Title

Differences in mutations across tumour sizes in clear-cell renal cell carcinoma.

Authors

Monda, Steven

Carney, Benjamin

May, Allison

et al.

Publication Date

2024-09-12

DOI

10.1111/bju.16527

Peer reviewed

BJU Int 2025; **135:** 269–278 doi:10.1111/bju.16527

Original Article

BJUI
BJU International

Differences in mutations across tumour sizes in clear-cell renal cell carcinoma

Steven M. Monda[1,2], Benjamin W. Carney[2], Allison M. May[4], Shuchi Gulati[3], Simpa S. Salami[4], Thenappan Chandrasekar[1], Evan T. Keller[4], Nicolai A. Huebner[1,5], Ganesh S. Palapattu[4] and Marc A. Dall'Era[1]

[1]Department of Urologic Surgery, [2]Department of Radiology, [3]Division of Hematology and Oncology, UC Davis, Sacramento, CA, [4]Department of Urology, University of Michigan, Ann Arbor, MI, USA, and [5]Department of Urology, Medical University of Vienna, Vienna, Austria

Objective

To assess the distribution of key mutations across tumour sizes in clear-cell renal cell carcinoma (ccRCC), and secondarily to examine the prognostic impact of aggressive mutations in smaller ccRCCs.

Patient and Methods

The distribution of mutations (*VHL*, *PBRM1*, *SETD2*, *BAP1* and *CDKN2A* loss) across tumour sizes was assessed in 1039 ccRCCs treated with nephrectomy in cohorts obtained from the Tracking Cancer Evolution (TRACERx), The Cancer Genome Atlas (TCGA) and the Cancer Genomics of the Kidney (CAGEKID) projects. Logistic regression was used to model the presence of each mutation against size. In our secondary analysis, we assessed a subset of ccRCCs ≤7 cm for associations of key aggressive mutations (*SETD2*, *BAP1*, and *CDKN2A* loss) with metastasis, invasive disease and overall survival, while controlling for size. A subset of localised tumours ≤7 cm was also used to assess associations with recurrence after nephrectomy.

Results

On logistic regression, each 1-cm increase in tumour size was associated with aggressive mutations, *SETD2*, *BAP1*, and *CDKN2A* loss, at odds ratios (ORs) of 1.09, 1.10 and 1.19 ($P < 0.001$), whereas no significant association was observed between tumour size and *PBRM1* (OR 1.02; $P = 0.23$). *VHL* was mildly negatively associated with a 1-cm increase in size (OR 0.95; $P = 0.01$). Among tumours ≤7 cm, *SETD2* and *CDKN2A* loss were associated with metastatic disease at ORs of 3.86 and 3.84 ($P < 0.05$) while controlling for tumour size. *CDKN2A* loss was associated with worse overall survival, with a hazard ratio (HR) of 2.19 ($P = 0.03$). Among localised tumours ≤7 cm, *SETD2* was associated with worse recurrence-free survival (HR 2.00; $P = 0.03$).

Conclusion

Large and small ccRCCs are genomically different. Aggressive mutations, namely, *SETD2*, *BAP1*, and *CDKN2A* loss, are rarely observed in small ccRCCs and are observed more frequently in larger tumours. However, when present in tumours ≤7 cm, *SETD2* mutations and *CDKN2A* loss were still independently associated with invasive disease, metastasis, worse survival, and recurrence after resection, after controlling for size.

Keywords

bap1, cdkn2a, clear-cell renal cell carcinoma, genomics, renal cell carcinoma, setd2, survival, tcga, tracerx, tumour size

Introduction

Clear-cell RCC (ccRCC) is a heterogeneous disease, with large tumours behaving aggressively through local invasion, high rates of recurrence after resection, and metastasis at presentation [1,2]. However small primary ccRCCs rarely metastasise [3,4]. Many of these small tumours are diagnosed incidentally in older adults, and some can be safely observed through active surveillance [4,5]. It is not known how different aggressive mutations are distributed across tumour sizes in ccRCC. Determining the distribution of mutations across different ccRCC tumour sizes may help clarify why small ccRCCs behave so differently from large ccRCCs. Further, determining whether certain mutations portend worse outcomes in small ccRCCs may help risk stratify both biopsied and resected tumours and assist in determining candidates for active surveillance and for adjuvant therapy after resection.

Substantial efforts have been made to delineate mutational pathways to progression in ccRCC and the prognostic importance of key mutations. Loss of chromosome 3p is almost invariably the earliest evolutionary event in the development of sporadic ccRCC and is followed by mutations in other key drivers [6–8]. Subsequent mutations in the remaining copy of *VHL* is the most frequent next event, 'the second hit'. Second most frequent are *PBRM1* mutations, which are also highly prevalent and often truncal in ccRCC evolution [9,10]. Neither *VHL* nor *PBRM1* have consistently been associated with worse early ccRCC prognosis across several studies [10–12]. However, other common subsequent mutations, namely, *BAP1*, *SETD2*, and *CDKN2A* copy number loss (CDKN2A loss), are associated with metastatic disease, sarcomatoid features, and poor prognosis [9,13–16]. In particular, *CDKN2A* loss, included in the loss of chromosome 9p21.3, is highly selected for in ccRCC metastases and has been shown to establish metastatic competence in murine models [9,17,18].

We hypothesise that mutations that have been suggested to be early and truncal, namely, *VHL* and *PBRM1*, would be approximately equally prevalent across tumour sizes, while mutations associated with aggressive disease, namely, *SETD2*, *BAP1* and *CDKN2A* loss, would be predominantly observed in larger tumours, possibly occurring later in tumour evolution. Our primary aim was to assess the distribution of these mutations across tumour sizes in a combined cohort of tumours from three series: Tracking Cancer Evolution (TRACERx), The Cancer Genome Atlas (TCGA) and Cancer Genomics of the Kidney (CAGEKID). Our secondary aim was to assess invasiveness, metastatic status, and overall survival in smaller tumours (≤7 cm) that harbour aggressive mutations, *SETD2*, *BAP1*, and *CDKN2A* loss, to determine if these mutations influence prognosis in small masses. Further, we sought to assess recurrence rates in smaller localised tumours (≤7 cm) with *SETD2* and *BAP1* mutations that were not metastatic at presentation. Finally, we sought to assess radiographic predictors of aggressive mutations, *SETD2*, *BAP1* and *CDKN2A* loss, as previously investigated in smaller TCGA and Memorial Sloan Kettering Cancer Center (MSKCC) cohorts [19,20].

Materials and Methods

Cohorts

Our cohort included 106 ccRCC tumours from TRACERx Renal, 227 ccRCC tumours from TCGA-Kidney Renal Clear Cell Carcinoma (TCGA-KIRC), and 706 ccRCC tumours from CAGEKID. All data were publicly available and de-identified. Full details of specimen acquisition and clinical data for these cohorts can be found in their respective publications [8,10,16,21].

TRACERx is a recent cohort of cytoreductive and definitive nephrectomy specimens beginning in 2012, with multiple genomic regions sampled (median of 7 biopsies) per primary tumour. In obtaining sequencing from multiple regions, TRACERx reported a mutation as clonal if a mutation was present in all regions. TRACERx Renal data were accessed as published by Turaljilic et al. [10] (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5938372/ Supplementary Data S1 and S2).

TCGA-KIRC is a genomic cohort of single-region biopsy data obtained from either cytoreductive or definitive nephrectomy specimens between 2002 and 2013. TCGA mutational data were accessed as reported by Ricketts et al. [13] (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6075733/; Supplementary Data S1). All imaging was accessed via The Cancer Imaging Archive (TCIA) portal (wiki.cancerimagingarchive.net). Of 481 patients in TCGA with available mutation data, 227 were included for analysis and had preoperative cross-sectional imaging (MRI or CT) through TCIA. Tumour size was measured as the largest diameter on cross-sectional imaging as reviewed by an abdominal radiologist who was blinded to both clinical and mutational data. Radiological assessment of necrosis, calcifications, tumour thrombus and ill-defined margins was obtained by this radiologist based on prior definitions in the literature [19].

CAGEKID is a European genomic cohort of single-region biopsy data from nephrectomy specimens obtained between 1998 to 2014 [16,21]. Individual mutations were assessed in CAGEKID but not copy number alterations and thus *CDKN2A* loss was not reported. CAGEKID data, including reported size and clinical variables, were accessed as published by Vasudev et al. [16] and tumours were included if their size was reported (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10068441/; Supplementary Data S1 and S2) [16].

Mutation Selection

VHL, *PBRM1*, *SETD2*, *BAP1* and *CDKN2A* loss were selected for analysis based on being both the most common mutations (*VHL*, *PBRM1*, *SETD2* and *BAP1* in decreasing order) and the most commonly associated with aggressive disease (*SETD2*, *BAP1* and *CDKN2A* loss) [22–24]. Although other mutations (*MTOR*, *TP53*) carry prognostic importance for ccRCC, they are much rarer, with incidence rates of approximately 10% and 4%, respectively. Finally, as 3p loss is present in almost all sporadic ccRCCs (91% in TCGA, 95% in TRACERx), it was not included for separate analysis. In targeting just these five mutations, *VHL*, *PBRM1*, *SETD2*, *BAP1*, and *CDKN2A* loss, we sought to minimise multiple comparisons. Only *SETD2*, *BAP1*, and *CDKN2A* copy number loss were assessed for clinical outcomes, given these mutations are consistently associated with worse outcomes

[11,25]. *CDKN2A* loss and loss of chromosome 9p21.3 were considered synonymous.

Mutation plots were generated for the TRACERx, TCGA and CAGEKID cohorts, sorted by tumour size, to visualise the presence of individual mutations in each included tumour. For consistent nomenclature between cohorts, non-frameshift insertions and deletions were grouped as 'in-frame indels', non-frameshift substitutions and nonsynonymous single nucleotide variants were grouped as 'missense', and stop-gain mutations were grouped as 'nonsense'. Synonymous single nucleotide variants were not considered.

Statistics and Outcomes

All statistical analyses were performed in R version 4.3.0 using packages ggplot2, survival, survminor, dplyr and cowplot.

Logistic regression was used to model the presence of each mutation as a binary outcome against tumour size as a continuous variable within the combined cohort of TCGA, CAGEKID and TRACERx patients ('glm'). Bonferroni correction was used for multiple comparisons. In reporting percentage incidence of mutations, we used known radiographic T-stage thresholds of 4 cm and 7 cm in Fig. 1 and expanded this to 4 cm (T1a–T1b threshold), 7 cm (T1b–T2a threshold) and 10 cm (T2a–T2b threshold) for Fig. 2a,b. In support of the size difference, the Wilcoxon rank sum test was used to compare sizes between tumours with and without each mutation ('wilcox.test').

Subsets of the combined cohort with tumours ≤7 cm was used to assess the association between clinical outcomes and *SETD2*, *BAP1*, and *CDKN2A* loss in small masses. The threshold of 7 cm was chosen for T1–T2 RCC. Although 4 cm is the commonly accepted clinical threshold for a small tumour, aggressive mutations were exceptionally rare in tumours ≤4 cm in size, therefore, 7 cm was chosen to expand our sample.

Logistic regression was used to model the presence of metastatic disease, as defined by pathological N or M status, against *SETD2*, *BAP1*, and *CDKN2A* loss as binary variables and tumour size as a continuous variable ('glm'). Similarly, logistic regression was used to model the presence of invasive disease as defined by pathological T3 or greater, against *SETD2*, *BAP1*, and *CDKN2A* loss as binary variables and tumour size as a continuous variable ('glm'). Cox proportional hazards was used to model overall survival against *SETD2*, *BAP1*, and *CDKN2A* loss as binary variables and tumour size as a continuous variable ('coxph'). Tumours from CAGEKID were excluded from our logistic regression on invasiveness and metastatic disease, and Cox models for overall survival, as copy number analysis was not available and we did not have data on *CDKN2A* status for the CAGEKID patients.

For our analysis of recurrence in localised tumours ≤7 cm, we used tumours from TRACERx and CAGEKID with no known metastatic disease at time of resection, that is, N0 M0, and known follow-up. TCGA did not have readily available data on time of recurrence and was excluded. As mentioned previously, CAGEKID does not have copy number analysis and thus we did not assess the association between *CDKN2A* loss and recurrence. Although no copy number analysis was available, it was necessary to include CAGEKID to ensure an adequate sample size to observe recurrence.

Radiological predictors of mutations, including invasiveness, were assessed using methods described in a previous TCGA-TCIA study in 103 patients [19]. Fisher's exact test was used to assess associations between binary radiological features (necrosis, ill-defined capsule, vein thrombus, and calcifications) with presence of *SETD2*, *BAP1*, and *CDKN2A* loss mutations ('fisher.test').

Results

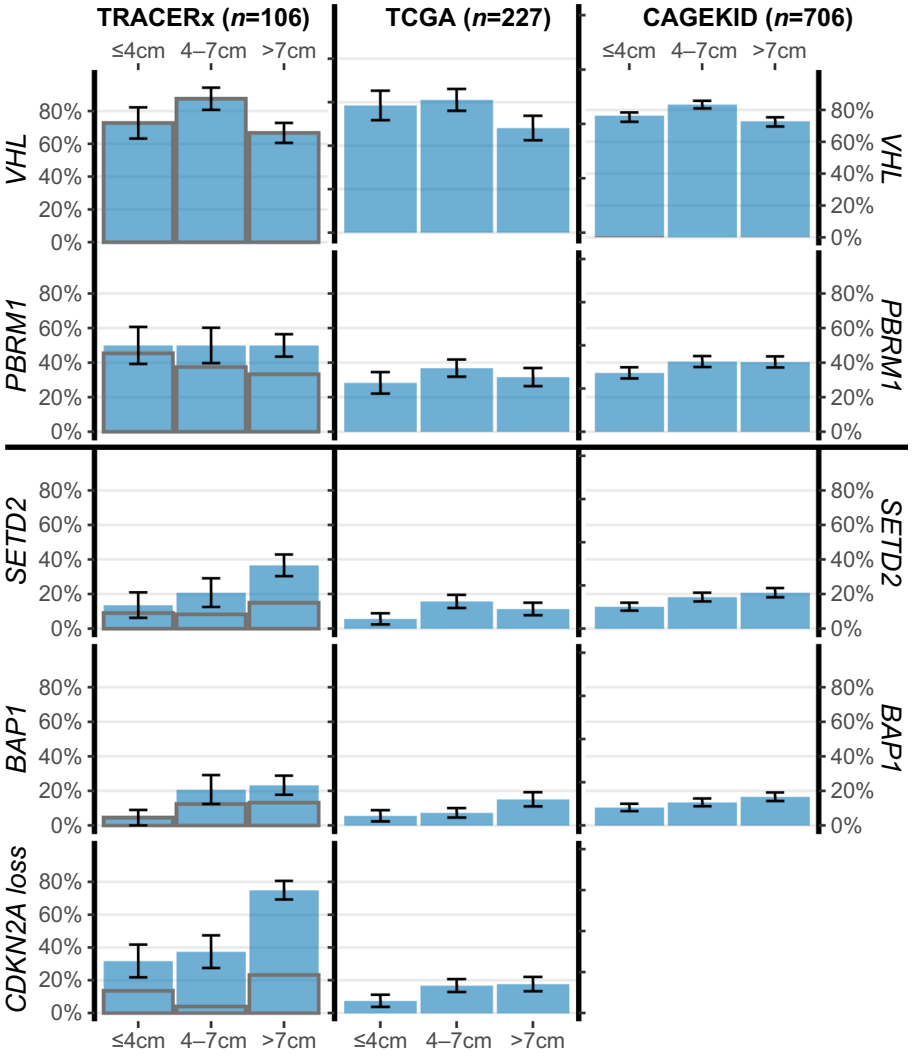
Associations of Size and Mutation Presence

Our final cohort included 1039 total patients, 106 from TRACERx, 227 from TCGA, and 706 from CAGEKID (Table 1). Figures 2a,b and S1 show the presence of mutations for each tumour stratified by tumour size in the TRACERx cohort, the TCGA cohort and the CAGEKID cohort, respectively. Clonal mutations in TRACERx, which were observed in all biopsied tumour regions, are further indicated in Fig. 2a with bolded outlines. Only six *SETD2* mutations and eight *BAP1* mutations were observed among 99 tumours that were <3 cm in all cohorts, 6% and 8%. In contrast, among 148 tumours >10 cm, 39 and 31, 26% and 21%, had *SETD2* and *BAP1* mutations, respectively. Similarly, in tumours with copy number data, 17% of tumours <3 cm had *CDKN2A* loss, while 49% of tumours >10 cm had *CDKN2A* loss. The incidences of *VHL* and *PBRM1* mutations were approximately 60% and 40% across different tumour sizes and across studies (Fig. 1 and Table 2).

On logistic regression for the combined cohort of 1039 tumours, each 1-cm increase in tumour size was associated with aggressive mutations, *SETD2*, *BAP1*, and *CDKN2A* loss, with odds ratios (ORs) of 1.09, 1.10 and 1.19 ($P < 0.001$), results which remained significant after Bonferroni correction, whereas no significant association was observed between tumour size and *PBRM1* (OR 1.02; $P = 0.23$). *VHL* was mildly negatively associated with 1-cm increase in size (OR 0.95; $P = 0.01$).

Tumours with *SETD2*, *BAP1* and *CDKN2A* loss mutations were larger than tumours without on Wilcoxon rank sum test ($P < 0.001$), with mean sizes of 7.5 cm vs 6.3 cm, 7.6 cm vs 6.4 cm, and 8.8 cm vs 6.4 cm. Tumours with *VHL* mutations were slightly smaller than those without (6.4 cm vs 7.0 cm),

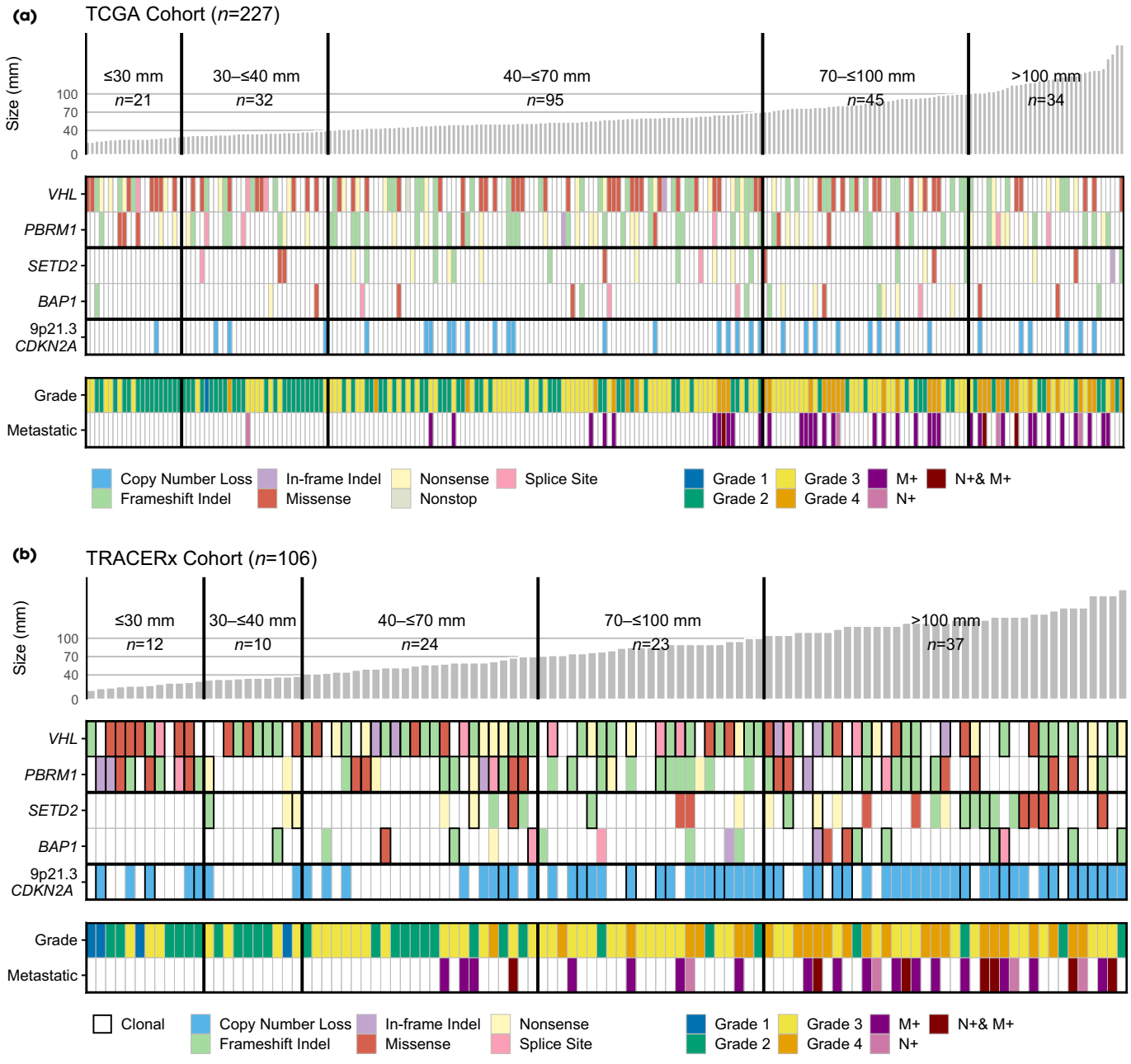
Fig. 1 The proportion of tumours with given mutations within each cohort among cT1a, cT1b, and cT2 size thresholds. Copy number and *CDKN2A* loss were not available for Cancer Genomics of the Kidney (CAGEKID). Within Tracking Cancer Evolution (TRACERx), the grey-outlined segment represents the proportion each mutation that was clonal- i.e. detected in every biopsied region.



but this was not significant after correction for multiple comparisons ($P = 0.029$, Bonferroni-corrected to $P = 0.15$). No significant tumour size difference was observed between tumours with and without *PBRM1* mutations on Wilcoxon rank sum test ($P = 0.12$), with mean sizes of 6.7 cm vs 6.4 cm.

To mitigate bias associated with combining cohorts with different sampling protocols, we also analysed size in the multi-region biopsy cohort (TRACERx) and the single-region biopsy cohorts (TCGA and CAGEKID) separately using a Wilcoxon rank sum test comparing mutated and unmutated tumours. *CDKN2A* loss tumours were 3.5 and 0.7 cm larger compared to wild-type tumours in the multi-region and single-region groups, respectively ($P < 0.001$ and $P = 0.12$). *SETD2* mutated tumours were 2.9 and 0.7 cm larger ($P < 0.001$ and $P = 0.01$) *BAP1* mutated tumours were 1.7 cm larger and 1.0 cm larger ($P = 0.08$ and $P = 0.003$). *PBRM1* mutated tumours were not significantly different in either group ($P = 0.77$ and $P = 0.17$). *VHL* mutated tumours were slightly smaller: 1.0 and 0.6 cm ($P = 0.23$ and $P = 0.05$). Similarly, to mitigate bias associated with the different sampling protocols for different cohorts, the logistic regression analysis was repeated after controlling for cohort (TRACERx, TCGA, CAGEKID). *CDKN2A* loss, *BAP1* and *SETD2* remained positively associated with increasing tumour size (hazard ratio [HR] 1.16, 1.09 and 1.08; $P < 0.001$) and *VHL* remained mildly negatively associated with increasing tumour size (HR 0.95; $P < 0.001$) after controlling for cohort.

Fig. 2 (**a, b**) Tumour size, mutational status, grade, and metastatic status at presentation among clear-cell RCC tumours in Tracking Cancer Evolution (TRACERx) and The Cancer Genome Atlas (TCGA) sorted by tumour size. Clonal mutations are outlined in bold for TRACERx tumours when a mutation was present in all biopsies.



Invasiveness and Metastasis in ccRCC ≤7 cm

Among the 194 tumours with copy number data (TRACERx and TCGA) that were ≤7 cm, 16 were metastatic at diagnosis (Fig. S2a–c). *SETD2* mutations and *CDKN2A* loss were associated with metastatic disease on logistic regression after controlling for tumour size, with ORs of 3.86 and 3.84 ($P < 0.05$; Table 3). Of the 16 tumours that were metastatic with a primary tumour ≤7 cm, seven had *SETD2* mutations, eight had *CDKN2A* loss, and only four had neither a *SETD2* nor *CDKN2A* loss. *SETD2* was associated with invasive disease in tumours ≤7 cm, with an OR of 3.17 ($P = 0.024$). *CDKN2A loss* was associated with worse overall survival on Cox proportional hazards after controlling for tumour size, with an HR of 2.19 ($P < 0.05$). Kaplan–Meier curves for overall survival for tumours ≤7 cm and >7 cm with and

Table 1 Cohort characteristics.

	TRACERx *N* = 106 *n* (%) or median (IQR)	TCGA *N* = 227 *n* (%) or median (IQR)	CAGEKID *N* = 706 *n* (%) or median (IQR)
Age, years	64 (57–69)	59 (51.5–70)	61 (54–68)
Race			
White	57 (56)	203 (90)	†
Black, African-American, African-Caribbean	37 (37)	21 (9)	
Asian	7 (7)	2 (1)	
Female	33 (33)	78 (34)	270 (38)
Tumour size, mm	79 (46–120)	58 (42–84)	55 (40–80)
Staging			
T1a	20 (19)	69 (30)	196 (28)
T1b	15 (14)	46 (20)	155 (22)
T2	7 (7)	23 (10)	78 (11)
≥T3	56 (53)	86 (38)	261 (37)
T4	8 (8)	3 (1)	10 (1)
N+	11 (10)	7 (3)	33 (5)
M+	24 (23)	38 (17)	60 (8)
Grade			
1	4 (4)	1 (0)	75 (11)
2	28 (26)	92 (41)	303 (43)
3	49 (46)	96 (42)	246 (35)
4	25 (24)	38 (17)	80 (11)

†*Not reported but for CAGEKID patients, 31% from the UK and 69% from Eastern Europe.*
CAGEKID, Cancer Genomics of the Kidney; IQR, interquartile range; TCGA, The Cancer Genome Atlas; TRACERx, Tracking Cancer Evolution.

Table 2 Percentage of clear-cell RCC tumours with mutations at different tumour sizes.

Cohort:	Incidence of mutations at different ccRCC tumour sizes											
	TRACERx (*n* = 106)			TCGA (*n* = 227)			CAGEKID (*n* = 706)			Combined (*n* = 1039)		
Tumour size	≤4 cm	>4–≤7 cm	>7 cm	≤4 cm	>4–≤7 cm	>7 cm	≤4 cm	>4–≤7 cm	>7 cm	≤4 cm	>4–≤7 cm	>7 cm
N	22	24	60	53	95	79	220	246	240	295	365	379
Early mutations												
*VHL**A, % (*n*)	73 (16)	88% (21)	67 (40)	58 (31)	61 (58)	48 (38)	76 (166)	83 (205)	73 (174)	72 (213)	78 (284)	66 (252)
PBRM1, % (*n*)	50 (11)	50 (12)	50 (30)	28 (15)	37 (35)	32 (25)	34 (75)	41 (100)	40 (97)	34 (101)	40 (147)	40 (152)
Aggressive mutations												
SETD2†, % (*n*)	14 (3)	21 (5)	37 (22)	6 (3)	16 (15)	11 (9)	13 (28)	18 (45)	21 (50)	12 (34)	18 (65)	21 (81)
BAP1†, % (*n*)	5 (1)	21 (5)	27 (16)	6 (3)	7 (7)	15 (12)	10 (23)	13 (33)	17 (40)	9 (27)	12 (45)	17 (66)
CDKN2A loss†, % (*n*)	32 (7)	38 (9)	75 (45)	8 (4)	17 (16)	18 (14)	–	–	–	15 (11)†	21 (25)†	42 (59)†

**P* < *0.05 for size negatively associated with presence of mutation on logistic regression (VHL).* †*P* < *0.001 for size positively associated with presence of mutation on logistic regression (SETD2, BAP1, CDKN2A loss).*
CAGEKID, Cancer Genomics of the Kidney; TCGA, The Cancer Genome Atlas; TRACERx, Tracking Cancer Evolution.

Table 3 Results of multivariable logistic regression in clear-cell RCC tumours ≤7 cm from TRACERx and TCGA (*n* = 194) for pathological invasive disease (≥pT3) and pathological metastatic disease (N+ and/or M+) at time of diagnosis, and results of the multivariable Cox model for overall survival.

	Invasive disease		Metastasis at presentation†		Overall survival		Disease-free survival‡	
	OR (95% CI)	*P*	OR (95% CI)	*P*	HR (95% CI)	*P*	HR (95% CI)	*P*
SETD2	3.17 (1.17–8.77)	0.024*	3.86 (1.10–13.25)	0.031*	1.09 (0.50–2.38)	0.83	2.00 (1.09–3.68)	0.026*
BAP1†	2.16 (0.60–7.58)	0.23	N/A	N/A	0.44 (0.10–1.87)	0.26	1.68 (0.84–3.37)	0.23
CDKN2A loss‡	1.56 (0.60–3.95)	0.35	3.84 (1.15–12.89)	0.027*	2.19 (1.09–4.41)	0.028*	N/A	N/A
Size, +1 cm	2.57 (1.84–3.78)	<0.001*	2.76 (1.61–5.43)	<0.001*	1.31 (1.01–1.71)	0.043*	1.32 (1.10–1.60)	0.004*

Also recurrence-free survival results of multivariable Cox regression in localised (N0 M0) ccRCC tumours ≤7 cm from Tracking Cancer Evolution (TRACERx) and Cancer Genomics of the Kidney (CAGEKID; n = 459) examining time to recurrence after nephrectomy. All models included each mutation and tumour size as predictors of the outcome of interest. **P* < *0.05.* †*No metastases observed for BAP1 mutated tumours ≤7 cm.*
‡*Recurrence free survival included CAGEKID cohort where CDKN2A loss data was not available.*
HR, hazard ratio; OR, odds ratio.

without *SETD2* mutations, *BAP1* mutations, and *CDKN2A* loss are shown in Fig. S2a–c, respectively.

Recurrence for Localised ccRCC ≤7 cm

Among 459 tumours ≤7 cm in size with disease-free survival data (TRACERx and CAGEKID) in patients with localised disease (no metastatic or nodal involvement) at time of nephrectomy, 56 recurred (Fig. S3a,b). *SETD2* was associated with worse disease-free survival, with an HR of 2.00 ($P = 0.03$) after controlling for tumour size (Table 3). Of patients with a *SETD2* mutation, 21% had recurrence compared to 11% of those without. Associations between *CDKN2A* loss and recurrence were not assessed given the lack of copy number data for CAGEKID. Kaplan–Meier curves for recurrence-free survival for tumours ≤7 cm and >7 cm with and without *SETD2* mutations and *BAP1* mutations are shown in Figs S3a and S3c. Kaplan–Meier curves for disease-free survival for tumours ≤7 and >7 cm with and without *SETD2* mutations and *BAP1* mutations are shown in Figs S3a and S3b.

Radiological Associations with Mutations

Among 217 TCGA patients with adequate cross-sectional imaging, tumours with *BAP1* mutations had a significantly higher proportion of radiographically observed ill-defined margins compared to those without (36.4% vs 14.9%; $P = 0.031$) and *SETD2* mutation was associated with a near significant higher proportion of ill-defined margins (30.8% vs 15.2%; $P = 0.056$) compared to tumours without *SETD2* mutations (Table S1). Of tumours with *CDKN2A* loss, 29.0% had a vein thrombus vs 14.0% of tumours without *CDKN2A* loss, but the difference did not reach statistical significance ($P = 0.060$).

Discussion

In this study, we established that certain aggressive mutations associated with metastatic and invasive ccRCC, specifically, *SETD2*, *BAP1*, and *CDKN2A* loss, are rare in small tumour sizes and are increasingly common at larger tumour sizes. In contrast, mutations in *VHL* and *PBRM1* were fairly evenly distributed across tumour sizes; h VHL and *PBRM1* have been shown to occur earlier and to be more truncal in TRACERx. Similar results were observed across all cohorts - TRACERx, CAGEKID and TCGA - with a trend towards higher rates of *SETD2*, *BAP1*, and *CDKN2A* loss in larger tumours . Although rarer in smaller tumours, we observed associations between *SETD2* and *CDKN2A* loss with worse prognosis across several outcomes (invasiveness, metastasis, overall survival and recurrence-free survival) in tumours ≤7 cm after controlling for size.

The TRACERx cohorts had a higher incidence of mutation detection compared to the TCGA and CAGEKID cohorts (Fig. 1 and Table 2). This was particularly striking for larger masses and aggressive mutations, with TRACERx detecting approximately twofold more *BAP1*, *SETD2* and *CDKN2A* loss mutations. This difference in mutation detection rates can also be inferred from Figs 2a,b and S1. This in part highlights the intra-tumoural heterogeneity of ccRCC and the limitations of a single-region biopsy protocol (TCGA and CAGEKID) to characterise a tumour's mutations compared to multiple biopsies, with a range of 4–8 biopsies needed to capture the majority of mutations [10]. Data from TCGA and CAGEKID are probably limited by false negatives, particularly for large heterogeneous tumours. Despite these shortcomings, TCGA and CAGEKID remain relevant as the largest available genomics databases for ccRCC, and we felt justified in their inclusion.

A limitation of this study was the combining of cohorts with different protocols, which can insert bias. As discussed above, multiple biopsies were taken in the TRACERx cohort, compared to a single-region biopsy in TCGA. Tissue processing and definitions of copy number loss differed between the studies. Nevertheless, we report a similar trend in increased *SETD2*, *BAP1*, and *CDKN2A* loss with tumour size in both cohorts (Fig. 1 and Table 2), and we feel that our study was strengthened by the inclusion of tumours from three large and established datasets.

TRACERx did report whether a mutation was 'clonal' if present in a high fraction of all biopsy regions from the tumour; a median of 7 biopsies were taken per tumour. In the TRACERx cohort, 100% of *VHL* mutations and 74% of *PBRM1* mutations were clonal, whereas only 43% and 29% of *SETD2* mutations and *CDKN2A* loss were clonal. In all, 60% of *BAP1* mutations were clonal. This suggests that *SETD2* and *CDKN2A* loss sub-clones develop later in tumour evolution, possibly when a tumour has reached a certain size. However, whether a mutation occurring later in tumour evolution also means it occurs at a larger tumour size warrants future investigation.

To our knowledge, the only previous study examining an association between tumour size and the presence of specific mutations was by Karlo et al. [20] Although no significant difference across size was observed, all mutations were rare, with 14 *BAP1* mutations and 17 *SETD2* mutations, suggesting the study may have been underpowered to observe a difference. Ueno et al. [26] did establish increased copy number variations, including chromosome 9p21.3 loss, in large vs small tumours, consistent with our results, but did not examine specific mutations. Turajlic et al. [10] did demonstrate that '*VHL* monodriver' tumours were smaller, with absent *SETD2* and *BAP1* mutations, consistent with our results.

Our exploratory results also suggest that, when present in small tumours, these aggressive mutations are associated with worse outcomes after controlling for size. In tumours ≤7 cm, we observed increased metastasis with *SETD2* mutations and *CDKN2A* loss, and correspondingly worse overall survival with *CDKN2A* copy number loss. We also observed an association between *SETD2* mutations and invasive disease in tumours ≤7 cm in size. Strikingly, 12 of the 16 patients with data on *CDKN2A* status who had metastases with tumours ≤7 cm, had either *SETD2* mutations or *CDKN2A* loss. Although *SETD2* mutations and *CDKN2A* loss were more common in large tumours, when present in smaller tumours, they still appeared to portend worse prognosis. Furthermore, in localised ccRCC tumours ≤7 cm, N0 and M0, *SETD2* was associated with recurrence after nephrectomy. In essence, smaller tumours with these mutations may behave more like large tumours; for instance, the overall survival curves of smaller tumours with an *SETD2* mutation or *CDKN2A* loss more closely paralleled larger tumours with or without these mutations (Fig. S2a,c). Similarly, recurrence-free survival for smaller tumours with *SETD2* mutations more closely paralleled those for larger tumours (Fig. S3a). Finally, overall survival and recurrence-free survival associations without accounting for stage may be of limited clinical applicability. However, when overall survival was examined for all tumours (including those >7 cm), *CDKN2a* loss remained significant (HR 1.63; $P = 0.035$) after controlling for nodal and metastatic status.

Manley et al. [27] also examined the association of individual mutations with recurrence in resected ccRCC ≤4 cm and found that patients who died from ccRCC had a *SETD2* mutation rate of 33.3%, compared to 9.9% among those with non-recurrent tumours. Although not statistically significant, this is consistent with our results. Vasudev et al. [16] examined the risk of recurrence in resected ccRCC using the same CAGEKID data and found a *SETD2* mutation rate of 24.7% in patients who had recurrence vs 15.2% in those without recurrence; the difference was not significant but again was consistent with our results. Copy number losses, and by extension *CDKN2A* loss, were not examined in the studies by Manley et al. or Vasudev et al.

BAP1 mutations were rare in ccRCC ≤7 cm in the TRACERx and TCGA cohorts. In a larger cohort of ccRCC ≤4 cm, Kapur et al. [28] demonstrated an association between *BAP1* and metastasis, although the overall incidence of *BAP1* mutations was still very low, at 55 of 715 tumours. Given the rarity of *BAP1* in smaller tumours, it is possible our study was underpowered to observe an association with overall survival or metastasis for small *BAP1* mutated tumours. Alternatively, *BAP1* may be more highly associated with large and fast growing tumours whose outcomes are more closely linked to their size.

A limitation of our study is the inclusion of tumours ≤7 cm in our analysis of 'smaller' masses. A small renal mass is generally considered to be ≤4 cm. However, *SETD2*, *BAP1*, and *CDKN2A* loss were very rare in tumours ≤4 cm, and metastasis at this size was exceptionally rare. We did not have an adequate sample size to examine outcomes association for ccRCC ≤4 cm and instead examined ≤7 cm as the T1–T2 stage threshold.

Even by including all tumours ≤7 cm, aggressive mutations and metastasis were rare. As such, we could only include so many predictor variables in our logistic regression and Cox models; we chose to focus on *SETD2*, *BAP1*, *CDKN2A* loss, and size, but omitted age, TNM staging, and grade in our models. This limitation would require a larger sample size to overcome. Further, the genomics of ccRCC have complex interactions, with some studies suggesting the mutual exclusivity of *PBRM1* and *BAP1* mutations and others suggesting worse prognosis when certain mutations are paired [10,16]. Similar to additional clinical variables, these complex interactions would probably require a larger sample size to tease out, but still represent a limitation of our analysis. Further, this study considered individual mutations but did not group patients based on evolutionary subtype ('multiple clonal driver', 'BAP1 driven', 'PBRM1 → SCNA', etc) as defined in TRACERx; indeed, it is not clear how these subtypes would be defined for single-region biopsy cohorts with copy number data, such as TRACERx and CAGEKID [9,10]. Finally, we performed multiple comparisons in our small sample of tumours ≤7 cm and false discoveries are a possibility.

We chose to focus on ccRCC and not to include papillary RCC or other subtypes as ccRCC is more common, more genomic data are available, and the common ccRCC mutations are better characterised. Additionally, we chose to focus on only a small subset of the most common mutations in ccRCC, and did not examine other rare mutations that are probably also prognostic, such as 14q loss, *TP53* and *KDM5C*. All of these are important areas of future work and limitations of the existing study.

In addition, we did not examine the role of *VHL* wild-type tumours where the remaining *VHL* copy remains unmutated, which have been shown to behave more aggressively, with higher frequency of sarcomatoid differentiation [10]. *VHL* was mildly negatively associated with tumour size, suggesting that *VHL* wild type might be more common in larger tumours. The association of *VHL* wild-type tumours with outcomes was beyond the scope of our study.

Importantly, our results do not establish a causal link between size and presence of aggressive *SETD2*, *BAP1*, and *CDKN2A* loss mutations. Tumours with these mutations may grow faster and attain a larger size, or these mutations may

arise once tumours reach a certain size. More complex genomic analysis may help delineate growth and evolutionary timing of these mutations, but was beyond the scope of this study.

We saw similar results to those obtained by others in the association of radiographic features with the presence of *SETD2* and *BAP1* mutations. *SETD2* and *BAP1* mutated tumours had an approximately twofold higher incidence of ill-defined margins [19,20]. *CKDN2A* copy number loss tumours also had a nearly twofold higher incidence of vein thrombus. Other imaging features associated with aggressive disease were more common in tumours with these mutations as well, but none reached statistical significance. Although many of these radiographic features are uncommon in small tumours (for instance, only five tumours ≤7 cm had ill-defined margins), our data support the importance of factoring in such radiographic findings when considering surveillance of smaller masses.

As adjuvant therapies established for patients with resected ccRCC [29], many of whom are cured with surgery alone, determining who is at higher risk of recurrence will allow more targeted initiation of therapy. Further, as more small ccRCC tumours are biopsied, a better understanding of genomic risk may help determine the need for surgery or ablation. Our results suggest that *SETD2* and *CDKN2A* loss may in the future be useful biomarkers to guide therapy. Future directions for our work include assessing whether *CDKN2A* loss may help identify who would benefit from adjuvant immune checkpoint inhibitors.

To conclude, aggressive mutations, *SETD2*, *BAP1*, and *CDKN2A* loss, are rare in small ccRCC tumours, and the incidence of these mutations increases with tumour size. Other truncal and early mutations (*VHL*, *PBRM1*) are frequently observed across all tumour sizes. In tumours ≤7 cm, *SETD2 mutations* were associated with invasive disease, and *SETD2* mutations and *CDKN2A* loss were both associated with metastatic disease. *CDKN2A* loss was associated with worse overall survival in tumours ≤7 cm. *SETD2* was associated with recurrence after nephrectomy in tumours ≤7 cm. Whether certain mutations are associated with faster growing tumours or whether tumours acquire these mutations once they reach larger sizes requires further investigation. This work suggests genomic differences may explain the more indolent behaviour of small ccRCC tumours compared to the aggressive nature of large tumours.

Acknowledgements

Steven Monda is supported by the National Institutes of Health/National Cancer Institute Advanced Training in Urologic Oncology Grant (T32 CA180984) and the Clark Family Fellowship in Kidney Cancer Research. The authors would also like to thank Nick Steinauer, whose coding insight helped make good figures better and Denis Whelan whose common-sense statistical advice was far from common.

Disclosure of Interests

The authors do not have conflicts of interest to report and have completed the requested ICMJE Disclosure Form.

References

1 **Bhindi B, Thompson RH, Lohse CM et al.** The probability of aggressive versus indolent histology based on renal tumor size: implications for surveillance and treatment. *Eur Urol* 2018; 74: 489–97
2 **Leibovich BC, Blute ML, Cheville JC et al.** Prediction of progression after radical nephrectomy for patients with clear cell renal cell carcinoma. *Cancer* 2003; 97: 1663–71
3 **Monda SM, Lui HT, Pratsinis MA, Chandrasekar T, Evans CP, Dall'Era MA.** The metastatic risk of renal cell carcinoma by primary tumor size and subtype. *Eur Urol Open Sci* 2023; 52: 137
4 **Sanchez A, Feldman AS, Hakimi AA.** Current management of small renal masses, including patient selection, renal tumor biopsy, active surveillance, and thermal ablation. *J Clin Oncol* 2018; 36: 3591–600
5 **Beisland C, Hjelle KM, Reisæter LAR, Bostad L.** Observation should be considered as an alternative in management of renal masses in older and comorbid patients. *Eur Urol* 2009; 55: 1419–29
6 **Mitchell TJ, Turajlic S, Rowan A, Swanton C, Campbell PJ.** Timing the landmark events in the evolution of clear cell renal cell cancer: TRACERx renal. *Cell* 2018; 173: 611–23.e17
7 **Hsieh JJ, Le VH, Oyama T, Ricketts CJ, Ho TH, Cheng EH.** Chromosome 3p loss–orchestrated VHL, HIF, and epigenetic deregulation in clear cell renal cell carcinoma. *J Clin Oncol* 2018; 36: 3533–9
8 **Creighton CJ, Morgan M, Gunaratne PH et al.** Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* 2013; 499: 43–9
9 **Turajlic S, Xu H, Litchfield K et al.** Tracking cancer evolution reveals constrained routes to metastases: TRACERx renal. *Cell* 2018; 173: 581
10 **Turajlic S, Xu H, Litchfield K et al.** Deterministic evolutionary trajectories influence primary tumor growth: TRACERx renal. *Cell* 2018; 173: 595–610.e11
11 **Kapur P, Peña-Llopis S, Christie A et al.** Effects on survival of BAP1 and PBRM1 mutations in sporadic clear-cell renal-cell carcinoma: a retrospective analysis with independent validation. *Lancet Oncol* 2013; 14: 159–67
12 **Motzer RJ, Martini JF, Mu XJ et al.** Molecular characterization of renal cell carcinoma tumors from a phase III anti-angiogenic adjuvant therapy trial. *Nat Commun* 2022; 13: 5959
13 **Ricketts CJ, De Cubas AA, Fan H et al.** The cancer genome atlas comprehensive molecular characterization of renal cell carcinoma. *Cell Rep* 2018; 23: 313–26.e5
14 **Bakouny Z, Braun DA, Shukla SA et al.** Integrative molecular characterization of sarcomatoid and rhabdoid renal cell carcinoma. *Nat Commun* 2021; 12(1): 1–14
15 **Lin E, Zhu P, Ye C et al.** Integrative analysis of the genomic and immune microenvironment characteristics associated with clear cell renal cell carcinoma progression: implications for prognosis and immunotherapy. *Front Immunol* 2022; 13: 830220
16 **Vasudev NS, Scelo G, Glennon KI et al.** Application of genomic sequencing to refine patient stratification for adjuvant therapy in renal cell carcinoma. *Clin Cancer Res* 2023; 29(7): 1220–31
17 **Carbone F, Huang JK, Perelli L et al.** 9p21 loss defines the evolutionary patterns of aggressive renal cell carcinomas. *bioRxiv* 2020

18 **Perelli L**, **Carbone F**, **Zhang L et al.** Interferon signaling promotes tolerance to chromosomal instability during metastatic evolution in renal cancer. *Nat Cancer* 2023; 4: 984–1000
19 **Shinagare AB**, **Vikram R**, **Jaffe C et al.** Radiogenomics of clear cell renal cell carcinoma: preliminary findings of the cancer genome atlas-renal cell carcinoma (TCGA-RCC) imaging research group. *Abdom Imaging* 2015; 40: 1684–92
20 **Karlo CA**, **Di Paolo PL**, **Chaim J et al.** Radiogenomics of clear cell renal cell carcinoma: associations between CT imaging features and mutations. *Radiology* 2014; 270: 464–71
21 **Scelo G**, **Riazalhosseini Y**, **Greger L et al.** ARTICLE variation in genomic landscape of clear cell renal cell carcinoma across Europe. *Nat Commun* 2014; 5: 5135
22 **Cotta BH**, **Choueiri TK**, **Cieslik M et al.** Current landscape of genomic biomarkers in clear cell renal cell carcinoma. *Eur Urol* 2023; 84: 166–75
23 **D'Avella C**, **Abbosh P**, **Pal SK**, **Geynisman DM**. Mutations in renal cell carcinoma. *Urol Oncol* 2020; 38: 763–73
24 **Bui TO**, **Dao VT**, **Nguyen VT**, **Feugeas JP**, **Pamoukdjian F**, **Bousquet G**. Genomics of clear-cell renal cell carcinoma: a systematic review and meta-analysis. *Eur Urol* 2022; 81: 349–61
25 **Kim BJ**, **Kim JH**, **Kim HS**, **Zang DY**. Prognostic and predictive value of VHL gene alteration in renal cell carcinoma: a meta-analysis and review. *Oncotarget* 2017; 8: 13979–85
26 **Ueno D**, **Xie Z**, **Boeke M et al.** Genomic heterogeneity and the small renal mass. *Clin Cancer Res* 2018; 24: 4137
27 **Manley BJ**, **Reznik E**, **Ghanaat M et al.** Characterizing recurrent and lethal small renal masses in clear cell renal cell carcinoma using recurrent somatic mutations. *Urol Oncol* 2019; 37: 12–7
28 **Kapur P**, **Zhong H**, **Araj E et al.** Predicting oncologic outcomes in small renal tumors. *Eur Urol Oncol* 2022; 5: 687–94
29 **Choueiri TK**, **Tomczak P**, **Park SH et al.** Adjuvant pembrolizumab after nephrectomy in renal-cell carcinoma. *N Engl J Med* 2021; 385: 683–94

Correspondence: Steven M. Monda, Michigan Department of Urology, 1500 E Medical Center Dr, Ann Arbor, MI 48109, USA.
e-mail: smonda@med.umich.edu

Abbreviations: CAGEKID, Cancer Genomics of the Kidney; ccRCC, clear-cell RCC; HR, hazard ratio; OR, odds ratio; TCGA, The Cancer Genome Atlas; TCIA, The Cancer Imaging Archive; TRACERx, Tracking Cancer Evolution.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Radiographic predictors of mutation status (n=217).

Figure S1. Tumour size, mutational status, grade, and metastatic status at presentation among ccRCC tumours in CAGEKID sorted by tumour size.

Figure S2. (a–c) Overall survival curves for renal masses larger and smaller than 7 cm by SETD2 mutation, BAP1 mutation, and CDKN2A copy-number loss status.

Figure S3. (a, b) Disease free survival curves for localised ccRCCs after nephrectomy by larger and smaller than 7 cm and by SETD2 mutation and BAP1 mutation status.

