



HAL
open science

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, et al.

► **To cite this version:**

Ming-Jun Shi, Xiang-Yu Meng, Philippe Lamy, A. Rouf Banday, Jie Yang, et al.. APOBEC-mediated Mutagenesis as a Likely Cause of FGFR3 S249C Mutation Over-representation in Bladder Cancer. *European Urology*, 2019, 76 (1), pp.9-13. <10.1016/j.eururo.2019.03.032>. <hal-02369388>

HAL Id: hal-02369388

<https://hal.science/hal-02369388v1>

Submitted on 18 Nov 2019

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

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



HAL Authorization

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



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 Dyrskjød^{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 \rightarrow TGC), representing 62% of all recurrent *FGFR3* mutations. We wondered whether this overrepresentation of *FGFR3*-S249C was associated with some specific mutational processes. Considering all mutational signatures [3], the S249C (TCC \rightarrow TGC) is most similar to an APOBEC-type mutation (TCN \rightarrow 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;

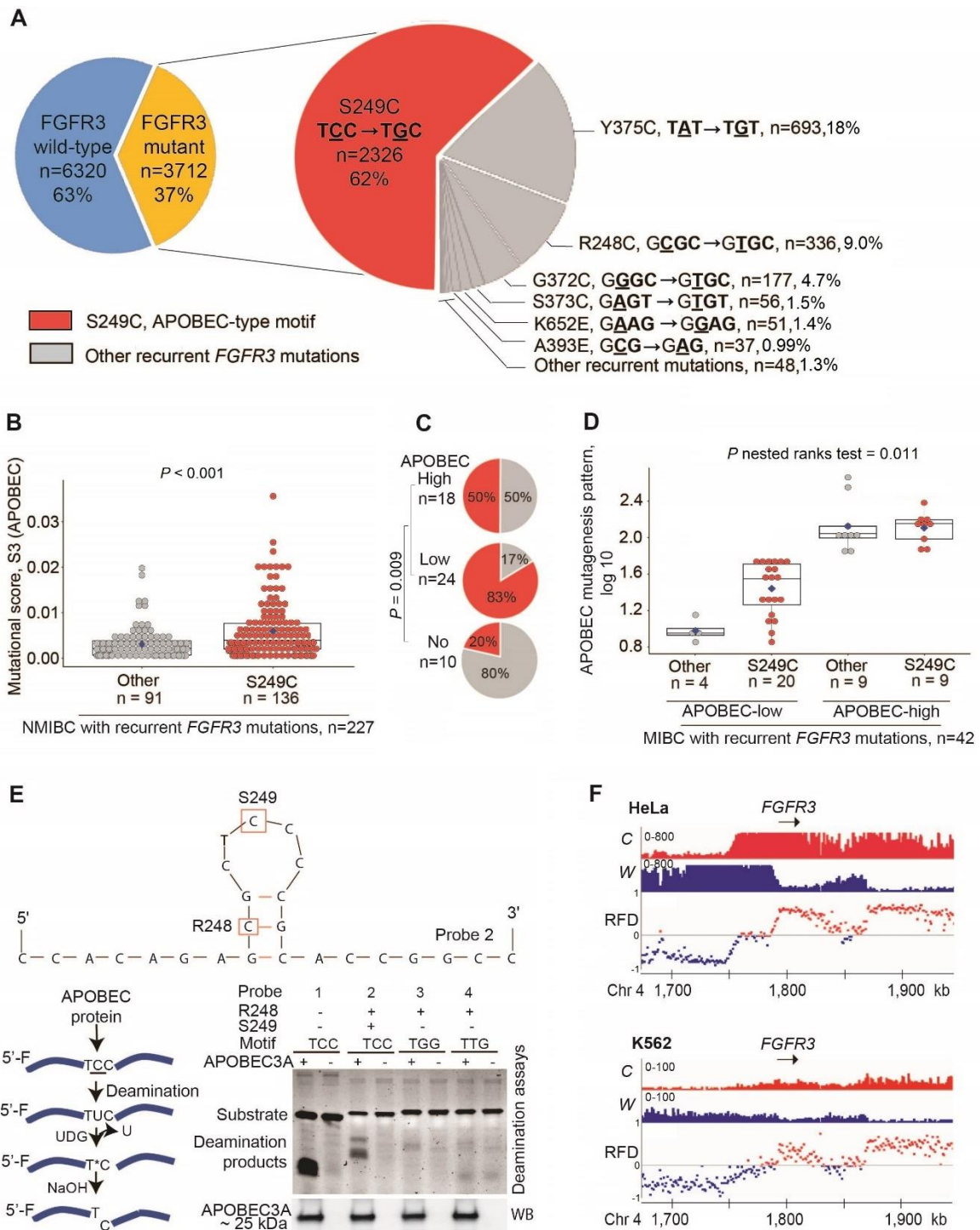


Figure 1. *FGFR3*-S249C mutation as a possible outcome of APOBEC-mediated mutagenesis.

(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).

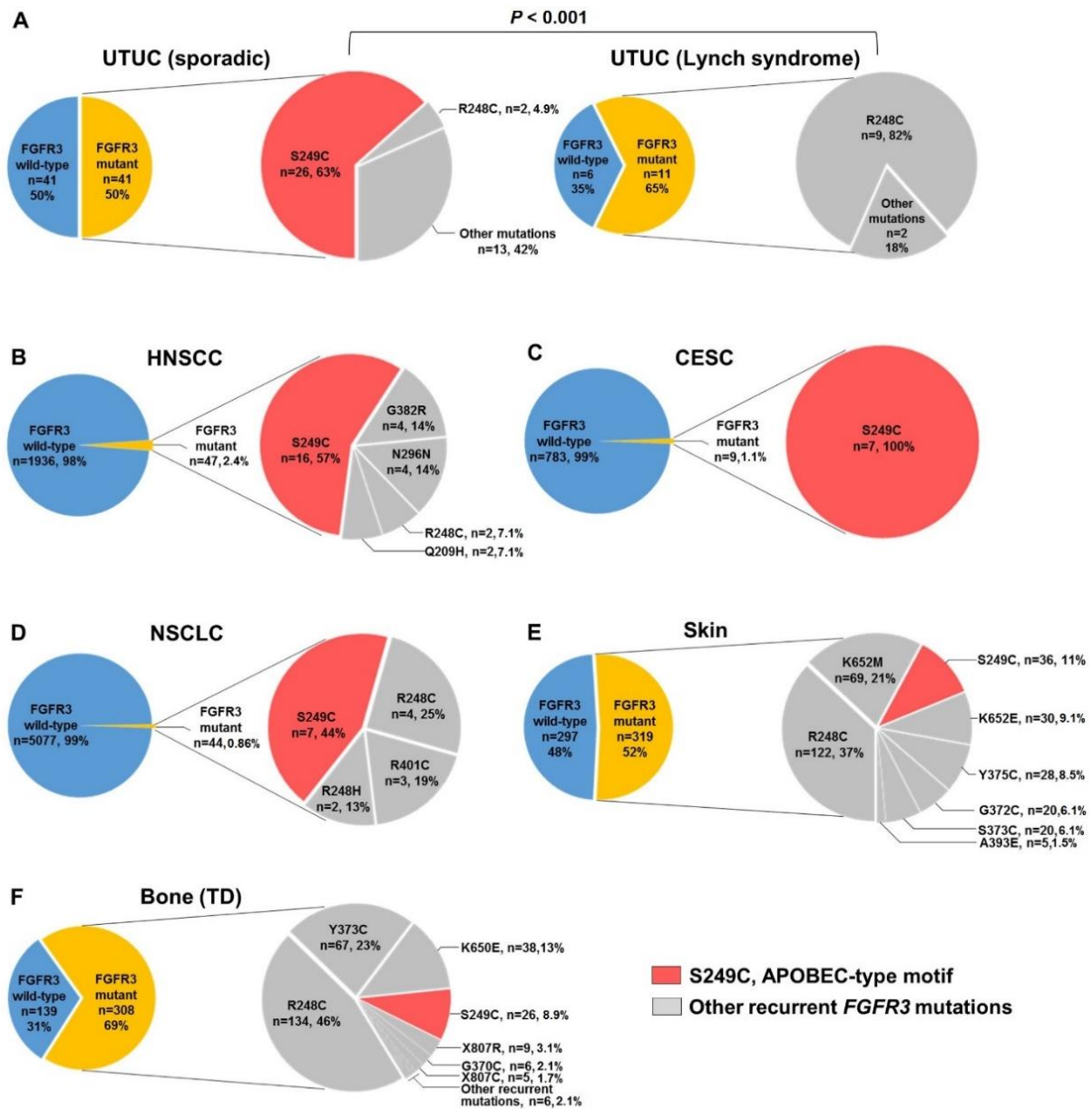


Figure 2. *FGFR3* mutation spectrum across several cancer types, benign skin tumors and bone disorders. (A-F) The rates and distribution of *FGFR3* mutations in patients with sporadic ($n = 82$) and Lynch syndrome-associated ($n = 17$) UTUC, HNSCC ($n = 1983$), CESC ($n = 792$), NSCLC ($n = 5121$), benign 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 seborrheic 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, Dyrskjöt, 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, Dyrskjöt, 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, Dyrskjöt.

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

Obtaining funding: Bernard-Pierrot, Radvanyi.

Administrative, technical, or material support: None.

Supervision: Dyrskjöt, 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**
2. **Supplementary Figures**
3. **Supplementary Tables**
3. **Supplementary references**

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°C 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

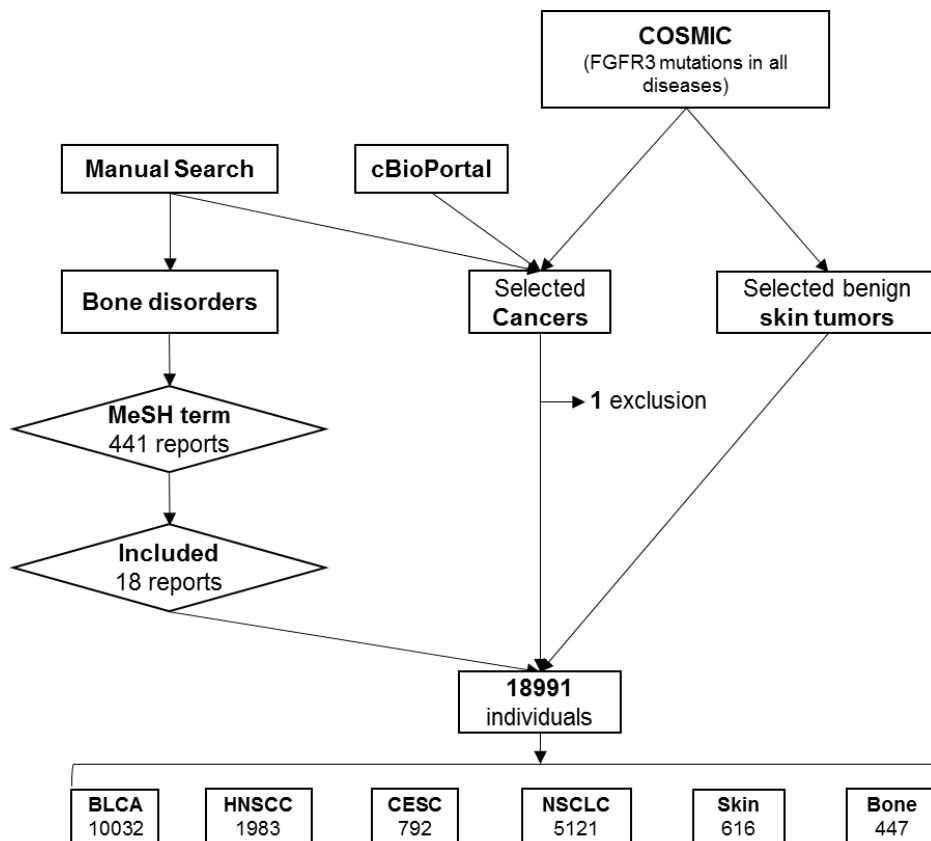


Fig. S1 Workflow of data collection for *FGFR3* mutation spectrum. MeSH terms can be found in **Table S1**. BLCA, bladder cancer; HNSCC, head and neck squamous cell carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; NSCLC, non-small cell lung cancer; Benign skin tumors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

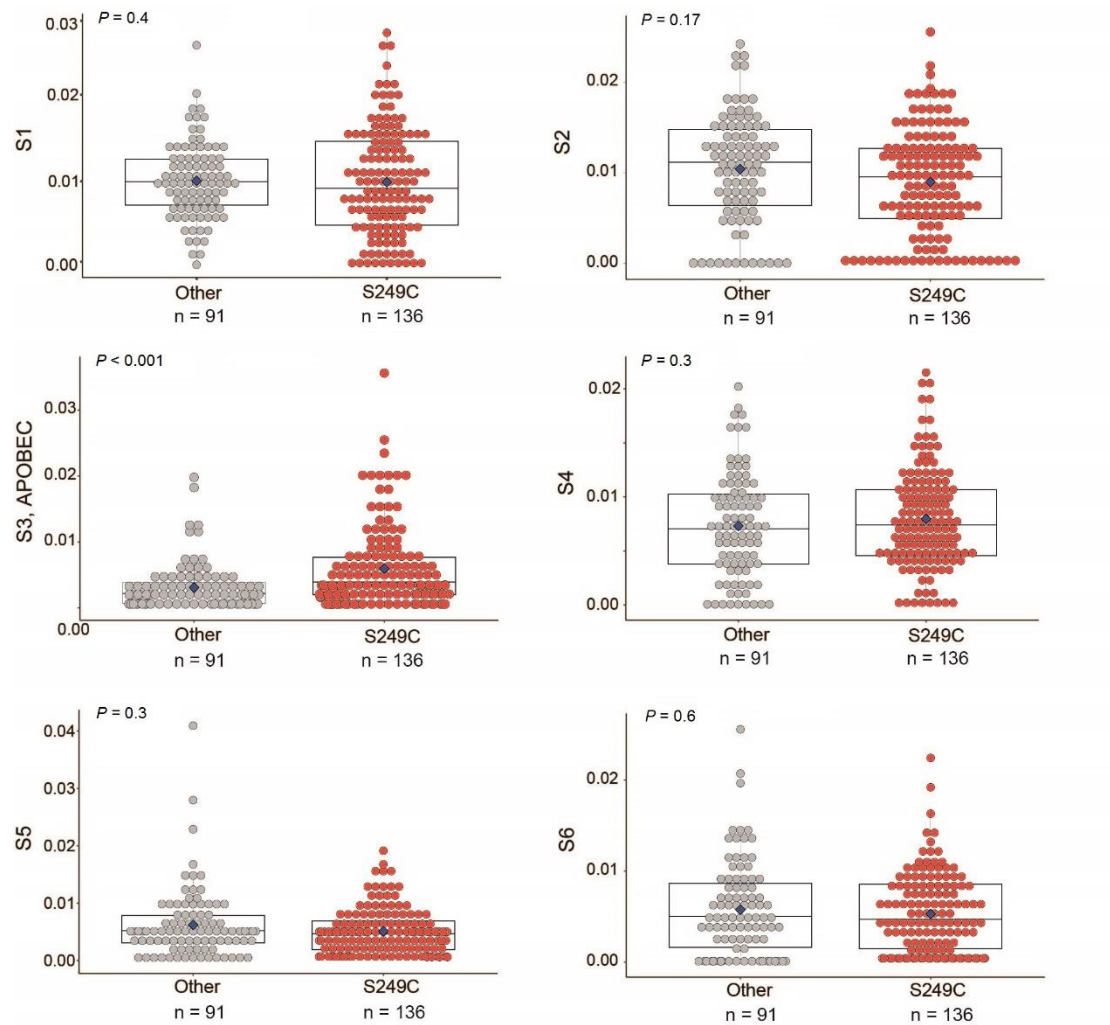


Fig. S2 Distribution of mutational signature scores in 227 tumors from patients with non-muscle-invasive bladder cancer (NMIBC [4]) with recurrent *FGFR3* mutations - S249C vs. other mutations. Recurrent mutations were defined as those found in at least 2 patients in analysis presented in **Fig.1A** 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.

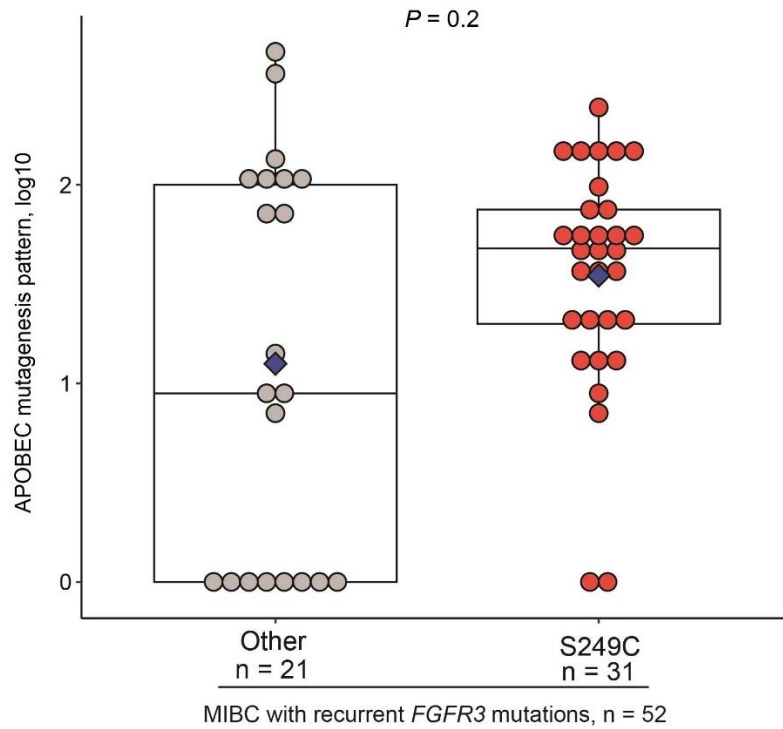


Fig. S3 Distribution of APOBEC mutational pattern in 52 MIBC [10] patients with recurrent *FGFR3* mutations - S249C vs. other mutations. Recurrent mutations were defined as those found in at least 2 patients in analysis presented in **Fig. 1A** and listed in **Table S2**. *P*-values are for Mann-Whitney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

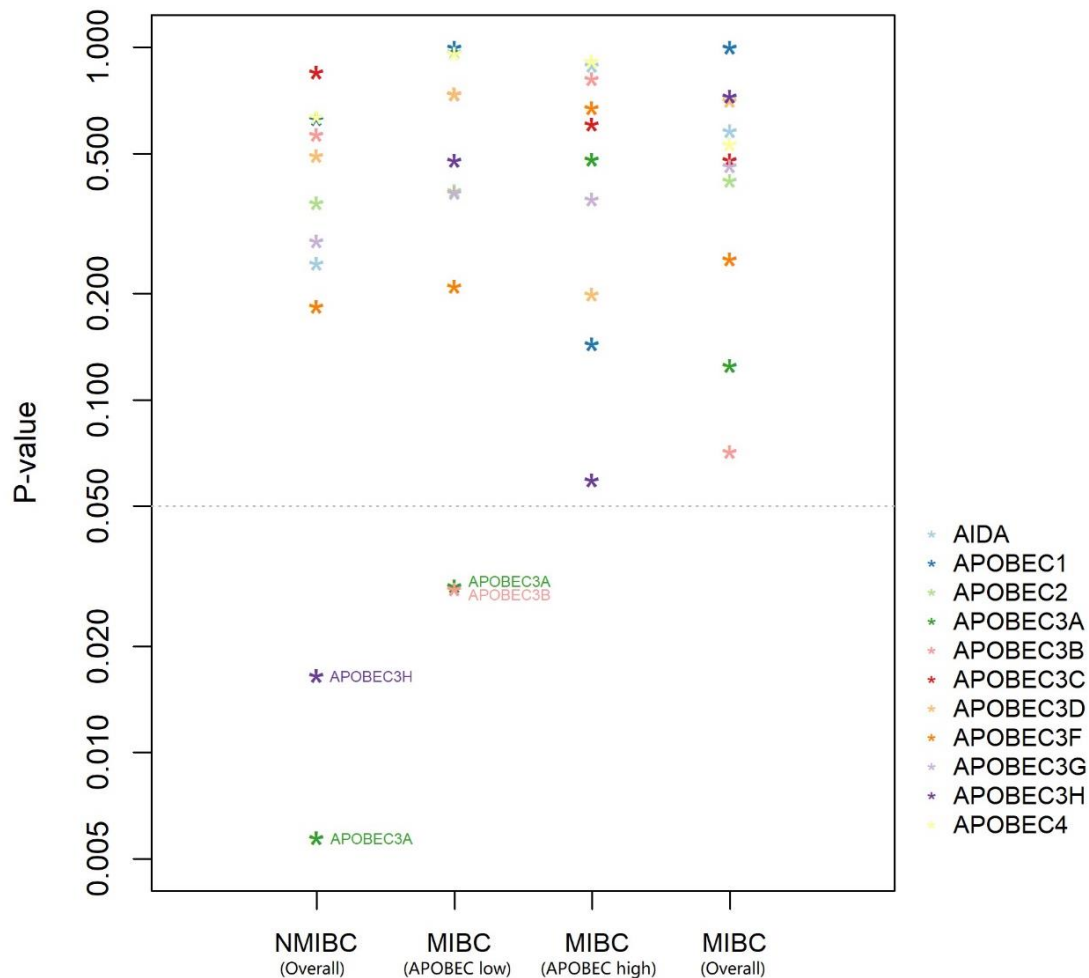


Fig. S4 Statistical significance for the association between *AID/APOBEC* gene expression (FPKM, log₂) and recurrent *FGFR3* mutations - S249C vs. other recurrent *FGFR3* mutations in 270 NMIBC [4] and 41 MIBC (RSEM, log₂) [10] patients (24 with APOBEC-low and 17 with APOBEC-high tumors). Low and high 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.

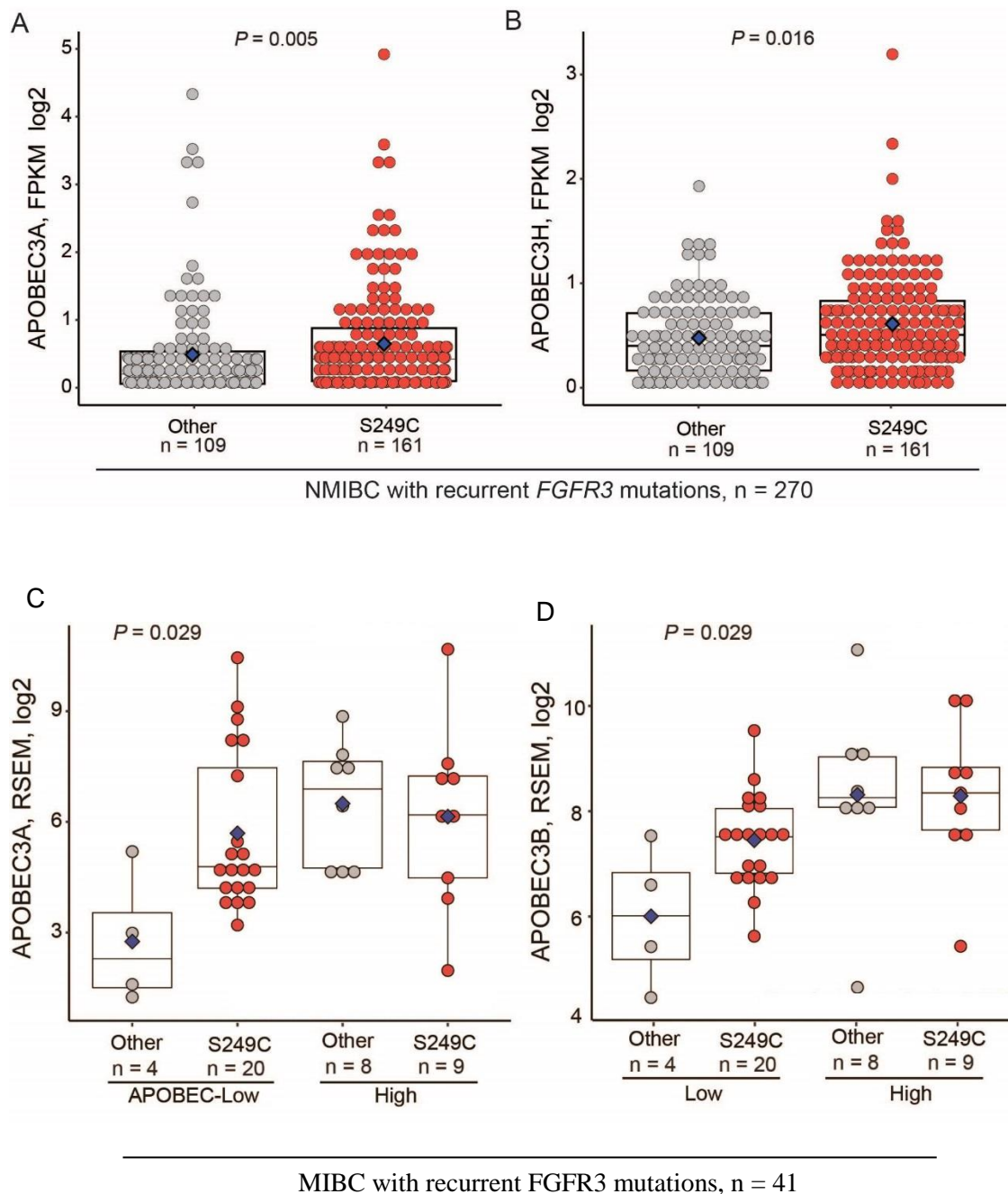


Fig. S5 Distribution of expression values for the *AID/APOBEC* genes significantly associated with recurrent *FGFR3* mutations - S249C vs. other mutations in 270 NMIBC [4] and 41 MIBC [10] patients in analysis presented in Fig. S4. **(A)** *APOBEC3A* in NMIBC tumors. **(B)** *APOBEC3H* in NMIBC tumors. **(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).

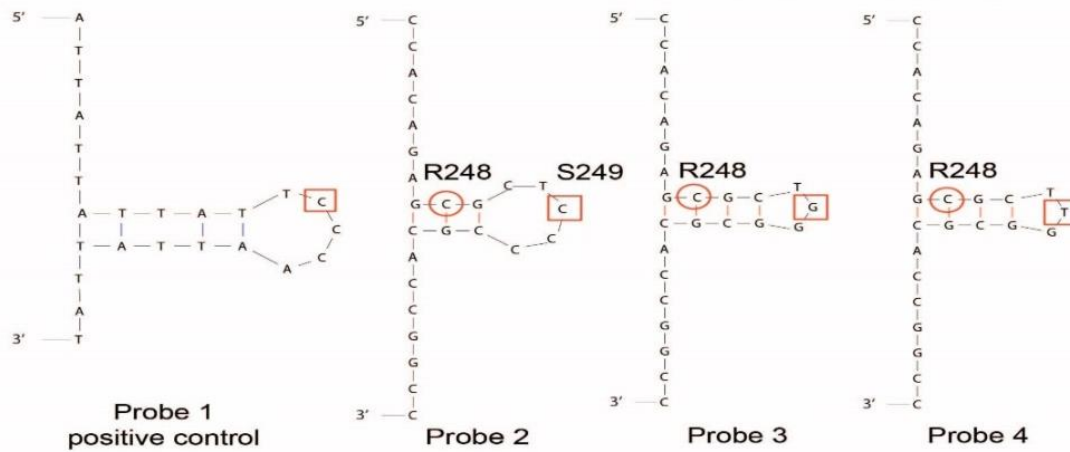


Fig. S6 Mfold analysis of secondary structures of all probes used for deamination assays - S249 is located within a single-stranded 5-nucleotide loop, while R248 is located within the double-stranded hairpin stem; shown are central 25 bp nucleotides of each probe. APOBEC-mediated mutagenesis is accumulated 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].

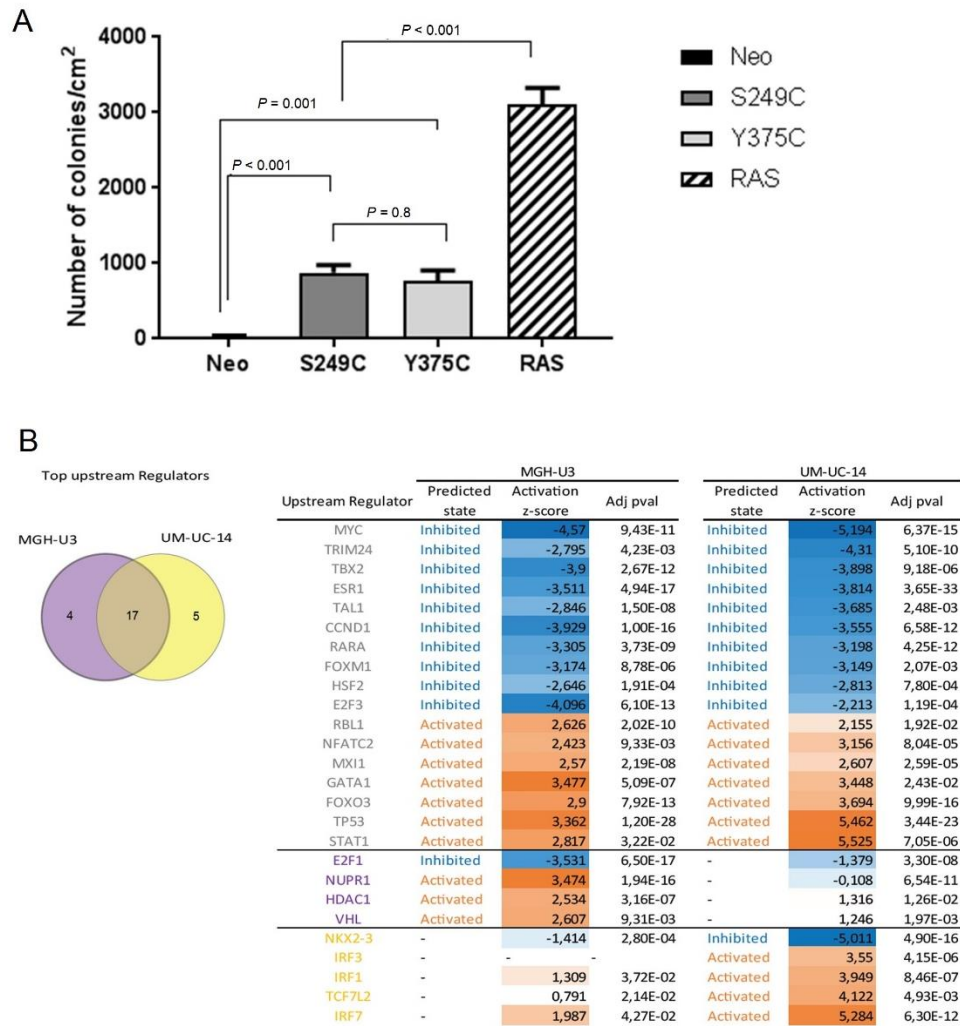


Fig. S7 Similar tumorigenic potential of FGFR3 with S249C and Y375C mutations. **(A)** Overexpression of FGFR3-S249C and FGFR3-Y375C in NIH-3T3 cells shows similar transformation potential and significantly lower compared to HRAS-Q61R (positive control), based on the number of soft agar colonies; Dunnett's multiple comparisons test. Shown are mean \pm 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 <i>FGFR3</i> 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)	GCGC	GTCG
c.1114G>T	p.G372C	5% (177)	/	/	/	6% (20)	2% (6)	GGGC	GTCG
c.1117A>T	p.S373C	1% (56)	/	/	/	6% (20)	1% (2)	GAGT	GTCG
c.1954A>G	p.K652E	1% (51)	/	/	/	9% (30)	13% (38)	G<u>A</u>AG	GGAG
c.1178C>A	p.A393E	1% (37)	/	/	/	2% (5)	/	GCG	GAG
c.1144G>C or c.1144G>A	p.G382R	1% (20)	14% (4)	/	/	/	/	CGGG	CC<u>AGG</u>
c.1955A>T	p.K652M	0.3% (12)	/	/	/	21% (69)	1% (2)	AAG	ATG
c.1954A>C	p.K652Q	0.1% (5)	/	/	/	/	/	GAAG	GCAG
c.1955A>C	p.K652T	0.1% (5)	/	/	/	/	/	AAG	ACG
c.1156T>C	p.F386L	0.1% (2)	/	/	/	/	/	CTTC	CCTC
c.1178C>T	p.A393V	0.1% (2)	/	/	/	/	/	GCG	GTCG
c.1927G>A	p.D643N	0.1% (2)	/	/	/	/	/	GGAC	GAAC

*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 Seborrheic 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.

1st round review: letter to editor and point by point response



Paris, January 22th, 2019

Ms. No: EURUROL-D-18-01506

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

Dear Dr. Catto,

We would like to thank you and the reviewers for considering our manuscript and providing constructive comments and suggestions. We have thoroughly revised the paper, included new results, and provided point-by-point responses reflecting changes in the manuscript.

In recent years, APOBEC mutagenesis has been identified as an important molecular feature of bladder tumors and disease aggressiveness. Our results on molecular etiology of *FGFR3*-S249C, the most common and a clinically relevant *FGFR3* mutation, will improve our understanding of bladder cancer.

We expect our paper will be of interest for the broad scientific community, and specifically for researchers working on bladder cancer, *FGFR3* and APOBEC mutagenesis.

We hope that after this thorough revision, our work meets the high standards for publication in *European Urology*.

Yours sincerely,

On behalf of the authors,

François Radvanyi

Point-by-point response to reviewers

We thank the reviewers for their constructive comments and suggestions that led to new analyses and helped to improve our manuscript. Overall, the data presented in our revised manuscript strengthened the hypothesis about the link between APOBEC mutagenic activity and over-representation of *FGFR3*-S249C mutation in BLCA and other conditions.

Comment from Reviewer # 1:

1. *“The authors have attempted to identify a correlation between APOBEC mutational signature and FGFR3-S249C mutations in bladder cancer. The authors identified a correlation between signature and mutation in patients with non-muscle invasive bladder cancer, but are unable to do so in patients with muscle invasive bladder cancer. The explanation to why this occurs is anemic and requires further analysis.”*

Response:

Initially, we focused our analysis on APOBEC signature 13 (TCW->TGW, W=A or T) as the closest match for the *FGFR3*-S249C mutation motif (TCC->TGC). Because this analysis is based on assigning statistical probabilities to different signatures and this deconvolution process is associated with uncertainties, we initially used only 66 NMIBC samples out of 227 with *FGFR3* mutations. In these 66 samples signature 13 was considered to be dominant. To use information from all 227 tumors with *FGFR3* mutations, we decided to use estimates for total APOBEC mutagenesis.

In the current version we made the following changes in the analysis:

1. In NMIBC, we used the total RNAseq-based APOBEC mutation signature score (S3) that reflects the global APOBEC activity without distinction between signatures 2 and 13, allowing us to use information from all 227 NMIBC tumors with *FGFR3* mutations [1].

2. For MIBC, a recent bioRxiv manuscript preprint (Alexandrov, <https://doi.org/10.1101/322859>) provided signature 2 and 13 scores for all TCGA samples. Our analysis using these scores showed similar trends for both signatures in relation to *FGFR3*-S249C mutation, justifying the use of the combined estimate for total APOBEC mutagenesis, which was available in Table S1 (APOBEC induced mutation load (P-MACD) variable) of the MIBC TCGA paper, Robertson et al [2].

3. We use an existing classification of all MIBC samples as “APOBEC-high, low or no” based on the presence of APOBEC-signature mutations provided in Table S1 of MIBC TCGA paper, Robertson et al [2]. APOBEC-no tumors have no mutation that can be confidently classified as APOBEC-type; the remaining tumors were assigned into two groups based on mutation load above and below the median. The use of this classification is justified by the striking distribution of *FGFR3*-S249C mutation in these groups (20% in APOBEC-no, 50% in APOBEC-high and 83% in APOBEC-low group, **Fig. 1C**).

We observed both differences and similarities between NMIBC and MIBC.

Similarities:

1. Similar rate of *FGFR3*-S249C mutation (59.6%) of all recurrent *FGFR3* mutations both in NMIBC and MIBC, despite much higher frequency of tumors with recurrent *FGFR3* mutations in NMIBC (66.7%) compared to MIBC (12.6%).
2. Analysis of expression of all 11 genes from the *AID/APOBEC* family identified *APOBEC3A* as a likely APOBEC enzyme responsible for the *FGFR3*-S249C mutation both in NMIBC and MIBC (**Fig.S4** and **Fig.S5**).
3. APOBEC-type mutagenesis is significantly higher in *FGFR3*-S249 tumors compared to other recurrent *FGFR3* mutations in NMIBC (**Fig.1B**) and APOBEC-low MIBC (**Fig.1D**).

Differences:

1. APOBEC-type mutagenesis is not significantly associated with *FGFR3*-S249 compared to other recurrent *FGFR3* mutations in MIBC overall and in APOBEC-high tumors (**Fig.S3** and **Fig.1D**). This difference could be due to the lower number of mutated tumors in MIBC compared to NMIBC (52 MIBC tumors with recurrent *FGFR3* mutations compared to 227 in NMIBC), or due to higher heterogeneity in MIBC.

These points are now presented in the manuscript: fewer MIBC tumors with recurrent *FGFR3* mutations compared to NMIBC; heterogeneity of *FGFR3* mutations in groups based on APOBEC mutation load (**Fig. 1C**); differences in *FGFR3* mutation pattern related to APOBEC activity (**Fig.1D**). These data suggest that APOBEC mutagenesis is necessary for generation of *FGFR3*-S249C mutation; this occurs at the low level of APOBEC mutagenesis and likely early in tumor development. In APOBEC-high tumors, *FGFR3*-S249C mutation might be masked by high background of other mutations that may appear later in tumor development as a result of multiple factors, including genomic instability and treatment.

2. *“Furthermore, attempt to validate the findings from the NMIC cohort in other cancer types has resulted in variable findings.”*

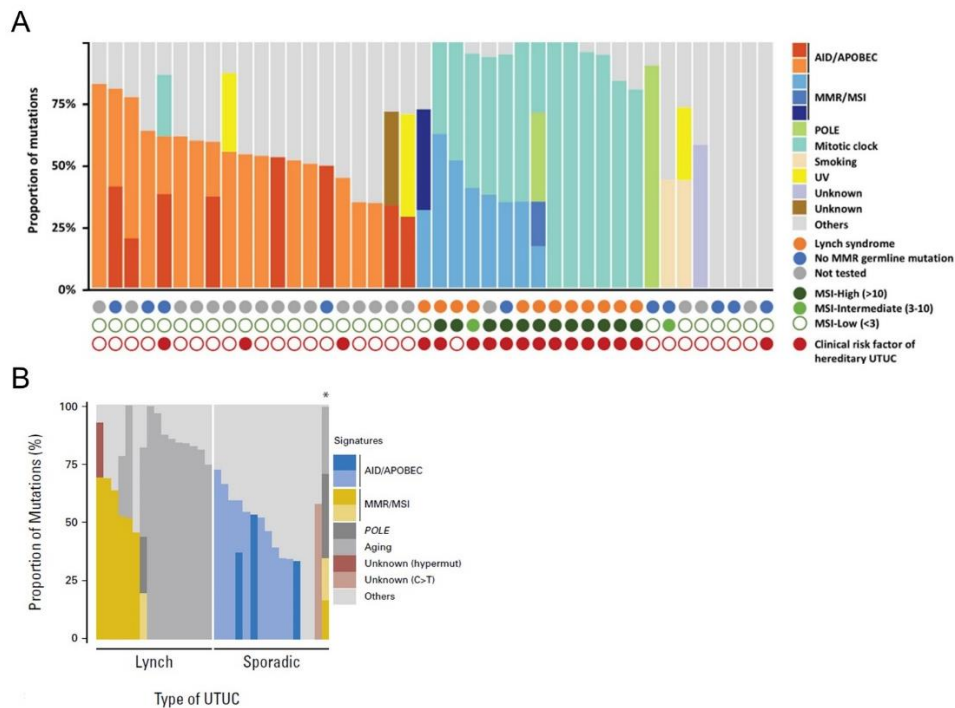
Response:

Our intent in this analysis was to link differential distribution of *FGFR3* mutations to their possible etiologies. Unlike mutations in *TP53* or *PIK3CA* genes, which are uniformly and frequently found in many cancer types, *FGFR3* mutations are more context-dependent. The *FGFR3* mutation rates range from 30% to 50% in bladder cancer, benign skin tumors, while being present in less than 3% in other tumor types (**Fig. 1A** and **Fig. 2A-F**).

We suggest that the over-representation of *FGFR3*-S249C mutation is common in conditions linked with APOBEC mutagenesis, such as bladder cancer, head and neck cancer and cervical cancer. In other conditions, such as benign skin tumors (seborrheic keratosis) and a

bone disorders (thanatophoric dysplasia) linked with other mutational processes (aging, UV exposure [3–5]), *FGFR3*-S249C is a rare *FGFR3* mutation.

Two recent publications [6,7] provided additional support for this trend (**Fig.2A**). In upper urinary tract tumors (UTUC) from patients with Lynch syndrome, which present no or low APOBEC mutagenesis, *FGFR3*-S249C mutation is rare, while it is common in sporadic UTUC that shows evidence of APOBEC mutagenesis (**Fig.2A** and **Reply Fig. 1A** and **1B**). As this difference was observed in cancers of the same tissue type (the urothelium of the upper urinary tract), this provides important support for APOBEC mutagenesis as an etiological cause of *FGFR3*-S249C mutation.



Reply-Figure 1. Distribution of *FGFR3* mutations and APOBEC mutational signature in UTUC (upper urinary tract urothelial cancer) in patients with and without Lynch syndrome. **(A)** Screenshot from the report of Audenet et al, 2018 CCR (original Fig. 4C); **(B)** Screenshot from the report of Donahue et al, 2018 JCO Precision Oncology (original Fig. 1B).

3. “Further analysis, *in vitro* or *in vivo*, would be important to help understand the significance of these findings for clinical use.”

Response:

We believe that identifying etiological mechanisms of this mutation that represents almost 60% of all recurrent *FGFR3* mutations both in NMIBC and MIBC, is of clinical significance. To improve our understanding of these mechanisms, we have performed several experiments and included additional analyses. Specifically, we show that: 1) The presence of S249C mutation is significantly associated with *APOBEC3A* expression both in NMIBC and MIBC (**Fig.S4** and **Fig.S5**); 2) *FGFR3* with S249C mutation has the same potential as *FGFR3* with Y375C mutation (the second most frequent mutation found in 18% of tumors, and not an APOBEC-type) to transform NIH-3T3 cells and activate the same regulatory pathway in

bladder cancer cells suggesting lack of functional advantage specifically provided by FGFR3-S249C (**Fig.S7**); 3) S249C mutation might be occurring at a high rate because of its position within a ssDNA loop of a hairpin, which makes it an efficient target of APOBEC enzymes (**Fig.1E**); 4) Cytosine in the S249 position (TCC) is efficiently deaminated by the activity of recombinant APOBEC3A enzyme (**Fig.1E**). Taken together, our results suggest that the over-representation of *FGFR3*-S249C mutation is more likely linked to APOBEC-mediated mutagenesis than to a particular functional advantage of this mutation compared to other recurrent *FGFR3* mutations.

Comment from Reviewer # 2:

1. *“However, the manuscript is very hard to understand and requires extensive re-writing.*

Response:

We have significantly revised our manuscript to improve presentation.

2. *Furthermore, the interpretation of the data may be flawed: admittedly, the conclusions may be correct but there is not sufficient evidence excluding other explanations for the S249C predominance in bladder cancer.”*

Response:

As provided in response to other questions and reflected in the updated manuscript, we now present several new lines of evidence to strengthen our conclusion that APOBEC-mediate mutagenesis is a significant contributor to the over-representation of the *FGFR3*-S249C mutation in bladder cancer and other conditions.

We also provide new functional evidence to exclude that the over-representation results only from a functional advantage specifically provided by *FGFR3*-S249C. *FGFR3* with S249C mutation has the same potential as *FGFR3* with Y375C mutation (the second most frequent mutation found in 18% of tumors, and not an APOBEC-type) to transform NIH-3T3 cells and activate the same regulatory pathway in bladder cancer cells (**Fig.S7**). Although we cannot confidently exclude other factors, these results reinforce the likelihood of APOBEC-mediated mutagenesis to cause the over-representation of *FGFR3*-S249C.

Specific comments

1. *“The authors should cite a recent report that relates signature 13 with PIK3CA and ERBB2 mutations in bladder cancer (Poulos et al PLOS Genet 2018). In this very broad analysis of associations between somatic mutations and mutational signatures in more than 7000 tumors, the S249C mutation is not identified as significantly associated with APOBEC signatures in bladder cancer. Can the authors speculate on the reasons for this discrepancy?”*

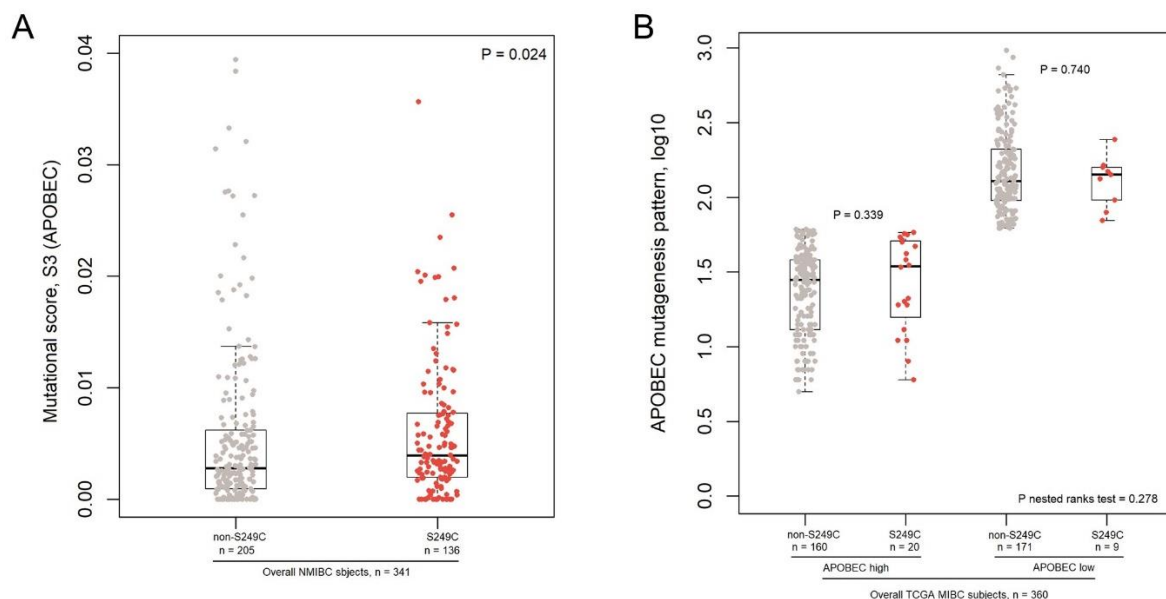
Response:

The paper by Poulos et al [8] did not identify association between *FGFR3*-S249C mutation and APOBEC signatures 2 and 13 by comparing 30 tumors with *FGFR3*-S249C mutation vs. 368 wild-type tumors of MIBC TCGA (Mann-Whitney *P*-value = 0.56 for signature 13 and

0.15 for signature 2, respectively, Table S3 of Poulos paper). This could be because of several important differences in our analyses:

1. Our strongest results are observed in 227 NMIBC with recurrent *FGFR3* mutations and APOBEC signature scores, while the analysis in Poulos paper [8] was limited by only 50 MIBC TCGA tumors with *FGFR3* mutations and high heterogeneity in APOBEC mutation loads in these tumors.
2. Because we wanted to understand the difference in etiology between *FGFR3*-S249C and other *FGFR3* mutations, we performed our analysis only in tumors with recurrent *FGFR3* mutations, defined as those observed in at least two out of more than 10,000 bladder tumors analyzed (**Fig 1A** and **Table S1** and **S2**). In contrast, the Poulos paper compared tumors with a given mutation vs. all other tumors. We performed similar analysis using APOBEC signature score (S3) in NMIBC tumors with *FGFR3*-S249C mutation vs. all other tumors. Although the difference was still detectable, it became attenuated (from $p=2.45E-05$ to 0.024, **Fig.1B** and **Reply-Figure 2A**). Similar analysis in the MIBC TCGA data showed that the difference became not significant within the APOBEC-low group (changed from $p=0.014$ to 0.339) and in all groups (changed from $p=0.011$ to 0.278) (**Reply-Figure 2B** and **Fig.1D**).

The comparison with Poulos paper has not been included in the manuscript due to space limitation and 10 references allowed in this Brief Communication format.



Reply-Figure 2. Distribution of APOBEC mutagenesis between tumors with *FGFR3*-S249C mutations vs. all other tumors. **(A)** RNA-seq derived APOBEC mutation score (S3) in 341 NMIBC tumors in relation to *FGFR3*-S249C mutation. **(B)** APOBEC mutagenesis pattern (log10) in 360 TCGA MIBC tumors in relation to *FGFR3*-S249C mutation in APOBEC-high and APOBEC-low groups. *P*-values are for Mann-Whitney U tests between two groups or for nested ranks test between all groups of samples. NMIBC, non-muscle-invasive bladder cancer; MIBC, muscle-invasive bladder cancer.

2. “The authors should comment on the differences in Figures 1B and 1C between NMIBC and MIBC.”

Response:

Please see our detailed response to Reviewer 1.

3. “It would also be important to correlate the occurrence of the signature and *FGFR3* mutation with the corresponding levels of *APOBEC* mRNA species in tumor tissue when that is available.”

Response:

Thank you for this suggestion. There are 11 *APOBEC* genes that could potentially contribute to *APOBEC* mutagenesis - activation-induced deaminase (*AICDA* (*AID*)), *APOBEC1* (*A1*), *APOBEC2* (*A2*), *APOBEC3* (*A3A*, *A3B*, *A3C*, *A3D*, *A3F*, *A3G* and *A3H*) and *APOBEC4* (*A4*) [9]. In NMIBC, only expression of *APOBEC3A* and *APOBEC3H* significantly differed, for both genes being higher in tumors with S249C compared to tumors with other recurrent *FGFR3* mutations (**Fig.S4** and **Fig.S5**).

Similar analysis in MIBC in The Cancer Genome Atlas (TCGA) showed association with increased expression of *APOBEC3A* and *APOBEC3B* in *APOBEC*-low tumors with S249C compared to tumors with other recurrent *FGFR3* mutations (**Fig.S4** and **Fig.S5**). Thus, expression of the *APOBEC3A* appeared to be associated with *APOBEC*-mediated mutagenesis in carriers of *FGFR3*-S249C mutations both in NMIBC and MIBC.

In line with this, our *in vitro* deamination assays with recombinant *APOBEC3A* protein experimentally confirmed S249 as a target of its deamination activity. *APOBEC3A* might be the strongest candidate of all *APOBEC3s* because it is a typical interferon-stimulated gene that is strongly induced by exogenous infectious stimuli [10]. Thus, generation of *FGFR3*-S249C mutation could be a collateral effect of the immune response inducing expression of *APOBEC3A* or other *APOBEC3s*.

4. “The paragraph describing the analysis of replication fork directionality (RFD) of the *FGFR3* gene (p.4, paragraph 1; Fig 1E-F) should be rewritten for clarity. For instance, the sentence “Indeed, *APOBEC* was found to be linked to a strong replicative asymmetry and specifically targeting lagging strand template in clinical samples [8]” is confusing: in ref. 8, the mutational bias is linked to the transcription sense/antisense strands, not to leading/lagging DNA replication strands. The authors should state in the main text that they have used available genome-wide RFD data generated by the O. Hyrien lab (current ref. 7 and Petryk et al, *Nat Commun* 7:10208; 2016) to analyze *in silico* whether *FGFR3* gene is replicated by the leading or lagging strand. The results of this test support replication by the lagging strand, which fits with the authors' hypothesis. In this context, it seems very important to cite previous studies that link *APOBEC* mutational signatures to lagging strand DNA synthesis: e.g. Haradhvala et al, *Cell* 2016 and Hoopes et al, *Cell Reports* 2016.”

Response:

This paragraph about replication fork directionality (RFD) was completely rewritten and references were cited accordingly. As the format of this Brief Correspondence allows only 10 references, we cited the article of Wu et al 2018 and Haradhvala et al 2016 in the main text, and the works of Petryk et al 2016 and Hoopes et al 2016 were cited in the Supplementary Methods.

5 - *“It is stated that the S249C mutation does not dominate in benign skin tumors and bone dysplasias and that in those conditions an APOBEC mutational signature has not been reported. Of course, this signature has not been reported because most likely it has not been investigated: in order to identify an APOBEC mutational signature, exome sequencing needs to be performed and very few exome sequencing data are available for the two conditions mentioned. Following this argument, one could say that the S249C mutation is again better selected during the evolution towards malignant diseases rather than in benign non-neoplastic conditions.”*

Response:

We agree that we cannot confidently state that there is no reported APOBEC mutational signature in seborrheic keratosis (with only one sample sequenced) or in bone dysplasia. However, there are several arguments suggesting that these conditions are caused by other factors. Aging and cumulative exposure to sunlight are independent risk factors for the development of seborrheic keratosis [4,5]. The one seborrheic keratosis tumor that has been sequenced [11] showed a UV signature and *FGFR3*-K652M mutation. *FGFR3* mutation in thanatophoric dysplasia has been linked to aging [3]. The manuscript has been modified accordingly.

Please also see responses to other questions that indicate that S249 might be an efficient target for APOBEC activity due to the secondary structure of the surrounding sequence and that *FGFR3* is being replicated from lagging strand template. Because the transforming properties of *FGFR3* with S249C were similar to those of the less frequent mutation, Y375C, we suggest that the frequency of *FGFR3*-S249C mutation is determined by the rate of its generation due to APOBEC activity, and not of its selection, at least when compared to Y375C mutation.

6. *“APOBEC-low tumors do present with the S249C mutation. Therefore, it appears that the APOBEC-dependent mechanism would be only one of the contributors to the S249C mutation. Alternatively, the author's proposal would not be substantiated by the data and - for yet unknown reasons - the S249C would be particularly efficient to transform urothelial cells rather than other cell types (i.e. skin keratinocytes).”*

Response:

In fact, this is exactly our point – that APOBEC-low tumors are a distinct group with just enough APOBEC expression (mutagenesis) to cause enrichment of *FGFR3*-S249C mutation without high background of mutational noise found in APOBEC-high MIBC tumors that would mask this enrichment. The striking distribution of *FGFR3*-S249C mutation in these

groups (20% in APOBEC-no, 50% in APOBEC-high and 83% in APOBEC-low group, **Fig. 1C**) supports this point.

In general, enrichment of *FGFR3*-S249C mutation was found in tumor types that can be considered APOBEC-low: NMIBC, APOBEC-low MIBC and sporadic UTUC. Other driver mutations in *FGFR3* and other genes might be responsible in APOBEC-no and APOBEC-high tumors.

Although *FGFR3* with S249C mutation is very potent in transforming cells and being selected in tumors, we found its potency comparable to that of another recurrent, but much less frequent mutation, Y375C (**Fig.S7**). Thus, we conclude that the initiation rate due to APOBEC activity defines the distribution and frequency of *FGFR3*-S249C mutation.

Minor comments

1. *“Figure 1A should cite the mutations at the nucleotide rather than the amino acid level in order to link with the corresponding explanations in the text.”*

Response:

Thank you for this suggestion. We have now added in **Fig. 1A** the mutation information at the nucleotide level, flanking the mutated nucleotide (underlined) and with codons marked in bold.

2. *“There are various places in the manuscript where the statements are not clear: APOBEC3 high vs. low and it is not properly qualified what this refers to.”*

Response:

We used an existing definition described in the Table S1 of the MIBC TCGA paper by Robertson et al [2]. This table provides classification of all MIBC samples as “APOBEC-high, low or no” based on the presence of APOBEC-signature mutations. APOBEC-no tumors have no mutations that can be confidently classified as APOBEC-type; the remaining tumors were assigned into two groups based on mutation load above and below the median. We now clearly explain the source of the data and the definitions used.

3. Comment from Reviewer # 3:

Major points:

1. - *While there is evidence of association between the specific *FGFR3*-S249C mutation and Sig.13 (the authors were able to produce a significant p-value), this association not exclusive to this specific mutation. In fact, as shown in Figure 1, there is still significant overlap in the APOBEC Sig.13 frequency and fraction between *FGFR3*-S249C and other *FGFR3* mutations in both NMIBC and MIBC.*

Response:

Mutational signatures represent a statistically defined probability (enrichment) of a mutation to occur at a specific motif. However, the utilization of these motifs is very much context-dependent. As we show by several lines of evidence, the *FGFR3*-S249C mutation might

represent a specific case due to an optimal location within an APOBEC-type motif, in the loop of a hairpin and within the gene replicated from a lagging strand. All these conditions create an opportunity for this site being mutated by APOBECs. Inducible expression of APOBEC3s, and particularly of *APOBEC3A* in the same samples, as well as evidence of overall APOBEC-mediated mutagenic activity, increases this chance.

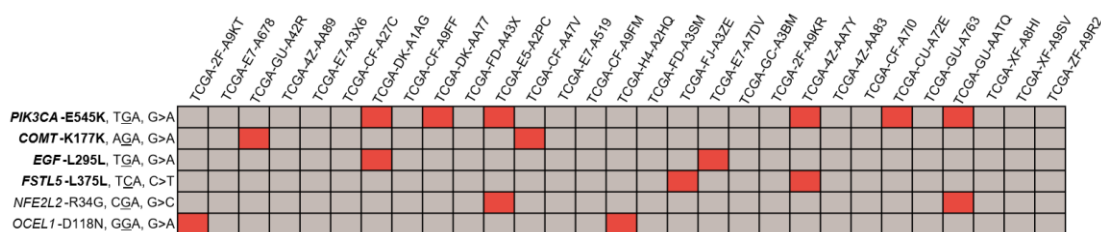
For this reason and to avoid interpreting Sig 2 and Sig 13 motifs too literally (as TCC->TGC mutation fits either of them only partially), we now used an estimate of total APOBEC activity as S3 RNA-seq based score in NMIBC and mutation APOBEC mutation load in MIBC. The re-analyses did not change our results and conclusions but we believe reduced a possible uncertainty associated with deconvolution into signatures 2 and 13 (**Fig. 1B, 1D and Fig.S3**).

2. - *How would the authors explain FGFR3 S249 mutagenesis in the absence of true APOBEC Sig.13? Is there is a unique co-mutation pattern that the authors can identify?*

Response: As we mentioned in response to previous questions, mutational signatures are derived computationally based on frequency metrics. APOBEC signatures 2 and 13 are enriched in specific mutation types (TCW → TTW or TGW), but they don't exclude other patterns, such as (TCN → TTN or TGN). The *FGFR3-S249C* (TCC → TGC) mutation is one of the possible events. We conclude that the high frequency of S249C mutation in APOBEC-low tumors could be explained by the high rate of its occurrence due to the S249 being an optimal target located within a ssDNA loop of the hairpin, availability of APOBEC3A, as well as the transforming potential of this mutation leading to its quick enrichment in tumors.

Using MIBC TCGA dataset described by Robertson et al [2], we identified 6 single nucleotide variations (SNVs) co-occurring with *FGFR3-S249C*, defined as SNVs significantly more frequent in *FGFR3-S249C* subjects (n = 29, with detailed mutation data) compared to other TCGA MIBC subjects (Fisher's exact test, *P* < 0.05): *COMT-K177K*, *EGF-L295L*, *FSTL5-L375L*, *NFE2L2-R34G*, *OCELI1-D118N*, and *PIK3CA-E545K*.

We mapped these co-mutations to the 29 *FGFR3-S249C* subjects (**Reply-Figure 3**). We observed no pattern in distribution of these mutations, which may be due to small sample set of only 29 tumors.



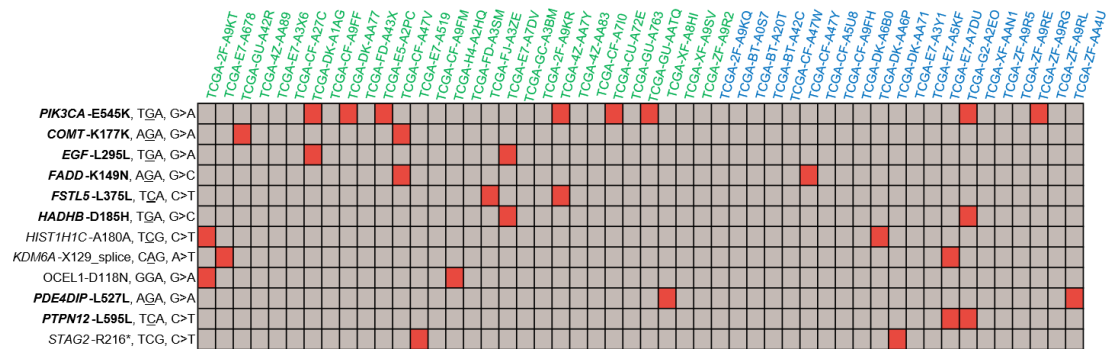
Reply-Figure 3. SNVs co-occurring with *FGFR3-S249C* in TCGA MIBC. The mutations in bold letters present a common APOBEC motif (TCW, W = T or A). Red and grey boxes mark presence and absence of the indicated mutations.

We did not perform the same analyses for NMIBC subjects because the RNA-seq derived mutations may contain more noise.

3. - Did the authors explore the mutational signature of other commonly co-mutated genes with *FGFR3* both S249C and others? For example *KDM6A*, *PIK3Ca*, *STAG2* to name a few.

Response:

Using MIBC TCGA dataset described by Robertson et al [2], we identified 12 SNVs significantly more frequent in subjects with recurrent *FGFR3* mutations compared to other MIBC subjects (Fisher’s exact test, $P < 0.05$): *COMT*-K177K, *EGF*-L295L, *FADD*-K149N, *FSTL5*-L375L, *HADHB*-D185H, *HIST1H1C*-A180A, *KDM6A*-X129_splice, *OCELI*-D118N, *PDE4DIP*-L527L, *PIK3CA*-E545K, *PTPN12*-L595L, and *STAG2*-R216* (**Reply-Figure 4**).



Reply-Figure 4. SNVs co-occurring with recurrent *FGFR3* mutations in TCGA MIBC. Mutations in bold letters present a common APOBEC motif (TCW, W = T or A). In green – 29 patients with *FGFR3*-S249C mutation, in blue - 21 patients with other recurrent *FGFR3* mutations. Red and grey boxes mark presence and absence of indicated mutations.

Among these co-mutations, we observed a total of 22 mutation events corresponding to a common APOBEC motif (TCW, W = A or T), including 15 events found in 29 *FGFR3*-S249C subjects (overall SNV load = 6144) and 7 events in 21 subjects with other recurrent *FGFR3* mutations (overall SNV load = 5041). The event frequency of *FGFR3* co-mutations related to APOBEC in *FGFR3*-S249C subjects was not statistically significant (*FGFR3*-S249C: 15/6144, 0.244%; other recurrent *FGFR3* mutations: 7/5041, 0.139%; Fisher’s exact test, $P = 0.28$), but there was a nearly two-fold increase in mutations within APOBEC motifs in S249C tumors over tumors with other recurrent *FGFR3* mutations.

We did not perform the same analyses for NMIBC subjects because the RNA-seq derived mutations may contain more noise.

As these results on co-mutations are not statistically significant (might due to limited sample size), they were not included in our updated manuscript.

Minor point:

- The data could have been presented more clearly to show the exact number of *FGFR3* S249 mutations that are associated with APOBEC Sig.13 and those that are not (instead of presenting % only).

Response:

Because both APOBEC signatures 2 and 13 showed similar trends for association with *FGFR3*-S249C mutation status (not shown) based on the data of recent bioRxiv manuscript preprint (Alexandrov, <https://doi.org/10.1101/322859>), we now use estimates of the total APOBEC mutagenesis for MIBC, without computational deconvolution into signature 2 and 13, which is not possible to do with confidence in all cases. For this analysis, we used an existing set of data defined as APOBEC induced mutation load (P-MACD), available in the Table S1 of the MIBC TCGA paper by Robertson et al [2].

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] 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.
- [4] Kwon OS, Hwang EJ, Bae JH, Park HE, Lee JC, Youn JI, et al. Seborrhic keratosis in the Korean males: Causative role of sunlight. *Photodermatol Photoimmunol Photomed* 2003;19:73–80. doi:10.1034/j.1600-0781.2003.00025.x.
- [5] Hafner C, Van Oers JMM, Hartmann A, Landthaler M, Stoehr R, Blaszyk H, et al. High frequency of *FGFR3* mutations in adenoid seborrhic keratoses. *J Invest Dermatol* 2006;126:2404–7. doi:10.1038/sj.jid.5700422.
- [6] Donahue TF. Genomic Characterization of Upper Tract Urothelial Carcinoma in Patients with Lynch Syndrome n.d.
- [7] Audenet F, Isharwal S, Cha EK, Donoghue MTA, Drill E, Ostrovnya 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.
- [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:e1007779. doi:10.1371/journal.pgen.1007779.
- [9] 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.
- [10] 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.

[11] Heidenreich B, Denisova E, Rachakonda S, Sanmartin O, Dereani T, Hosen I, et al. Genetic alterations in seborrheic keratoses. *Oncotarget* 2017;8:36639–49. doi:10.18632/oncotarget.16698.

2nd round review: letter to editor and point by point response



Paris, March 8th, 2019

Ms. No: EURUROL-D-18-01506

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

Dear Dr. Catto,

We would like to thank you and the reviewers for positive comments and providing new constructive suggestions. We have revised the paper accordingly and provided point-by-point responses reflecting changes in the manuscript. The modifications are highlighted in the manuscript.

We hope that after this second revision, our work meets the high standards for publication in *European Urology*.

Yours sincerely,

On behalf of the authors,

François Radvanyi

Point-by-point response to reviewers

We thank the reviewers for their comments and suggestions that helped to improve our manuscript.

Comment from Reviewer # 1:

“The authors once again present a brief communication looking at the correlation between APOBEC and FGFR3-249c mutations in patients with bladder cancer. They have addressed the majority of the reviewers concerns. The in vitro data does add strength to the manuscript. Still confusing is the differences found between NMI and MI analysis. The text regarding this portion is very confusing. Would recommend revisions to simplify what the authors are trying to state or just focus on the NMI samples.”

Response:

We thank the reviewer for this encouraging comment. We think that presenting the data for both NMIBC and MIBC is important and we did our best to simplify the presentation for MIBC.

Although *FGFR3* is less frequently mutated in MIBC compared to NMIBC (13% vs. 67%, respectively), we found the same proportion of *FGFR3*-S249C mutation (60%) among all recurrent *FGFR3* mutations in both NMIBC and MIBC. This suggests the existence of a potential common mechanism between the two tumor types that might be related to S249C mutation, specifically. We propose exposure to APOBEC mutagenesis, which generates S249C mutation, as a common mechanism. First, we observed an overall significant association between APOBEC-mediated mutagenesis and over-representation of S249C mutation, even if we had to consider APOBEC stratification as a confounder in MIBC. In addition, we found consistent results in both NMIBC and MIBC when looking for APOBEC mutagen through gene expression analysis.

However, we agree that our conclusions made in MIBC are based on a small number of *FGFR3* mutated samples and still need to be confirmed with a larger sample size in the future.

We modified this section (manuscript pages 3-4) as:

“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; 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.”

Comment from Reviewer # 4:

*“Please see the European Urology guidelines for the presentation of statistics: <https://doi.org/10.1016/j.eururo.2018.12.014>. In particular, please follow guideline 4.1 for precision. That includes figures (for example, many of the figures include *p* values to inappropriate precision).*

*Moreover, don't report *p* values for between comparisons and then for the difference between comparisons (e.g. figure 1D), report only the latter (i.e. nested ranks test). Please use exact statistics for comparisons where there are low cell counts (e.g. figure 1c).”*

Response:

We have carefully reviewed the recommended European Urology guidelines [1] for the presentation of statistics and modified our paper accordingly.

Detailed changes are stated below:

- a). Report *P*-value: All the *P*-values reported in the main figures (Figure 1B, 1C and Figure 2A) as well as supplementary figures (Figure S2, S3, S5 and Figure S7) were modified with the recommended appropriate precision, for instance, < 0.001, 0.004, 0.045, 0.13, 0.3, 1.
- b). Report percentage: All the percentages presented in our manuscript, Figure 1A and Figure 2A-F were reported to two significant figures, for example, 75%, 3.4%, 0.13%.

According to guideline section 3.1 about accepting a null hypothesis, we modified one statement interpreting result of Figure S3. We changed the sentence “Analysis of APOBEC mutation load in all tumors with recurrent *FGFR3* mutations did not show association with S249C mutation status (Fig.S3).” to the current one (manuscript page 3) “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)”.

We followed the reviewer’s suggestions, and reported only nested rank test *P*-value in Figure 1D to make it clear.

As suggested, for Figure 1C and Figure 2A we now present results for Fisher’s exact test instead of Chi-square test, with *P*-values remaining significant.

Reference

- [1] Assel M, Sjoberg D, Elders A, Wang X, Huo D, Botchway A, et al. Guidelines for Reporting of Statistics for Clinical Research in Urology. *Eur Urol* 2018;75:358–67. doi:10.1016/j.eururo.2018.12.014.