation of PP2A *in vitro* occurred exclusively on Y307 and this phosphorylation can be catalysed by SRC, LCK, EGFR, and IGFRs (not reported yet for FGFRs) (Chen et al., 1992). Interestingly, phosphorylation at Y307 makes PP2A inactive. Therefore, I hypothesized that PP2A may be phosphorylated at Y307 by FGFR3 and being inactivated during the activating status of FGFR3; oppositely, upon inhibition of FGFR3, even though PP2A complex remained constitutive, they can change conformations and became active. Further, the activated PP2A can dephosphorylate multiple downstream effectors of FGFR3, like MAPK and PI3K/AKT signaling in which PP2A has been reported as negative regulators (Kiely and Kiely, 2015). We detected the Y307 phosphorylation level of PP2A with specific antibody by comparing FGFR3 activation to five-time-point kinetic inhibition (treated with 100 nM PD173074 for 5 min, 15 min, 30 min, 60 min and 120 min, respectively)

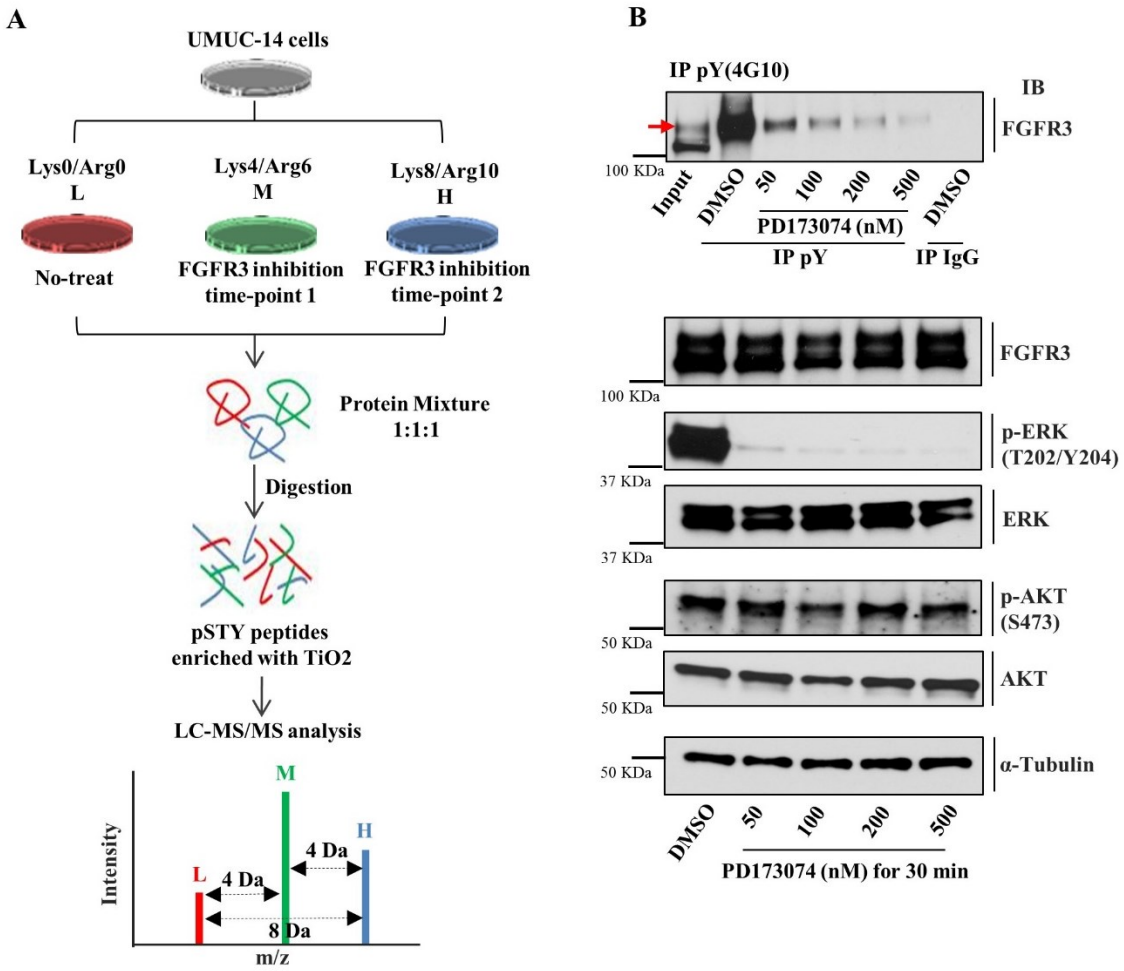
conditions in UMUC-14 cells. However, we did not see any change for phospho-PP2A (Y307) levels (data not shown). In line with literature, we found MAPK and PI3K/AKT signaling can be significantly abolished by FGFR3 inhibition in UMUC-14 cells. We asked whether the PP2A activation that released from FGFR3 inhibition was responsible for this dephosphorylation/inactivation of MAPK and PI3K/AKT signaling. For this purpose, we applied FGFR3 and PP2A inhibitors (100 nM PD173074 for FGFR3 and 5 nM OA for PP2A) as concurrent treatment for 5 min, 30 min and 24 h, and expected to observe a rescued activation of ERK compared to FGFR3 inhibitor along. Nevertheless, phospho-ERK was inhibited equally in both conditions, indicating there was no effect by adding PP2A inhibitors (data not shown). **Hence, for the importance of FGFR3/PP2A complex, except for their confirmed interaction, I cannot achieve any clear conclusion regarding the regulation mechanism behind.** Of note, several points need to be noted: 1) it is still unclear whether phospho-PP2A (Y307) levels reflect the enzymatic activity itself; 2) only one commercialized anti-phospho-PP2A (Y307) antibody is available, thus the antibody quality is undetermined; 3) it is a technical challenge and uncertainty in terms of correct order of combined drugs administration, which may affect largely the efficiency of inhibitors.

Other partners that were failed to be validated. I was also interested in two other partners within the list: 1) CaMKII, it is a constitutive partner and a well-studied serine/threonine protein kinase; 2) HMGB1, it is a detached binder and a well-known cancerdriver. CaMKII is activated by the Ca^{2+}/calmodulin (CaM) complex, consisting of catalytic, regulatory (autoinhibitory/CaM binding) and association domains. It not only can function as scaffold for PPIs, but can also act as a hub where multiple signaling pathways crosstalk. CaMKII has been well accepted as master upstream kinase that activates cAMP-response element-binding protein (CREB1) at S133 residue, which are critical feature for long-term potentiation (Kim et al., 2016; Ma et al., 2014). It is also shown that CaMKII can stabilize MYC protein by direct phosphorylation at S62 and promotes T Cell lymphoma (Gu et al., 2017). High mobility group box 1 (HMGB1) is a DNA-binding nuclear protein, released actively following cytokine stimulation as well as functioning directly as transcription factors. HMGB1 is implicated in many biological processes, such as inflammatory response, cell differentiation, cell migration and tumor metastasis (Sims et al., 2010; Tang et al., 2010). Similar to PP2A complex, CaMKII and HMGB1 lack of SH2/PTB domain, thus not supporting direct bindings with FGFR3. Because the molecular sizes of CaMKII and HMGB1 are around 50 kDa and 25 kDa, respectively, we cannot recognize specific binding after IP of FGFR3 as the heavy and light

(h/l) chains of antibodies were located in the same molecular size as well. I struggled with optimizing experimental conditions to avoid the interference of h/l chains during IP, such as eluting proteins in non-reducing buffer, applying specific secondary antibodies reducing h/l chains and utilizing primary/secondary antibodies of different species, but everything failed. I finally tried reversed IP to pull down CaMKII and HMGB1 proteins, nevertheless, I failed again as those antibodies were not suitable for IP. Much effort still need to be done to overcome this technical challenge. Lastly, I asked whether the activity of CaMKII and protein expression level of HMGB1 was regulated by FGFR3 even though they were not validated yet as FGFR3 binders due to technical limitations. Disappointedly, in my short time point kinetic experiment described before, I did not observe a significant modulation of phosphorylation of CaMKII (T286). In addition, after knockdown of FGFR3 for 40 h, the HMGB1 protein level was not regulated as well (data not shown). **In brief, the hypotheses that the FGFR3-CaMKII or FGFR3-HMGB1 regulation axes are important but could not be validated.**

MS-based FGFR3 phosphoproteome

Workflow for MS-based FGFR3 phosphoproteome. Triple SILAC strategy was applied for UMUC-14 cells. All cells were starved (0% FBS) for 16 h before treatment, 'Light (L)' cells were treated with DMSO, 'Medium (M)' and 'Heavy (H)' cells were treated with 50 nM of FGFR3 inhibitor (PD173074) for 5 min and 30 min, respectively. Following steps for sample preparation and MS data acquisition are described in the Methods section. A simplified workflow is presented in Figure 3-6A, page 194. PD173074 is a selective pan-FGFRs inhibitor with IC50 of ~25 nM and also inhibits KDR (VEGFR2) with IC50 of 100-200 nM, ~1000-fold more selective for FGFRs than for SRC and other RTKs, like PDGFR, EGFR and MEK (Mohammadi et al., 1998). In fact, in BCa, the gene expression for *FGFR1/2/4* and *KDR* are relatively low compare to *FGFR3*. Despite that, I tried to minimize the concentration of PD173074 to reduce interference from potential drug off-targets when exploring downstream phosphorylation events induced by FGFR3. In UMUC-14 cells, I found even with the lowest PD173074 concentration of 50 nM, that the phosphorylation level of FGFR3 itself as well as downstream p-ERK were clearly attenuated whereas only a slight decrease for another downstream p-AKT was observed (Figure 3-6B, page 194). Considering that the downstream phosphoproteome profile of FGFR3 can undergo rapid and dynamical changes upon inhibition time duration, I conducted multiple time-points kinetic assay in UMUC-14 cells with two different inhibitor concentrations (50 nM and 100 nM, respectively). Interestingly, I found that


A
UMUC-14 cells
Lys0/Arg0
L
Lys4/Arg6
M
Lys8/Arg10
H
No-treat
FGFR3 inhibition
time-point 1
FGFR3 inhibition
time-point 2
Protein Mixture
1:1:1
Digestion
pSTY peptides
enriched with TiO2
LC-MS/MS analysis
Intensity
M
H
L
4 Da
4 Da
8 Da
m/z
B
IP pY(4G10)
IB
FGFR3
100 KDa
Input
DMSO
50
100
200
500
DMSO
PD173074 (nM)
IP pY
IP IgG
FGFR3
100 KDa
p-ERK
(T202/Y204)
37 KDa
ERK
37 KDa
p-AKT
(S473)
50 KDa
AKT
50 KDa
α-Tubulin
50 KDa
DMSO
50
100
200
500
PD173074 (nM) for 30 min

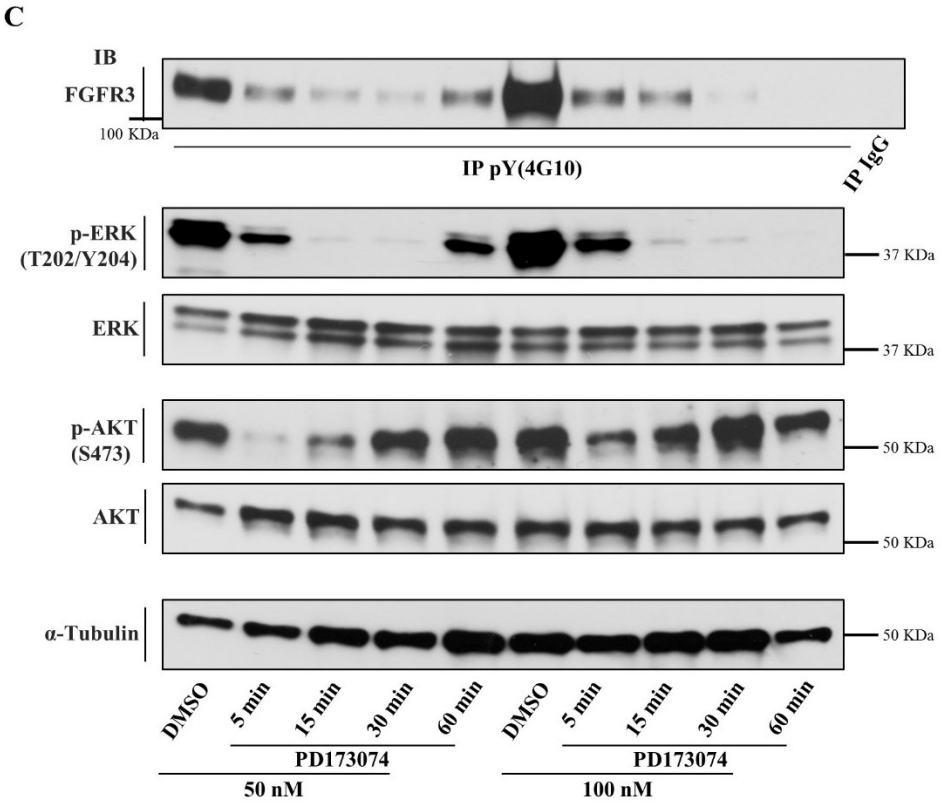

C
IB
FGFR3
100 KDa
IP pY(4G10)
IP IgG
p-ERK
(T202/Y204)
37 KDa
ERK
37 KDa
p-AKT
(S473)
50 KDa
AKT
50 KDa
α-Tubulin
50 KDa
DMSO
5 min
15 min
30 min
60 min
DMSO
5 min
15 min
30 min
60 min
PD173074
50 nM
PD173074
100 nM

Figure 3-6: *Workflow and optimized conditions for MS-based FGFR3 phosphoproteome.* ***A)*** *Graphical workflow for FGFR3 phosphoproteome. No-treat, DMSO; FGFR3 inhibition, using PD173074 inhibitor at 50 nM for 5 min and 30 min, respectively.* ***B)*** *Optimization for the concentration of PD173074 inhibitor. UMUC-14 cells were treated with different concentrations for 30 min.* ***C)*** *Optimization for the time points. UMUC-14 cells were treated with 50 nM or 100 nM PD173074 for different time courses.*

the activation of FGFR3 and ERK proteins was continually diminished from 5 min inhibition whereas the activation of AKT was only shortly inhibited at 5 min inhibition and rapidly recovered to basic level after 30 min inhibition (Figure 3-6C, page 194). The time courses results between 50 nM inhibition and 100 nM inhibition were consistent. Therefore, we selected two time points (5 and 30 min) that potentially represented the most significant protein phosphorylation changes, and fixed a minimum inhibitor concentration of 50 nM as optimal conditions to accomplish the workflow for MS-based FGFR3 phosphoproteome.

459 regulated phosphosites by FGFR3. According to data processing criteria described in the method section, a total of 2306 phosphosites (phospho-serine, threonine and tyrosine, pSTY) were identified and quantified from MS in two out of three replicates (with 1216 pSTY sites were common in triplicate). The PCA analysis with 1216 pSTY sites showed two distinct groups which corresponded to profiles reflecting the treatment of two time points (Figure 3-7A; page 195). The correlation analysis with 2306 pSTY sites indicated data reproducibility among replicates (Figure 3-7B; page 195). To highlight reliably regulated pSTY sites, I considered the ratio of either '|log2 (M/L)|' or '|log2 (H/L)|' over than '1' as cut-off threshold (L, no-treat; M, 5 min inhibition and H, 30 min inhibition) and eventually obtained 459 pSTY sites (corresponded to 310 proteins) responding to FGFR3 inhibition. These regulated pSTY sites were followed dynamic changes and mainly classified to five groups as demonstrated in Figure 3-7C; page 195. Except for the group 2 where we only listed top 10 regulated pSTY sites, all other regulated sites and their corresponding groups are indicated in Figure 3-7D; page 195.

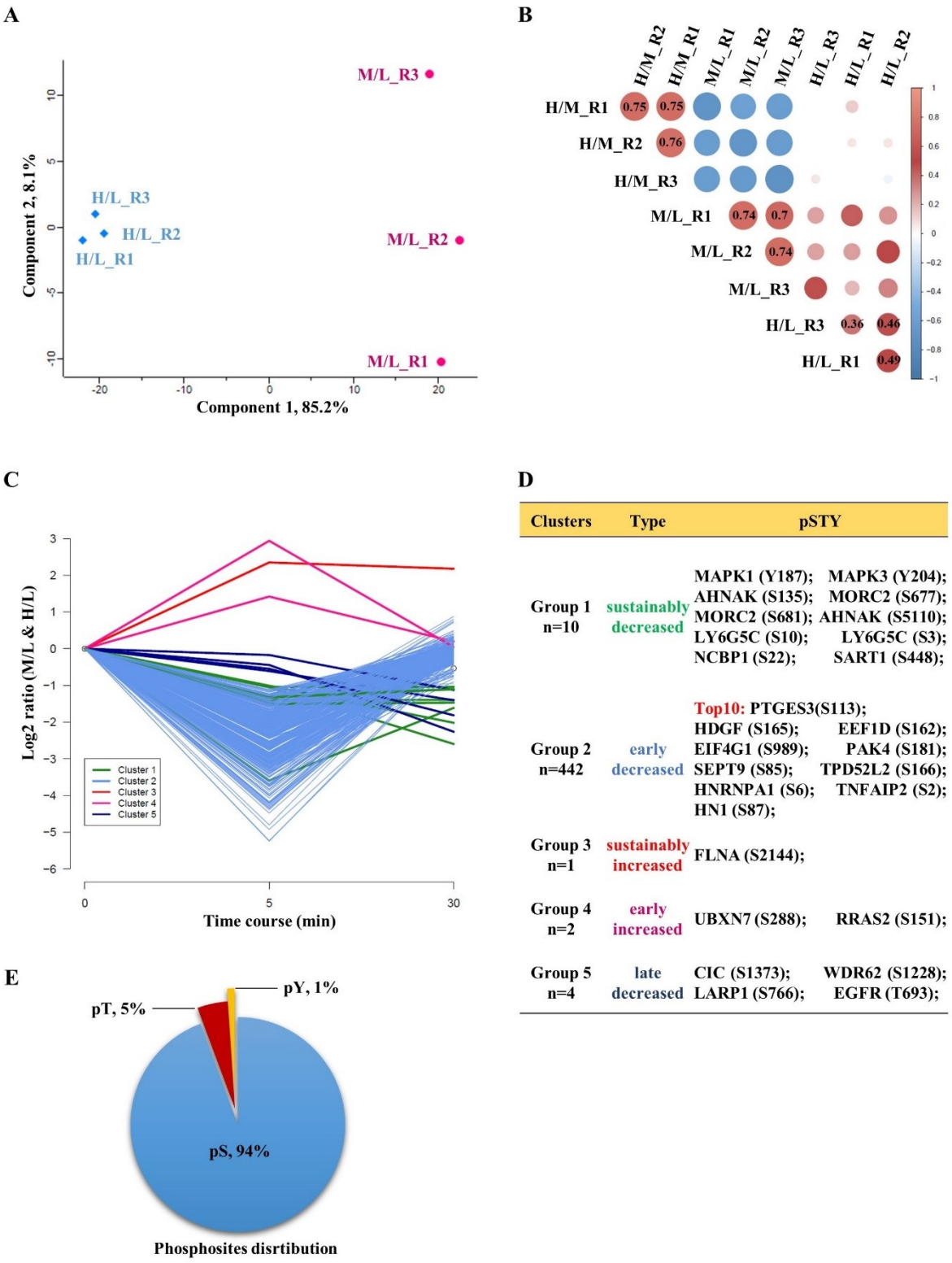


Figure 3-7: *Identification of 459 regulated phosphosites conferred by FGFR3 inhibiton.*

A) *Principal component analysis with 1216 commonly regulated pSTY sites in all three replicates. L, no-treat; M, 5 min inhibition and H, 30 min inhibition. R, replicate.* ***B)*** *Correlations among triplicate reflect experimental reproducibility (2306 identified and*

quantified pSTY sites in at least 2 of 3 replicates). ***C)*** *459 regulated pSTY sites showed dynamic changing profiles.* ***D)*** *List of regulated pSTY sites in each group. Only top 10 regulated pSTY sites was listed out for group 2.* ***E)*** *Distribution of pSTY. S, serine; T, thereonine; Y, tyrosine.*

Consistently, we found p-ERK was ranked in the top that sustainably decreased at both 5 min and 30 min inhibition of FGFR3 (group 1). However, p-AKT was not included in my list. Strikingly, majority of these 459 pSTY sites were strongly inhibited at 5 minutes but rapidly recovered after 30 minutes. The reason for that was still not understandable. Only 1.1% of these 459 regulated pSTY sites were pY sites (including *MAPK1* Y187, *MAPK3* Y204, *ATP1A1* Y511, *LARP1* Y633 and *SLC4A1AP* Y464) and the remaining were pS/T sites (Figure 3-7E; page 195). This bias may be a result of generally less abundance of pY than pS/T events in cells, but also due to the use of TiO2 that favour enrichment of pS/T containing peptides.

Pathways enrichment analysis. To evaluate the rationality of the 459 regulated pSTYs dataset that I generated above, I performed KEGG pathways enrichment analysis through online application of Enrichr. MAPK signaling was significantly enriched as expected. Interestingly, many of these significantly enriched pathways were also enriched previously with the data of 68 highly confident partners of FGFR3, including ERBB signaling, HIF1 signaling and cell adhesion, etc (Table 3-4; page 196). We also noted that a number of endocytosis-related proteins were highlighted, which may be important for the route of FGFR3 trafficking. Additionally, a clear crosstalk between mRNA transport and membrane trafficking has been reviewed (Jansen et al., 2014). In short, I supposed that the pinpointed 459 regulated pSTY sites were rational and reflected potential downstream signaling induced by FGFR3 in BCa.

Table 3-4: *Pathways enrichment analysis with FGFR3 regulated phosphosites.*
Pathways enrichment analysis in Enrichr web application (input n = 310 regulated phosphorylated proteins).

Term	Overlap	Adj P-val	Z-score	Genes
ErbB signaling	**11/87**	**2.01E-05**	**-1.90**	*PAK1;RPS6KB1;BAD;SRC;EIF4EBP1;MAPK1;PAK2;EGFR;PAK4;MAPK3;NCK1*
RNA transport	**13/172**	**2.93E-04**	**-1.84**	*RANBP2;NCBP1;CASC3;THOC5;NUP160;SRRM1;CLNS1A;EIF3J;EIF4EBP1;ACIN1;SMN1;EIF4G3;EIF4G1*
MAPK signaling	**15/255**	**5.62E-04**	**-1.92**	*JUND;MAX;RRAS2;HSPB1;DUSP6;EGFR;IL1A;PAK1;FLNA;MAPK1;TP53;PAK2;MAP3K7;FGFR3;MAPK3*
Spliceosome	**11/134**	**5.17E-04**	**-1.75**	*SART1;PRPF38B;RBM25;HNRNPK;NCBP1;SRSF2;DDX42;SRSF3;ACIN1;HNRNPA1;SF3B1*
Thyroid hormone signaling	**10/118**	**6.07E-04**	**-1.75**	*HDAC2;BAD;SRC;HDAC1;TSC2;MAPK1;ATP1A1;TP53;SLC9A1;MAPK3*

Acute myeloid leukemia	**7/57**	**8.98E-04**	**-1.84**	*JUP;RPS6KB1;BAD;EIF4EBP1;MAPK1;PML;MAPK3*
Bladder cancer	**6/41**	**1.06E-03**	**-1.61**	*SRC;MAPK1;TP53;FGFR3;EGFR;MAPK3*
Endocytosis	**13/259**	**4.24E-03**	**-1.74**	*USP8;SRC;SNX12;EPS15L1;PML;EGFR;SNX1;RABEP1; NEDD4;FAM21A;FGFR3;SNX5;GIT1*
Adherens junction	**7/74**	**3.66E-03**	**-1.64**	*TJP1;SRC;CTNND1;MAPK1;MAP3K7;EGFR;MAPK3*
HIF-1 signaling	**8/103**	**4.24E-03**	**-1.63**	*RPS6KB1;EIF4EBP1;PGK1;MAPK1;ALDOA;GAPDH;EGFR; MAPK3*
Proteoglycans in cancer	**11/203**	**5.62E-03**	**-1.73**	*PAK1;RPS6KB1;SRC;RRAS2;FLNA;MAPK1;ITPR3;TP53; EGFR;SLC9A1;MAPK3*
Insulin signaling	**9/139**	**5.62E-03**	**-1.53**	*RPS6KB1;BAD;FASN;PRKAR2A;EIF4EBP1;TSC2;MAPK1; ACACA;MAPK3*
Regulation of actin cytoskeleton	**11/214**	**7.22E-03**	**-1.62**	*PAK1;SRC;RRAS2;MAPK1;PAK2;FGFR3;GIT1;EGFR; SLC9A1;PAK4;MAPK3*
…	…	…	…	…
Focal adhesion	**10/202**	**1.13E-02**	**-1.39**	*PAK1;BAD;SRC;FLNA;MAPK1;ARHGAP5;PAK2;EGFR; PAK4;MAPK3*
Tight junction	**7/139**	**3.29E-02**	**-0.86**	*TJP1;EPB41L1;SRC;PRKCD;RRAS2;SYMPK;TJP2*
AMPK signaling	**7/124**	**2.19E-02**	**-1.08**	*RPS6KB1;FASN;EIF4EBP1;TSC2;PPP2R5D;MAP3K7; ACACA*
Estrogen signaling	**6/99**	**2.65E-02**	**-1.02**	*SRC;PRKCD;MAPK1;ITPR3;EGFR;MAPK3*
PI3K-Akt signaling	**10/341**	**1.23E-01**	**-0.41**	*RPS6KB1;BAD;EIF4EBP1;TSC2;MAPK1;PPP2R5D;TP53; FGFR3;EGFR;MAPK3*
cAMP signaling	**7/199**	**1.18E-01**	**-0.42**	*PAK1;BAD;RRAS2;MAPK1;ATP1A1;SLC9A1;MAPK3*

Common pathways enriched with 68 highly confident partners **$P < 0.05$**

Upstream phosphatases prediction. Phosphatases are crucial regulators to balance proper signaling dissemination by either activating or inactivating substrates through dephosphorylation. I asked what phosphatases can be involved in the regulation of the 459 pSTY sites (310 proteins). Enrichr application was used to predict upstream phosphatases. I found a dozen of phosphatases predicted, including serine/threonine phosphatases, tyrosine phosphatases as well as dual specificity phosphatases (Table 3-5; page 198). Some of these phosphatases have been reported to modulate FGFRs signaling, such as PTPN1 and PTPN11 (Neben et al., 2019; St-Germain et al., 2015). Even though all of the listed phosphatases were statistically significant, many of substrates were shared by multiple phosphatases apart from their limited numbers for each prediction. Thus, these results should be treated with caution. However, PPP2CA (catalytic subunit of PP2A complex), the top second prediction, might be a 'bona fide' candidate because it has been validated as FGFR3 binding partners previously. This result also reflected the coordination between our interactome and phosphoproteome data.

Table 3-5: *Upstream phosphatases prediction.*

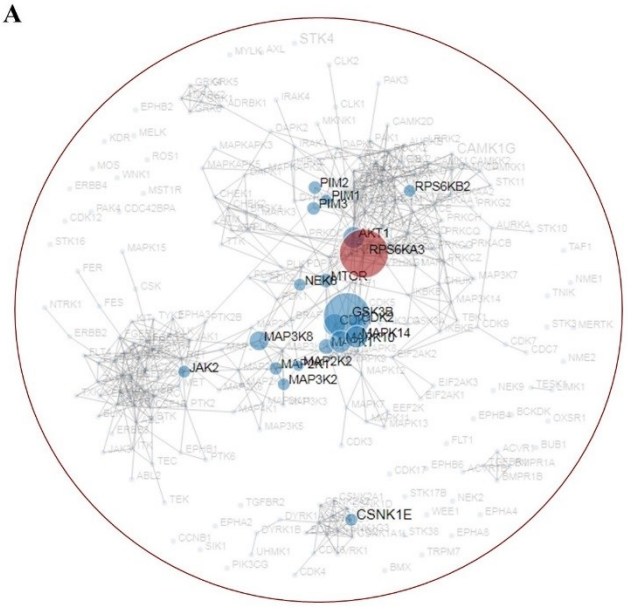
Predicted upstream phosphatases in Enrichr web application (input n = 310 regulated phosphorylated proteins).

Phosphatase	Overlap	Adj P-val	Genes
PTPRJ	**6/17**	**5.38E-06**	*TJP1;SRC;CTNND1;MAPK1;EGFR;MAPK3*
PPP2CA	**9/88**	**1.71E-04**	*PAK1;RPS6KB1;SRC;BAD;CAD;MAPK1;HSPB1;MAP3K7;TP53*
PPP3CA	**5/25**	**4.52E-04**	*BAD;PRKAR2A;CANX;HSPB1;FLNA*
PTPRE	**3/7**	**1.17E-03**	*SRC;MAPK1;MAPK3*
PTPN6	**4/19**	**1.23E-03**	*RBM39;SRC;CTNND1;EGFR*
CDC25A	**3/8**	**1.23E-03**	*MAPK1;EGFR;MAPK3*
PPP1CA	**5/46**	**2.70E-03**	*POLD3;BAD;CAD;TP53;LMNB1*
PTPN7	**2/3**	**2.70E-03**	*MAPK1;MAPK3*
DUPD1	**2/3**	**2.70E-03**	*MAPK1;MAPK3*
DUSP4	**2/3**	**2.70E-03**	*MAPK1;MAPK3*
DUSP3	**3/13**	**3.25E-03**	*MAPK1;EGFR;MAPK3*
DUSP9	**2/5**	**6.79E-03**	*MAPK1;MAPK3*
DUSP1	**2/5**	**6.79E-03**	*MAPK1;MAPK3*
DUSP22	**2/6**	**9.36E-03**	*RBM39;MAPK1*
PTPN11	**3/21**	**1.01E-02**	*SRC;ARHGAP5;EGFR*
PPM1D	**2/9**	**1.79E-02**	*TP53;UNG*
PPM1B	**2/9**	**1.79E-02**	*BAD;MAP3K7*
CDC14A	**2/12**	**3.01E-02**	*KIF23;TP53*
PTPN1	**3/38**	**4.17E-02**	*RBM39;SRC;EGFR*
ACP1	**2/15**	**4.18E-02**	*SRC;EGFR*

Highly confident interactor of FGFR3 **$P < 0.05$**

Upstream kinases prediction. Compared to phosphatases, the kinases are much more abundant and diverse. In order to identify upstream kinases which could be responsible for phosphorylating aforementioned pSTY sites, I applied three different available tools to predict those master regulators. First, the KEA tool developed by Ma'ayan's lab was used. Even though only 58 out of 459 regulated pSTY sites were recognized by KEA database, I found the prediction based on that was encouraging since the majority of the significantly highlighted kinases were already well-known to be involved in FGFR3 signaling, such as RSK2, p38, GSK-3β and AKT, etc (Figure 3-8A; page 199). The potential involvement of cyclin-dependent kinase 1 (CDK1) and JNK3 (*MPAK10*) kinases in FGFR3 signaling were new and have not been reported before. Second, NetworKIN tool is another commonly used tool to explore phosphoproteomic data, which combines the NetPhorest (a phylogenetic tree–based algorithm to classify phosphorylation sites in terms of kinases and phosphobinding domains) probability

A



Kinase	Adj. P-val	Total pSTY set	Total pSTY in input	Total pSTY intersected
RPS6KA3 (RSK2)	3.19E-06	493	58	17
GSK3B	5.35E-06	992	58	23
MAPK14 (p38)	8.95E-03	630	58	13
CDK1	8.98E-03	737	58	14
MAPK10 (JNK3)	8.98E-03	153	58	6
AKT1	9.71E-03	222	58	7

Well-known kinases involved in FGFR3 signaling

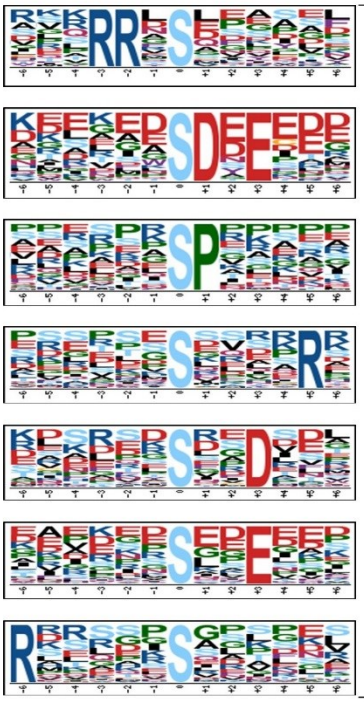
$P < 0.05$

B

Kinase	pSTY target count		#pooled NetworKIN score
	pSTYs	*Stringent pSTYs	
CSNK2A1 (CK2)	178	51	341.39
CDK1	149	81	266.57
MAPK3 (ERK1)	150	116	244.47
MAPK1 (ERK2)	150	65	216.07
BTRC	165	165	161.58
FBXW7	165	165	161.58
PIN1	142	142	151.09
CDK2	149	54	146.28
CSNK1A1	153	94	112.98
CSNK2A2	178	104	79.28
CLK1	248	35	60.90
MAPK8 (JNK1)	142	24	58.60
MAPK11 (p38)	144	22	51.22
TGFBR2	208	23	19.69

Well-known kinases involved in FGFR3 signaling

C



Motif-X score	Kinase*
30.21	PKA
22.56	CK2
16	CDK1 & MAPK1/3
10	Unknown
8.57	Unknown
6.56	CK2
6.27	PKA & PKC

* literature matched

Figure 3-8: *Upstream kinases prediction.*

A) *KEA tool predicted kinases containing an input of 58 regulated pSTY sites.* ***B)*** *NetworKIN tool predicted kinases with 459 regulated pSTY sites. *Stringent filtering criteria: 1) pSTY count for a given kinase was over than median count of all kinases within the matrix; 2) NetworKIN score of a given kinase-substrate pair was over than average score of all pairs within the matrix. #, the score for each kinase was the pooled from all of its substrates that*

filtered by stringent criteria. ***C)*** *Motif-X predicted motifs with 274 mono-pS sites and kinases were matched according to literature records. Motif-X score was computed by inner algorithm with defaulting parameters.*

and the STRING-derived proximity (PPI database) score using the naive Bayes method (Horn et al., 2014). With the input of 459 pSTY sites, an initial list of 269 kinases were predicted by NetworKIN. I removed potential fake kinases by considering only kinases with wider coverage of targets (pSTY counts were over than median count of all kinases within the matrix) and with higher NetworKIN score (NetworKIN score of each kinase-substrate pair was over than average score of all pairs within the matrix). Finally, there were 14 kinases were pinpointed as confident candidates and were ranked by pooled NetworKIN score (Figure 3-8B; page 199). Again, some known kinases relevant to FGFR3 signaling were included, such as ERK1/2 and p38. CDK1 kinase, which was also predicted by KEA, was ranked in the top, whereas most of other predicted kinases were distinct from KEA prediction. Third, the motif-X tool was used to discover enriched motif based on input pSTY sites and then potential kinases that can bind those consensus motifs were matched through literature review. The limitation of motif-X is that it recognizes only mono-phosphosites. Hence, with an input of 274 mono-phosphosites of serine (mono-pS), seven phosphorylation-specific motifs were enriched and their corresponding enrich scores were showed (Figure 3-8C; page 199). R-R-x-S/T motifs are identical for protein kinase A (PKA) substrate (Smith et al., 2011). Several motifs are known to be phosphorylated by casein kinase 2 (CK2), such as S-D-x-E and S-x-x-E (Villen et al., 2007). Both CDK1 and ERK1/2 (*MAPK3/1*) are proline-directed kinases and are known to phosphorylate substrates on the consensus motif S/T-P (Shah et al., 2003; Suzuki et al., 2015; Villen et al., 2007). One of the remarkable substrate enriched with such motif was S6K (*RPS6KB1*), which showed up to six S/T-P sites (i.e. Ser411, Ser418, Thr421, Ser424, Ser429, and Thr447) and was known to be phosphorylated by CDK1 and ERK1/2 (Shah et al., 2003). **Considering all these evidence above, except for those classically FGFR3-related kinases, I proposed three kinases – CDK1, CK2 and PKA – were interesting to follow as new candidates involved in FGFR3 signaling.**

Candidate kinases exploration & validation. As mentioned above, three kinases – CK2, PKA and CDK1 – might be involved in FGFR3 signaling, thus I tried to validate this hypothesis *in vitro*. Phosphatase and tensin homolog (PTEN) is a confirmed negative regulator of AKT activation in many context. However, PTEN overexpression was observed in human T cell

acute lymphoblastic leukemia (T-ALL) where PI3K/AKT pathway had been constitutively activated. In fact, this PTEN overexpression was associated with decreased PTEN phosphatase activity, resulting from CK2 overexpression and hyperactivation (Silva et al., 2008). Similarly, in BCa, our team has also noted that there was a positive and moderate correlation between p-AKT and PTEN protein expressions in both Ta and T1 tumors (Calderaro et al., 2014). Frequent mutations and overactivation of FGFR3 were common in Ta/T1 tumors. Therefore, the highlighted CK2 kinase from FGFR3 phosphoproteome data may be rational. Unlike other kinase, a certain posttranslational modification that represents as a surrogate marker for activation or inhibition of CK2 has not yet determined. Alternatively, p-AKT at S129 and p-PTEN at S380 have been identified as CK2 substrates and could partially reflect CK2 activation (Melão et al., 2016; Silva et al., 2008). Phosphorylation of PTEN at S380 results in PTEN inactivation. First, I treated three FGFR3-dependent cancer cell lines with CK2 inhibitor (CX-4945) for 40 h and validated that p-PTEN (S380) was indeed downregulated in all models (Figure 3-9A; page 202). Then, I checked whether CK2 inhibitor has an impact on cell proliferation. Unexpectedly, CX-4549 did not show any effect on cell viability in all of three cell models tested (Figure 3-9B; page 202). Lastly, I assessed the activation of CK2 after FGFR3 inhibition for 40 h. There were no protein changes for both p-PTEN (S380) and overall CK2 level (Figure 3-9C; page 202). **As a result, our results did not support the hypothesis that CK2 was regulated by FGFR3 as a downstream kinase.**

PKA exists as a tetramer of two catalytic and two regulatory subunits. Binding of cAMP by the regulatory subunits results in the dissociation of the tetramer complex and release of the active catalytic subunit (Ould Amer and Hebert-Chatelain, 2018). One of activating phosphorylation site of PKA is T197. It has been known that EGFR activated MAPK signaling was inhibited by PKA activation via phosphorylation of C-Raf (RAF1) (Cook and Mccormick, 1993) whereas FGFR1 activated MAPK signaling was independent of PKA activation (Cross et al., 2002). Experiments from growth plate chondrocytes revealed that *FGFR3* gene expression was repressed by PKA activation though binding a transcriptional regulatory element – CSRh – located in the promoter region of *FGFR3* (McEwen et al., 1999; Ornitz and Marie, 2015). I performed a preliminary test to see whether PKA kinase itself was regulated by FGFR3 through short time point inhibition in UMUC-14 cells, but I did not observe any changes for PKA activation (Figure 3-9D; page 202). We previously found that the gene expression of *FGFR3* itself was downregulated after long-term FGFR3 inhibition which was interpreted by FGFR3/MYC positive feedback loop (Mahe et al., 2018). Considering a

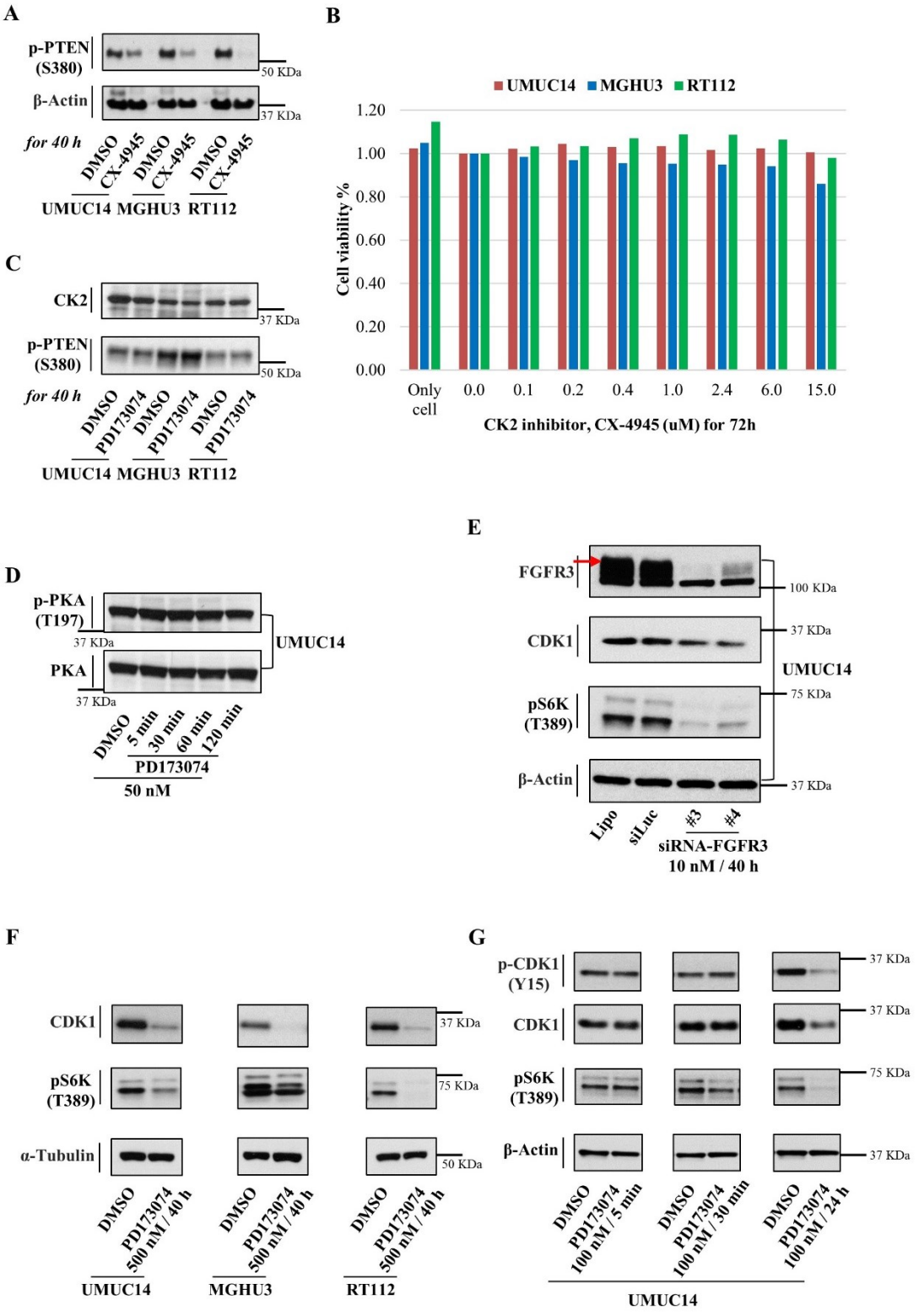


Figure 3-9: *Candidate kinases exploration & validation.*

A) *Effect of CK2 inhibitor (CX4945) on p-PTEN (S380) in three FGFR3-dependent cell lines. CX4945 was used at 20 uM for 40 h.* ***B)*** *Cell viability assay of three FGFR3-dependent models treated with different concentrations of CK2 inhibitor (CX4945) for 72h.* ***C)*** *WB of CK2 and p-PTEN (S380) in three FGFR3-dependent cell lines treated with a FGFR3 inhibitor*

(PD173074, 500 nM) for 40 h. ***D)*** *WB of PKA and p-PKA (T197) in UMUC-14 cells treated with short time point kinetic inhibition of FGFR3 (PD173074, 50 nM). Cells were starved (0% FBS) for 16 h prior to treatment.* ***E)*** *Regulation of CDK1 and p-S6K (T389) after knockdown of FGFR3 (40 h) by two siRNAs in UMUC-14 cells. F) Regulation of CDK1 and p-S6K (T389) upon long-term inhibition of FGFR3 (PD173074, 500 nM /40 h) in three FGFR3-dependent cell lines.* ***G)*** *Regulation of CDK1 and p-S6K (T389) upon short time course inhibition of FGFR3 (PD173074, 100 nM for 5 min – 30 min – 24 h) in UMUC-14 cells. Cells were starved (0% FBS) 16 h prior to treatment for 5 and 30 min.*

potential involvement of PKA regulation in FGFRs/RTKs signaling, I am wondering whether PKA activation would also participate into *FGFR3* gene downregulation conferred by long-term FGFR3 inhibition. Experiments regarding this part are still ongoing.

CDKs, a family of serine/threonine, can control the cell cycle progression and transcription. CDK4/cyclin D, CDK6/cyclin D and CDK2/cyclin E facilitate the G1-S phase transition by sequentially p-RB, while CDK1/cyclin A, CDK2/cyclin A and CDK1/cyclin B are essential for S-phase progression and G2-M transition, respectively (Varma et al., 2013). Dysregulated activity of CDKs results in loss of cell-cycle checkpoint function and has been linked to carcinogenesis (Geleta et al., 2016). Very recently, *in vitro* and *in vivo* evidence in BCa showing encouraging preclinical efficacy of CDK4/6 dual inhibitors (Palbociclib), which was even more dramatic when combing with cisplatin. The main mechanism was partially independent on *RB1* gene status but rely on FOXM1 phosphorylation and activation (Rubio et al., 2019). Of note, the regulation between FGFRs and CDKs is not unexpected, in lung cancer, it has been shown that FGF2-FGFR1-MAPK-mTOR-p27^{Kip1} but not FGFR1-AKT pathway conferred the emerging resistance of palbociclib by reactivating CDK6/cyclin D (Haines et al., 2018). In line with that, through the analysis of FGFR3 phosphoproteome data, CDK1 was the strongest candidate kinase that was predicted by all of three software used. CDK1/cyclin B activity depends on phosphorylation. CDK1 is inactivated through phosphorylation on T14 and Y15 by WEE1 and MYT1 kinases in interphase, and activated by phosphorylating T161 in M phase (Petrone et al., 2016). CDC25C is a crucial phosphatase to remove p-Y15 site of CDK1 and enable subsequential activation at the G2/M transition, and degradation of cyclin B is important to enable CDK1 to be switched off during mitosis exit. In our MS-based list, a well-known CDK1 substrate – S6K (*RPS6KB1*) – was also identified as one of 310 FGFR3-regulated phosphoproteins. Phosphorylation on T389 is crucial to enable full activation of S6K.

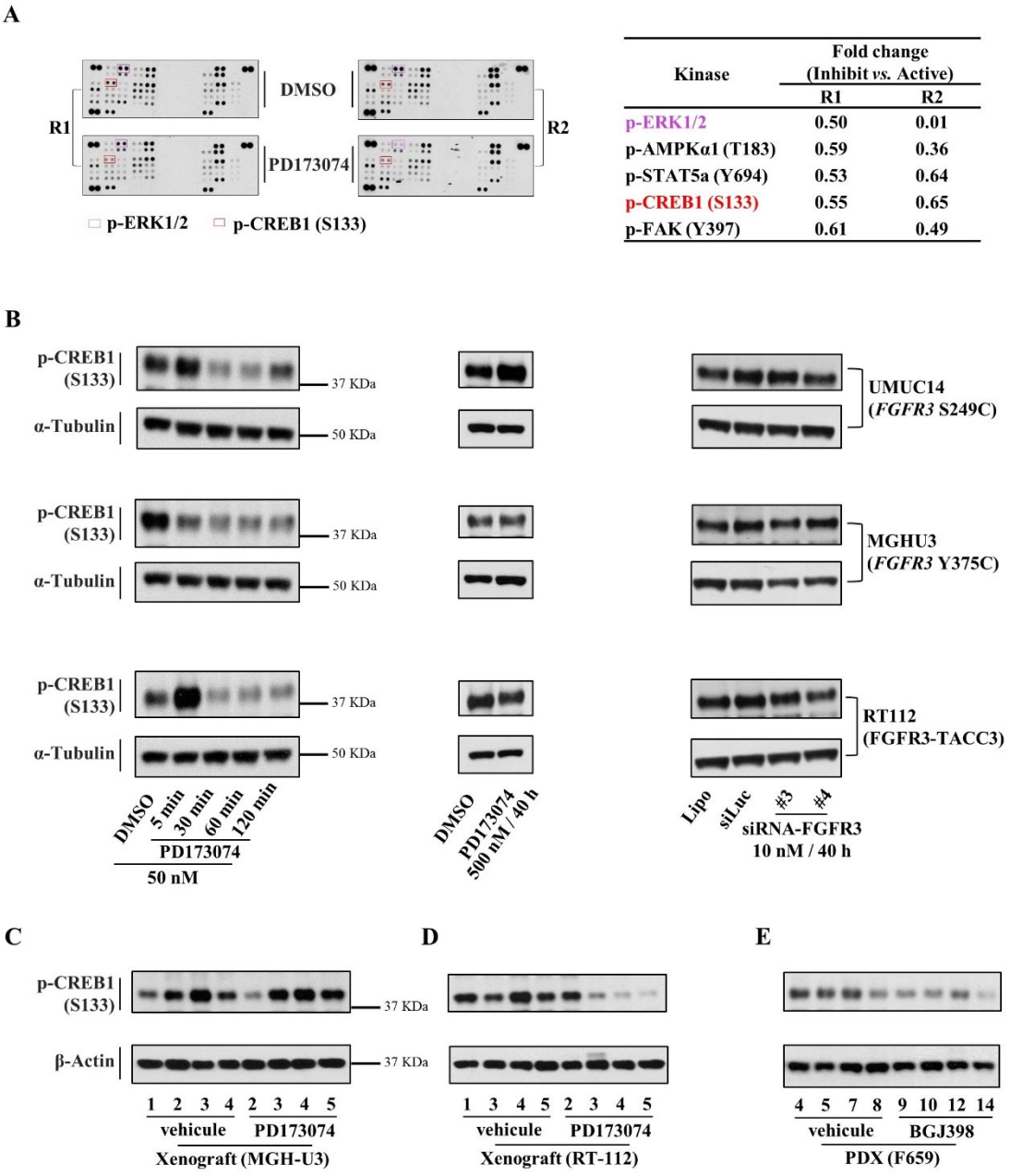
p-S6K (T389) is not a S/T-P sites, thus cannot be targeted by CDK1 kinase directly. However, it has been shown that CDK1 negatively and indirectly regulated T389 to make S6K inactive in mitosis in spite of its positive phosphorylation on multiple S/T-P sites of S6K at meantime (Shah et al., 2003). To see whether FGFR3 indeed regulated CDK1 as well as its substrate, I performed experiments by either knockdown or inhibition of FGFR3. I found that overall CDK1 protein was slightly decreased by FGFR3 knockdown while p-S6K (T389) was strongly attenuated (Figure 3-9E; page 202). However, by using FGFR3 inhibitor for 40 h, I found a strong inhibition for both overall CDK1 and p-S6K (T389) after treatment in all of three FGFR3-dependent models (Figure 3-9F; page 202). In fact, FGFR3 inhibitor, PD173074 – the one that I used throughout my experiments, has been reported to induce upregulation of cyclin-dependent kinase inhibitor 1B ($p27^{Kip1}$, *CDKN1B*) and G1/G0 arrest of the cell cycle in FGFR3 mutated BCa cell lines (UMUC-14 and MGH-U3 cells) (Miyake et al., 2009). $p27^{Kip1}$ has been known as endogenous CDKs inhibitors (Bencivenga et al., 2014; Le Sage et al., 2007). Further, in UMUC-14 cells, I showed both overall CDK1 and p-CDK1 (Y15) downregulation as early as 24 h after FGFR3 inhibition whereas the downregulation of p-S6K (T389) was observed more earlier, 30 minutes after the start of the treatment (Figure 3-9G; page 202).

Here, I failed to observe a negative regulation of p-S6K (T389) by CDK1 as described in the literature. I guess this may be due to the multifaceted regulation of S6K conferred by multiple kinases, such as PDK1 (the best known kinase directly phosphorylating T389), mTOR, ERK1/2 and JNK, etc (Shah et al., 2003), and some of which were also involved in FGFR3 signaling. **Briefly, I validated the regulation axis FGFR3-CDK1/S6K, but the regulation between CDK1 and S6K need further investigation. In addition, the biological effect of this axis need to be explored as well in future steps.**

FGFR3 regulated phospho-kinases based on antibody array

Human Phospho-Kinase Array & Candidate validation. Nearly half of the candidates included in the Human Phospho-Kinase Array (R&D Systems, cat.no. ARY003B) are tyrosine phosphorylated kinases, hence I thought this array would provide me complementary candidates in addition to the already identified pSTY sites from MS. Meanwhile, some of the regulated phosphoptoteins based on MS were also included in this array, such as ERK1/2 and S6K, thus the results from array could be considered as an additional validation. UMUC-14 cells were treated with either DMSO or a FGFR3 inhibitor (PD173074, 50 nM for 30 min) and the experiment was repeated twice. Representative film exposures for each replicate are shown

(Figure 3-10A, left; page 205). I compared treated versus control condition for each protein by quantifying the dots' intensity and considered proteins with a fold change less than 0.7 in duplicate as significant candidates (listed in Figure 3-10A, right; page 205). As expected, p-ERK1/2 were strongly downregulated. I focused on another downregulated candidate – CREB1,



Kinase	Fold change (Inhibit *vs.* Active)	
	R1	R2
p-ERK1/2	0.50	0.01
p-AMPKα1 (T183)	0.59	0.36
p-STAT5a (Y694)	0.53	0.64
p-CREB1 (S133)	0.55	0.65
p-FAK (Y397)	0.61	0.49

Figure 3-10: *Phospho-kinase array & validation for CREB1.*

A) *Identify phospho-kinases regulated by FGFR3 in UMUC-14 cells with human phospho-kinase array (R&D Systems, cat.no. ARY003B). Duplicate experiments. Candidates with significant fold change were listed (fold change < 0.7).* ***B)*** *In vitro evidence (three FGFR3-depedent cell lines) showed CREB1 was regulated by FGFR3 in short time but without long-*

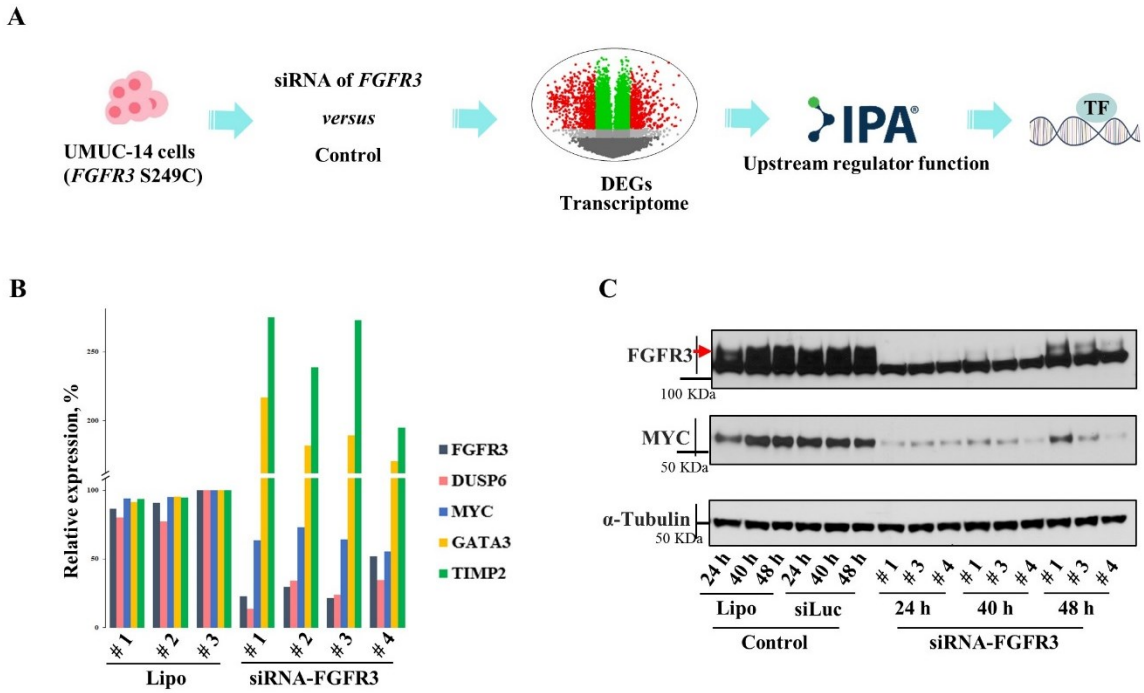
term effect. UMUC-14 cells were starved 16 h before treatment for short time kinetics. siRNA transfection as described in method section. ***C)*** *No regulation of CREB1 by FGFR3 in MGH-U3 xenograft model (9 days treatment). Four control mice versus four mice treated FGFR3 inhibtion (PD173074).* ***D)*** *CREB1 was regulated by FGFR3 in RT-112 xenograft model (9 days treatment).* ***E)*** *CREB1 was regulated by FGFR3 in PDX model (4 days treatment). BGJ398, FGFR3 inhibitor. PDX (F659) model was grafted from a patient with bladder cancer containing FGFR3 S249C mutation.*

because it has been reported that nuclear FGFR1 can directly bind to and activate CREB1 (Dunham-Ems et al., 2009). In addition, one of the most well-known upstream kinase of CREB1 is CaMKII protein, which has been identified as highly confident partners of FGFR3 previously. Pooling all these evidence, I hypothesized that CREB1 might be a downstream effector of FGFR3 signaling. Phosphorylation of CREB1 at serine 133 is crucial for its activation. To test my hypothesis, first, I checked activation status of CREB1 in different conditions, like after short term (up to 2 h) or long-term (40 h) inhibition and knockdown (40 h) of FGFR3 in three *in vitro* FGFR3-dependent models. I found p-CREB1 (S133) was indeed downregulated after short time inhibition of FGFR3 in all models but not modulated in neither long-term inhibition nor knockdown of FGFR3 (Figure 3-10B, from left to right; page 205). Of note, I previously found that CaMKII activation (p-T286) was not modulated by FGFR3 after short-term inhibition, hence the regulation of CREB1 by FGFR3 may be due to other mechanism instead of being mediated by CaMKII kinase. Second, I also checked the activation status of CREB1 responding to FGFR3 inhibition in different *in vivo* models (two cell lines xenograft models and one PDX model, all of which were FGFR3-dependent (Mahe et al., 2018)). A clear diminution of p-CREB1 (S133) level upon FGFR3 inhibition was observed in RT-112 cells xenograft and PDX model but not found in MGH-U3 cells xenograft (Figure 3-10C-E; page 205). It was still difficult to understand the inconsistency among these *in vivo* models. Considering our xenograft models were treated by FGFR3 inhibitor for 9 days and PDX was treated for 4 days, the results from *in vivo* models was also conflict to what I found *in vitro* after long-term inhibition or knockdown of FGFR3. Therefore, I found CREB1 was indeed a bona fide candidate weakly regulated by FGFR3, and I suspected that another compensatory mechanism could be responsible for the rapid reactivation of CREB1 after FGFR3 inhibition in BCa.

DNA array-based FGFR3 transcriptome

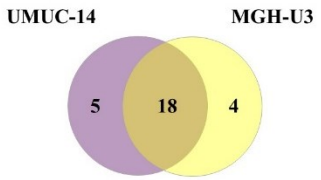
Workflow & condition optimization & results evaluation. I compared the transcriptome profiles between knockdowns of FGFR3 *vs.* non-treated conditions in UMUC-14 cells. With the differentially expressed genes ($|\log2(\text{fold change})| \geq 0.58$, adjusted P-value < 0.05), I predicted upstream transcription factors by IPA software for further investigation. A simplified workflow is shown in Figure 3-11A, page 208. I optimized experimental conditions for sample preparation. Marker genes, such as *MYC*, *DUSP6*, *GATA3* and *TIMP2*, that were known to be downregulated or upregulated by *FGFR3* knockdown from previous laboratory's work, were detected for verifying the efficiency of FGFR3 interference in UMUC-14. All four siRNAs of FGFR3 showed similar significant efficiency (Figure 3-11B, page 208). Next, I performed a time course knockdown of FGFR3 in UMUC-14 cells to figure out the optimal time point. I found that both MYC and FGFR3 were strongly reduced from 24 h and there was a tendency of protein expression recovery in the longest time point – 48 h (Figure 3-11C, page 208). Hence, I considered the intermediate time point – 40 h as the optimal treatment duration. Cells treated with three siRNAs of FGFR3 *versus* three lipofectamine controls were prepared for DNA array. After data processing as presented in workflow, a dozens of transcription factors (TF) were predicted by IPA tool. To evaluate the data rationality, I compared with the data that were similarly generated from MGH-U3 cells (harbouring *FGFR3* Y375C mutation) by previous lab work. I found that the majority of the predicted upstream regulators were shared between UMUC-14 and MGH-U3 cells as well as the consistency for predicted activation status of these TFs (Figure 3-11D, page 208).

HIF1A was predicted as downstream TF and regulated by FGFR3 in vitro. Our transgenic mouse model based on UII-*FGFR3* S249C have already highlighted the FGFR3/HIF1A regulation axis. Interestingly, from aforementioned proteomic data, I noted that HIF1A signaling was enriched using KEGG pathways analysis in both FGFR3 interactome and regulated phosphoproteome datasets. Therefore, when exploring the transcriptome data from *in vitro* FGFR3-dependent models, I was wondering whether HIF1A can be pinpointed as a downstream TF. As expected, HIF1A was predicted significantly as a downstream TF in all of



D

Upstream Regulator



Upstream Regulator	UMUC-14			MGH-U3		
	Predicted state	Activation z-score	Adj *P*-val	Predicted state	Activation z-score	Adj *P*-val
MYC	Inhibited	-5.194	6.37E-15	Inhibited	-4.570	9.43E-11
TRIM24	Inhibited	-4.310	5.10E-10	Inhibited	-2.795	4.23E-03
TBX2	Inhibited	-3.898	9.18E-06	Inhibited	-3.900	2.67E-12
ESR1	Inhibited	-3.814	3.65E-33	Inhibited	-3.511	4.94E-17
TAL1	Inhibited	-3.685	2.48E-03	Inhibited	-2.846	1.50E-08
CCND1	Inhibited	-3.555	6.58E-12	Inhibited	-3.929	1.00E-16
RARA	Inhibited	-3.198	4.25E-12	Inhibited	-3.305	3.73E-09
FOXM1	Inhibited	-3.149	2.07E-03	Inhibited	-3.174	8.78E-06
HSF2	Inhibited	-2.813	7.80E-04	Inhibited	-2.646	1.91E-04
E2F3	Inhibited	-2.213	1.19E-04	Inhibited	-4.096	6.10E-13
HIF1A	Inhibited	-1.000	2.09E-05	Inhibited	-0.750	5.29E-05
RBL1	Activated	2.155	1.92E-02	Activated	2.626	2.02E-10
NFATC2	Activated	3.156	8.04E-05	Activated	2.423	9.33E-03
MXI1	Activated	2.607	2.59E-05	Activated	2.570	2.19E-08
GATA1	Activated	3.448	2.43E-02	Activated	3.477	5.09E-07
FOXO3	Activated	3.694	9.99E-16	Activated	2.900	7.92E-13
TP53	Activated	5.462	3.44E-23	Activated	3.362	1.20E-28
STAT1	Activated	5.525	7.05E-06	Activated	2.817	3.22E-02
E2F1	NA	-1.379	3.30E-08	Inhibited	-3.531	6.50E-17
NUPR1	NA	-0.108	6.54E-11	Activated	3.474	1.94E-16
HDAC1	NA	1.316	1.26E-02	Activated	2.534	3.16E-07
VHL	NA	1.246	1.97E-03	Activated	2.607	9.31E-03
NKX2-3	Inhibited	-5.011	4.90E-16	NA	-1.414	2.80E-04
IRF3	Activated	3.550	4.15E-06	NA	-	-
IRF1	Activated	3.949	8.46E-07	NA	1.309	3.72E-02
TCF7L2	Activated	4.122	4.93E-03	NA	0.791	2.14E-02
IRF7	Activated	5.284	6.30E-12	NA	1.987	4.27E-02

Figure 3-11: *Workflow for FGFR3 transcriptome data collection.*

A) *Simplified data processing flow to acquire transcription factors regulated by FGFR3 in UMUC-14 cells.* ***B)*** *Detection of FGFR3 knockdown efficiency by qPCR for several known FGFR3 regulated markers. siRNA #1 and #2 were used at 20 nM and #3 and #4 were used at 10 nM. UMUC-14 cells were treated for 40 h.* ***C)*** *Optimization for time duration of siRNA treatment. Same concentrations were used as above. MYC protein was loaded as positive*

control because it has been shown as downstream effector of FGFR3. ***D)*** *Comparison of upstream regulators identified from UMUC-14 cells with those previously defined from MGH-U3 cells (Mahe et al., 2018). This data table was cited from Shi et al (Shi et al., 2019).*

three cell models (data for UMUC-14 was generated in current project while the ones for MGH-U3 and RT-112 cells were public data from the team) even though it was not ranked in the top. HIF1A was moderately inhibited after knockdowns of FGFR3 (IPA predicted activation Z score were -1, -0.75 and -1.56 for UMUC-14, MGH-U3 and RT-112 cells respectively, and all showed significant P-values).

Human tissues do not make use of oxygen (O2) equally, but O2 gradients are formed and show organ / cell type-specificity. Thus, we need to consider hypoxia as a relative term rather than a fixed O2 concentration (Wenger et al., 2015). Cell culture under 20.9% O2 in incubator are usually referred to as 'normoxic', in physiological terms, they are rather 'hyperoxic' because not even lung alveolar cells (containing highest O2 concentration) are ever exposed to 20.9% O2. I found in many BCa cell lines cultured in 'normoxic / hyperoxic' condition, regardless of FGFR3 mutation status, an enhanced expression of HIF1A protein was observed Figure 3-12A, page 210. The explanation for these results is challenging because HIF1A is usually induced in hypoxic conditions. However, it has also been reported that HIF1A protein expression can be induced in 'normoxic / hyperoxic' condition by specific mechanism, such as FGF2 stimuli in lung cancer or MYC gene overexpression in breast cancer (Doe et al., 2012; Fumarola et al., 2017). I thereby investigated the regulation of HIF1A by FGFR3 *in vitro* in normal cultured conditions. I found that HIF1A protein was strongly attenuated by knocking down of FGFR3 in the three FGFR3-dependent cell models (Figure 3-12B, from top to down, page 210). Consistent results were also observed by FGFR3 inhibition in these cell lines (Figure 3-12C, page 210, DUSP6 was loaded as positive control as it has been described as a marker for FGFR3 inhibition (Nakanishi et al., 2015b)). Thus, HIF1A was validated to be regulated by FGFR3. My next plan is to explore this regulation axis by culturing cells under hypoxic conditions. Interestingly, there is evidence showing that HIF1A activation itself can induce FGFR3 expression by supressing miR-100 expression (a known negative regulator of FGFR3) in RT-112 cells (Blick et al., 2013), implying the existence of positive feedback loop between FGFR3 and HIF1A. This will also be future direction to follow.

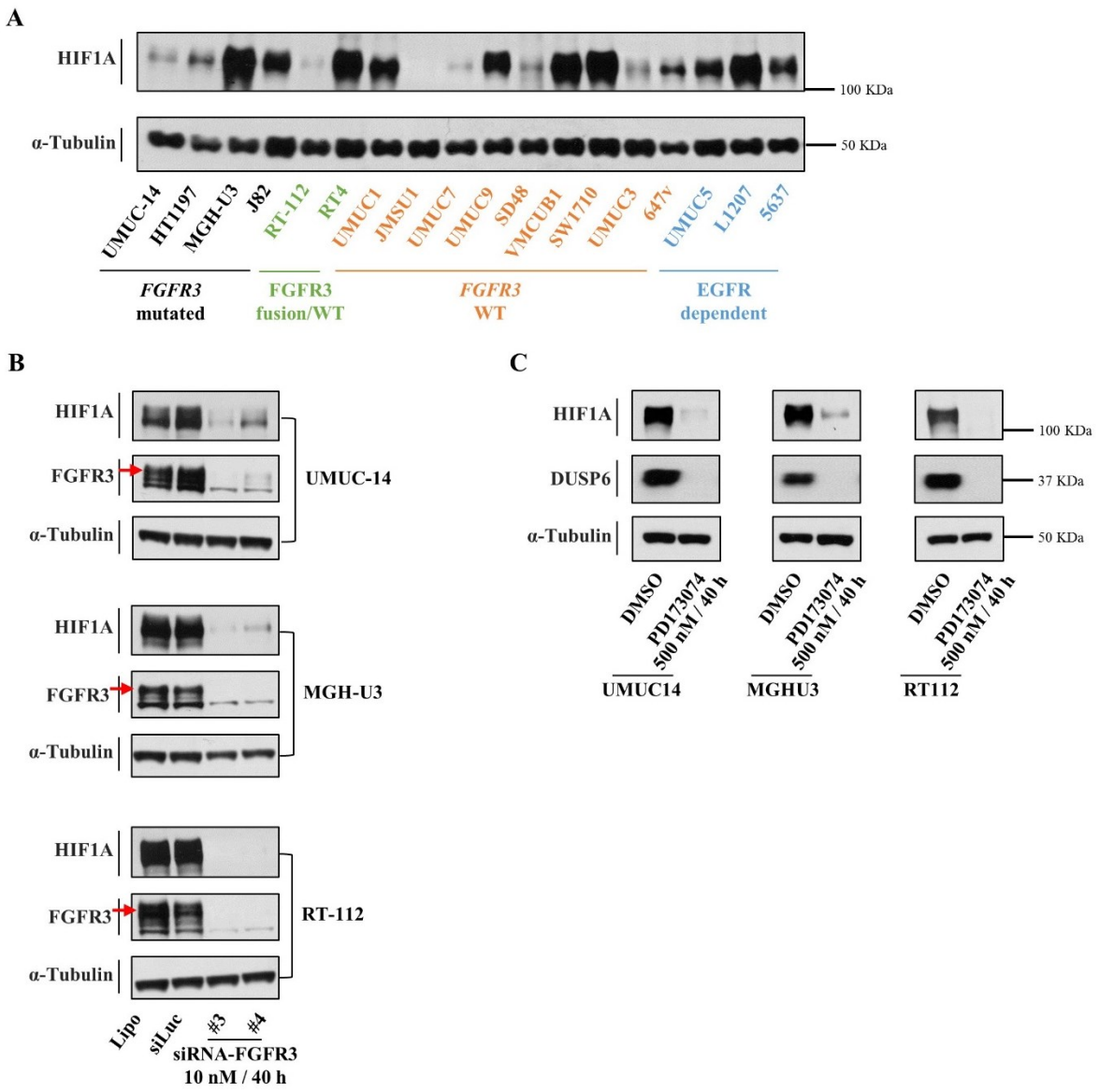


Figure 3-12: *HIF1A was regulated by FGFR3 in vitro.*

A) *HIF1A protein expression in various bladder cancer cell lines cultured in 'normoxic' condition.* ***B)*** *HIF1A was downregulated after knockdowns of FGFR3 in three FGFR3-dependent cell models.* ***C)*** *HIF1A was downregulated after long-term inhibition of FGFR3 in three FGFR3-dependent cell models.*

Summary –omic data analysis & ongoing experiments

PP2A complex were the only validated partners of FGFR3 currently. As discussed previously, I failed to show the impact of PP2A on inactivating pAKT and pERK, which were known to be dephosphorylated by PP2A (Kiely and Kiely, 2015). No link was found between PP2A

complex and currently investigated FGFR3 transcriptome data. CREB1, a target identified from phospho-kinase array, was regulated by FGFR3 in a time-dependent manner and seem not be a strong downstream effector. Initially I was interested in it because I found two well-known upstream kinases of CREB1, CaMKII and PKA (Zheng and Keifer, 2009), identified from the interactome and phosphoproteome data. However, except for the reason that I failed to validate the interaction between FGFR3 and CaMKII, the activity of both CaMKII (p-T286) and PKA (p-T197) were not modulated upon FGFR3 inhibition as expected. Therefore, this piece will not be the major direction for further exploration. FGFR3 is identified as the one of most strong client of HSP90/Cdc37 complex to date, and even different mutations of FGFR3 can alter binding affinity to this complex (Bunney et al., 2018; Taipale et al., 2012). In line with literature, we also identified HSP90/Cdc37 complex as highly confident partners. Of note, Cdc37 is obligatory phosphorylated on S13 by CK2 to maintain the stability of HSP90/Cdc37 ternary complex (Bunney et al., 2018), implying the rationality that CK2 was predicted as upstream kinase by FGFR3 phosphoproteome in our model. This point needs to be further elucidated.

Our transgenic mouse model brings new insight into FGFR3-driven signaling pathways and highlighted the involvement of angiogenesis that mediated by transcription factor – HIF1A. However, a direct regulation between FGFR3 and HIF1A remains questionable although a positive feedback loop may exist regarding HIF1A which can upregulate FGFR3 expression by repressing miR-100 in BCa cell lines (Blick et al., 2013). Interestingly, the HIF1A signaling can also be enriched from both interactome and phosphoproteome data, and the regulation of HIF1A by FGFR3 was also predicted from transcriptomic data as well as validated from *in vitro* models cultured in normoxia conditions. One of the next task will be the validation of this regulation axis in hypoxic cell culture conditions. With the aim trying to integrate similar observations from these omics datasets, we expected to figure out potential downstream kinase or effector of FGFR3 mediating the regulation between FGFR3 and HIF1A. From the three candidate upstream kinases predicted from FGFR3 phosphoproteome, CDK1 is a plausible candidate that may mediate this regulation. This is an unexpected link, because it is common that HIF1A is a major regulator of G1 cell cycle arrest by upregulating expression of $p27^{Kip1}$ during hypoxia whereas CDK1/cyclin B activation exclusively occurs in mitosis (Gardner et al., 2001; Goda et al., 2003). However, consistent with my hypothesis, there are emerging evidence showing that both CDK1 and CDK2 can directly interact with HIF1A, but only CDK1/cyclin B can stabilize HIF1A and promotes cell cycle progression and tumor growth,

even under 'normoxic' condition (Hubbi et al., 2014; Warfel et al., 2013). Part of current work is to validating whether CDK1 is involved in the regulation between FGFR3 and HIF1A in BCa. Nevertheless, the regulation axis of FGFR3/CDK1/HIF1A may be not that simple considering the involvement of MYC protein, but rather a complex network. There are many publications consistently reporting that MYC can either directly upregulated CDK1/cyclin B expressions or inhibit $p27^{Kip1}$ protein to promote CDK1 activation indirectly, and CDK1 inhibition (but not CDK4/6 inhibition) is selectively lethal to MYC-dependent tumors (Garcia-Gutierrez et al., 2016; Goga et al., 2007; Kang et al., 2014; Yang et al., 2018). Therefore, future validation steps should not overlook the role of MYC.

In summary, we generated the first FGFR3 proteomic dataset with several unique features: 1) utilized BCa cell line bearing FGFR3 mutation and with endogenously high activation of FGFR3; 2) applied SILAC coupled MS technique for precise protein quantification. We also generated corresponding transcriptome dataset for the same cell line by knocking down of FGFR3. The integration of these different datasets centred on FGFR3 enable us to propose potential regulatory axis towards FGFR3/CDK1/HIF1A/angiogenesis, where MYC may also played a role.

Chapter 4 Conclusions & perspectives

4.1 Conclusions

In my PhD, I focused on exploring the association between APOBEC mutagenesis and frequent mutations (including *FGFR3* S249C) in BCa to identify the aetiology of certain frequent mutations. I also participated in generating a transgenic mouse model of non-muscle-invasive BCa by targeting the expression of *FGFR3* S249C in the urothelium. Mostly, I made effort to reveal FGFR3-driven signaling pathways with proteomic strategy in BCa.

4.1.1 APOBEC mutagenesis & frequent mutations

We propose that APOBEC mutagenesis can generate frequent mutations in BCa and shape landscape of frequent mutations for each targeted gene, such as in the case of *FGFR3* S249C, one of the most frequent mutations in BCa. These APOBEC-associated mutations induce not solely gain-of-function mutations and are not confined to canonical APOBEC motifs. Location of mutations on lagging strand templates during DNA replication and loops with hairpin structure facilitate spatial accessibility to APOBEC enzymes. Hopefully, these APOBEC-associated mutations provide potency to select patients with BCa for optimal clinical treatment by evaluating the level of APOBEC mutagenesis in tumor samples.

4.1.2 Transgenic mouse model based on UII-*FGFR3* S249C

We generate the first GEM model of spontaneous low-grade papillary BCa formation and demonstrate that a mutated *FGFR3* alone can alter urothelium differentiation leading to hyperplasia and act as an oncogene leading to neoplastic transformation *in vivo*. Similar to human luminal papillary BCa, our model is also non-T-cell-inflamed and presents poorly enriched immune-environment. We also show that the transformation process is associated with increased genome instability and enhanced angiogenesis that probably mediated by hypoxia-inducing factor (*HIF1A*).

4.1.3 FGFR3-driven signaling pathways

To our knowledge, this is the first proteomic study, coupling SILAC labelling system with LC-MS/MS technology, investigating FGFR3 binding partners and downstream phosphoproteome in BCa cells that harboured endogenously mutated/activated FGFR3. I identified new FGFR3 interacting protein complex such as PP2A/STRN, and validated several FGFR3 regulating downstream kinases such as CDK1, but still there are many validation experiments are ongoing.

We also generated transcriptome data by knocking down FGFR3 within the same cell line, by which HIF1A was pinpointed as a downstream TF and validated *in vitro*. This is consistent to what we have observed from our GEM model. Interestingly, HIF1 signaling can also be enriched from both interactome and phosphoproteome data, thus we found this is a good sign to follow this pathway in-depth. It has been reported that CDK1 can regulate HIF1A directly. Part of current work is to validating whether CDK1 is involved in the FGFR3/HIF1A regulation axis in BCa. To make a long story short, I summarized the background knowledge, the main findings and the future perspectives in Figure 4-1; page 216. Perspectives will be described in details in the next section.

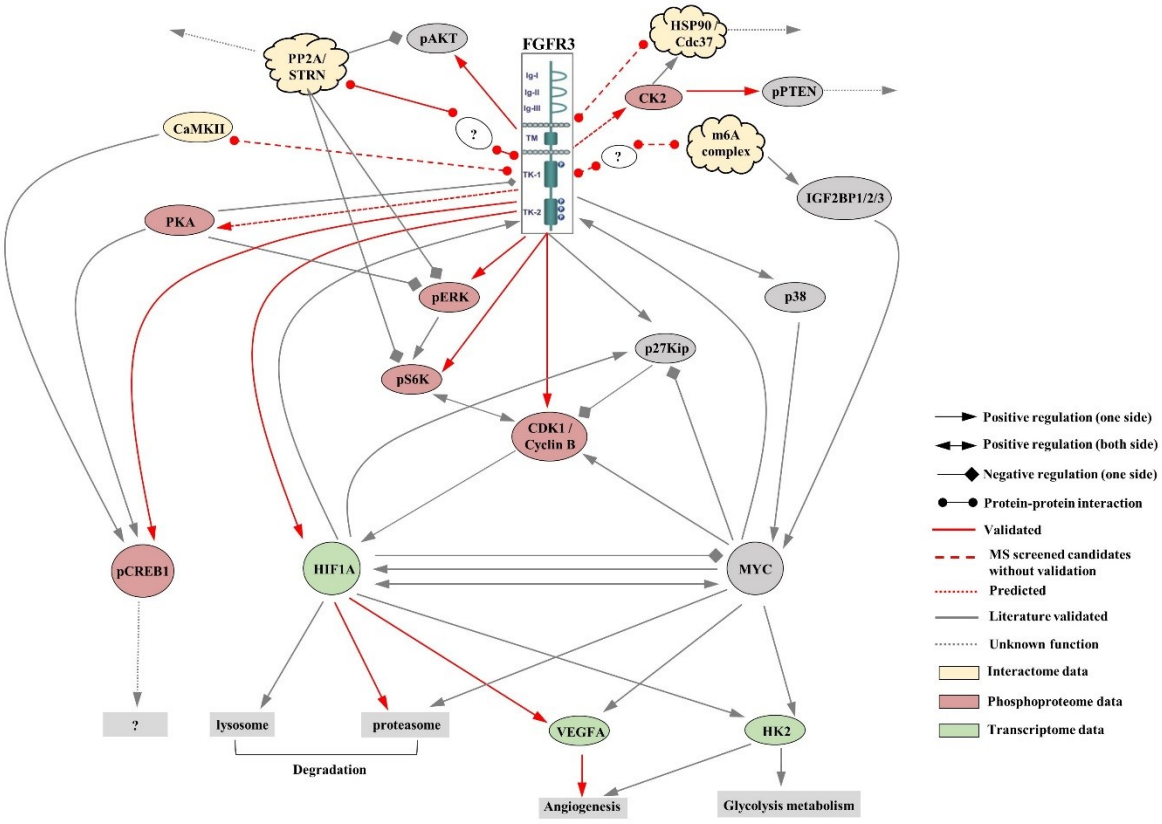


Figure 4-1: *Overview of –omics data integration and perspectives.*

4.2 Perspectives

4.2.1 APOBEC mutagenesis & frequent mutations

In BCa, APOBEC-mediated mutagenesis seems to be an early event and is associated with many clonal mutations. Although several APOBEC enzymes have been proposed as responsible mutagenesis origins in other cancer types, there is still not yet such clear evidence in BCa. To explore exactly which enzyme will be the main contributor in BCa is critical for a better understanding of etiologies. In addition we should not overlook other enzymes not belong to APOBEC family that could also co-contribute to the process of APOBEC mutagenesis.

The observation that APOBEC enzymes not only show trinucleotide context-specificity but also the loop within DNA-hairpin, gives us a lesson that in addition to context-dependency, the similar loop or other unknown features may exist for other mutational signatures as well.

Our results challenge the dogma that hotspot mutations are always associated with gain-of-function affecting oncogenes since they can also be associated with passenger mutations. Thus, future studies aiming to screen cancer driver mutations based on their prevalence should be treated with caution. Our model predicts new driver hotspot mutations in BCa that could be potential new therapeutic targets but need to be validated experimentally in the future.

From our preliminary findings, patients bearing APOBEC-associated hotspot mutations may be susceptibile to immunotherapy or anti-ATR treatment. It should be confirmed in BCa cell lines with specific anti-ATR inhibitors first; and then being validated in a prospective clinical cohort if results are optimising *in vitro*.

4.2.2 Transgenic mouse model based on UII-*FGFR3* S249C

Considering the life span of the mouse, the time (18 months) we observed the tumor formation is too tartive and limited to many possible following experimental designs. This result also indicates the second or prior event that couples with FGFR3 mutation is necessarily to accelerate tumorigenesis. Herein, future models combined UII-*FGFR3* S249C mutation with other mutational events / environmental exposures are useful to understand such crucial events.

Our transgenic mouse model show an immune-suppression, thus probably be less responsive to immunotherapy. This hypothesis should be further validated and if so, this is better evidence to show the counteraction between FGFR3 mutation and immunotherapy than the undergoing clinical trial with the same aim.

We show clear regulation of HIF1A by FGFR3 and associated this axis with angiogenesis. However, HIF1A has been reported to upregulate FGFR3 expression by repressing miR-100 in BCa cell lines (Blick et al., 2013). Herein, a positive feedback loop may exist between FGFR3 and HIF1A and which should be validated in the next step.

Of note, HIF1A does not only regulate angiogenesis but also glycolysis metabolism as well as others. In addition, MYC proteins may be potentially involved in the regulation of FGFR3/HIF1A/angiogenesis. This complex regulatory network and interaction between different biological processes should be addressed in the future.

4.2.3 FGFR3-driven signaling pathways

We show clear interaction between PP2A complex and FGFR3. However, it is still undetermined about the biological influence of PP2A complex in BCa and not clear how PP2A complex is regulated by FGFR3? m6A complex is highlighted in our interactome, validating the interaction between FGFR3 and this complex may be interesting should be explored in the future, as emerging evidence shows that m6A complex stabilised MYC protein (Huang et al., 2018), which is a master downstream hub of FGFR3 in BCa (Mahe et al., 2018). Another identified interactor of FGFR3 is HSP90/Cdc37 complex, whose inhibition alone or combined with other inhibitors, such as RTKs inhibition, have achieved synergistic and optimising anticancer activity (Dai and Whitesell, 2005; Li et al., 2018). Thus, it may be interesting to explore combined inhibition of FGFR3 and HSP90/Cdc37 to see their anticancer efficacy in BCa. I fail to trap classical partners of FGFRs/FGFR3, such as FRS2, PLCG1, GRB2 and GAB1, etc. This is may be due to distinct PPI profiles between mutated FGFR3 versus wild-type FGFR3 or FGFRs, and need to be explored in the future. Lastly, many of these partners are localized in the nuclei, implying a function of nuclear FGFR3. Thus, future experimental optimization should pay attention to compartment-specific interactome.

In our phosphoproteome experiments, only two time points of treatment are used for for drawing dynamic regulation profile, which seems not be sufficient. I propose both short and

long-term treatments should be included in future experimental design to achieve a more representative and dynamic picture. One limitation of current study is the lack of phosphotyrosine phosphoproteome, which would be also interesting regarding RTK signaling. This should be considered to be supplemented. As discussed previously, CDK1 is a plausible candidate that may mediate the regulation of FGFR3/HIF1A in BCa. Current work is ongoing to test this hypothesis. Interestingly, CDK1 inhibition (but not CDK4/6 inhibition) is selectively lethal to MYC-dependent hematologic or breast tumors (Garcia-Gutierrez et al., 2016; Goga et al., 2007; Kang et al., 2014; Yang et al., 2018). Considering the important positive feedback loop between FGFR3 and MYC in BCa, a synergistic effect can be expected by combining FGFR3 and CDK1 inhibitors, which needs to be further validated in our model and if so, will give new hope for BCa treatment.

Current *in vitro* validation for FGFR3/HIF1A regulation is performed in cell lines cultured under normoxic conditions. Future experiments should try to see whether we could reproduce the same results in cell lines cultured under hypoxic conditions. As summarized previously, MYC-CDK1-FGFR3/HIF1A proteins may play as an orchestra. The involvement of MYC in this regulatory network should be investigated.

Bibliography

BIBLIOGRAPHY

Bibliography

Abern, M.R., Dude, A.M., Tsivian, M., and Coogan, C.L. (2013). The characteristics of bladder cancer after radiotherapy for prostate cancer. Urol. Oncol. *31*, 1628–1634.

Adar, R., Monsonego-Ornan, E., David, P., and Yayon, A. (2002). Differential activation of cysteine-substitution mutants of fibroblast growth factor receptor 3 is determined by cysteine localization. J. Bone Miner. Res. *17*, 860–868.

Adolph, M.B., Love, R.P., Feng, Y., and Chelico, L. (2017). Enzyme cycling contributes to efficient induction of genome mutagenesis by the cytidine deaminase APOBEC3B. Nucleic Acids Res. *45*, 11925–11940.

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198–207.

Ahmad, I., Singh, L.B., Foth, M., Morris, C.-A., Taketo, M.M., Wu, X.-R., Leung, H.Y., Sansom, O.J., and Iwata, T. (2011). K-Ras and b-catenin mutations cooperate with Fgfr3 mutations in mice to promote tumorigenesis in the skin and lung, but not in the bladder. Dis. Model. Mech. *4*, 548–555.

Al-Ahmadie, H.A., Iyer, G., Janakiraman, M., Lin, O., Heguy, A., Tickoo, S.K., Fine, S.W., Gopalan, A., Chen, Y.B., Balar, A., et al. (2011). Somatic mutation of fibroblast growth factor receptor-3 (FGFR3) defines a distinct morphological subtype of high-grade urothelial carcinoma. J. Pathol. *224*, 270–279.

Al-Zalabani, A.H., Stewart, K.F.J., Wesselius, A., Schols, A.M.W.J., and Zeegers, M.P. (2016). Modifiable risk factors for the prevention of bladder cancer: a systematic review of meta-analyses (Springer Netherlands).

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Campbell, P.J., and Stratton, M.R. (2013a). Deciphering Signatures of Mutational Processes Operative in Human Cancer. Cell Rep. *3*, 246–259.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A. V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.-L., et al. (2013b). Signatures of mutational processes in human cancer. Nature *500*, 415–421.

Alfred Witjes, J., Lebret, T., Compérat, E.M., Cowan, N.C., De Santis, M., Bruins, H.M., Hernández, V., Espinós, E.L., Dunn, J., Rouanne, M., et al. (2017). Updated 2016 EAU Guidelines on Muscle-invasive and Metastatic Bladder Cancer. Eur. Urol. *71*, 462–475.

Allen, J.M., Schrock, A.B., Erlich, R.L., Miller, V.A., Stephens, P.J., Ross, J.S., Ou, S.H.I., Ali, S.M., and Vafai, D. (2017). Genomic Profiling of Circulating Tumor DNA in Relapsed EGFR-mutated Lung Adenocarcinoma Reveals an Acquired FGFR3-TACC3 Fusion. Clin. Lung Cancer *18*, e219–e222.

Alonso-López, D., Gutiérrez, M.A., Lopes, K.P., Prieto, C., Santamaría, R., and De Las Rivas, J. (2016). APID interactomes: Providing proteome-based interactomes with controlled quality for multiple species and derived networks. Nucleic Acids Res. *44*, W529–W535.

Awasthi, P., Foiani, M., and Kumar, A. (2016). ATM and ATR signaling at a glance. J. Cell Sci. *129*, 1285–1285.

Babina, I.S., and Turner, N.C. (2017). Advances and challenges in targeting FGFR signalling

in cancer. Nat. Rev. Cancer *17*, 318–332.

Babjuk, M., Böhle, A., Burger, M., Capoun, O., Cohen, D., Compérat, E.M., Hernández, V., Kaasinen, E., Palou, J., Rouprêt, M., et al. (2017). EAU Guidelines on Non–Muscle-invasive Urothelial Carcinoma of the Bladder: Update 2016. Eur. Urol. *71*, 447–461.

Baillat, G., Moqrich, A., Castets, F., Baude, A., Bailly, Y., Benmerah, A., and Monneron, A. (2001). Molecular Cloning and Characterization of Phocein, a Protein Found from the Golgi Complex to Dendritic Spines. Mol. Biol. Cell *12*, 663–673.

Balar, A. V., Galsky, M.D., Rosenberg, J.E., Powles, T., Petrylak, D.P., Bellmunt, J., Loriot, Y., Necchi, A., Hoffman-Censits, J., Perez-Gracia, J.L., et al. (2017). Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. Lancet *389*, 67–76.

Balek, L., Nemec, P., Konik, P., Kunova Bosakova, M., Varecha, M., Gudernova, I., Medalova, J., Krakow, D., and Krejci, P. (2018). Proteomic analyses of signalling complexes associated with receptor tyrosine kinase identify novel members of fibroblast growth factor receptor 3 interactome. Cell. Signal. *42*, 144–154.

Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007). Quantitative mass spectrometry in proteomics : a critical review. Anal Bioanal Chem *389*, 1017–1031.

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G. V, Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603–607.

Bellmunt, J., de Wit, R., Vaughn, D.J., Fradet, Y., Lee, J.-L., Fong, L., Vogelzang, N.J., Climent, M.A., Petrylak, D.P., Choueiri, T.K., et al. (2017). Pembrolizumab as Second-Line Therapy for Advanced Urothelial Carcinoma. N. Engl. J. Med. *376*, 1015–1026.

Bellot, F., Crumley, G., Kaplow, J.M., Schlessinger, J., Jaye, M., and Dionne, C.A. (1991). Ligand-induced transphosphorylation between different FGF receptors. EMBO J. *10*, 2849–2854.

Bencivenga, D., Tramontano, A., Borgia, A., Negri, A., Caldarelli, I., Oliva, A., Perrotta, S., Ragione, F. Della, and Borriello, A. (2014). P27Kip1 serine 10 phosphorylation determines its metabolism and interaction with cyclin-dependent kinases. Cell Cycle *13*, 3768–3782.

Bergeron, J.J.M., Di Guglielmo, G.M., Dahan, S., Dominguez, M., and Posner, B.I. (2016). Spatial and Temporal Regulation of Receptor Tyrosine Kinase Activation and Intracellular Signal Transduction. Annu. Rev. Biochem. *85*, 573–597.

Bernard-Pierrot, I., Brams, A., Dunois-Lardé, C., Caillault, A., Diez de Medina, S.G., Cappellen, D., Graff, G., Thiery, J.P., Chopin, D., Ricol, D., et al. (2006). Oncogenic properties of the mutated forms of fibroblast growth factor receptor 3b. Carcinogenesis *27*, 740–747.

Bertz, S., Abeé, C., Schwarz-Furlan, S., Alfer, J., Hofstädter, F., Stoehr, R., Hartmann, A., and Gaumann, A.K.A. (2014). Increased angiogenesis and FGFR protein expression indicate a favourable prognosis in bladder cancer. Virchows Arch. *465*, 687–695.

Best, S.A., Harapas, C.R., Kersbergen, A., Rathi, V., Asselin-Labat, M.L., and Sutherland, K.D. (2018). FGFR3-TACC3 is an oncogenic fusion protein in respiratory epithelium. Oncogene *37*, 6096–6104.

Billerey, C., Chopin, D., Aubriot-Lorton, M.H., Ricol, D., De Medina, S.G.D., Van Rhijn, B., Bralet, M.P., Lefrere-Belda, M.A., Lahaye, J.B., Abbou, C.C., et al. (2001). Frequent FGFR3 mutations in papillary non-invasive bladder (pTa) tumors. Am. J. Pathol. *158*, 1955–1959.

Black, P.C., Brown, G.A., Dinney, C.P., Kassouf, W., Inamoto, T., Arora, A., Gallagher, D., Munsell, M.F., Bar-Eli, M., McConkey, D.J., et al. (2011). Receptor heterodimerization: A new mechanism for platelet-derived growth factor induced resistance to anti-epidermal growth factor receptor therapy for bladder cancer. J. Urol. *185*, 693–700.

Blagoev, B., Kratchmarova, I., Ong, S.E., Nielsen, M., Foster, L.J., and Mann, M. (2003). A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. Nat. Biotechnol. *21*, 315–318.

Blencke, S., Zech, B., Engkvist, O., Greff, Z., Orfi, L., Horvath, Z., Keri, G., Ullrich, A., and Daub, H. (2004). Characterization of a Conserved Structural Determinant Controlling Protein Kinase Sensitivity to Selective Inhibitors. Chem. Biol. *11*, 691–701.

Blick, C., Ramachandran, A., Wigfield, S., McCormick, R., Jubb, A., Buffa, F.M., Turley, H., Knowles, M.A., Cranston, D., Catto, J., et al. (2013). Hypoxia regulates FGFR3 expression via HIF-1α and miR-100 and contributes to cell survival in non-muscle invasive bladder cancer. Br. J. Cancer *109*, 50–59.

Bocharov, E. V., Lesovoy, D.M., Goncharuk, S.A., Goncharuk, M. V., Hristova, K., and Arseniev, A.S. (2013). Structure of FGFR3 transmembrane domain dimer: Implications for signaling and human pathologies. Structure *21*, 2087–2093.

Bonneville, R., Krook, M.A., Kautto, E.A., Miya, J., Wing, M.R., Chen, H.-Z., Reeser, J.W., Yu, L., and Roychowdhury, S. (2017). Landscape of Microsatellite Instability Across 39 Cancer Types. JCO Precis. Oncol. 1–15.

Botteman, M.F., Pashos, C.L., Redaelli, A., Laskin, B., and Hauser, R. (2003). The Health Economics of Bladder Cancer: A Comprehensive Review of the Published Literature. Pharmacoeconomics *21*, 1315–1330.

De Bruin, E.C., McGranahan, N., Mitter, R., Salm, M., Wedge, D.C., Yates, L., Jamal-Hanjani, M., Shafi, S., Murugaesu, N., Rowan, A.J., et al. (2014). Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science (80-.). *346*, 251–256.

Buisson, R., Lawrence, M.S., Benes, C.H., and Zou, L. (2017). APOBEC3A and APOBEC3B activities render cancer cells susceptible to ATR inhibition. Cancer Res. *77*, 4567–4578.

Bunney, T.D., Inglis, A.J., Sanfelice, D., Farrell, B., Kerr, C.J., Thompson, G.S., Masson, G.R., Thiyagarajan, N., Svergun, D.I., Williams, R.L., et al. (2018). Disease Variants of FGFR3 Reveal Molecular Basis for the Recognition and Additional Roles for Cdc37 in Hsp90 Chaperone System. Structure *26*, 446-458.e8.

Burger, M., Catto, J.W.F., Dalbagni, G., Grossman, H.B., Herr, H., Karakiewicz, P., Kassouf, W., Kiemeney, L.A., Vecchia, C. La, Shariat, S., et al. (2013). Epidemiology and Risk Factors of Urothelial Bladder Cancer. Eur. Urol. *63*, 234–241.

Byron, S.A., Chen, H., Wortmann, A., Loch, D., Gartside, M.G., Dehkhoda, F., Blais, S.P., Neubert, T.A., Mohammadi, M., and Pollock, P.M. (2015). The N550K/H Mutations in FGFR2 Confer Differential Resistance to PD173074, Dovitinib, and Ponatinib ATP-Competitive Inhibitors. Neoplasia *15*, 975-IN30.

Calderaro, J., Rebouissou, S., De Koning, L., Masmoudi, A., Hérault, A., Dubois, T., Maille, P., Soyeux, P., Sibony, M., De La Taille, A., et al. (2014). PI3K/AKT pathway activation in bladder carcinogenesis. Int. J. Cancer *134*, 1776–1784.

Cannataro, V.L., Gaffney, S.G., Sasaki, T., Issaeva, N., Grewal, N.K.S., Grandis, J.R., Yarbrough, W.G., Burtness, B., Anderson, K.S., and Townsend, J.P. (2019). APOBEC-

induced mutations and their cancer effect size in head and neck squamous cell carcinoma. Oncogene *38*, 3475–3487.

Cappellen, D., Oliveira, C. De, Ricol, D., Medina, S.G.D. de, Bourdin, J., Sastre-Garau, X., Chopin, D., Thiery, J.P., and Radvanyi, F. (1999). Frequent activating mutations of FGFR3 in human bladder. Nat. Genet. *23*, 18–20.

Casci, T., Vinós, J., and Freeman, M. (1999). Sprouty, an intracellular inhibitor of Ras signaling. Cell *96*, 655–665.

Chaffee, B.R., Hoang, T. V., Leonard, M.R., Bruney, D.G., Wagner, B.D., Dowd, J.R., Leone, G., Ostrowski, M.C., and Robinson, M.L. (2016). FGFR and PTEN signaling interact during lens development to regulate cell survival. Dev. Biol. *410*, 150–163.

Chan, K., Roberts, S.A., Klimczak, L.J., Sterling, J.F., Saini, N., Malc, E.P., Kim, J., Kwiatkowski, D.J., Fargo, D.C., Mieczkowski, P.A., et al. (2015). An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. Nat. Genet. *47*, 1067–1072.

Chell, V., Balmanno, K., Little, A.S., Wilson, M., Andrews, S., Blockley, L., Hampson, M., Gavine, P.R., and Cook, S.J. (2013). Tumour cell responses to new fibroblast growth factor receptor tyrosine kinase inhibitors and identification of a gatekeeper mutation in FGFR3 as a mechanism of acquired resistance. Oncogene *32*, 3059–3070.

Chen, M.K., and Hung, M.C. (2015). Proteolytic cleavage, trafficking, and functions of nuclear receptor tyrosine kinases. FEBS J. *282*, 3693–3721.

Chen, F., Degnin, C., Laederich, M., Horton, W.A., and Hristova, K. (2011). The A391E mutation enhances FGFR3 activation in the absence of ligand. Biochim. Biophys. Acta - Biomembr. *1808*, 2045–2050.

Chen, H., Ma, J., Li, W., Eliseenkova, A. V., Xu, C., Neubert, T.A., Miller, W.T., and Mohammadi, M. (2007). A Molecular Brake in the Kinase Hinge Region Regulates the Activity of Receptor Tyrosine Kinases. Mol. Cell *27*, 717–730.

Chen, J., Martin, B.L., and Brautigan, D.L. (1992). Regulation of Protein Serine-Threonine Phosphatase Type-2A by Tyrosine Phosphorylation. Science (80-.). *257*, 1261–1264.

Chen, J., Williams, I.R., Lee, B.H., Duclos, N., Huntly, B.J.P., Donoghue, D.J., and Gilliland, D.G. (2005). Constitutively activated FGFR3 mutants signal through PLCr-dependent and -independent pathways for hematopoietic transformation. Blood *106*, 328–337.

Chen, T.W., Lee, C.C., Liu, H., Wu, C.S., Pickering, C.R., Huang, P.J., Wang, J., Chang, I.Y.F., Yeh, Y.M., Chen, C. De, et al. (2017). APOBEC3A is an oral cancer prognostic biomarker in Taiwanese carriers of an APOBEC deletion polymorphism. Nat. Commun. *8*, 465.

Chen, W., Zheng, R., Baade, P.D., Zhang, S., Zeng, H., Bray, F., Jemal, A., Yu, X.Q., and He, J. (2016). Cancer statistics in China, 2015. CA. Cancer J. Clin. *66*, 115–132.

Chioni, A.M., and Grose, R. (2012). FGFR1 cleavage and nuclear translocation regulates breast cancer cell behavior. J. Cell Biol. *197*, 801–817.

Chou, M.F., and Schwartz, D. (2011). Biological sequence motif discovery using motif-x. Curr Protoc Bioinforma. *Chapter 13*, 15–24.

Cook, S.J., and Mccormick, F. (1993). Inhibition by cAMP of Ras-Dependent Activation of Raf. Science (80-.). *262*, 1069–1072.

Cooper, D.N., and Krawczak, M. (1989). Cytosine methylation and the fate of CpG

dinucleotides in vertebrate genomes. Hum. Genet. *83*, 181–188.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Cross, M.J., Lu, L., Magnusson, P., Nyqvist, D., Holmqvist, K., Welsh, M., and Claesson-Welsh, L. (2002). The Shb Adaptor Protein Binds to Tyrosine 766 in the FGFR-1 and Regulates the Ras/MEK/MAPK Pathway via FRS2 Phosphorylation in Endothelial Cells. Mol. Biol. Cell *13*, 2881–2893.

Van Cutsem, E., Bang, Y.J., Mansoor, W., Petty, R.D., Chao, Y., Cunningham, D., Ferry, D.R., Smith, N.R., Frewer, P., Ratnayake, J., et al. (2017). A randomized, open-label study of the efficacy and safety of AZD4547 monotherapy versus paclitaxel for the treatment of advanced gastric adenocarcinoma with FGFR2 polysomy or gene amplification. Ann. Oncol. *28*, 1316–1324.

Dai, C., and Whitesell, L. (2005). HSP90: A rising star on the horizon of anticancer targets. Futur. Oncol. *1*, 529–540.

Daly, C., Castanaro, C., Zhang, W., Zhang, Q., Wei, Y., Ni, M., Young, T.M., Zhang, L., Burova, E., and Thurston, G. (2017). FGFR3-TACC3 fusion proteins act as naturally occurring drivers of tumor resistance by functionally substituting for EGFR/ERK signaling. Oncogene *36*, 471–481.

Davies, H., Glodzik, D., Morganella, S., Yates, L.R., Staaf, J., Zou, X., Ramakrishna, M., Martin, S., Boyault, S., Sieuwerts, A.M., et al. (2017). HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. Nat. Med. *23*, 517–525.

Degnin, C.R., Laederich, M.B., and Horton, W.A. (2011). Ligand activation leads to regulated intramembrane proteolysis of fibroblast growth factor receptor 3. Mol. Biol. Cell *22*, 3861–3873.

Deng, X., Su, R., Weng, H., Huang, H., Li, Z., and Chen, J. (2018). RNA N 6 -methyladenosine modification in cancers: Current status and perspectives. Cell Res. *28*, 507–517.

Dimitroff, C.J., Klohs, W., Sharma, A., Pera, P., Driscoll, D., Veith, J., Steinkampf, R., Schroeder, M., Klutchko, S., Sumlin, A., et al. (1999). Anti-angiogenic activity of selected receptor tyrosine kinase inhibitors, PD166285 and PD173074: Implications for combination treatment with photodynamic therapy. Invest. New Drugs *17*, 121–135.

Dobruch, J., Daneshmand, S., Fisch, M., Lotan, Y., Noon, A.P., Resnick, M.J., Shariat, S.F., Zlotta, A.R., and Boorjian, S.A. (2016). Gender and Bladder Cancer: A Collaborative Review of Etiology, Biology, and Outcomes. Eur. Urol. *69*, 300–310.

Doe, M.R., Ascano, J.M., Kaur, M., and Cole, M.D. (2012). Myc posttranscriptionally induces HIF1 protein and target gene expression in normal and cancer cells. Cancer Res. *72*, 949–957.

Dudka, A.A., Sweet, S.M.M., and Heath, J.K. (2010). Signal transducers and activators of transcription-3 binding to the fibroblast growth factor receptor is activated by receptor amplification. Cancer Res. *70*, 3391–3401.

Dunham-Ems, S.M., Lee, Y.-W., Stachowiak, E.K., Pudavar, H., Claus, P., Prasad, P.N., and Stachowiak, M.K. (2009). Fibroblast Growth Factor Receptor-1 (FGFR1) Nuclear Dynamics Reveal a Novel Mechanism in Transcription Control. Mol. Biol. Cell *20*, 2401–00.

Earl, J., Rico, D., Carrillo-de-Santa-Pau, E., Rodríguez-Santiago, B., Méndez-Pertuz, M., Auer, H., Gómez, G., Grossman, H.B., Pisano, D.G., Schulz, W.A., et al. (2015). The UBC-40

Urothelial Bladder Cancer cell line index: A genomic resource for functional studies. BMC Genomics *16*.

Eble, J.N., Sauter, G., Epstein, J.I., and Sesterhenn, I.A. (2004). World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of the Urinary System and Male Genital Organs. IARC Press Lyon.

Eichhorn, P.J.A., Creyghton, M.P., and Bernards, R. (2009). Protein phosphatase 2A regulatory subunits and cancer. Biochim. Biophys. Acta - Rev. Cancer *1795*, 1–15.

Fantini, D., Glaser, A.P., Rimar, K.J., Wang, Y., Schipma, M., Varghese, N., Rademaker, A., Behdad, A., Yellapa, A., Yu, Y., et al. (2018). A Carcinogen-induced mouse model recapitulates the molecular alterations of human muscle invasive bladder cancer. Oncogene *37*, 1911–1925.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D., and Bray, F. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer *136*, E359-86.

Fichter, C.D., Timme, S., Braun, J.A., Gudernatsch, V., Schöpflin, A., Bogatyreva, L., Geddert, H., Faller, G., Klimstra, D., Tang, L., et al. (2014). EGFR, HER2 and HER3 dimerization patterns guide targeted inhibition in two histotypes of esophageal cancer. Int. J. Cancer *135*, 1517–1530.

Fioriti, D., Pietropaolo, V., Dal Forno, S., Laurenti, C., Chiarini, F., and Degener, A.M. (2003). Urothelial bladder carcinoma and viral infections: Different association with human polyomaviruses and papillomaviruses. Int. J. Immunopathol. Pharmacol. *16*, 283–288.

Foldynova-Trantirkova, S., Wilcox, W.R., and Krejci, P. (2012). Sixteen years and counting: The current understanding of fibroblast growth factor receptor 3 (FGFR3) signaling in skeletal dysplasias. Hum. Mutat. *33*, 29–41.

Forbes, S.A., Beare, D., Boutselakis, H., Bamford, S., Bindal, N., Tate, J., Cole, C.G., Ward, S., Dawson, E., Ponting, L., et al. (2017). COSMIC: Somatic cancer genetics at high-resolution. Nucleic Acids Res. *45*, D777–D783.

Forget, A., Martignetti, L., Puget, S., Calzone, L., Brabetz, S., Picard, D., Montagud, A., Liva, S., Sta, A., Dingli, F., et al. (2018). Aberrant ERBB4-SRC Signaling as a Hallmark of Group 4 Medulloblastoma Revealed by Integrative Phosphoproteomic Profiling. Cancer Cell *34*, 379-395.e7.

Foth, M., Ismail, N.F.B., Kung, J.S.C., Tomlinson, D., Knowles, M.A., Eriksson, P., Sjödahl, G., Salmond, J.M., Sansom, O.J., and Iwata, T. (2018). FGFR3 mutation increases bladder tumourigenesis by suppressing acute inflammation. J. Pathol. *246*, 331–343.

Freedman, N.D., Silverman, D.T., Hollenbeck, A.R., Schatzkin, A., and Abnet, C.C. (2011). Association Between Smoking and Risk of Bladder Cancer Among Men and Women. JAMA *306*, 737.

Fujiki, H., and Suganuma, M. (1993). Tumor Promotion by Inhibitors of ProteinZ Phosphatases 1 and 2A: The Okadaic Acid Class of Compounds. Adv. Cancer Res. *61*, 143–194.

Fujiki, H., Sueoka, E., Watanabe, T., and Suganuma, M. (2018). The concept of the okadaic acid class of tumor promoters is revived in endogenous protein inhibitors of protein phosphatase 2A, SET and CIP2A, in human cancers. J. Cancer Res. Clin. Oncol. *144*, 2339–2349.

Fumarola, C., Cretella, D., La Monica, S., Bonelli, M.A., Alfieri, R., Caffarra, C., Quaini, F.,

Madeddu, D., Falco, A., Cavazzoni, A., et al. (2017). Enhancement of the anti-tumor activity of FGFR1 inhibition in squamous cell lung cancer by targeting downstream signaling involved in glucose metabolism. Oncotarget *8*, 91841–91859.

Furdui, C.M., Lew, E.D., Schlessinger, J., and Anderson, K.S. (2006). Autophosphorylation of FGFR1 kinase is mediated by a sequential and precisely ordered reaction. Mol. Cell *21*, 711–717.

Galsky, M.D., Hahn, N.M., Rosenberg, J., Sonpavde, G., Hutson, T., Oh, W.K., Dreicer, R., Vogelzang, N., Sternberg, C.N., Bajorin, D.F., et al. (2011). Treatment of patients with metastatic urothelial cancer "Unfit" for cisplatin-based chemotherapy. J. Clin. Oncol. *29*, 2432–2438.

Garcia-Gutierrez, L., Bretones, G., Arechaga, I., Santamaria, D., Barbacid, M., and Leon, J. (2016). Myc-dependent cell cycle progression through the activation of CDK1 and phosphorylation of p27. BioRxiv 1–15.

Gardner, L.B., Li, Q., Park, M.S., Flanagan, W.M., Semenza, G.L., and Dang, C. V. (2001). Hypoxia Inhibits G1/S Transition through Regulation of p27 Expression. J. Biol. Chem. *276*, 7919–7926.

Geleta, B., Makonnen, E., and Abay, S.M. (2016). Cyclic Dependent Kinase (CDK): Role in Cancer Pathogenesis and as Drug Target in Cancer Therapeutics. J. Cancer Sci. Ther. *8*, 160–167.

Giacomelli, A.O., Yang, X., Lintner, R.E., McFarland, J.M., Duby, M., Kim, J., Howard, T.P., Takeda, D.Y., Ly, S.H., Kim, E., et al. (2018). Mutational processes shape the landscape of TP53 mutations in human cancer. Nat. Genet. *50*, 1381–1387.

Gibbs, L., and Legeai-Mallet, L. (2007). FGFR3 intracellular mutations induce tyrosine phosphorylation in the Golgi and defective glycosylation. Biochim. Biophys. Acta - Mol. Cell Res. *1773*, 502–512.

Goda, N., Goda, N., Ryan, H.E., Ryan, H.E., Khadivi, B., Khadivi, B., McNulty, W., McNulty, W., Rickert, R.C., Rickert, R.C., et al. (2003). Hypoxia-Inducible Factor 1 Is Essential for Cell Cycle Arrest during Hypoxia. Mol. Cell. Biol. *23*, 359–369.

Goga, A., Yang, D., Tward, A.D., Morgan, D.O., and Bishop, J.M. (2007). Inhibition of CDK1 as a potential therapy for tumors over-expressing MYC. Nat. Med. *13*, 820–827.

Gordan, J.D., Thompson, C.B., and Simon, M.C. (2007). HIF and c-Myc: Sibling Rivals for Control of Cancer Cell Metabolism and Proliferation. Cancer Cell *12*, 108–113.

Gordon, J., Hwang, J., Carrier, K.J., Jones, C.A., Kern, Q.L., Moreno, C.S., Karas, R.H., and Pallas, D.C. (2011). Protein phosphatase 2a (PP2A) binds within the oligomerization domain of striatin and regulates the phosphorylation and activation of the mammalian Ste20-Like kinase Mst3. BMC Biochem. *12*, 1–18.

Goudreault, M., D'Ambrosio, L.M., Kean, M.J., Mullin, M.J., Larsen, B.G., Sanchez, A., Chaudhry, S., Chen, G.I., Sicheri, F., Nesvizhskii, A.I., et al. (2008). A PP2A Phosphatase High Density Interaction Network Identifies a Novel Striatin-interacting Phosphatase and Kinase Complex Linked to the Cerebral Cavernous Malformation 3 (CCM3) Protein. Mol. Cell. Proteomics *8*, 157–171.

Granvogl, B., Plöscher, M., and Eichacker, L.A. (2007). Sample preparation by in-gel digestion for mass spectrometry-based proteomics. Anal. Bioanal. Chem. *389*, 991–1002.

Green, A.M., Budagyan, K., Hayer, K.E., Reed, M.A., Savani, M.R., Wertheim, G.B., and

Weitzman, M.D. (2017). Cytosine deaminase APOBEC3A sensitizes leukemia cells to inhibition of the DNA replication checkpoint. Cancer Res. *77*, 4579–4588.

Greenman, C., Stephens, P., Smith, R., Dalgliesh, G.L., Hunter, C., Bignell, G., Davies, H., Teague, J., Butler, A., Stevens, C., et al. (2007). Patterns of somatic mutation in human cancer genomes. Nature *446*, 153–158.

Gu, Y., Zhang, J., Ma, X., Kim, B., Wang, H., and Li, J. (2017). Supplemental Information Stabilization of the c-Myc Protein by CAMKII g Promotes T Cell Lymphoma. Cancer Cell *32*, 115–128.

Gulhan, D.C., Lee, J.J.-K., Melloni, G.E.M., Cortés-Ciriano, I., and Park, P.J. (2019). Detecting the mutational signature of homologous recombination deficiency in clinical samples. Nat. Genet.

Gust, K.M., McConkey, D.J., Awrey, S., Hegarty, P.K., Qing, J., Bondaruk, J., Ashkenazi, A., Czerniak, B., Dinney, C.P., and Black, P.C. (2013). Fibroblast Growth Factor Receptor 3 Is a Rational Therapeutic Target in Bladder Cancer. Mol. Cancer Ther. *12*, 1245–1254.

Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., and Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat. Biotechnol. *17*, 1–6.

Hafner, C., López-Knowles, E., Luis, N.M., Toll, A., Baselga, E., Fernández-Casado, A., Hernández, S., Ribé, A., Mentzel, T., Stoehr, R., et al. (2007). Oncogenic PIK3CA mutations occur in epidermal nevi and seborrheic keratoses with a characteristic mutation pattern. Proc. Natl. Acad. Sci. *104*, 13450–13454.

Hafner, C., Toll, A., Fernandez-Casado, A., Earl, J., Marques, M., Acquadro, F., Mendez-Pertuz, M., Urioste, M., Malats, N., Burns, J.E., et al. (2010). Multiple oncogenic mutations and clonal relationship in spatially distinct benign human epidermal tumors. Proc. Natl. Acad. Sci. *107*, 20780–20785.

Haines, E., Chen, T., Kommajosyula, N., Chen, Z., Sprie, G.S.H.-, Cornell, L., Wong, K.-K., and Shapiro, G.I. (2018). Palbociclib resistance confers dependence on an FGFR-MAP kinase-mTOR-driven pathway in KRAS-mutant non-small cell lung cancer. Oncotarget *9*, 31572–31589.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Haradhvala, N.J., Polak, P., Stojanov, P., Covington, K.R., Shinbrot, E., Hess, J.M., Rheinbay, E., Kim, J., Maruvka, Y.E., Braunstein, L.Z., et al. (2016). Mutational Strand Asymmetries in Cancer Genomes Reveal Mechanisms of DNA Damage and Repair. Cell *164*, 538–549.

Harper, J.W., and Bennett, E.J. (2016). Proteome complexity and the forces that drive proteome imbalance. Nature *537*, 328–338.

Harris, K., and Pritchard, J.K. (2017). Rapid evolution of the human mutation spectrum. Elife *6*, 1–17.

Harris, B.H.L., Barberis, A., West, C.M.L., and Buffa, F.M. (2015). Gene Expression Signatures as Biomarkers of Tumour Hypoxia. Clin. Oncol. *27*, 547–560.

Hart, K.C., Robertson, S.C., and Donoghue, D.J. (2001). Identification of tyrosine residues in constitutively activated fibroblast growth factor receptor 3 involved in mitogenesis, Stat activation, and phosphatidylinositol 3-kinase activation. Mol. Biol. Cell *12*, 931–942.

Hedegaard, J., Lamy, P., Nordentoft, I., Algaba, F., Høyer, S., Ulhøi, B.P., Vang, S., Reinert,

T., Hermann, G.G., Mogensen, K., et al. (2016). Comprehensive Transcriptional Analysis of Early-Stage Urothelial Carcinoma. Cancer Cell *30*, 27–42.

Helleday, T., Eshtad, S., and Nik-Zainal, S. (2014). Mechanisms underlying mutational signatures in human cancers. Nat. Rev. Genet. *15*, 585–598.

Henderson, S., Chakravarthy, A., Su, X., Boshoff, C., and Fenton, T.R. (2014). APOBEC-Mediated Cytosine Deamination Links PIK3CA Helical Domain Mutations to Human Papillomavirus-Driven Tumor Development. Cell Rep. *7*, 1833–1841.

Hernández, S., López-Knowles, E., Lloreta, J., Kogevinas, M., Amorós, A., Tardón, A., Carrato, A., Serra, C., Malats, N., and Real, F.X. (2006). Prospective study of FGFR3 mutations as a prognostic factor in nonmuscle invasive urothelial bladder carcinomas. J. Clin. Oncol. *24*, 3664–3671.

Herrera-Abreu, M.T., Pearson, A., Campbell, J., Shnyder, S.D., Knowles, M.A., Ashworth, A., and Turner, N.C. (2013). Parallel RNA interference screens identify EGFR activation as an escape mechanism in FGFR3-mutant cancer. Cancer Discov. *3*, 1058–1071.

Hodgkinson, A., and Eyre-Walker, A. (2011). Variation in the mutation rate across mammalian genomes. Nat. Rev. Genet. *12*, 756–766.

Holmang, S., Andius, P., Hedelin, H., Wester, K., Busch, C., and Johansson, S.L. (2004). Stage Progression in Ta Papillary Urothelial Tumors: Relationship To Grade, Immunohistochemical Expression of Tumor Markers, Mitotic Frequency and Dna Ploidy. J. Urol. *165*, 1124–1130.

Holtz, C.M., Sadler, H.A., and Mansky, L.M. (2013). APOBEC3G cytosine deamination hotspots are defined by both sequence context and single-stranded DNA secondary structure. Nucleic Acids Res. *41*, 6139–6148.

Horn, H., Schoof, E.M., Kim, J., Robin, X., Miller, M.L., Diella, F., Palma, A., Cesareni, G., Jensen, L.J., and Linding, R. (2014). KinomeXplorer: an integrated platform for kinome biology studies. Nat. Methods *11*, 603–604.

Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B.S., Mesquita, A., Liu, C., Yuan, C.L., et al. (2018). Recognition of RNA N 6 -methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat. Cell Biol. *20*, 285–295.

Hubbi, M.E., Gilkes, D.M., Hu, H., Kshitiz, Ahmed, I., and Semenza, G.L. (2014). Cyclin-dependent kinases regulate lysosomal degradation of hypoxia-inducible factor 1 to promote cell-cycle progression. Proc. Natl. Acad. Sci. *111*, E3325–E3334.

Hunter, T., and Sefton, B.M. (1980). Transforming gene product of Rous sarcoma virus phosphorylates tyrosine (phosphotyrosine/protein kinase/src gene/phosphoproteins). Proc. Natl. Acad. Sci. *77*, 1311–1315.

Hurst, C.D., Alder, O., Platt, F.M., Droop, A., Stead, L.F., Burns, J.E., Burghel, G.J., Jain, S., Klimczak, L.J., Lindsay, H., et al. (2017). Genomic Subtypes of Non-invasive Bladder Cancer with Distinct Metabolic Profile and Female Gender Bias in KDM6A Mutation Frequency. Cancer Cell *32*, 701-715.e7.

Hwang, J., and Pallas, D.C. (2014). STRIPAK complexes: Structure, biological function, and involvement in human diseases. Int. J. Biochem. Cell Biol. *47*, 118–148.

Iioka, H., Saito, K., Sakaguchi, M., Tachibana, T., Homma, K., and Kondo, E. (2019). Crumbs3 is a critical factor that regulates invasion and metastasis of colon adenocarcinoma via the specific interaction with FGFR1. Int. J. Cancer 1–15.

Iyer, G., and Milowsky, M.I. (2013). Fibroblast growth factor receptor-3 in urothelial

tumorigenesis. Urol. Oncol. Semin. Orig. Investig. *31*, 303–311.

Jäger, W., Moskalev, I., Janssen, C., Hayashi, T., Awrey, S., Gust, K.M., So, A.I., Zhang, K., Fazli, L., Li, E., et al. (2013). Ultrasound-Guided Intramural Inoculation of Orthotopic Bladder Cancer Xenografts: A Novel High-Precision Approach. PLoS One *8*, 1–11.

Jansen, R.-P., Niessing, D., Baumann, S., and Feldbrügge, M. (2014). mRNA transport meets membrane traffic. Trends Genet. *30*, 408–417.

Jean, S., Mikryukov, A., Tremblay, M.G., Baril, J., Guillou, F., Bellenfant, S., and Moss, T. (2010). Extended-synaptotagmin-2 Mediates FGF receptor endocytosis and ERK activation in vivo. Dev. Cell *19*, 426–439.

Jensen, O.N. (2004). Modification-specific proteomics: Characterization of post-translational modifications by mass spectrometry. Curr. Opin. Chem. Biol. *8*, 33–41.

Jhamb, M., Lin, J., Ballow, R., Kamat, A.M., Grossman, H.B., and Wu, X. (2007). Urinary tract diseases and bladder cancer risk: A case-control study. Cancer Causes Control *18*, 839–845.

Jia, G., Fu, Y., and He, C. (2013). Reversible RNA adenosine methylation in biological regulation. Trends Genet. *29*, 108–115.

Jiang, X., Castelao, J.E., Groshen, S., Cortessis, V.K., Shibata, D., Conti, D. V., Yuan, J.M., Pike, M.C., and Gago-Dominguez, M. (2009). Urinary tract infections and reduced risk of bladder cancer in Los Angeles. Br. J. Cancer *100*, 834–839.

John, B.A., and Said, N. (2017). Insights from animal models of bladder cancer: recent advances, challenges, and opportunities. Oncotarget *8*, 57766–57781.

Johnson, S.B., and Yu, J.B. (2018). Bladder Preserving Trimodality Therapy for Muscle-Invasive Bladder Cancer. Curr. Oncol. Rep. *20*, 66.

Jung, J. (2014). Human tumor xenograft models for preclinical assessment of anticancer drug development. Toxicol. Res. *30*, 1–5.

Kang, J., Sergio, C.M., Sutherland, R.L., and Musgrove, E.A. (2014). Targeting cyclin-dependent kinase 1 (CDK1) but not CDK4/6 or CDK2 is selectively lethal to MYC-dependent human breast cancer cells. BMC Cancer *14*, 1–13.

Kardos, J., Chai, S., Mose, L.E., Selitsky, S.R., Krishnan, B., Saito, R., Iglesia, M.D., Milowsky, M.I., Parker, J.S., Kim, W.Y., et al. (2016). Claudin-low bladder tumors are immune infiltrated and actively immune suppressed. JCI Insight *1*, e85902.

Kates, M., Nirschl, T., Sopko, N.A., Matsui, H., Kochel, C.M., Reis, L.O., Netto, G.J., Hoque, M.O., Hahn, N.M., McConkey, D.J., et al. (2017). Intravesical BCG Induces CD4 + T-Cell Expansion in an Immune Competent Model of Bladder Cancer . Cancer Immunol. Res. *5*, 594–603.

Kaur, A., and Westermarck, J. (2016). Regulation of protein phosphatase 2A (PP2A) tumor suppressor function by PME-1. Biochem. Soc. Trans. *44*, 1683–1693.

Kiely, M., and Kiely, P.A. (2015). PP2A: The wolf in sheep's clothing? Cancers (Basel). *7*, 648–669.

Kim, J. -w., Gao, P., Liu, Y.-C., Semenza, G.L., and Dang, C. V. (2007). Hypoxia-Inducible Factor 1 and Dysregulated c-Myc Cooperatively Induce Vascular Endothelial Growth Factor and Metabolic Switches Hexokinase 2 and Pyruvate Dehydrogenase Kinase 1. Mol. Cell. Biol. *27*, 7381–7393.

Kim, K., Saneyoshi, T., Hosokawa, T., Okamoto, K., and Hayashi, Y. (2016). Interplay of enzymatic and structural functions of CaMKII in long-term potentiation. J. Neurochem. *139*, 959–972.

Knight, A., Askling, J., Granath, F., Sparen, P., and Ekbom, A. (2004). Urinary bladder cancer in Wegener's granulomatosis: Risks and relation to cyclophosphamide. Ann. Rheum. Dis. *63*, 1307–1311.

Knowles, W.A. (2006). Discovery and Epidemiology of the Human Polyomaviruses BK Virus (BKV) and JC Virus (JCV). Adv. Exp. Med. Biol. *577*, 19–45.

Knowles, M.A., and Hurst, C.D. (2015). Molecular biology of bladder cancer: New insights into pathogenesis and clinical diversity. Nat. Rev. Cancer *15*, 25–41.

Kobayashi, T., Owczarek, T.B., McKiernan, J.M., and Abate-Shen, C. (2015). Modelling bladder cancer in mice: Opportunities and challenges. Nat. Rev. Cancer *15*, 42–54.

Kong, M., Wang, C.S., and Donoghue, D.J. (2002). Interaction of fibroblast growth factor receptor 3 and the adapter protein SH2-B. A role in Stat5 activation. J. Biol. Chem. *277*, 15962–15970.

Koshiji, M., Kageyama, Y., Pete, E.A., Horikawa, I., Barrett, J.C., and Huang, L.E. (2004). HIF-1α induces cell cycle arrest by functionally counteracting Myc. EMBO J. *23*, 1949–1956.

Kostas, M., Haugsten, E.M., Zhen, Y., Sorensen, V., Szybowska, P., Fiorito, E., Lorenz, S., de Souza, G.A., Wiedlocha, A., and Wesche, J. (2018). Protein Tyrosine Phosphatase Receptor Type G (PTPRG) Controls Fibroblast Growth Factor Receptor (FGFR) 1 Activity and Influences Sensitivity to FGFR Kinase Inhibitors. Mol. Cell. Proteomics *17*, 850–870.

Kovalenko, D., Yang, X., Nadeau, R.J., Harkins, L.K., and Friesel, R. (2003). Sef inhibits fibroblast growth factor signaling by inhibiting FGFR1 tyrosine phosphorylation and subsequent ERK activation. J. Biol. Chem. *278*, 14087–14091.

Kreft, M.E., Sterle, M., Veranič, P., and Jezernik, K. (2005). Urothelial injuries and the early wound healing response: Tight junctions and urothelial cytodifferentiation. Histochem. Cell Biol. *123*, 529–539.

Krejci, P., Salazar, L., Goodridge, H.S., Kashiwada, T.A., Schibler, M.J., Jelinkova, P., Thompson, L.M., and Wilcox, W.R. (2008). STAT1 and STAT3 do not participate in FGF-mediated growth arrest in chondrocytes. J. Cell Sci. *121*, 272–281.

Kuleshov, M. V, Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. *44*, W90-7.

Kunova Bosakova, M., Nita, A., Gregor, T., Varecha, M., Gudernova, I., Fafilek, B., Barta, T., Basheer, N., Abraham, S.P., Balek, L., et al. (2019). Fibroblast growth factor receptor influences primary cilium length through an interaction with intestinal cell kinase. Proc. Natl. Acad. Sci. *116*, 4316–4325.

Kvach, M. V., Barzak, F.M., Harjes, S., Schares, H.A.M., Jameson, G.B., Ayoub, A.M., Moorthy, R., Aihara, H., Harris, R.S., Filichev, V. V., et al. (2019). Inhibiting APOBEC3 Activity with Single-Stranded DNA Containing 2′-Deoxyzebularine Analogues. Biochemistry *58*, 391–400.

L'Hôte, C.G.M., and Knowles, M.A. (2005). Cell responses to FGFR3 signalling: Growth, differentiation and apoptosis. Exp. Cell Res. *304*, 417–431.

Lachmann, A., and Ma'ayan, A. (2009). KEA: Kinase enrichment analysis. Bioinformatics *25*,

684–686.

Lackey, L., Law, E.K., Brown, W.L., and Harris, R.S. (2013). Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. Cell Cycle *12*, 762–772.

Lamy, P., Nordentoft, I., Birkenkamp-Demtroder, K., Thomsen, M.B.H., Villesen, P., Vang, S., Hedegaard, J., Borre, M., Jensen, J.B., Høyer, S., et al. (2016). Paired Exome Analysis Reveals Clonal Evolution and Potential Therapeutic Targets in Urothelial Carcinoma. Cancer Res. *76*, 5894–5906.

Lao, D.H., Yusoff, P., Chandramouli, S., Philp, R.J., Chee, W.F., Jackson, R.A., Tzuen, Y.S., Chye, Y.Y., and Guy, G.R. (2007). Direct binding of PP2A to Sprouty2 and phosphorylation changes are a prerequisite for ERK inhibition downstream of fibroblast growth factor receptor stimulation. J. Biol. Chem. *282*, 9117–9126.

Larsen, M.R., Thingholm, T.E., Jensen, O.N., Roepstorff, P., and Jørgensen, T.J.D. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol. Cell. Proteomics *4*, 873–886.

Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G. V, Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature *499*, 214–218.

Le, D.T., Durham, J.N., Smith, K.N., Wang, H., Bartlett, B.R., Aulakh, L.K., Lu, S., Kemberling, H., Wilt, C., Luber, B.S., et al. (2017). Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science (80-.). *357*, 409–413.

Ledda, F., and Paratcha, G. (2007). Negative Regulation of Receptor Tyrosine Kinase (RTK) Signaling: A Developing Field. Biomark. Insights *2*, 45–58.

Lee, H.J., Zhuang, G., Cao, Y., Du, P., Kim, H.J., and Settleman, J. (2014). Drug resistance via feedback activation of stat3 in oncogene-addicted cancer cells. Cancer Cell *26*, 207–221.

Lee, S.H., Hu, W., Matulay, J.T., Silva, M. V., Owczarek, T.B., Kim, K., Chua, C.W., Barlow, L.M.J., Kandoth, C., Williams, A.B., et al. (2018). Tumor Evolution and Drug Response in Patient-Derived Organoid Models of Bladder Cancer. Cell *173*, 515-528.e17.

Lee, W., Jiang, Z., Liu, J., Haverty, P.M., Guan, Y., Stinson, J., Yue, P., Zhang, Y., Pant, K.P., Bhatt, D., et al. (2010). The mutation spectrum revealed by paired genome sequences from a lung cancer patient. Nature *465*, 473–477.

Lemmon, M.A., and Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. Cell *141*, 1117–1134.

Letouzé, E., Shinde, J., Renault, V., Couchy, G., Blanc, J.F., Tubacher, E., Bayard, Q., Bacq, D., Meyer, V., Semhoun, J., et al. (2017). Mutational signatures reveal the dynamic interplay of risk factors and cellular processes during liver tumorigenesis. Nat. Commun. *8*, 1315.

Lew, E.D., Furdui, C.M., Anderson, K.S., and Schlessinger, J. (2009). The precise sequence of FGF receptor autophosphorylation is kinetically driven and is disrupted by oncogenic mutations. Sci. Signal. *2*, 1–11.

Lewis, S.A. (2000). Everything you wanted to know about the bladder epithelium but were afraid to ask. Am J Physiol Ren. Physiol *278*, F867-74.

Lewis, J.D., Ferrara, A., Peng, T., Hedderson, M., Bilker, W.B., Quesenberry, C.P., Vaughn, D.J., Nessel, L., Selby, J., and Strom, B.L. (2011). Risk of bladder cancer among diabetic patients treated with pioglitazone: Interim report of a longitudinal cohort study. Diabetes Care

34, 916–922.

Li, M., Shandilya, S.M.D., Carpenter, M.A., Rathore, A., Brown, W.L., Perkins, A.L., Harki, D.A., Solberg, J., Hook, D.J., Pandey, K.K., et al. (2012a). First-in-class small molecule inhibitors of the single-strand DNA cytosine deaminase APOBEC3G. ACS Chem. Biol. *7*, 506–517.

Li, T., Jiang, H.L., Tong, Y.G., and Lu, J.J. (2018). Targeting the Hsp90-Cdc37-client protein interaction to disrupt Hsp90 chaperone machinery. J. Hematol. Oncol. *11*, 1–10.

Li, Z., Adams, R.M., Chourey, K., Hurst, G.B., Hettich, R.L., and Pan, C. (2012b). Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ orbitrap velos. J. Proteome Res. *11*, 1582–1590.

Lim, B., Mun, J., and Kim, S.Y. (2017). Intrinsic Molecular Processes: Impact on Mutagenesis. Trends in Cancer *3*, 357–371.

Liu, X., Jing, X., Cheng, X., Ma, D., Jin, Z., Yang, W., and Qiu, W. (2016). FGFR3 promotes angiogenesis-dependent metastasis of hepatocellular carcinoma via facilitating MCP-1-mediated vascular formation. Med. Oncol. *33*, 46.

Logié, A., Dunois-Lardé, C., Rosty, C., Levrel, O., Blanche, M., Ribeiro, A., Gasc, J.M., Jorcano, J., Werner, S., Satre-Garau, X., et al. (2005). Activating mutations of the tyrosine kinase receptor FGFR3 are associated with benign skin tumors in mice and humans. Hum. Mol. Genet. *14*, 1153–1160.

Lombardi, B., Ashford, P., Moya-Garcia, A.A., Rust, A., Crawford, M., Williams, S. V., Knowles, M.A., Katan, M., Orengo, C., and Godovac-Zimmermann, J. (2017). Unique signalling connectivity of FGFR3-TACC3 oncoprotein revealed by quantitative phosphoproteomics and differential network analysis. Oncotarget *8*, 102898–102911.

Ma, H., Groth, R.D., Cohen, S.M., Emery, J.F., Li, B., Hoedt, E., Zhang, G., Neubert, T.A., and Tsien, R.W. (2014). γcaMKII shuttles Ca 2+ /CaM to the nucleus to trigger CREB phosphorylation and gene expression. Cell *159*, 281–294.

Mahe, M., Dufour, F., Neyret-Kahn, H., Moreno-Vega, A., Beraud, C., Shi, M., Hamaidi, I., Sanchez-Quiles, V., Krucker, C., Dorland-Galliot, M., et al. (2018). An FGFR3/MYC positive feedback loop provides new opportunities for targeted therapies in bladder cancers. EMBO Mol. Med. *10*, pii: e8163.

Makova, K.D., and Hardison, R.C. (2015). The effects of chromatin organization on variation in mutation rates in the genome. Nat. Rev. Genet. *16*, 213–223.

Di Martino, E., Tomlinson, D.C., and Knowles, M.A. (2012). A decade of FGF receptor research in bladder cancer: Past, present, and future challenges. Adv. Urol. *2012*, 429213.

McEwen, D.G., Green, R.P., Naski, M.C., Towler, D.A., and Ornitz, D.M. (1999). Fibroblast Growth Factor Receptor 3 Gene Transcription Is Suppressed by Cyclic Adenosine 3′,5′-Monophosphate. J. Biol. Chem. *274*, 30934–30942.

Melão, A., Spit, M., Cardoso, B.A., and Barata, J.T. (2016). Optimal interleukin-7 receptor-mediated signaling, cell cycle progression and viability of t-cell acute lymphoblastic leukemia cells rely on casein kinase 2 activity. Haematologica *101*, 1368–1379.

Mering, C. von, Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S., and Bork, P. (2002). Comparative assessment of large-scale data sets of protein–protein interactions. Nature *417*, 399–403.

Meyer, K.D., and Jaffrey, S.R. (2014). The dynamic epitranscriptome: N6-methyladenosine

and gene expression control. Nat. Rev. Mol. Cell Biol. *15*, 313–326.

Miaczynska, M. (2013). Effects of membrane trafficking on signaling by receptor tyrosine kinases. Cold Spring Harb. Perspect. Biol. *5*, a009035.

Miao, D., Margolis, C.A., Vokes, N.I., Liu, D., Taylor-Weiner, A., Wankowicz, S.M., Adeegbe, D., Keliher, D., Schilling, B., Tracy, A., et al. (2018). Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. Nat. Genet. *50*, 1271–1281.

Middlebrooks, C.D., Banday, A.R., Matsuda, K., Udquim, K.I., Onabajo, O.O., Paquin, A., Figueroa, J.D., Zhu, B., Koutros, S., Kubo, M., et al. (2016). Association of germline variants in the APOBEC3 region with cancer risk and enrichment with APOBEC-signature mutations in tumors. Nat. Genet. *48*, 1330–1338.

Miyake, M., Ishii, M., Koyama, N., Kawashima, K., Kodama, T., Anai, S., Fujimoto, K., Hirao, Y., and Sugano, K. (2009). 1-tert-Butyl-3-[6-(3,5-dimethoxy-phenyl)-2-(4-diethylamino-butylamino)-pyrido[2,3-d]pyrimidin-7-yl]-urea (PD173074), a Selective Tyrosine Kinase Inhibitor of Fibroblast Growth Factor Receptor-3 (FGFR3), Inhibits Cell Proliferation of Bladder Cancer Carryi. J. Pharmacol. Exp. Ther. *332*, 795–802.

Miyazaki, J., and Nishiyama, H. (2017). Epidemiology of urothelial carcinoma. Int. J. Urol. *24*, 730–734.

Mohammadi, M., Honegger, A.M., Rotin, D., Fischer, R., Bellot, F., Li, W., Dionne, C.A., Jaye, M., Rubinstein, M., and Schlessinger, J. (1991). A tyrosine-phosphorylated carboxy-terminal peptide of the fibroblast growth factor receptor (Flg) is a binding site for the SH2 domain of phospholipase C-gamma 1. Mol. Cell. Biol. *11*, 5068–5078.

Mohammadi, M., Froum, S., Hamby, J.M., Schroeder, M.C., Panek, R.L., Lu, G.H., Eiiseenkova, A. V., Green, D., Schlessinger, J., and Hubbard, S.R. (1998). Crystal structure of an angiogenesis inhibitor bound to the FGF receptor tyrosine kinase domain. EMBO J. *17*, 5896–5904.

Moreno, C.S., Park, S., Nelson, K., Ashby, D.G., Hubalek, F., Lane, W.S., and Pallas, D.C. (2000). The WD40 Repeat Proteins Striatin and SG2NA are Members of a Novel Family of Calmodulin-Binding Proteins that Associate with PP2A. J. Biol. Chem. *275*, 5257–5263.

Mostafa, M.H., Sheweita, S.A., and O'Connor, P.J. (1999). Relationship between schistosomiasis and bladder cancer. Clin. Microbiol. Rev. *12*, 97–111.

Mostofi, F.K., DavisJr, C.J., Sesterhenn, I.A., and Sobin, L.H. (1974). World Health Organization: Histological Typing of Urinary Bladder Tumours. Int. Histol. Classif. Tumours.

Mullenders, J., de Jongh, E., Brousali, A., Roosen, M., Blom, J.P.A., Begthel, H., Korving, J., Jonges, T., Kranenburg, O., Meijer, R., et al. (2019). Mouse and human urothelial cancer organoids: A tool for bladder cancer research. Proc. Natl. Acad. Sci. pii: 201803595.

Nabel, C.S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H.L., Stivers, J.T., Zhang, Y., and Kohli, R.M. (2012). AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. Nat. Chem. Biol. *8*, 751–758.

Nadratowska-Wesolowska, B., Haugsten, E.M., Zakrzewska, M., Jakimowicz, P., Zhen, Y., Pajdzik, D., Wesche, J., and Wiedlochad, A. (2014). RSK2 regulates endocytosis of FGF receptor 1 by phosphorylation on serine 789. Oncogene *33*, 4823–4836.

Nakanishi, Y., Akiyama, N., Tsukaguchi, T., Fujii, T., Satoh, Y., Mizuno, H., Ishii, N., and Aoki, M. (2015a). Mechanism of oncogenic signal activation by the novel fusion kinase FGFR3-BAIAP2L1. Cancer Res. *75*, 123–123.

Nakanishi, Y., Mizuno, H., Sase, H., Fujii, T., Sakata, K., Akiyama, N., Aoki, Y., Aoki, M., and Ishii, N. (2015b). ERK Signal Suppression and Sensitivity to CH5183284/Debio 1347, a Selective FGFR Inhibitor. Mol. Cancer Ther. *14*, 2831–2839.

Naski, M.C., Wang, Q., Xu, J., and Ornitz, D.M. (1996). Graded activation of fibroblast growth factor receptor 3 by mutations causing achondroplasia and thanatophoric dysplasia. Nat. Genet. *13*, 233–237.

Neben, C.L., Lo, M., Jura, N., and Klein, O.D. (2019). Feedback regulation of RTK signaling in development. Dev. Biol. *447*, 71–89.

Nelson, K.N., Meyer, A.N., Siari, A., Campos, A.R., Motamedchaboki, K., and Donoghue, D.J. (2016). Oncogenic Gene Fusion FGFR3-TACC3 Is Regulated by Tyrosine Phosphorylation. Mol. Cancer Res. *14*, 458–469.

Ng, J.C.F., Quist, J., Grigoriadis, A., Malim, M.H., and Fraternali, F. (2019). Pan-cancer transcriptomic analysis dissects immune and proliferative functions of APOBEC3 cytidine deaminases. Nucleic Acids Res. *47*, 1178–1194.

Nik-Zainal, S., Alexandrov, L.B., Wedge, D.C., Van Loo, P., Greenman, C.D., Raine, K., Jones, D., Hinton, J., Marshall, J., Stebbings, L.A., et al. (2012). Mutational processes molding the genomes of 21 breast cancers. Cell *149*, 979–993.

Nik-Zainal, S., Wedge, D.C., Alexandrov, L.B., Petljak, M., Butler, A.P., Bolli, N., Davies, H.R., Knappskog, S., Martin, S., Papaemmanuil, E., et al. (2014). Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. Nat. Genet. *46*, 487–491.

Nik-Zainal, S., Davies, H., Staaf, J., Ramakrishna, M., Glodzik, D., Zou, X., Martincorena, I., Alexandrov, L.B., Martin, S., Wedge, D.C., et al. (2016). Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature *534*, 47–54.

Nordentoft, I., Lamy, P., Birkenkamp-Demtröder, K., Shumansky, K., Vang, S., Hornshøj, H., Juul, M., Villesen, P., Hedegaard, J., Roth, A., et al. (2014). Mutational context and diverse clonal development in early and late bladder cancer. Cell Rep. *7*, 1649–1663.

Van Oers, J.M.M., Lurkin, I., Van Exsel, A.J.A., Nijsen, Y., Van Rhijn, B.W.G., Van Der Aa, M.N.M., and Zwarthoff, E.C. (2005). A simple and fast method for the simultaneous detection of nine fibroblast growth factor receptor 3 mutations in bladder cancer and voided urine. Clin. Cancer Res. *11*, 7743–7748.

Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. Cell *127*, 635–648.

Olsen, S.K., Ibrahimi, O.A., Raucci, A., Zhang, F., Eliseenkova, A. V., Yayon, A., Basilico, C., Linhardt, R.J., Schlessinger, J., and Mohammadi, M. (2004). Insights into the molecular basis for fibroblast growth factor receptor autoinhibition and ligand-binding promiscuity. Proc. Natl. Acad. Sci. *101*, 935–940.

Ong, S., and Mann, M. (2005). Mass spectrometry – based proteomics turns quantitative. Nat. Chem. Biol. *1*, 252–262.

Ong, S., and Mann, M. (2006). A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat. Protoc. *1*, 2650–2660.

Ong, S.H., Guy, G.R., Hadari, Y.R., Laks, S., Gotoh, N., Schlessinger, J., and Lax, I. (2000). FRS2 Proteins Recruit Intracellular Signaling Pathways by Binding to Diverse Targets on

Fibroblast Growth Factor and Nerve Growth Factor Receptors. Mol. Cell. Biol. *20*, 979–989.

Ong, S.H., Hadari, Y.R., Gotoh, N., Guy, G.R., Schlessinger, J., and Lax, I. (2001). Stimulation of phosphatidylinositol 3-kinase by fibroblast growth factor receptors is mediated by coordinated recruitment of multiple docking proteins. Proc. Natl. Acad. Sci. *98*, 6074–6079.

Ornitz, D.M., and Itoh, N. (2015). The fibroblast growth factor signaling pathway. Wiley Interdiscip Rev Dev Biol *4*, 215–266.

Ornitz, D.M., and Marie, P.J. (2015). Fibroblast growth factor signaling in skeletal development and disease. Genes Dev. *29*, 1463–1486.

Ould Amer, Y., and Hebert-Chatelain, E. (2018). Mitochondrial cAMP-PKA signaling: What do we really know? Biochim. Biophys. Acta - Bioenerg. *1859*, 868–877.

Pan, C.X., Zhang, H., Tepper, C.G., Lin, T.Y., Davis, R.R., Keck, J., Ghosh, P.M., Gill, P., Airhart, S., Bult, C., et al. (2015). Development and characterization of bladder cancer patient-derived xenografts for molecularly guided targeted therapy. PLoS One *10*, 1–22.

Pardo, O.E., Latigo, J., Jeffery, R.E., Nye, E., Poulsom, R., Spencer-Dene, B., Lemoine, N.R., Stamp, G.W., Aboagye, E.O., and Seckl, M.J. (2009). The fibroblast growth factor receptor inhibitor PD173074 blocks small cell lung cancer growth in vitro and in vivo. Cancer Res. *69*, 8645–8651.

Passos-Bueno, M.R., Sertié, A.L., Jehee, F.S., Fanganiello, R., and Yeh, E. (2008). Genetics of craniosynostosis: Genes, syndromes, mutations and genotype-phenotype correlations. Front. Oral Biol. *12*, 107–143.

Pawson, T. (2004). Specificity in Signal Transduction: From Phosphotyrosine-SH2 Domain Interactions to Complex Cellular Systems. Cell *116*, 191–203.

Pearson, A., Smyth, E., Babina, I.S., Herrera-Abreu, M.T., Tarazona, N., Peckitt, C., Kilgour, E., Smith, N.R., Geh, C., Rooney, C., et al. (2016). High-level clonal FGFR amplification and response to FGFR inhibition in a translational clinical trial. Cancer Discov. *6*, 838–851.

Petersen-Mahrt, S.K., Harris, R.S., and Neuberger, M.S. (2002). AID mutates E . coli suggesting a DNA deamination mechanism for antibody diversification. Nature *418*, 99–103.

Petrone, A., Adamo, M.E., Cheng, C., and Kettenbach, A.N. (2016). Identification of Candidate Cyclin-dependent kinase 1 (Cdk1) Substrates in Mitosis by Quantitative Phosphoproteomics. Mol. Cell. Proteomics *15*, 2448–2461.

Pfeifer, G.P., You, Y.H., and Besaratinia, A. (2005). Mutations induced by ultraviolet light. Mutat. Res. - Fundam. Mol. Mech. Mutagen. *571*, 19–31.

Del Piccolo, N., Sarabipour, S., and Hristova, K. (2017). A new method to study heterodimerization of membrane proteins and its application to fibroblast growth factor receptors. J. Biol. Chem. *292*, 1188–1301.

Poon, L.S., Huang, N.M., Choo, Y., McPherson, J.R., Yu, W., Heng, L.H., Gan, A., Myint, S.S., Siew, Y.E., Ler, D.L., et al. (2015). Mutation signatures implicate aristolochic acid in bladder cancer development. Genome Med. *7*, 38.

Porębska, N., Latko, M., Kucińska, M., Zakrzewska, M., Otlewski, J., and Opaliński, Ł. (2018). Targeting Cellular Trafficking of Fibroblast Growth Factor Receptors as a Strategy for Selective Cancer Treatment. J. Clin. Med. *8*, 7.

Poulos, R.C., Wong, Y.T., Ryan, R., Pang, H., and Wong, J.W.H. (2018). Analysis of 7,815 cancer exomes reveals associations between mutational processes and somatic driver mutations.

PLoS Genet. *14*, 1–20.

Prado, J., Monezi, T., Amorim, A., Lino, V., Paladino, A., and Boccardo, E. (2018). Human polyomaviruses and cancer: an overview. Clinics *73*.

Raynaud, F., Mina, M., Tavernari, D., and Ciriello, G. (2018). Pan-cancer inference of intra-tumor heterogeneity reveals associations with different forms of genomic instability. PLoS Genet. *14*, 1–18.

Rebouissou, S., Hérault, A., Letouzé, E., Neuzillet, Y., Laplanche, A., Ofualuka, K., Maillé, P., Leroy, K., Riou, A., Lepage, M.L., et al. (2012). CDKN2A homozygous deletion is associated with muscle invasion in FGFR3-mutated urothelial bladder carcinoma. J. Pathol. *227*, 315–324.

Rebouissou, S., Bernard-Pierrot, I., Reyniès, A. de, Lepage, M.-L., Krucker, C., Chapeaublanc, E., Hérault, A., Kamoun, A., Caillault, A., Letouzé, E., et al. (2014). EGFR as a potential therapeutic target for a subset of muscle-invasive bladder cancers presenting a basal-like phenotype. Sci. Transl. Med. *6*, 244ra91-244ra91.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

Roberts, S.A., Lawrence, M.S., Klimczak, L.J., Grimm, S.A., Fargo, D., Stojanov, P., Kiezun, A., Kryukov, G. V., Carter, S.L., Saksena, G., et al. (2013). An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat. Genet. *45*, 970–976.

Robertson, A.G., Kim, J., Al-Ahmadie, H., Bellmunt, J., Guo, G., Cherniack, A.D., Hinoue, T., Laird, P.W., Hoadley, K.A., Akbani, R., et al. (2017). Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell *171*, 540-556.e25.

Rodin, S.N., and Rodin, A.S. (2000). Human lung cancer and p53: The interplay between mutagenesis and selection. Proc. Natl. Acad. Sci. *97*, 12244–12249.

Roper, N., Gao, S., Maity, T.K., Banday, A.R., Zhang, X., Venugopalan, A., Cultraro, C.M., Patidar, R., Sindiri, S., Brown, A.L., et al. (2019). APOBEC Mutagenesis and Copy-Number Alterations Are Drivers of Proteogenomic Tumor Evolution and Heterogeneity in Metastatic Thoracic Tumors. Cell Rep. *26*, 2651-2666.e6.

Rosenberg, J.E., Hoffman-Censits, J., Powles, T., Van Der Heijden, M.S., Balar, A. V., Necchi, A., Dawson, N., O'Donnell, P.H., Balmanoukian, A., Loriot, Y., et al. (2016). Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: A single-arm, multicentre, phase 2 trial. Lancet *387*, 1909–1920.

Ross, P.L., Huang, Y.N., Marchese, J.N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., et al. (2004). Multiplexed Protein Quantitation in Saccharomyces cerevisiae Using Amine-reactive Isobaric Tagging Reagents. Mol. Cell. Proteomics *3*, 1154–1169.

Rubio, C., Martínez-Fernandez, M., Segovia, C., Lodewijk, I., Suarez-Cabrera, C., Segrelles, C., Lopez-Calder, F., Munera-Maravilla, E., Santos, M., Bernardini, A., et al. (2019). CDK4/6 inhibitor as a novel therapeutic approach for advanced bladder cancer independently of RB1 status. Clin. Cancer Res. *25*, 390–402.

Le Sage, C., Nagel, R., and Agami, R. (2007). Diverse ways to control p27Kip1 function: miRNAs come into play. Cell Cycle *6*, 2742–2749.

Saint-Jacques, N., Parker, L., Brown, P., and Dummer, T.J. (2014). Arsenic in drinking water and urinary tract cancers: A systematic review of 30 years of epidemiological evidence. Environ. Heal. *13*, 44.

Saito, R., Smith, C.C., Utsumi, T., Bixby, L.M., Kardos, J., Wobker, S.E., Stewart, K.G., Chai, S., Manocha, U., Byrd, K.M., et al. (2018). Molecular subtype-specific immunocompetent models of high-grade urothelial carcinoma reveal differential neoantigen expression and response to immunotherapy. Cancer Res. *78*, 3954–3968.

Salazar, L., Kashiwada, T., Krejci, P., Muchowski, P., Donoghue, D., Wilcox, W.R., and Thompson, L.M. (2009). A novel interaction between fibroblast growth factor receptor 3 and the p85 subunit of phosphoinositide 3-kinase: Activation-dependent regulation of ERK by p85 in multiple myeloma cells. Hum. Mol. Genet. *18*, 1951–1961.

Salter, J.D., Bennett, R.P., and Smith, H.C. (2016). The APOBEC Protein Family: United by Structure, Divergent in Function. Trends Biochem. Sci. *41*, 578–594.

Scherer, W., Syverton, J., and Gey, G. (1953). Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. J Exp Med *97*, 695–710.

Schlessinger, J., Plotnikov, A.N., Ibrahimi, O.A., Eliseenkova, A. V., Yeh, B.K., Yayon, A., Linhardt, R.J., and Mohammadi, M. (2000). Crystal Structure of a Ternary FGF-FGFR-Heparin Complex Reveals a Dual Role for Heparin in FGFR Binding and Dimerization. Mol. Cell *6*, 743–750.

Semenza, G.L. (2010). HIF-1: upstream and downstream of cancer metabolism. Curr. Opin. Genet. Dev. *20*, 51–56.

Seplyarskiy, V.B., Soldatov, R.A., Popadin, K.Y., Antonarakis, S.E., Bazykin, G.A., and Nikolaev, S.I. (2016). APOBEC-induced mutations in human cancers are strongly enriched on the lagging DNA strand during replication. Genome Res. *26*, 174–182.

Shah, O.J., Ghosh, S., and Hunter, T. (2003). Mitotic regulation of ribosomal S6 kinase 1 involves Ser/Thr, Pro phosphorylation of consensus and non-consensus sites by Cdc2. J. Biol. Chem. *278*, 16433–16442.

Sharma, S., and Baysal, B.E. (2017). Stem-loop structure preference for site-specific RNA editing by APOBEC3A and APOBEC3G. PeerJ *5*, e4136.

Sheehy, A.M., Gaddis, N.C., Choi, J.D., and Malim, M.H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature *418*, 4–8.

Shenoy, A., and Geiger, T. (2014). Super-SILAC: Current trends and future perspectives. Expert Rev. Proteomics *12*, 13–19.

Shi, M.J., Meng, X.Y., Lamy, P., Banday, A.R., Yang, J., Moreno-Vega, A., Chen, C.-L., Dyrskjøt, L., Bernard-Pierrot, I., Prokunina-Olsson, L., et al. (2019). APOBEC-mediated Mutagenesis as a Likely Cause of FGFR3 S249C Mutation Over-representation in Bladder Cancer. Eur. Urol. *76*, 9–13.

Silva, A., Yunes, J.A., Cardoso, B.A., Martins, L.R., Jotta, P.Y., Abecasis, M., Nowill, A.E., Leslie, N.R., Cardoso, A.A., and Barata, J.T. (2008). PTEN posttranslational inactivation and hyperactivation of the PI3K/Akt pathway sustain primary T cell leukemia viability. J. Clin. Invest. *118*, 3762–3774.

Silvas, T. V., Hou, S., Myint, W., Nalivaika, E., Somasundaran, M., Kelch, B.A., Matsuo, H.,

Kurt Yilmaz, N., and Schiffer, C.A. (2018). Substrate sequence selectivity of APOBEC3A implicates intra-DNA interactions. Sci. Rep. *8*, 7511.

Sims, G.P., Rowe, D.C., Rietdijk, S.T., Herbst, R., and Coyle, A.J. (2010). HMGB1 and RAGE in inflammation and cancer. Annu. Rev. Immunol. *28*, 367–388.

Sirvent, A., Vigy, O., Orsetti, B., Urbach, S., and Roche, S. (2012). Analysis of SRC Oncogenic Signaling in Colorectal Cancer by Stable Isotope Labeling with Heavy Amino Acids in Mouse Xenografts. Mol. Cell. Proteomics *11*, 1937–1950.

Smith, F.D., Samelson, B.K., and Scott, J.D. (2011). Discovery of cellular substrates for protein kinase A using a peptide array screening protocol. Biochem. J. *438*, 103–110.

Spruck, C.H., Ohneseit, P.F., Gonzalez-zulueta, M., Esrig, D., Miyao, N., Thai, Y.C., Lerner, S.P., Schmutte, C., Yang, A.S., Cote, R., et al. (1994). Two Molecular Pathways to Transitional Cell Carcinoma of the Bladder. Cancer Res. *54*, 784–788.

St-Germain, J.R., Taylor, P., Tong, J., Jin, L.L., Nikolic, A., Stewart, I.I., Ewing, R.M., Dharsee, M., Li, Z., Trudel, S., et al. (2009). Multiple myeloma phosphotyrosine proteomic profile associated with FGFR3 expression, ligand activation, and drug inhibition. Proc. Natl. Acad. Sci. *106*, 20127–20132.

St-Germain, J.R., Taylor, P., Zhang, W., Li, Z., Ketela, T., Moffat, J., Neel, B.G., Trudel, S., and Moran, M.F. (2015). Differential regulation of FGFR3 by PTPN1 and PTPN2. Proteomics *15*, 419–433.

Stamatoyannopoulos, J.A., Adzhubei, I., Thurman, R.E., Kryukov, G. V., Mirkin, S.M., and Sunyaev, S.R. (2009). Human mutation rate associated with DNA replication timing. Nat. Genet. *41*, 393–395.

Starrett, G.J., Luengas, E.M., McCann, J.L., Ebrahimi, D., Temiz, N.A., Love, R.P., Feng, Y., Adolph, M.B., Chelico, L., Law, E.K., et al. (2016). The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. Nat. Commun. *7*, 12918.

Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F.H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., et al. (2005). A human protein-protein interaction network: a resource for annotating the proteome. Cell *122*, 957–968.

Stratton, M.R., Campbell, P.J., and Futreal, P.A. (2009). The cancer genome. Nature *458*, 719–724.

Suganuma, M., Fujiki, H., Furuya-Suguri, H., Yoshizawa, S., Yasumoto, S., Kato, Y., Fusetani, N., and Sugimura, T. (1990). Calyculin A, an Inhibitor of Protein Phosphatases, a Potent Tumor Promoter on CD-I Mouse Skin. Cancer Res. *50*, 3521–3525.

Sun, L.M., Lin, C.L., Liang, J.A., Liu, S.H., Sung, F.C., Chang, Y.J., and Kao, C.H. (2013). Urinary tract infection increases subsequent urinary tract cancer risk: A population-based cohort study. Cancer Sci. *104*, 619–623.

Sutandy, F.R., Qian, J., Chen, C.-S., and Zhu, H. (2013). Overview of Protein Microarrays. Curr Protoc Protein Sci *Chapter 27*.

Suzuki, K., Sako, K., Akiyama, K., Isoda, M., Senoo, C., Nakajo, N., and Sagata, N. (2015). Identification of non-Ser/Thr-Pro consensus motifs for Cdk1 and their roles in mitotic regulation of C2H2 zinc finger proteins and Ect2. Sci. Rep. *5*, 1–9.

Swanton, C., McGranahan, N., Starrett, G.J., and Harris, R.S. (2015). APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity. Cancer Discov. *5*, 704–712.

Sweis, R.F., Spranger, S., Bao, R., Paner, G.P., Stadler, W.M., Steinberg, G., and Gajewski, T. (2016). Molecular drivers of the non-T cell-inflamed tumor microenvironment in urothelial bladder cancer. Cancer Immunol. Res. *4*, 563–568.

Taipale, M., Krykbaeva, I., Koeva, M., Kayatekin, C., Westover, K.D., Karras, G.I., and Lindquist, S. (2012). Quantitative analysis of Hsp90-client interactions reveals principles of substrate recognition. Cell *150*, 987–1001.

Tan, W.S., and Kelly, J.D. (2018). Intravesical device-assisted therapies for non-muscle-invasive bladder cancer. Nat. Rev. Urol. *15*, 667–685.

Tang, D., Kang, R., Zeh, H.J., and Lotze, M.T. (2010). High-mobility group box 1 and cancer. Biochim. Biophys. Acta *1799*, 131–140.

Tanizaki, J., Okamoto, I., Sakai, K., and Nakagawa, K. (2011). Differential roles of trans-phosphorylated EGFR, HER2, HER3, and RET as heterodimerisation partners of MET in lung cancer with MET amplification. Br. J. Cancer *105*, 807–813.

Temko, D., Tomlinson, I.P.M., Severini, S., Schuster-Böckler, B., and Graham, T.A. (2018). The effects of mutational processes and selection on driver mutations across cancer types. Nat. Commun. *9*, 1857.

Thompson, A., Kuhn, K., Kienle, S., Schwarz, J., Neumann, T., and Hamon, C. (2003). Tandem Mass Tags : A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS / MS. Anal Chem *75*, 1895–1904.

Toll, A., and Real, F.X. (2008). Somatic oncogenic mutations, benign skin lesions and cancer progression: Where to look next? Cell Cycle *7*, 2674–2681.

Turner, N., and Grose, R. (2010). Fibroblast growth factor signalling: From development to cancer. Nat. Rev. Cancer *10*, 116–129.

Tyanova, S., and Cox, J. (2018). Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. Methods Mol. Biol. *1711*, 133–148.

Ullrich, A., and Schlessinger, J. (1990). Signal transduction by receptors with tyrosine kinase activity. Cell *61*, 203–212.

Vallot, C., Stransky, N., Bernard-Pierrot, I., Hérault, A., Zucman-Rossi, J., Chapeaublanc, E., Vordos, D., Laplanche, A., Benhamou, S., Lebret, T., et al. (2011). A novel epigenetic phenotype associated with the most aggressive pathway of bladder tumor progression. J. Natl. Cancer Inst. *103*, 47–60.

Varma, R.K., Garg, V.K., Singh, L., and Kumar, R. (2013). Role of cyclin-dependent kinase 5 in neurodegenerative disorders. African J. Pharm. Pharmacol. *7*, 478–487.

Vasyutin, I., Zerihun, L., Ivan, C., and Atala, A. (2019). Bladder Organoids and Spheroids: Potential Tools for Normal and Diseased Tissue Modelling. Anticancer Res. *39*, 1105–1118.

Venkatesan, S., Rosenthal, R., Kanu, N., McGranahan, N., Bartek, J., Quezada, S.A., Hare, J., Harris, R.S., and Swanton, C. (2018). Perspective: APOBEC mutagenesis in drug resistance and immune escape in HIV and cancer evolution. Ann. Oncol. *29*, 563–572.

Vermeulen, S.H., Hanum, N., Grotenhuis, A.J., Castaño-Vinyals, G., Van Der Heijden, A.G., Aben, K.K., Mysorekar, I.U., and Kiemeney, L.A. (2015). Recurrent urinary tract infection and risk of bladder cancer in the Nijmegen bladder cancer study. Br. J. Cancer *112*, 594–600.

Villen, J., Beausoleil, S.A., Gerber, S.A., and Gygi, S.P. (2007). Large-scale phosphorylation analysis of mouse liver. Proc. Natl. Acad. Sci. *104*, 1488–1493.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Jr., L.A.D., and Kinzler, K.W. (2013). Cancer Genome Landscapes. Science (80-.). *339*, 1546–1558.

Wang, J., Mikse, O., Liao, R.G., Li, Y., Tan, L., Janne, P.A., Gray, N.S., Wong, K.K., and Hammerman, P.S. (2015). Ligand-associated ERBB2/3 activation confers acquired resistance to FGFR inhibition in FGFR3-dependent cancer cells. Oncogene *34*, 2167–2177.

Wang, L., Šuštić, T., Leite de Oliveira, R., Lieftink, C., Halonen, P., van de Ven, M., Beijersbergen, R.L., van den Heuvel, M.M., Bernards, R., and van der Heijden, M.S. (2017). A Functional Genetic Screen Identifies the Phosphoinositide 3-kinase Pathway as a Determinant of Resistance to Fibroblast Growth Factor Receptor Inhibitors in FGFR Mutant Urothelial Cell Carcinoma [Figure presented]. Eur. Urol. *71*, 858–862.

Warfel, N.A., Dolloff, N.G., Dicker, D.T., Malysz, J., and El-Deiry, W.S. (2013). CDK1 stabilizes HIF-1α via direct phosphorylation of Ser668 to promote tumor growth. Cell Cycle *12*, 3689–3701.

Webster, M.K., and Donoghue, D.J. (1996). Constitutive activation of fibroblast growth factor receptor 3 by the transmembrane domain point mutation found in achondroplasia. EMBO Journalournal *15*, 520–527.

Wenger, R., Kurtcuoglu, V., Scholz, C., Marti, H., and Hoogewijs, D. (2015). Frequently asked questions in hypoxia research. Hypoxia *3*, 35–43.

Wiedemann, M., and Trueb, B. (2000). Characterization of a novel protein (FGFRL1) from human cartilage related to FGF receptors. Genomics *69*, 275–279.

Williams, M.J., Werner, B., Barnes, C.P., Graham, T.A., and Sottoriva, A. (2016). Identification of neutral tumor evolution across cancer types. Nat. Genet. *48*, 238–244.

Williams, S. V., Hurst, C.D., and Knowles, M.A. (2013). Oncogenic FGFR3 gene fusions in bladder cancer. Hum. Mol. Genet. *22*, 795–803.

Wogan, G.N. (1992). Aflatoxins as risk factors for hepatocellular carcinoma in humans. Cancer Res. *52*, 2114s-2118s.

Wu, X.R. (2005). Urothelial tumorigenesis: A tale of divergent pathways. Nat. Rev. Cancer *5*, 713–725.

Wu, Y.M., Su, F., Kalyana-Sundaram, S., Khazanov, N., Ateeq, B., Cao, X., Lonigro, R.J., Vats, P., Wang, R., Lin, S.F., et al. (2013). Identification of targetable FGFR gene fusions in diverse cancers. Cancer Discov. *3*, 636–647.

Yang, Y., Xue, K., Li, Z., Zheng, W., Dong, W., Song, J., Sun, S., Ma, T., and Li, W. (2018). c-Myc regulates the CDK1/cyclin B1 dependent-G 2 /M cell cycle progression by histone H4 acetylation in Raji cells. Int. J. Mol. Med. *41*, 3366–3378.

Yokote, H., Fujita, K., Jing, X., Sawada, T., Liang, S., Yao, L., Yan, X., Zhang, Y., Schlessinger, J., and Sakaguchi, K. (2005). Trans-activation of EphA4 and FGF receptors mediated by direct interactions between their cytoplasmic domains. Proc. Natl. Acad. Sci. *102*, 18866–18871.

Yu, P., Wilhelm, K., Dubrac, A., Tung, J.K., Alves, T.C., Fang, J.S., Xie, Y., Zhu, J., Chen, Z., De Smet, F., et al. (2017). FGF-dependent metabolic control of vascular development. Nature *545*, 224–241.

Zanivan, S., Krueger, M., and Mann, M. (2012). In vivo SILAC labeling for proteomics. Methods Mol Biol *757*, 435–450.

Zeegers, M.P.A., Kellen, E., Buntinx, F., and van den Brandt, P.A. (2004). The association between smoking, beverage consumption, diet and bladder cancer: a systematic literature review. World J. Urol. *21*, 392–401.

Zhang, Z.T., Pak, J., Huang, H.Y., Shapiro, E., Sun, T.T., Pellicer, A., and Wu, X.R. (2001). Role of Ha-ras activation in superficial papillary pathway of urothelial tumor formation. Oncogene *20*, 1973–1980.

Zheng, Z., and Keifer, J. (2009). PKA Has a Critical Role in Synaptic Delivery of GluR1- and GluR4-Containing AMPARs During Initial Stages of Acquisition of In Vitro Classical Conditioning. J. Neurophysiol. *101*, 2539–2549.

Zhou, H., Ye, M., Dong, J., Corradini, E., Cristobal, A., Heck, A.J.R., Zou, H., and Mohammed, S. (2013). Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. Nat. Protoc. *8*, 461–480.

Titre: Identification des mutations fréquentes et associées avec APOBEC et caractérisation des voies de signalisation contrôlées par FGFR3 dans le cancer de la vessie

Mots clés: Cancer de la vessie, mutation FGFR3 S249C, mutagenèse due à APOBEC, mutations conductrices, voies de signalisation, protéomique, phospho-protéomique, souris transgénique

Résumé: Le cancer de la vessie (BCa), est une tumeur maligne de l'urothélium, fréquente dans le monde entier, dont le traitement particulièrement coûteux ne permet cependant pas d'éviter les récidives et les progressions. *FGFR3* est l'un des gènes les plus fréquemment mutés dans le BCa et les cellules tumorales sont dépendantes de son expression pour leur prolifération. La mutation *FGFR3* S249C est fortement surreprésentée (62% des mutations récurrentes de *FGFR3*). Dans la première partie de ma thèse, en réalisant une étude de la signature de mutation, nous avons montré que cette surreprésentation de la mutation *FGFR3* était liée à une pression sélective induite par la mutagenèse APOBEC et non due à un gain de fonction plus important induit par cette mutation. En plus de *FGFR3* S249C, 44 mutations fréquentes (représentant près de la moitié des mutations fréquentes du BCa) ont été identifiées comme étant associées à la signature mutationnelle APOBEC et la plupart d'entre elles étaient surreprésentées par rapport à d'autres mutations au sein du même gène. Il est intéressant de noter que ces mutations associées à APOBEC incluaient à la fois de nouveaux 'conducteurs' et des 'passagers' fréquents potentiels et qu'elles pouvaient potentiellement sélectionner des répondeurs à l'immunothérapie. Dans la deuxième partie de cette thèse, nous nous sommes intéressés aux effets fonctionnels du gène *FGFR3* dans le BCa. En utilisant un modèle de souris transgénique, nous avons apporté la première preuve *in vivo* selon laquelle cette mutation *FGFR3* S249C conférait un pouvoir de transformation maligne. Ce processus était associé à une instabilité accrue du génome, activation de MYC et à une angiogenèse accrue, probablement induites par le facteur induisant l'hypoxie (*HIF1A*). En outre, nous avons caractérisé le réseau de régulation contrôlé par FGFR3 en analysant des données protéomiques obtenues par spectrométrie de masse à partir d'une lignée de cellules cancéreuses du cancer de la vessie portant la mutation *FGFR3* S249C - UMUC14. Plusieurs voies de signalisation bien connues comme étant régulées par FGFR3 ont été identifiées. Nous avons également mis en évidence de nouvelles cascades de signalisation suite à l'activation de FGFR3 pouvant être jouer un rôle dans la progression tumorale, notamment un axe FGFR3 / HIF1A / angiogenèse qui a été validé dans certains modèles de BCa *in vitro* et *in vivo*.

Title: Identification of APOBEC-associated hotspot mutations and characterization of FGFR3-driven signaling pathways in bladder cancer

Keywords: Bladder cancer, FGFR3 S249C mutation, APOBEC mutagenesis, driver mutations, passenger mutations, signaling pathways, proteomics, phospho-proteomics, transgenic mice

Abstract: Bladder cancer (BCa) is a worldwide frequent and costly urothelial malignancy. *FGFR3* is one of the most frequently mutated genes in BCa and a driver of an oncogenic dependency. Here, we systematically catalogued the *FGFR3* point mutation spectrum in BCa and identified 14 recurrent residues (frequency ≥ 2). One hotspot mutation - *FGFR3* S249C - was strongly over-represented compared to other recurrent *FGFR3* mutations (62% of all recurrent mutations). Based on in-depth investigation of mutational signature, we revealed that this over-representation of *FGFR3* S249C mutation was merely favoured by APOBEC mutagenesis rather than a stronger functional selection compared to other oncodriver mutations on *FGFR3*. Similarly, together with *FGFR3* S249C, 44 hotspot mutations (accounts for nearly half of all hotspot mutations in BCa) were pinpointed to be associated with APOBEC mutational signature and most of them were over-represented compared to other mutations within the same gene. Interestingly, these APOBEC-associated mutations included both novel potential 'drivers' as well as 'frequent passengers', and had a potential to select responders for immunotherapy. On the other hand, we were interested in functional effects of FGFR3 activation in BCa. We provided the first *in vivo* evidence that *FGFR3* S249C mutation conferred potency to BCa transformation using a transgenic mice model. This process was associated with increased genome instability, MYC activation and enhanced angiogenesis probably mediated by hypoxia-inducing factor (*HIF1A*). Further, we tried to characterize FGFR3-driven regulatory network through mass spectrometry based proteomic data generated in a BCa cell line bearing *FGFR3* S249C mutation – UMUC14. As expected, several well-known FGFR3 regulated signaling pathways could be identified. Of note, we also highlighted some novel signaling cascades that may be relevant to FGFR3 activation, including a FGFR3/HIF1A/angiogenesis signaling axis that we validated in several *in vitro* and *in vivo* BCa models.