Supplementary Information for:

BAP1 loss defines a new class of renal cell carcinoma

Samuel Peña-Llopis[1,2,3], Silvia Vega-Rubin-de-Celis[1,2,3], Arnold Liao[4], Nan Leng[4], Andrea Pavía-Jiménez[1,2,3], Shanshan Wang[1,2,3], Toshinari Yamasaki[1,2,3], Leah Zhrebker[1,2,3], Sharanya Sivanand[1,2,3], Patrick Spence[1,2,3], Lisa Kinch[5], Tina Hambuch[4], Suneer Jain[4], Yair Lotan[6], Vitaly Margulis[6], Arthur I. Sagalowsky[6], Pia Banerji Summerour[3,7], Wareef Kabbani[7], S. W. Wendy Wong[9], Nick Grishin[5], Marc Laurent[4], Xian-Jin Xie[3], Christian Haudenschild[4], Mark T. Ross[9], David R. Bentley[9], Payal Kapur[8], and James Brugarolas[1,2,3†]

Departments of [1]Internal Medicine, [2]Developmental Biology, [5]Biochemistry, [6]Urology, [7]Clinical genetics, and [8]Pathology, UT Southwestern Medical Center, Dallas, TX

[3]Simmons Comprehensive Cancer Center, UT Southwestern Medical Center, Dallas, TX

[4]Illumina Inc., San Diego, CA

[9]Illumina Cambridge Ltd., Essex, UK

[†]To whom correspondence should be addressed (james.brugarolas@utsouthwestern.edu)

This PDF file includes:

- **Supplementary Note**
- **Supplementary Figures 1-10**
- **Supplementary Tables 1-10**

Other online materials:

- **Supplementary Data 1-4**

Supplementary Note

Tissue processing and nucleic acid extraction. Tissues were excised while on dry ice. Flanking sections were oriented using pathology dyes (StatLab Medical Products), fixed in formalin, embedded in paraffin, stained with hematoxilin/eosin and reviewed by a pathologist, who scored histology, Fuhrman grade, percentage tumor (or normal), and necrosis. Tumors with necrosis were discarded. A procedure was developed to simultaneously purify high-quality genomic DNA, RNA, low-molecular-weight RNA (enriched in miRNA) and protein from the same sample (**Supplementary Fig. 2c**). Tissues were homogenized in Lysis/binding buffer (Ambion) using an RNase-free pestle (VWR) with alternating freeze/thaw cycles (dry and wet ice) and subsequently passed through a QIAshredder column (Qiagen). Lysates were passed through AllPrep DNA spin column (Qiagen) to isolate DNA and the flow-through was mixed with 1 volume of acid-phenol:chloroform and centrifuged. Large and low-molecular-weight RNA were extracted from the aqueous phase using *mir*Vana columns (Ambion). The proteins in the organic fraction were precipitated with 2 volumes of acetone, washed with acetone and subsequently ethanol, and resuspended with protein lysis buffer (50 mM Tris-HCl [pH 7.4], 250 mM NaCl, 0.5% Igepal supplemented with protease and phosphatase inhibitors)[1]. Nucleic acid yield and quality was assessed using a Nanodrop ND1000 spectrophotometer. RNA quality (large and low molecular weight) was further inspected by quantifying the abundance of ribosomal RNA fractions with Experion (Bio-Rad) and/or Agilent 2100 Bioanalyzer. PBMCs were isolated with BD Vacutainer CPT and the DNA was extracted with QIAamp DNA Blood Mini (Qiagen).

Library construction, exome enrichment, and sequencing. Prior to enrichment, indexed libraries were prepared starting with 3 µg of genomic DNA using the Illumina TruSeq DNA Sample Prep Kit according to manufacturer instructions with an added gel purification step for proper insert size selection (250-300 bp inserts). The libraries were then pooled in equimolar fashion and enriched using the Illumina Exome enrichment kit, which contains a pool of DNA-based 95-mer capture oligomers (or baits) targeting 62 Mb representing 201,121 genomic loci including coding exons, their flanking regions, and miRNA genes. Libraries were constructed for 7 pairs of tumor and matched normal samples including one additional metastasis sample. Each pool of 6 indexed samples was subjected to two liquid phase hybridizations for enrichment process followed by a single PCR amplification. Each pool of libraries was then sequenced in 6 lanes of an Illumina Hiseq 2000 sequencer using V3 chemistry and HCS/RTA version 1.1.37/1.7.48. All sequencing was run with paired-end 75-bp reads and was performed according to Illumina's standard protocol. For each sample, on average, ~78 million purity-filtered read pairs were generated and ~67 million read pairs mapped to UCSC hg19 reference genome. The mean percentage of duplicate read pairs due to PCR and optical artifacts average was ~27% in the data set. After removing these duplicate read pairs, ~49 million uniquely mapped read pairs were obtained for each sample.

Comparison of CNA in index patient to publicly-available datasets. Genomic DNA from the primary tumor and PBMC were hybridized to a Human1M-Duo DNA Analysis Beadchip (Illumina). Signal intensities and genotypes were determined by GenomeStudio (Illumina) and

exported to Partek Genomics Suite 6.5, where paired copy number was calculated from intensities. Copy-number profile was segmented with CBS. Paired copy numbers of 317K markers for 78 ccRCC patients from M.D. Anderson Cancer Center[2] were obtained from GEO, transformed to raw copy numbers, and segmented using the Gain and Loss of DNA (GLAD)[3] module of GenePattern[4]. Segmentation data for 99 primary ccRCC tumors and 27 renal cancer cell lines from the Dana-Farber Cancer Institute[5], which were analyzed with Affymetrix 238K StyI arrays and largely published previously[6], were obtained from the Tumorscape portal (http://www.broad.mit.edu/tumorscape). Segmentation files were visualized with IGV (**Supplementary Fig. 2d**).

Primer design. Primers listed in **Supplementary Table 8** were either taken from the literature (where indicated)[7,8] or designed using the NCBI Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast) and checked for specificity using electronic PCR. Primers not listed were designed by Beckman Coulter Genomics (and are proprietary) using a Linux-based amplicon design software set to produce 400-600 bp amplicons and to include at least 50 bp from intron/exon boundaries. Where necessary, overlapping amplicons were designed to assure that all exon bases were covered with at least one high quality read. Repeatmasker, homopolymer detection, and electronic PCR were performed to optimize primer specificity and robustness. M13 priming sites were added to the 5' end of each PCR primer to provide universal priming sites for sequencing. Primer pairs were validated against control DNA (Coriell) and Beckman Coulter Genomics production conditions to ensure they produced high quality sequence.

Sanger sequencing. Genomic DNA was amplified in 384-well format PCR setup. Each PCR reaction contained 10 ng DNA, Thermo-StartReddyMix PCR Master Mix (ThermoScientific) and 0.2 µM forward and reverse PCR primers in a 10 µl reaction. PCR cycling parameters were: one cycle of 95 °C for 15 min, 35 cycles of 95 °C for 20 sec, 60 °C for 30 sec and 72 °C for 1 min, followed by one cycle of 72 °C for 3 min. The resultant PCR products were purified using solid phase reversible immobilization (SPRI) chemistry (AMPure - Beckman Coulter Genomics) followed by dye-terminator fluorescent sequencing with universal M13 primers. Thermocycling conditions: 95 °C for 15 min followed by 40 cycles of 95 °C for 10 sec, 50 °C for 5 sec, 60 °C for 2 min 30 sec. Dye-terminator removal using SPRI (CleanSEQ – Agencourt Bioscience Corporation); Sequencing fragments were detected via capillary electrophoresis using ABI Prism 3730xl DNA analyzer (Applied Biosystems, Foster City, CA). Overall, >95% amplicons (232/239) corresponding to the coding regions and splice sites of 22 genes (inclusive of *VHL*) were successfully sequenced in the Discovery Set. All 106 amplicons analyzed for the validation of the mutations in the exomes were successfully sequenced (**Supplementary Table 6** and **7**). In addition, all 44 amplicons for *BAP1* and *PBRM1* in the Discovery and Validation Sets were successfully sequenced.

For the primers listed in **Supplementary Table 9**, PCR was performed using *Pfu* polymerase in 20 µl reactions with 10 ng of genomic DNA. Cycling parameters were 94 °C for 5 min, followed by 35-40 cycles of 94 °C for 20 sec, 56-62 °C for 20 sec, 72 °C for 1–1.5 min, and a final step of 72°C for 7 min. The amplified samples were subjected to agarose gel electrophoresis and purified using the QIAquick PCR purification kit (Qiagen). Samples were

sequenced at the UTSW McDermott Sequencing Center using Big Dye Terminator 3.1 chemistry (Applied Biosystems) with the listed primers and run on a 3730x1 DNA analyzer (Applied Biosystems).

Immunohistochemistry. Tissue was fixed in formalin for up to 24 h, dehydrated, and paraffinized in a Microm STP 120, and embedded in paraffin blocks. Standard immunohistochemistry staining procedure for BAP1, PBRM1, Ⓟ-S6 ribosomal protein (Ser235/236), and Ⓟ-4E-BP1 (Thr37/46) was conducted using the Benchmark XT automated stainer (Ventana) (**Supplementary Table 10**). Briefly, formalin-fixed, paraffin-embedded tissue sections were cut at 3-4 µm and air-dried overnight. The sections were deparaffinized, rehydrated, and subjected to heat-induced epitope retrieval with Tris-based cell conditioning 1 buffer solution (Ventana). Sections were then incubated with appropriate primary antibody. For signal detection the following were used: ultraView universal detection system (Ventana) for PBRM1, Ⓟ-S6, and Ⓟ-4E-BP1; indirect biotin streptavidin based iView DAB detection systems (Ventana) for BAP1. The slides were developed using 3-3'-diaminobenzidine chromogen and counterstained with hematoxylin. Appropriate positive and negative controls were utilized for each run of immunostains.

Nuclear reactivity was considered positive for BAP1 and PBRM1, and in each tumor sections lymphocytes, stromal fibroblasts and endothelial cells served as internal positive control cells. Phospho-S6 positivity was evaluated as cytoplasmic pattern and Ⓟ-4E-BP1 positivity was evaluated as nuclear and/or cytoplasmic pattern. For Ⓟ-S6 and Ⓟ-4E-BP1 a *H* score was assigned as the product of intensity of staining (0 for negative, 1 for weakly positive, 2 for moderately positive, and 3 for strongly positive) and extent of immunoexpression (0 for no positive cells, 1 for less than 25%, 2 for 26-69%, and 3 for more than 70% cells staining).

Plasmids and cell culture. pcDNA3-FLAG-BAP1 (Laboratory Database ID p744), pBabe-Puro-FLAG-BAP1 (p758) and BAP1 shRNA plasmids (p759 [pENTRmir-shControl], p760 [pENTRmir-shBAP1-A], p761 [pENTRmir-shBAP1-B]) were a generous gift of Dr. Y. Machida (Mayo Clinic). BAP1 cDNA was subcloned into a pBabe-hygro vector using EcoRI and SalI sites (p755), and the different mutations were introduced by site-directed mutagenesis (p756 [HBM mutant ^{363}NHNY^{366} to ^{363}AAAA^{366}], p757 [Y33D mutant]).

769-P cells were obtained from ATCC and maintained in RPMI medium supplemented with 10% FBS and 1% Pen/Strep. UMRC cells were obtained from Dr. Grossman and maintained in DMEM medium supplemented with 10% FBS and 1% P/S. Retroviral infection was performed as previously reported[9] and cells were selected in 2 µg/ml of puromycin (Sigma) or 250 µg/ml hygromycin (Invitrogen). Where indicated, 769-P cells were transiently transfected with an empty vector (EV) or pcDNA3-FLAG-BAP1 (BAP1) using Lipofectamine 2000 (Invitrogen), according to the manufacturer's instructions. Confocal analysis of cells was performed as described elsewhere[9].

For growth curves, cells were plated in triplicate wells and counted every 24 h. UMRC6 growth curves represent two independent experiments each involving 3 wells per cell type and timepoint. Data are represented as fold change over counts at the beginning of the experiment. For olaparib treatment, cells were plated, 24 h after medium was changed and vehicle (DMSO) or olaparib were added at the indicated concentrations. Cells were counted after four days on

olaparib. For irradiation, cells were plated and 24 h later were irradiated and counted for the indicated periods of time.

Colony formation assays. For colony formation assays, cells were irradiated (or not) and grown (with olaparib or vehicle), washed with PBS and stained with crystal violet solution (0.005% in water) for 2 h at room temperature under gentle agitation. Plates were washed thrice in water, scanned with a Microtek Scanmaker i800 scanner at 1,200 dpi, and average pixel intensity for every well was quantified. Colony formation was evaluated 7 days after treatment.

Irradiation of cells. Cells were irradiated using a ^{137}Cs source (JL Shepherd and Associates, San Fernando, CA) and evaluated 4 days after irradiation.

Immunoprecipitation, western blot, and antibodies. Immunoprecipitations were performed as previously described[1] using the immunoprecipitation (IP) buffer (50 mM Tris [pH 7.5], 150 mM NaCl, 5 mM EDTA, 1% Triton X-100 containing protease and phosphatases inhibitors)[10]. Where indicated, cells were fractionated as described[9]. Tumorgrafts were homogenized with 10 volumes of the IP buffer with a pestle and passed through a QIAshredder column (Qiagen). HCF-1 and BAP1 immunoprecipitations from nuclear fractions of 769-P cells were performed using 2 µg of antibody/mg of nuclear protein. Western blot was performed as previously described[1]. Antibodies are listed in **Supplementary Table 10**.

Gel filtration. Tumorgrafts or cells were lysed in IP buffer (see above) and run in a Superose 6 column (GE Healthcare) using an ÄKTA *FPLC* machine (GE Healthcare). 1 ml fractions were collected at a 0.5 ml/min flow rate. Molecular weight markers were from Bio-Rad Laboratories. Collected fractions were subjected to 10% trichloroacetic acid precipitation (where indicated) and analyzed by western blot.

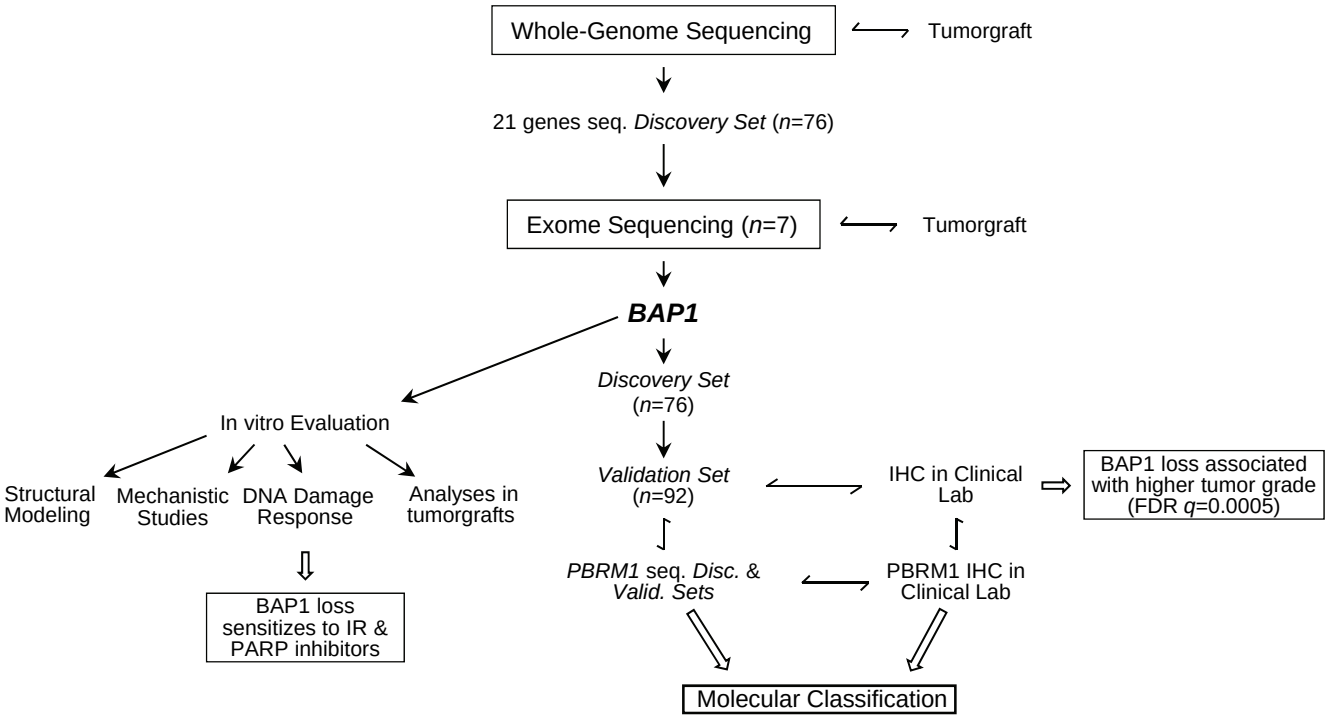
Histone purification. Histones were isolated by acid extraction[11]. Cells were washed with PBS, scraped, spun at 1,000·*g* for 5 min and cell pellets were resuspended in 400 µl of cold 0.4 M HCl and rocked at 4 °C for at least 30 min. Tumorgrafts were homogenized with 10 volumes of 0.4 M HCl with a pestle. Extracts were pelleted after spinning 10 min at 16,000·*g*, and supernatants were precipitated with 10% trichloroacetic acid and analyzed by western blot.

BAP1 and PBRM1 structural modeling. Structure templates were identified using the HHPRED server[12] with the human BAP1 sequence as a query. BAP1 sequence corresponding to the N-terminal UCH domain (residues 1 to 250) confidently identified the structure templates Uch-L3 (PDB: 1xd3) with 100% probability over the entire template sequence (1xd3: residues 1-230). BAP1 sequence limited to the C-terminal ULD (residues 618 to 694) identified Uch37 (3ihr) with 99.83% probability over the template ULD (3ihr: residues 246 to 314). A structure model of the BAP1 UCH domain was generated using the SWISS-MODEL workspace[13] based on the Uch-L3 template alignment from HHPRED. To illustrate DUB domain interactions, the BAP1 UCH domain was superimposed with DUB domains from the ubiquitin-bound Uch-L3 template and ULH-containing Uch37 structure with DaliLite[14]. *PBRM1* missense mutations from this study (isoform NP_060635.2, magenta spheres) and from the Sanger Institute (isoform

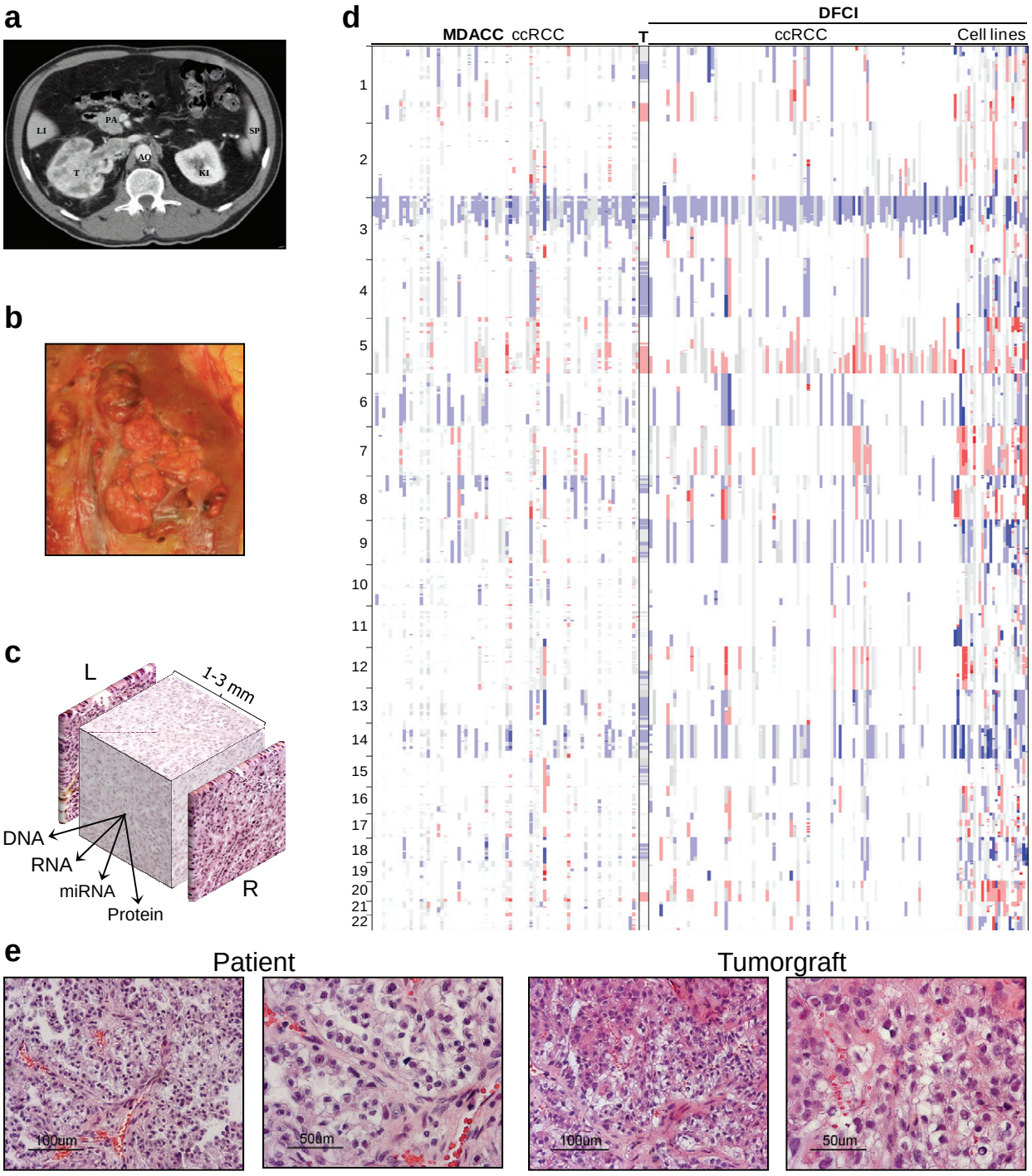
NP_851385.1, gray spheres) are numbered according to NP_060635.2 and displayed in the following structures of individual PBRM1 domains: BR1 (PDB: 3iu5); BR2 (PDB:1ljw), BR4 (PDB: 3tlp), BR5 (PDB:3mb4), BR6 (PDB:3iu6), and BAH1 (PDB: 1w4s).

Supplementary References

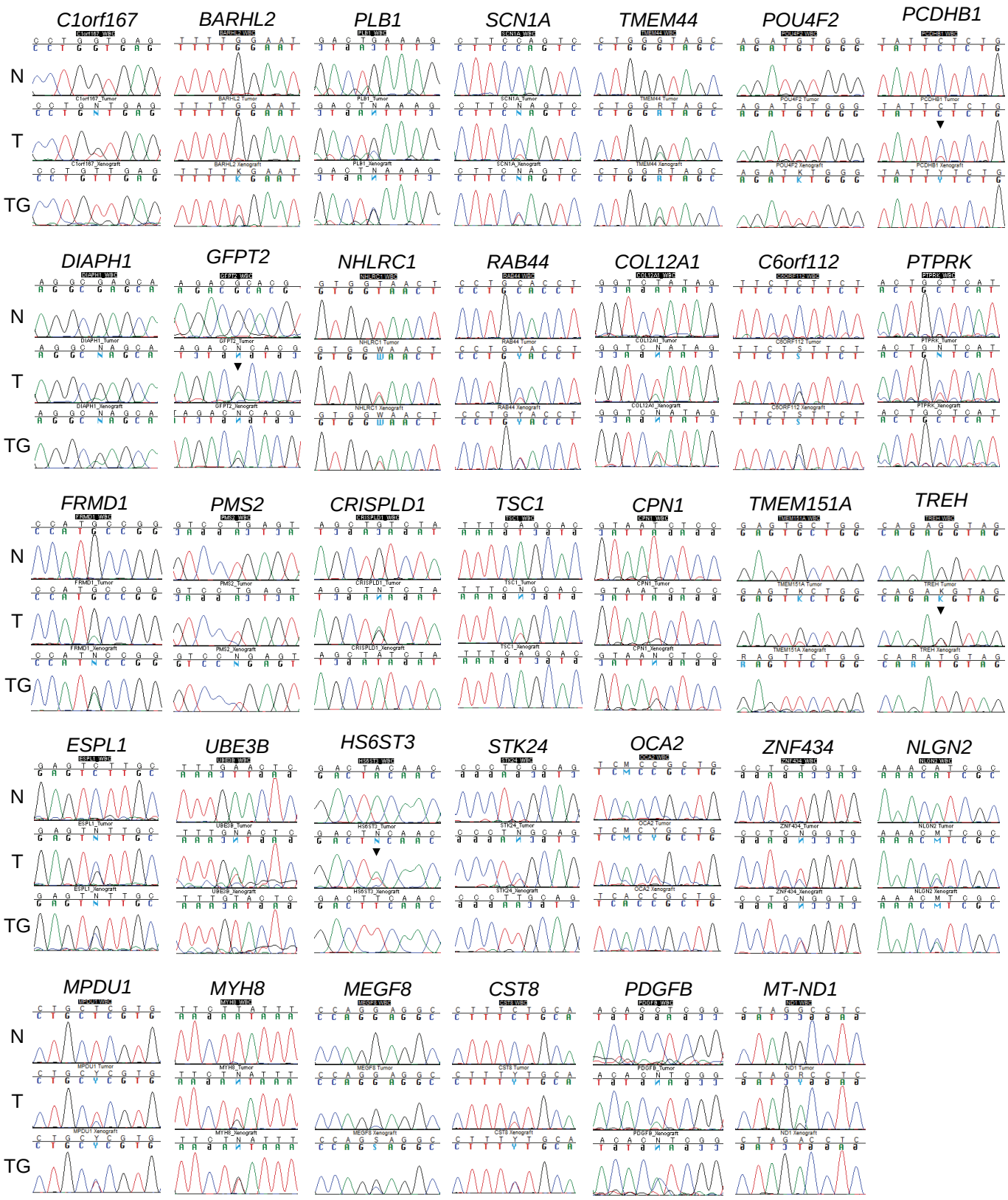
1. Vega-Rubin-de-Celis, S. *et al.* Structural analysis and functional implications of the negative mTORC1 regulator REDD1. *Biochemistry* **49**, 2491-2501 (2010).
2. Chen, M. *et al.* Genome-wide profiling of chromosomal alterations in renal cell carcinoma using high-density single nucleotide polymorphism arrays. *Int J Cancer* **125**, 2342-2348 (2009).
3. Hupé, P., Stransky, N., Thiery, J.P., Radvanyi, F. & Barillot, E. Analysis of array CGH data: from signal ratio to gain and loss of DNA regions. *Bioinformatics* **20**, 3413-3422 (2004).
4. Reich, M. *et al.* GenePattern 2.0. *Nat Genet* **38**, 500-501 (2006).
5. Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899-905 (2010).
6. Beroukhim, R. *et al.* Patterns of gene expression and copy-number alterations in von-hippel lindau disease-associated and sporadic clear cell carcinoma of the kidney. *Cancer Res* **69**, 4674-4681 (2009).
7. Sjöblom, T. *et al.* The consensus coding sequences of human breast and colorectal cancers. *Science* **314**, 268-274 (2006).
8. Dalgliesh, G.L. *et al.* Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**, 360-363 (2010).
9. Peña-Llopis, S. *et al.* Regulation of TFEB and V-ATPases by mTORC1. *EMBO J* **30**, 3242-3258 (2011).
10. Yu, H. *et al.* The ubiquitin carboxyl hydrolase BAP1 forms a ternary complex with YY1 and HCF-1 and is a critical regulator of gene expression. *Mol Cell Biol* **30**, 5071-5085 (2010).
11. Shechter, D., Dormann, H.L., Allis, C.D. & Hake, S.B. Extraction, purification and analysis of histones. *Nat Protoc* **2**, 1445-1457 (2007).
12. Hildebrand, A., Remmert, M., Biegert, A. & Soding, J. Fast and accurate automatic structure prediction with HHpred. *Proteins* **77 Suppl 9**, 128-132 (2009).
13. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195-201 (2006).
14. Holm, L. & Park, J. DaliLite workbench for protein structure comparison. *Bioinformatics* **16**, 566-567 (2000).



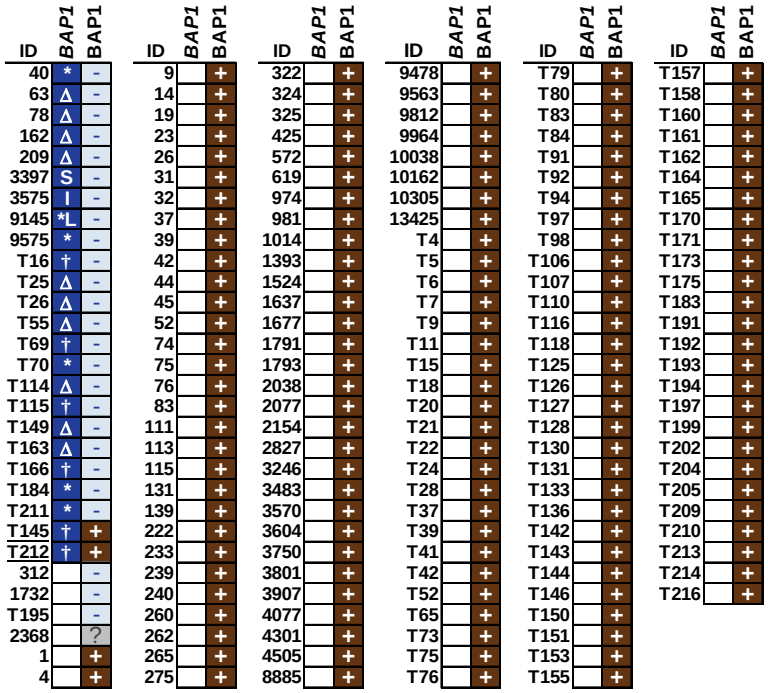
Supplementary Figure 1. Research outline.



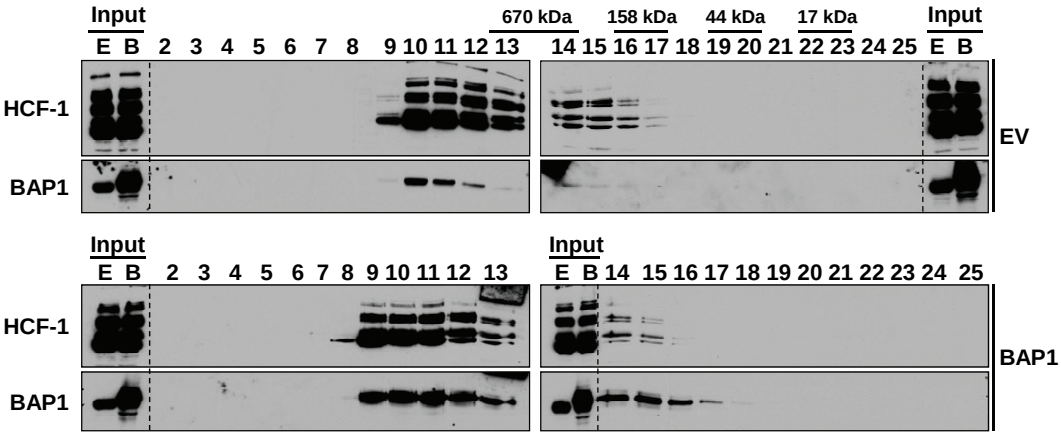
Supplementary Figure 2. Characterization of the primary tumor used for whole-genome sequencing and corresponding tumorgraft. (**a**) Abdominal CT scan showing heterogeneously enhancing central tumor in the right kidney extending into the lumen of the renal vein (Li, liver; T, tumor in right kidney, Pa, pancreas; Ao, aorta; Ki, left kidney; Sp, spleen). (**b**) Photograph of the bisected kidney with embedded multifocal ccRCC (pT3b pN0 M0). (**c**) Schematic illustrating tissue selection and processing (L, left; R, right). (**d**) Comparative analysis of copy number alterations in the patient's tumor vs. other ccRCCs and tumor cell lines (T, tumor; MDACC, MD Anderson Cancer Center; DFCI, Dana-Farber Cancer Institute; see methods). Red and blue refer to segmented areas of chromosomal amplification and deletion respectively. (**e**) Hematoxylin/eosin sections of the primary tumor and corresponding tumorgraft.



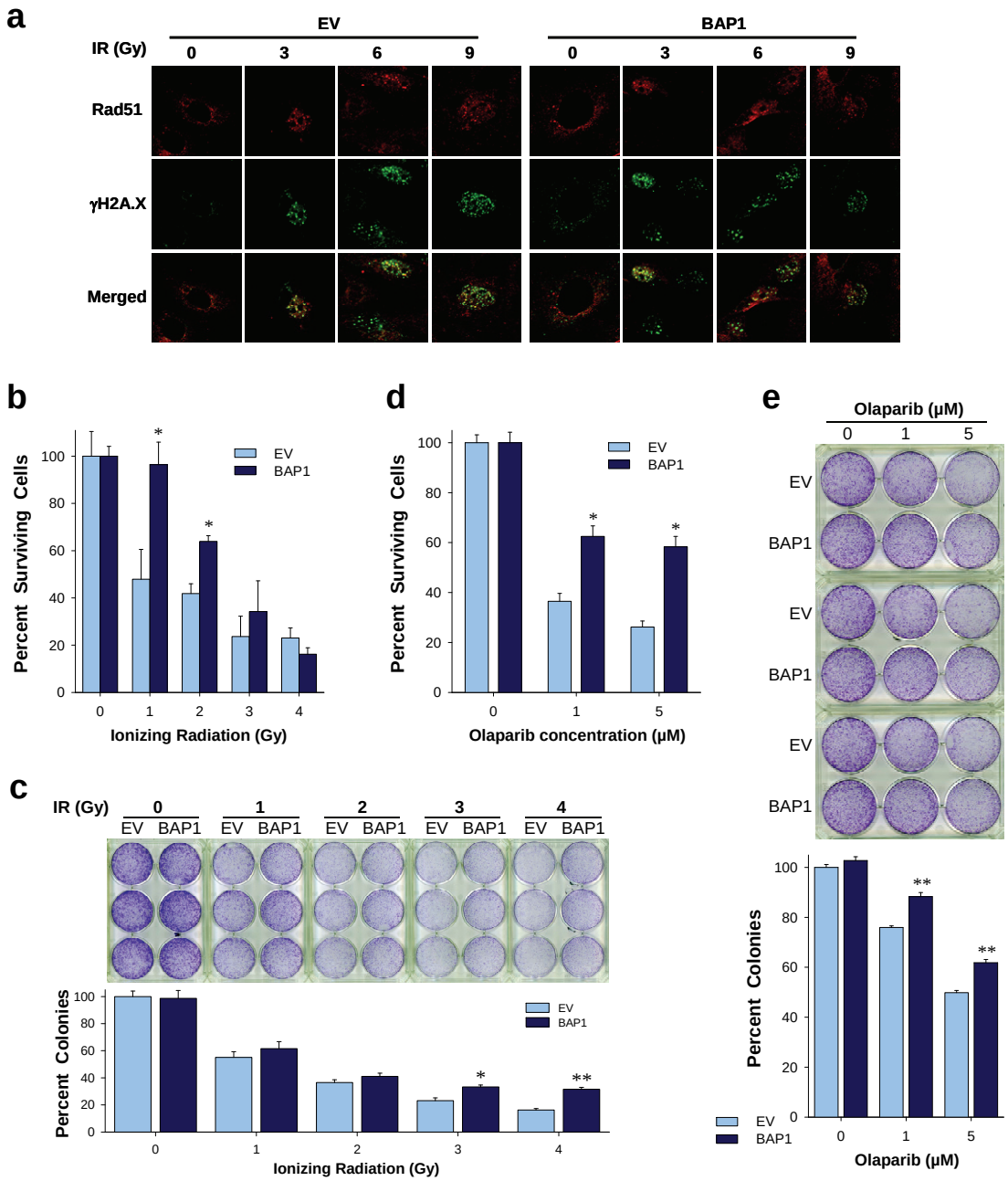
Supplementary Figure 3. Chromatograms illustrating point mutations in the tumor and corresponding tumorgraft. N, normal; T, tumor from patient; TG, tumorgraft.



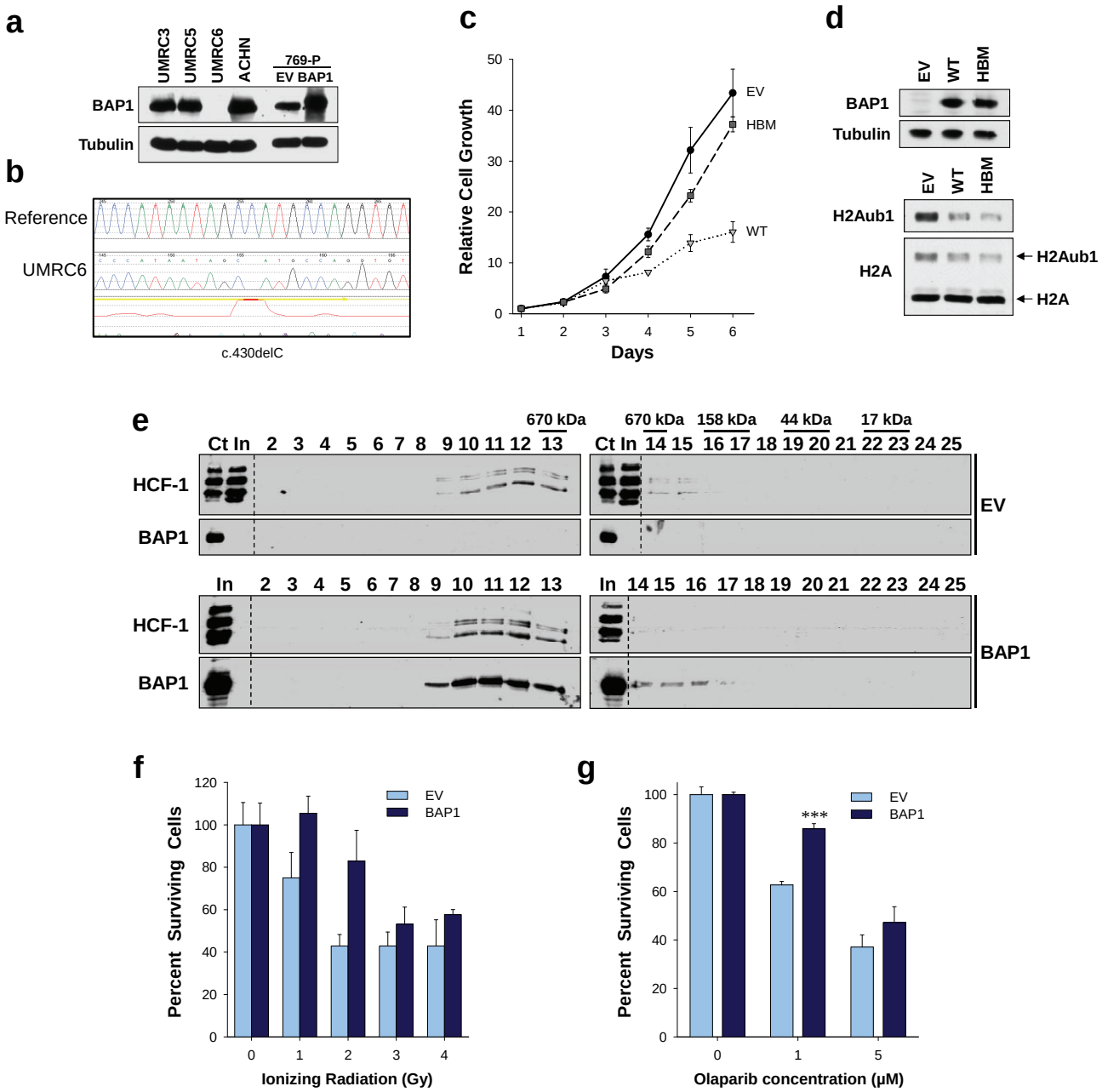
Supplementary Figure 4. Correlation of BAP1 mutation status and IHC in tumors. IHC: +, positive; –, negative; ?, unclear. Mutation: I, insertion; Δ, deletion; †, missense; *, non-sense; S, splice site; *L, stop codon lost; blank, no somatic mutation detected.



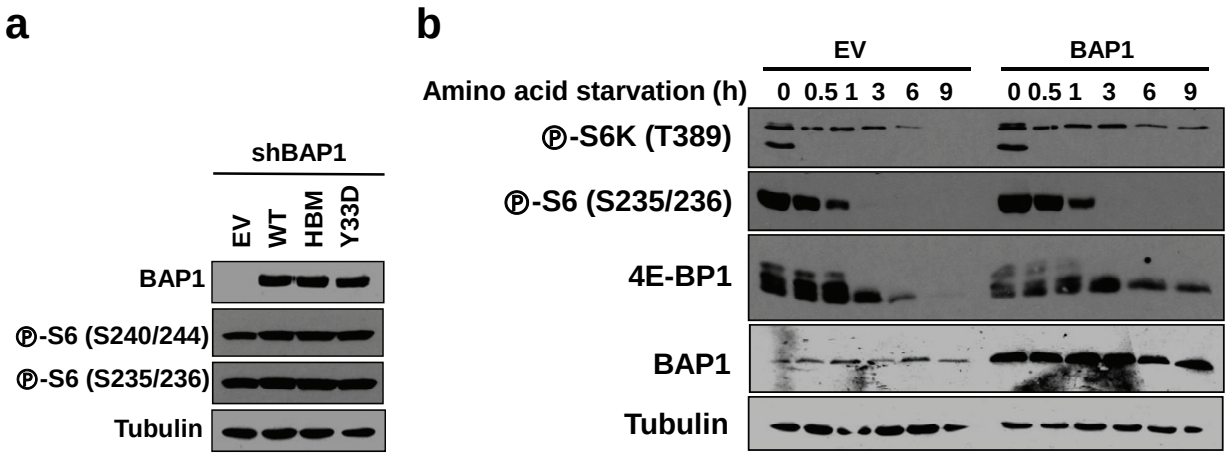
Supplementary Figure 5. Co-fractionation of BAP1 with HCF-1. Western blot of gel filtration fractions of 769-P cells expressing endogenous mutant BAP1 and transduced with an empty vector (EV) control or a BAP1 expression vector (BAP1). Both endogenous mutant as well as ectopically expressed wild-type BAP1 co-fractionate with HCF-1.



Supplementary Figure 6. BAP1 loss sensitizes 769-P cells to ionizing radiation and PARP inhibitors. (**a**) Confocal microscopy images of cells reconstituted as indicated 4 hours after irradiation. (**b**) Histogram of surviving fractions of cells reconstituted as indicated 4 days after ionizing radiation. (**c**) Colony formation assay and histogram quantitation of cells reconstituted as indicated and treated with the stated amounts of radiation. (**d**) Histogram of surviving fractions of 769-P cells reconstituted as indicated and treated with the stated concentrations of olaparib. (**e**) Colony formation assays and histogram quantitation of 769-P cells reconstituted as indicated and treated with olaparib. Error bars represent SEM (n=3). *, p<0.05; **, p<0.01.

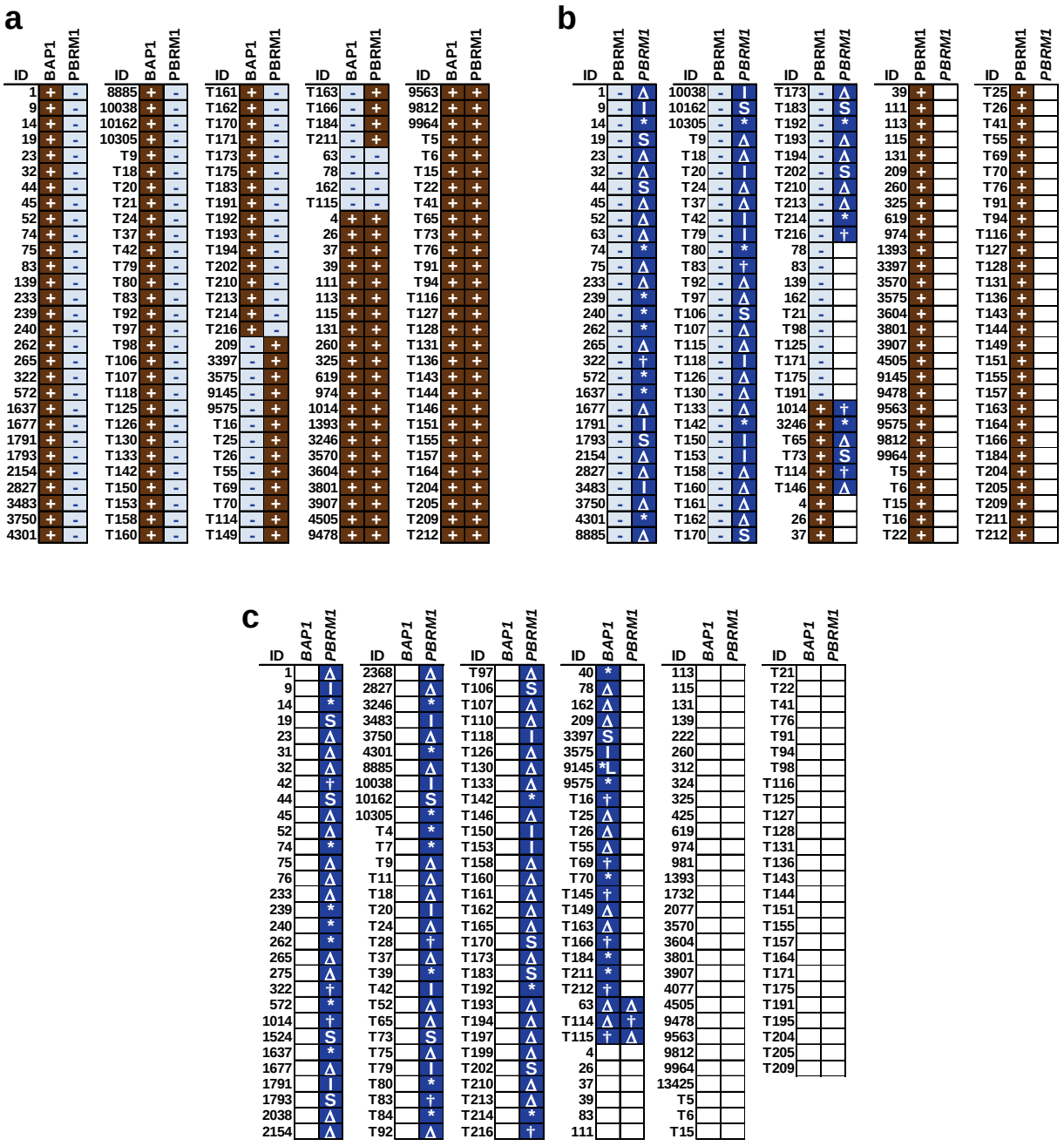


Supplementary Figure 7. HCF-1-dependent suppression of cell proliferation by BAP1 and sensitivity to genotoxic stress in UMRC6 cells. (**a**) Single direction *BAP1* sequence chromatogram of UMRC6 cells. (**b**) Western blot of the indicated cell lines or 769-P control cells transduced as indicated with an empty vector (EV) or BAP1. (**c**) Proliferation curves of UMRC6 cells transduced with an empty vector (EV), wild-type BAP1 or an HBM mutant. (**d**) Western blot of whole cell lysates (top) and partially purified histone fractions (bottom) of UMRC6 cells transduced with an empty expression vector (EV), wild-type *BAP1* (WT), or an *HMB* mutant (HBM). (**e**) Western blot of gel filtration fractions of cells transduced with a BAP1 expression vector (BAP1) or empty vector (EV) control. Histogram of surviving fractions of cells reconstituted as indicated and treated with ionizing radiation (**f**) or olaparib (**g**). Error bars represent SEM (*n*=3-6). ***, $p<0.001$.

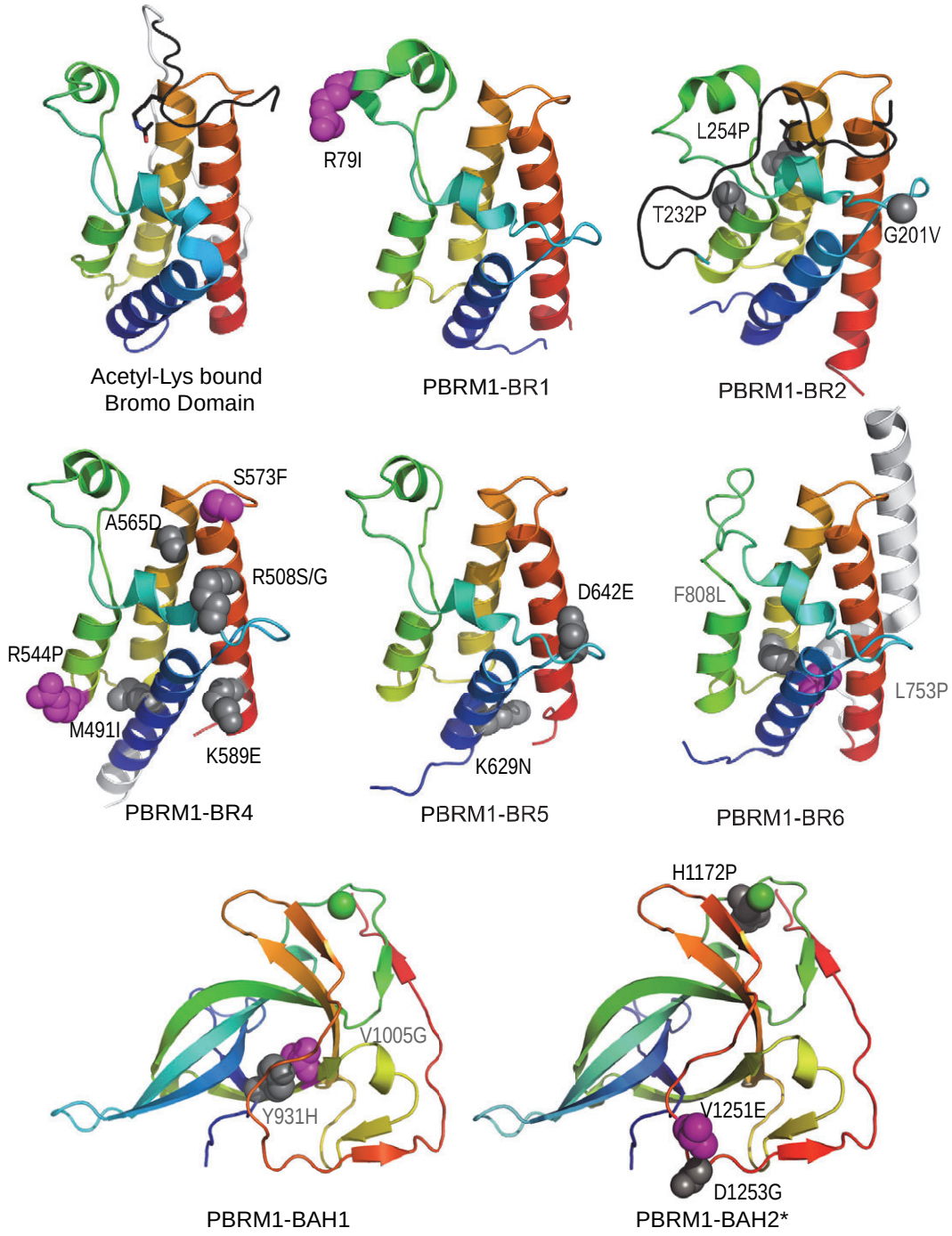


Supplementary Figure 8. mTORC1 activity in 769-P cells. (**a**) Western blot of 769-P cells depleted of endogenous BAP1 and reconstituted with wild type (WT), HCF-1 binding site mutant (HBM), BAP1^{Y33D} or an empty vector (EV). (**b**) Western blot of 769-P cells reconstituted as indicated and starved for amino acids.

Nature Genetics: doi:10.1038/ng.2323



Supplementary Figure 9. Mutation and IHC correlations for BAP1 and PBRM1. (**a**-**c**) +, positive IHC staining; –, negative IHC staining. I, insertion; Δ, deletion; †, missense; *, nonsense; S, splice site; *L, stop codon lost.



Supplementary Figure 10. PBRM1 models. PBRM1 missense mutations in RCC from this study (isoform NP_060635.2, magenta spheres) and from the COSMIC database (isoform NP_851385.1, gray spheres) are numbered according to NP_060635.2 and displayed in the structures of individual PBRM1 bromodomains. PBRM1 comprises six N-terminal bromodomains (BR1-BR6) that bind acetylated lysine residues in histones, followed by two bromo-adjacent homology domains (BAH1 and BAH2), and a HMG DNA-binding domain. All mutations fell within the BR (none on BR3) and BAH domains. Two of the mutations abrogated protein expression (L753P and V1005G) and were buried within the hydrophobic core of BR6 (L753P) or BAH1 (V1005G). Interestingly, the side chain of L^{753} interacted with F^{808} (which was also mutated) and the side chain of V^{1005} interacted with Y^{931}, which was also mutated. Many of the remaining BR mutations (9 out of 13) disrupted residues within or near acetylated lysine binding sites (R79I, T232P, L253P, G201V, S573F, A565D, R508S/G, and D642E) suggesting that binding to acetylated lysine residues is essential for PBRM1 tumor suppressor function.

Supplementary Table 1. Distribution of somatic point mutations and indels in ccRCC genome

	SNV	Indels
Total mutations	5,891	680
Intragenic mutations	2,106	292
Coding	50	9
Synonymous	16	
Missense	32	
Nonsense	2	
5'UTR	9	1
3'UTR	36	9
Intronic	2,011	273
Splice site	4	
miRNA	1	
ncRNA	1	1

SNV, single nucleotide variant (somatically acquired).

Supplementary Table 2. Integrated analysis of somatic mutations and DNA copy number alterations in tumor and tumorgraft

		Chr	Position[§]	Nucleotide change	Gene	Mutant Allele Ratios			T			TG			RefSeq	Protein
							Sanger Seq.			ASCN			ASCN			
						Illumina	T	TG	PCN	Min	Max	PCN	Min	Max		
Single Nucleotide Mutations	Protein Coding	1	36,593,565	C>A	*STK40*	0.30	0.32	**1.00**	**1.39**	**0.43**	1.00	**1.00**	**0.003**	1.04	NM_032017	p.Met133Ile
		1	90,955,307	C>A	*BARHL2*	0.48	<0.10	0.68	**1.45**	**0.43**	1.09	**0.99**	**0.003**	1.04	NM_020063	p.Gly12*
		2	28,602,280	G>C	*PLB1*	0.38	0.30	0.53	1.99	0.89	1.02	2.03	0.85	1.05	NM_153021	p.Glu97Gln
		2	166,616,498	C>T	*SCN1A*	0.37	0.22	0.47	2.06	0.89	1.06	2.01	0.85	1.05	NM_006920	p.Trp314*
		2	228,554,698	G>T	*SPHKAP*	0.26	0.29	0.48	2.01	0.89	1.03	2.01	0.85	1.05	NM_001142644	p.Asp1694Glu
		3	10,166,479	C>G	*VHL*	0.37	0.52	**1.00**	**1.39**	**0.43**	1.07	**0.98**	**0.003**	1.05	NM_000551	p.Leu158Val
		3	195,825,297	C>T	*TMEM44*	0.33	0.34	0.68	1.92	0.91	1.07	2.01	0.86	1.06	NM_138399	p.Gly186Asp
		4	147,781,015	G>T	*POU4F2*	0.47	<0.10	0.28	**1.44**	**0.44**	1.08	**0.97**	**0.004**	1.06	NM_004575	p.Val279Leu
		5	140,412,708	C>T	*PCDHB1*	0.32	0.17	0.28	**2.53**	0.97	**1.55**	**3.05**	0.97	**2.00**	NM_013340	p.Ser490Phe
		5	140,885,872	C>T	*DIAPH1*	0.26	0.20	0.31	**2.53**	0.97	**1.55**	**3.05**	0.97	**2.00**	NM_005219	p.Arg1164Gln
		6	18,230,125	A>T	*NHLRC1*	0.31	0.41	0.62	1.97	0.88	1.15	1.99	0.83	1.07	NM_198586	p.Val231Glu
		6	36,798,440	C>T	*RAB44*	0.38	0.27	0.51	1.95	0.88	1.04	2.00	0.87	1.07	ENST00000457893	p.His554Tyr
		6	105,701,902	G>C	*C6orf112*	0.28	0.36	0.64	1.99	0.92	1.08	2.01	0.87	1.07	XM_098536	p.Lys34Asn
		6	128,430,370	C>T	*PTPRK*	0.28	0.30	0.00	2.07	0.92	1.08	1.99	0.84	1.07	NM_002844	p.Ser715Asn
		6	168,208,441	C>T	*FRMD1*	0.24	0.21	0.43	1.96	0.88	1.03	1.99	0.84	1.03	NM_024919	p.Ala203Thr
		7	5,993,129	T>C	*PMS2*	0.41	0.23	0.48	1.94	0.89	1.03	2.05	0.85	1.05	NM_000535	p.Gln598Arg
		8	76,088,864	G>A	*CRISPLD1*	0.57	0.56	**1.00**	2.02	**0.41**	**1.63**	1.98	**0.003**	**2.00**	NM_031461	p.Val200Ile
		10	101,813,414	T>G	*CPN1*	0.33	0.16	0.32	1.97	0.87	1.05	2.00	0.85	1.04	NM_001308	p.Asp273Ala
		11	65,818,643	G>T	*TMEM151A*	0.36	0.18	**1.00**	1.62	**0.41**	1.16	1.93	**0.003**	**1.96**	NM_153266	p.Cys117Phe
		11	118,035,289	C>A	*TREH*	0.54	0.50	**1.00**	1.62	**0.41**	1.19	1.89	**0.003**	**1.96**	NM_007180	p.Gly478Cys
		12	51,964,210	C>G	*ESPL1*	0.35	0.38	0.59	1.96	0.90	1.08	2.03	0.86	1.06	NM_012291	p.Ser1060Cys
		12	108,456,960	A>T	*UBE3B*	0.37	0.40	**1.00**	**1.39**	**0.43**	1.01	**1.00**	**0.003**	1.06	NM_130466	p.Glu1066Tyr
		13	96,283,428	A>T	*HS6ST3*	0.38	0.38	**1.00**	**1.39**	**0.43**	1.06	**0.98**	**0.004**	1.04	NM_153456	p.Tyr464Phe
		13	97,907,504	C>T	*STK24*	0.28	0.38	**1.00**	**1.39**	**0.43**	1.06	**0.98**	**0.004**	1.04	NM_003576	p.Arg405Gln
		15	25,941,256	G>A	*OCA2*	0.40	0.37	0.00	**1.39**	**0.43**	1.05	**0.98**	**0.003**	1.06	NM_000275	p.Pro211Leu
		16	3,373,160	T>C	*ZNF434*	0.29	0.30	0.55	1.95	**0.41**	**1.57**	1.98	**0.003**	**1.99**	NM_017810	p.Gln384Arg
		17	7,258,792	A>C	*NLGN2*	0.21	0.38	0.59	1.96	0.88	1.03	2.04	0.86	1.05	NM_020795	p.Ile249Leu
		17	7,427,942	T>C	*MPDU1*	0.40	0.28	0.55	1.96	0.88	1.03	2.04	0.86	1.05	NM_004870	p.Leu13Pro
		17	10,240,607	T>G	*MYH8*	0.37	0.30	0.54	1.96	0.88	1.03	2.04	0.86	1.05	NM_002472	p.Lys1506Gln
		17	77,788,470	G>A	*SLC16A3*	0.29	-	-	1.93	0.88	1.03	2.09	0.86	1.19	NM_004207	p.Gly179Ser
		19	47,545,684	G>C	*MEGF8*	0.29	0.12	0.33	1.96	0.89	1.03	2.06	0.89	1.08	NM_001410	p.Gly764Ala
		20	23,421,650	G>A	*CST8*	0.27	0.27	0.50	1.97	0.89	1.05	1.98	0.86	1.05	NM_005492	p.Arg96Lys
		22	37,957,723	C>A	*PDGFB*	0.30	0.26	0.52	1.95	0.85	1.01	2.03	0.96	1.08	NM_002608	p.Glu102Asp
		MT	3,609	G>A	*MT-ND1*	0.77	0.82	**1.00**	1.95	-	-	1.53	-	-	ENST00000361390	p.Gly101Asp
	Splice	1	11,767,238	G>T	*C1orf167*	0.38	0.36	**1.00**	**1.39**	**0.43**	1.00	**1.02**	**0.003**	1.04	XM_001715322	
		5	179,662,025	G>A	*GFPT2*	0.43	0.38	0.65	**2.51**	0.97	**1.54**	**3.05**	0.97	**2.00**	NM_005110	
		6	75,894,863	T>A	*COL12A1*	0.27	0.26	0.53	2.07	0.92	1.09	2.05	0.87	1.07	NM_004370	
		9	134,790,949	T>A	*TSC1*	0.48	0.28	0.00	**1.36**	**0.44**	1.03	**1.02**	**0.004**	1.06	NM_000368	
Indels	Protein Coding	2	48,727,373	delA	*GTF2A1L*	0.20	0.32	0.54	2.08	0.89	1.06	2.03	0.85	1.05	NM_172196	
		2	203,338,446	delA	*FAM117B*	0.19	0.25	0.45	1.98	0.89	1.06	2.01	0.85	1.05	NM_173511	
		2	212,520,459	insAAA	*ERBB4*	0.10	0.30	0.51	2.08	0.89	1.06	2.01	0.85	1.05	NM_005235	
		5	137,482,465	insT	*NME5*	0.13	0.20	0.33	**2.53**	0.97	**1.55**	**3.05**	0.96	**2.00**	NM_003551	
		7	150,185,054	insC	*ABP1*	0.15	0.28	0.49	1.96	0.88	1.03	2.01	0.85	1.05	NM_001091	
		9	18,767,566	delGGAGCACCA	*ADAMTSL1*	0.25	0.34	**1.00**	**1.42**	**0.44**	1.12	**1.01**	**0.004**	1.06	NM_001040272	
		10	49,674,424	delG	*WDFY4*	0.17	0.29	0.51	1.94	0.88	1.03	2.00	0.85	1.04	NM_020945	
		11	57,333,402	delG	*CTNND1*	0.36	0.38	**1.00**	1.74	**0.41**	1.16	1.95	**0.003**	**1.96**	NM_001085458	
		14	73,275,194	delTCTCGCTCCCG GCCATTGGGTGCTAGTCTCT CCCCATCAGGCAGTAGCTG	*C14orf43*	0.32	0.35	**1.00**	**1.39**	**0.44**	1.07	**0.99**	**0.004**	1.05	NM_194278	

Mutation analyses of whole-genome sequenced tumor/normal pair and corresponding tumorgraft. [§], Annotated with NCBI36.1 and Ensembl build 54. Orange gradient intensity is proportional to the ratio of mutant to wild-type alleles. DNA copy numbers were inferred from segmented data at mutation sites; PCN, paired copy number; ASCN, allele-specific copy number. Min and Max represent the minimum and maximum ASCN for heterozygous SNPs; blue denotes deletion (PCN<1.5 or ASCN<0.5) and red amplification (PCN>2.5 or ASCN>1.5). T, patient tumor; TG, tumorgraft.

Supplementary Table 3. Patient and tumor characteristics

	Discovery Set (*n* =76)	Exomes (*n* =7)	Validation Set (*n* =92)
Mean age - yr (95% CI)	62 (59-65)	58 (45-70)	60 (58-63)
Race - No. (%)			
White	57 (76)	4 (66)	63 (80)
Hispanic	12 (16)	1 (17)	8 (10)
African American	6 (8)		4 (5)
Indian			3 (4)
Asian			1 (1)
Native American		1 (17)	
Sex - No. (%)			
F	34 (45)	1 (14)	41 (45)
M	42 (55)	6 (86)	50 (55)
BMI - No. (%)			
<25	9 (19)	1 (14)	15 (17)
≥25 - <30	17 (34)	5 (72)	35 (39)
≥30	23 (47)	1 (14)	39 (44)
Smoking - No. (%)			
N	34 (63)	4 (57)	37 (42)
Y	20 (37)	3 (43)	52 (58)
Family History of RCC - No. (%)			
N	64 (98)	5 (100)	73 (92)
Y	1 (2)		6 (8)
Fuhrman Grade - No. (%)			
1	4 (5)	0 (0)	3 (3)
2	45 (59)	1 (14)	44 (48)
3	20 (27)	2 (29)	34 (37)
4	7 (9)	4 (57)	11 (12)
pT Stage - No. (%)			
T1	37 (49)	2 (29)	45 (50)
T2	11 (14)	1 (14)	12 (14)
T3	27 (36)	3 (43)	30 (33)
T4	1 (1)	1 (14)	3 (3)
pN - No. (%)			
0	37 (95)	1 (33)	26 (96)
1	2 (5)	2 (67)	1 (4)
M - No. (%)			
0	46 (87)	6 (86)	80 (91)
1	7 (13)	1 (14)	8 (9)
Stage - No. (%)			
I	12 (35)	1 (20)	9 (26)
II	5 (15)		4 (12)
III	10 (29)	2 (40)	11 (32)
IV	7 (21)	2 (40)	10 (29)

Percentages refer to the total number of samples with information available. Staging based on TNM classification. pT, pathological T stage; pN, pathological N stage; M, clinical metastases.

Supplementary Table 4. Mutated genes in whole-genome sequence evaluated in Discovery Set

Gene	CDS	Protein	ID
BARHL2	**c.34G>T**	**p.G12***	**T22**
C6orf112	**c.102G>C**	**p.K34N**	**T22**
C14orf43	c.734C>A	p.P245Q	T7
	c.1222_1271del(50 Nt)	**Fs**	**T22**
CRISPLD1	c.461C>G	p.P154R	1637
	c.463T>A	p.Y155N	324
	c.598G>A	**p.V200I**	**T22**
CST8	**c.287G>A**	**p.R96K**	**T22**
FRMD1	**c.607G>A**	**p.A203T**	**T22**
	c.1120A>G[§]	p.S374G	T55
MEGF8	**c.2291G>C**	**p.G764A**	**T22**
MPDU1	**c.38T>C**	**p.L13P**	**T22**
	IVS170-5C>T	Sp	115
	IVS303-6C>A	Sp	T9
MT-ND1	c.162delA	Fs	37
	c.265delC	**Fs**	**T22**
	c.302G>A	**p.G101D**	**T22**
	c.745G>A	p.A249T	T16
NHLRC1	**c.692T>A**	**p.V231E**	**T22**
NLGN2	**c.745A>C**	**p.I249L**	**T22**
	c.1663A>G	p.T555A	T52
OCA2	**c.632C>T**	**p.P211L**	**T22**
	c.1327G>A	p.V443I	322
PCDHB1	**c.1469C>T**	**p.S490F**	**T22**
	c.2216A>G	p.N739S	4077
POU4F2	**c.835G>T**	**p.V279L**	**T22**
	c.886C>T	p.Q296*	265
RAB44	**c.1660C>T**	**p.H554Y**	**T22**
	c.1559C>T[§]	p.P520L	T55
STK40	**c.399G>T**	**p.M133I**	**T22**
TMEM151A	**c.350G>T**	**p.C117F**	**T22**
	c.880G>C	p.A294P	76
	c.1030A>G	p.T344A	325
	c.1093G>A	p.V365I	240
TMEM44	**c.557G>A**	**p.G186D**	**T22**
TREH	**c.1432G>T**	**p.G478C**	**T22**
TSC1	**IVS211-2A>T**	**Sp**	**T22**
	c.1342C>T[§]	p.P448S	T55
	c.1546C>T	p.Q516*	3246
	c.2459dupA	Fs	9
ZNF434	**c.1151A>G**	**p.Q384R**	**T22**
	c.1184G>A[§]	p.C395Y	T11

Mutations within non-coding areas are prefixed with "IVS" and its position is identified by the number of nucleotides away from the closest coding nucleotide. Bold denotes mutation in whole-genome sequence from index patient. [§], no paired normal tissue available; Fs, frameshift mutation; Sp, splice-site mutation; Nt, nucleotide; ID, tumor identifier. Mutations in *TSC1* previously reported in Kucejova *et al.* 2011.

Supplementary Table 5. Distribution of somatic point mutations and indels in ccRCC exomes

		T127		T142		T144	
	Average	Normal	Tumor	Normal	Tumor	Normal	Tumor
Bases in target region	59,415,126	59,776,368	59,804,879	59,236,262	58,788,223	59,294,311	59,582,386
Average # of reads per targeted base	64.2	80.5	114.7	45.8	53.8	49.2	81.5
Percent positions covered	95.7	96.3	96.3	95.4	94.7	95.5	96.0
Targeted bases with at least 10 reads (%)	89.5	91.4	92.4	88.5	88.3	88.6	90.8
Mutations identified by CASAVA1.8	20,021	21,067	21,292	20,240	20,221	21,188	21,362
Predicted somatic mutations	779	682		769		903	
Non-synon., splice-site, and indel mutations	463	401		446		556	
Somatic mutations validated visually	43	21		40		62	

T163		T164		T166		T183			
Normal	Tumor	Normal	Tumor	Normal	Tumor	Normal	Tumor	Metastasis	
59,253,978	59,527,984	59,385,244	59,676,549	59,030,748	59,188,224	59,587,283	59,602,362	59,492,084	
45.9	75.2	45.2	65.5	59.2	51.3	58.6	77.1	59.7	
95.4	95.9	95.7	96.1	95.1	95.3	96.0	96.0	95.8	
88.5	89.8	88.2	90.0	88.9	88.4	89.4	90.7	88.6	
									Total
20,403	20,572	20,183	17,235	20,747	20,349	20,860	20,814	19,483	306,016
818		907		700		759		692	6,230
474		515		437		453		421	3,703
36		42		68		37 (5 unique)		39 (7 unique)	345

Supplementary Table 6. List of mutated genes randomly selected from exome project for validation by Sanger sequencing with corresponding mutant allele ratios and status in tumorgrafts

							Illumina Sequencing Reads			MAR (Sanger Seq.)						Paired Copy Number					
							Normal	Tumor	Metastasis				Tumorgrafts						Tumorgrafts		
	Gene	ID	Chr	Position[§]	CDS	Protein	Mut/wt (%)	Mut/wt (%)	Mut/wt (%)	N	T	M	(1)	(2)	(3)	T	M	(1)	(2)	(3)	
1	*ATF7IP*	T127	12	14,577,562	c.A713G	p.D238G	0/141 (0)	47/138 (0.25)		0.03	0.24		0.45	0.47		1.98		2.14	2.02		
2	*ITPR3*	T127	6	33,659,443	c.G7277T	p.G2426V	0/51 (0)	28/69 (0.29)		0.00	0.52		0.00	0.05		1.98		1.10	1.11		
3	*KCNN2*	T127	5	113,698,690	c.218_229del	fs	0/28 (0)	6/13 (0.32)		0.00	<0.10		0.31	0.41		1.98		2.02	2.00		
4	*SLC16A11*	T127	17	6,945,095	c.C1319T	p.A440V	0/37 (0)	19/36 (0.35)		0.00	0.30		0.55	0.54		1.98		2.02	1.99		
5	*SLC41A2*	T127	12	105,303,531	c.565delG	p.V189fs	0/81 (0)	20/122 (0.14)		0.00	0.27		0.00	0.00		1.98		2.14	2.02		
6	*CFH*	T142	1	196,694,326	c.A1772T	p.E591V	0/86 (0)	32/80 (0.29)		0.00	0.32		0.30	0.32		1.99		1.57	1.55		
7	*CTCF*	T142	16	67,660,619	IVS1518+1G>T	Sp	0/36 (0)	16/33 (0.33)		0.01	0.25		0.41	0.40		1.99		2.02	2.01		
8	*FAM84A*	T142	2	14,774,146	c.G43A	p.E15K	0/36 (0)	13/31 (0.3)		0.03	0.25		0.03	0.05		1.99		2.01	2.03		
9	*MCM3AP*	T142	21	47,676,880	c.C3755A	p.A1252D	0/13 (0)	2/8 (0.2)		0.00	0.00		0.00	0.00		2.12		1.93	1.55		
10	*MPP4*	T142	2	202,545,715	c.C775T	p.P259S	0/40 (0)	16/50 (0.24)		0.00	0.39		0.57	0.59		1.99		2.01	2.03		
11	*NDFIP2*	T142	13	80,107,515	c.C545T	p.S182F	0/66 (0)	33/71 (0.32)		0.00	0.34		0.57	0.57		1.99		2.01	2.02		
12	*PRKAR2B*	T142	7	106,685,610	c.G258T	p.E86D	0/5 (0)	2/5 (0.29)		0.04	0.00		0.04	0.17		2.59		2.42	2.50		
13	*SNED1*	T142	2	241,987,756	c.G1298A	p.C433Y	0/11 (0)	2/6 (0.25)		0.00	0.00		0.00	0.00		1.46		1.13	1.09		
14	*TMPRSS11B*	T142	4	69,093,666	c.C1214T	p.S405F	0/46 (0)	25/55 (0.31)		0.05	0.42		**1.00**	**1.00**		1.98		1.09	1.04		
15	*UNC13B*	T142	9	35,397,188	c.C3310G	p.L1104V	0/40 (0)	21/35 (0.38)		0.00	0.29		0.00	0.00		1.99		1.12	1.09		
16	*ZC3H4*	T142	19	47,570,581	c.G2944A	p.E982K	0/18 (0)	6/14 (0.3)		0.05	0.41		0.37	0.40		2.00		1.60	1.61		
17	*ANKRD35*	T144	1	145,562,188	c.A1876T	p.N626Y	0/74 (0)	22/75 (0.23)		0.03	0.07		0.53	0.60	0.52	2.04		1.97	2.01	1.94	
18	*BCL3*	T144	19	45,260,646	c.G787C	p.E263Q	0/22 (0)	12/23 (0.34)		0.00	0.19		0.00	0.03	0.00	2.03		1.97	1.97	1.52	
19	*BDKRB2*	T144	14	96,707,683	c.C1018G	p.R340G	0/45 (0)	13/30 (0.3)		0.00	0.34		**1.00**	**1.00**	**1.00**	1.53		1.15	1.09	1.08	
20	*CBS*	T144	21	44,486,486	c.G318T	p.L106F	0/12 (0)	8/14 (0.36)		0.00	0.25		0.40	0.40	0.46	2.02		1.96	0.61	0.53	
21	*DYNC2H1*	T144	11	103,128,394	c.T10519A	p.S3507T	0/85 (0)	42/55 (0.43)		0.00	0.28		**0.95**	**1.00**	**1.00**	1.53		1.13	1.10	1.06	
22	*GATAD1*	T144	7	92,079,890	c.376_404del	fs	0/102 (0)	12/150 (0.07)		0.00	0.43		0.62	0.62	0.87	2.04		1.99	2.03	1.53	
23	*HGSNAT*	T144	8	43,024,320	c.G568T	p.D190Y	0/82 (0)	44/105 (0.3)		0.00	0.30		0.51	0.55	0.38	2.02		1.92	1.88	1.58	
24	*LRP2*	T144	2	169,985,191	c.13950delA	p.K4650fs	0/52 (0)	28/67 (0.29)		0.00	0.18		0.37	0.49	0.49	2.07		2.00	2.02	1.96	
25	*PAPPA*	T144	9	119,065,182	c.C3100A	p.Q1034K	1/48 (0.02)	24/52 (0.32)		0.00	0.25		0.45	0.44	0.40	2.01		1.98	2.00	1.98	
26	*PRPF8*	T144	17	1,576,715	c.C3593T	p.P1198L	0/39 (0)	14/24 (0.37)		0.00	0.46		**1.00**	**1.00**	**1.00**	1.53		1.19	1.13	1.09	
27	*SPHK1*	T144	17	74,383,398	c.C928T	p.R310C	0/37 (0)	20/62 (0.24)		0.03	0.26		0.38	0.42	0.56	2.50		2.58	2.63	1.92	
28	*SPTLC1*	T144	9	94,812,274	c.G856A	p.G286S	0/68 (0)	33/86 (0.28)		0.00	0.18		0.45	0.41	0.46	2.01		1.92	1.99	1.93	
29	*TAS1R2*	T144	1	19,175,898	c.C1404G	p.Y468*	0/25 (0)	16/15 (0.52)		0.00	0.34		0.00	0.00	0.00	2.25		1.96	2.01	1.95	
30	*UGT2B7*	T144	4	69,973,880	c.1150_1151insTC	p.I384fs	0/69 (0)	22/74 (0.23)		0.01	0.46		**1.00**	**1.00**	**1.00**	1.55		1.07	1.08	1.11	
31	*WWC1*	T144	5	167,868,745	c.G2339A	p.W780*	0/39 (0)	22/39 (0.36)		0.00	0.32		**0.97**	**0.97**	**1.00**	1.51		1.18	1.10	1.48	
32	*ZNF213*	T144	16	3,187,364	c.G83T	p.W28L	0/17 (0)	10/13 (0.43)		0.03	0.18		0.40	0.42	0.47	2.00		1.93	1.95	1.93	
33	*ATRNL1*	T163	10	117,309,044	c.G3793T	p.E1265*	0/86 (0)	30/84 (0.26)		0.00	0.32					1.97					
34	*CCDC155*	T163	19	49,899,051	c.G361A	p.A121T	0/56 (0)	18/53 (0.25)		0.00	0.29					2.02					
35	*COL7A1*	T163	3	48,608,117	c.A7299T	p.E2433D	0/26 (0)	10/13 (0.43)		0.00	0.35					1.46					
36	*ELP3*	T163	8	27,989,919	c.904delA	p.T302fs	0/34 (0)	11/23 (0.32)		0.00	0.22					1.99					
37	*GTPBP4*	T163	10	1,045,035	c.G654T	p.Q218H	0/53 (0)	18/53 (0.25)		0.00	0.24					1.97					
38	*KIF1B*	T163	1	10,386,341	c.G2710T	p.E904*	0/12 (0)	8/8 (0.5)		0.00	0.40					1.49					
39	*KLC1*	T163	14	104,139,370	c.C1007T	p.P336L	0/37 (0)	13/16 (0.45)		0.00	0.46					1.45					
40	*KLHL22*	T163	22	20,812,232	c.G1168A	p.A390T	0/29 (0)	10/21 (0.32)		0.00	0.29					2.00					
41	*RALGPS1*	T163	9	129,958,783	c.T1068A	p.S356R	0/63 (0)	14/51 (0.22)		0.00	0.20					1.90					
42	*AFF2*	T164	X	148,055,040	c.G2230A	p.A744T	0/12 (0)	19/4 (0.83)		0.00	**0.90**		**1.00**	**1.00**		1.19		1.14	1.33		
43	*AFM*	T164	4	74,357,644	c.A899G	p.K300R	0/80 (0)	59/13 (0.82)		0.00	0.70		**0.94**	**0.89**		1.17		1.09	0.87		
44	*AMOTL2*	T164	3	134,085,209	c.1102_1103insC	p.K368fs	0/42 (0)	31/49 (0.39)		0.06	0.27		**1.00**	**1.00**		1.17		1.04	1.15		
45	*COPG2*	T164	7	130,337,759	c.269_273del,	fs	0/50 (0)	34/100 (0.25)		0.01	0.44		0.50	0.50		1.96		1.97	2.13		
46	*CREBBP*	T164	16	3,801,726	IVS3779+1delG	Sp	0/48 (0)	35/34 (0.51)		0.00	0.39		0.30	0.44		1.95		1.67	2.33		
47	*DLST*	T164	14	75,357,775	c.C347T	p.P116L	0/42 (0)	50/17 (0.75)		0.00	0.72		**0.98**	**1.00**		1.18		1.05	1.23		
48	*HECTD1*	T164	14	31,598,054	c.T4523G	p.L1508R	0/48 (0)	37/14 (0.73)		0.00	0.76		**1.00**	**1.00**		1.19		1.10	1.02		
49	*MARCH6*	T164	5	10,402,517	c.T1075C	p.F359L	0/96 (0)	174/181 (0.49)		0.00	0.43		0.44	0.46		3.47		3.52	3.60		
50	*RAD54L2*	T164	3	51,575,667	IVS1-48770G>GT	Sp	0/47 (0)	43/11 (0.8)		0.01	0.69		0.03	0.03		1.17		1.00	1.24		
51	*SETD2*	T164	3	47,144,892	c.G4861A	p.G1621R	0/68 (0)	17/54 (0.24)		0.00	0.73		**0.98**	**0.97**		1.17		1.05	1.24		
52	*SOS1*	T164	2	39,222,350	c.3232_3260del	fs	0/31 (0)	22/38 (0.37)		0.00	0.63		0.68	0.49		2.67		1.55	1.99		
53	*TLX2*	T164	2	74,741,964	c.C31T	p.L11F	0/9 (0)	16/21 (0.43)		0.02	0.46		0.50	0.44		1.96		1.12	2.04		
54	*BCL9L*	T166	11	118,772,682	c.G1770C	p.M590I	0/23 (0)	7/7 (0.5)		0.02	0.52		0.70	0.69	0.67	1.67		1.53	1.51	1.52	
55	*BPHL*	T166	6	3,129,277	IVS328-2A>AT	Sp	1/57 (0.02)	9/23 (0.28)		0.00	0.21		0.34	0.29	0.30	1.63		1.52	1.51	1.53	
56	*C17orf57*	T166	17	45,517,797	c.2639delG	p.R880fs	0/51 (0)	14/37 (0.27)		0.00	0.29		0.48	0.50	0.45	2.01		1.99	2.00	1.99	
57	*CSNK1D*	T166	17	80,207,463	c.G901T	p.A301S	0/7 (0)	2/4 (0.33)		0.00	0.00		0.00	0.00	0.00	2.03		1.98	1.94	1.95	
58	*FAM126B*	T166	2	201,846,181	c.A1405G	p.I469V	0/33 (0)	21/27 (0.44)		0.00	0.39		0.66	0.62	0.63	1.61		1.53	1.57	1.52	
59	*FZD2*	T166	17	42,636,527	c.G1471C	p.E491Q	0/55 (0)	19/31 (0.38)		0.00	0.30		0.36	0.50	0.32	2.01		1.99	2.00	1.99	
60	*GRIN2A*	T166	16	9,934,641	c.C1514A	p.A505E	0/31 (0)	19/26 (0.42)		0.00	0.27		0.55	0.57	0.56	1.86		1.53	1.46	1.47	
61	*HSDL1*	T166	16	84,163,953	c.G304T	p.E102*	0/107 (0)	32/76 (0.3)		0.00	0.35		0.64	0.66	0.66	1.90		1.51	1.50	1.47	
62	*KIAA2022*	T166	X	73,960,664	c.G3728A	p.R1243H	0/34 (0)	24/22 (0.52)		0.00	0.65		**1.00**	**1.00**	**1.00**	1.83		1.98	1.82	1.88	
63	*LAMB1*	T166	7	107,616,265	c.C1058T	p.A353V	0/48 (0)	12/23 (0.34)		0.04	0.52		0.78	0.77	0.76	1.84		2.00	1.97	2.00	
64	*MSL1*	T166	17	38,285,755	c.C461G	p.S154*	0/76 (0)	31/42 (0.42)		0.00	0.36		0.53	0.52	0.53	1.97		1.99	1.97	1.99	
65	*MTOR*	T166	1	11,168,338	c.G7534C	p.D2512H	0/46 (0)	16/17 (0.48)		0.00	0.48		**1.00**	**1.00**	**1.00**	1.37		1.13	1.08	1.09	
66	*NBEAL2*	T166	3	47,043,451	c.G4824C	p.Q1608H	0/27 (0)	9/7 (0.56)		0.00	0.75		**1.00**	**1.00**	**1.00**	1.46		1.10	1.50	1.51	
67	*PIK3C2A*	T166	11	17,143,877	c.G2515C	p.D839H	0/78 (0)	13/42 (0.24)		0.00	0.24		0.35	0.36	0.32	1.65		1.53	1.51	1.52	
68	*SIN3A*	T166	15	75,693,171	c.T1637C	p.I546T	0/116 (0)	21/68 (0.24)		0.05	0.36		0.59	0.55	0.60	1.99		1.99	1.99	1.95	
69	*STARD13*	T166	13	33,692,292	c.G2191T	p.V731L	0/118 (0)	20/49 (0.29)		0.00	0.35		0.51	0.52	0.52	1.94		1.97	1.90	1.95	
70	*TET1*	T166	10	70,404,686	c.2200delG	p.E734fs	0/70 (0)	24/33 (0.42)		0.00	0.45		0.31	0.35	0.34	1.39		1.52	1.55	1.51	
71	*XAB2*	T166	19	7,687,448	c.G1471A	p.A491T	0/55 (0)	15/38 (0.28)		0.00	0.42		0.65	0.62	0.62	1.69		0.59	1.49	1.50	
72	*GRK4*	T183	4	3,021,416	c.C790G	p.L264V	0/45 (0)	14/39 (0.26)		0.00	0.27	0.02	0.03	0.02		1.66	1.18	1.12	1.11		
73	*NDST4*	T183	4	115,997,643	c.C550G	p.L184V	0/58 (0)	25/50 (0.33)		0.00	0.24	0.00	0.00	0.03		1.64	1.13	1.06	1.07		
74	*PCNX*	T183	14	71,443,934	c.880_916del	fs	0/83 (0)	24/54 (0.31)		0.01	0.48	0.01	0.01	0.00		1.66	1.16	1.09	1.10		
75	*PSMA5*	T183	1	109,957,979	c.T103A	p.S35T	0/69 (0)	29/41 (0.41)		0.04	0.34	0.02	0.02	0.00		2.01	1.16	1.09	1.09		
76	*AHCYL1*	T183	1	110,561,085	c.A1214C	p.D405A	0/134 (0)	67/101 (0.4)	55/18 (0.75)	0.00	0.28	0.73	**0.92**	**1.00**		2.01	1.16	1.11	1.09		
77	*ANKS3*	T183	16	4,747,097	c.C1903A	p.L635M	1/46 (0.02)	25/40 (0.38)	12/20 (0.38)	0.00	0.45	0.53	0.56	0.61		2.02	2.02	2.02	2.02		
78	*CLPX*	T183	15	65,449,241	c.A1087G	p.I363V	0/103 (0)	49/86 (0.36)	46/70 (0.4)	0.00	0.40	0.48	0.52	0.52		2.02	2.02	2.03	2.01		
79	*COL21A1*	T183	6	56,031,778	c.G1204T	p.D402Y	0/93 (0)	46/79 (0.37)	49/54 (0.48)	0.01	0.30	0.39	0.41	0.47		2.02	2.04	2.04	2.02		
80	*CUL9*	T183	6	43,155,650	c.C1781T	p.P594L	0/57 (0)	23/38 (0.38)	19/18 (0.51)	0.00	0.28	0.35	0.40	0.39		2.02	2.04	2.04	2.02		
81	*DHX33*	M183	17	5,356,916	c.G1380A	p.M460I	0/55 (0)		10/28 (0.26)	0.00	0.00	0.41	0.48	0.46		2.01	2.02	2.05	2.01		
82	*IER5L*	M183	9	131,939,425	c.907_908insC	p.K303fs	0/14 (0)		6/13 (0.32)	0.00	0.01	0.50	0.48	0.58		2.02	2.02	2.03	2.02		

[§], Annotated with hg19 reference genome. M/wt, mutant vs. wild-type calls in Illumina reads; MAR, mutant allele ratio. Orange fill is proportional to MAR; DNA copy numbers were inferred from segmented data at mutation sites and are proportional to the degree of deletion (blue) or amplification (red).

Supplementary Table 7. List of recurrently mutated genes in exomes and their corresponding status in tumorgrafts

						MAR (Sanger Seq.)						Paired Copy Number				
									Tumorgrafts					Tumorgrafts		
Gene	ID	RefSeq	Chr	Position[§]	Mutation	N	T	M	(1)	(2)	(3)	T	M	(1)	(2)	(3)
BAP1	T163	NM_004656	3	52,436,628	c.2028_2046delCACCTTTATCTCCATGCTG	0.00	0.43					**1.46**				
BAP1	T166	NM_004656	3	52,442,066	c.283G>C,p.A95P	0.00	0.51		**1.00**	**1.00**	**1.00**	**1.46**		**1.10**	**1.50**	**1.51**
CEP68	T144	NM_015147	2	65,299,963	c.1733A>T,p.Q578L	0.00	0.31		0.53	0.43	0.00	2.05		1.98	2.00	1.85
CEP68	T183	NM_015147	2	65,301,495	c.1964G>A,p.R655K & c.1966T>G,p.S656A	0.01	0.40	0.46	0.51	0.50		2.02	2.03	2.02	2.03	
COL17A1	T142	NM_000494	10	105,809,175	c.2219C>T,p.G740E	0.02	0.31		0.03	0.04		1.98		2.01	2.01	
COL17A1	T166	NM_000494	10	105,813,716	c.1796G>C,p.P599R	0.00	0.34		0.57	0.57	0.55	**1.50**		**1.52**	**1.52**	**1.53**
LPIN1	T144	NM_145693	2	11,913,805	c.656_661delTGGCTG	0.00	0.39		0.55	0.57	0.54	2.05		1.97	1.95	1.83
LPIN1	T163	NM_145693	2	11,943,157	IVS1905+0dupT	0.00	0.31					2.00				
LRRK2	T144	NM_198578	12	40,760,823	c.7406T>C,p.V2469A	0.00	0.30		0.49	0.43	0.32	2.04		1.95	1.91	1.53
LRRK2	T183	NM_198578	12	40,740,662	c.6217A>G,p.I2073V	0.00	0.28	0.38	0.41	0.43		2.02	2.03	2.02	2.01	
NFE2L1	T144	NM_003204	17	46,128,958	c.485dupC	0.00	0.29		0.60	0.51	0.37	2.03		1.91	1.90	1.51
NFE2L1	T163	NM_003204	17	46,133,961	IVS723+1G>T	0.00	0.26					2.01				
OR11L1	T127	NM_001001959	1	248,004,360	c.830_839delTCTACACTGT	0.00	0.33		0.54	0.55		1.99		2.02	2.03	
OR11L1	T183	NM_001001959	1	248,004,978	c.221C>A,p.T74K	0.00	0.40	0.49	0.54	0.54		2.03	2.03	2.02	2.02	
PBRM1	T142	NM_018313	3	52,595,948	c.4027C>A,p.E1343*	0.00	0.48		**1.00**	**1.00**		**1.40**		**1.10**	**1.07**	
PBRM1	T183	NM_018313	3	52,621,368	c.3027+1C>T	0.02	0.51	0.80	**0.91**	**0.94**		**1.20**	**1.13**	**1.10**	**1.30**	
SPEF2	T144	NM_024867	5	35,654,709	c.859G>A,p.D287N	0.03	0.24		0.46	0.44	0.48	2.06		1.96	2.00	1.98
SPEF2	T166	NM_024867	5	35,659,271	c.1129C>T,p.R377*	0.00	0.44		0.53	0.52	0.51	1.95		1.99	1.94	1.99
VHL	T127	NM_000551	3	10,183,763	c.232_233delAA	0.00	0.20		**1.00**	**1.00**		**1.51**		**1.10**	**1.24**	
VHL	T142	NM_000551	3	10,191,532	c.525_533delCAGGAGACT	0.00	0.23		**1.00**	**1.00**		**1.38**		**1.10**	**1.07**	
VHL	T144	NM_000551	3	10,191,513	c.506T>C,p.L169P	0.02	0.46		**1.00**	**0.98**	**1.00**	**1.53**		**1.12**	**1.09**	**1.07**
VHL	T163	NM_000551	3	10,188,200	c.343C>A,p.H115N	0.00	0.42					**1.42**				
VHL	T166	NM_000551	3	10,183,755	c.224_226delTCT	0.01	0.48		**1.00**	**1.00**	**1.00**	**1.47**		**1.07**	**1.49**	**1.51**
VHL	T183	NM_000551	3	10,188,271	c.414_421delATCTCTCA	0.00	0.55	0.76	**1.00**	**1.00**		**1.15**	**1.11**	**1.08**	**1.32**	

[§], Annotated with hg19 reference genome. MAR, mutant allele ratios. Orange fill is proportional to MAR. DNA copy numbers were inferred from segmented data at mutation sites; blue denotes deletion (PCN≤1.5). N, normal from patient; T, tumor from patient; M, metastasis from patient; Tumorgraft series.

Supplementary Table 8. Tumor cell lines examined for *BAP1* mutation

Cell Line	CDS	Protein
A498	wild type	
A704	wild type	
ACHN	wild type	
Caki-1	wild type	
Caki-2	wild type	
CCF-RC1	wild type	
CCF-RC2	wild type	
NC65	wild type	
OSRC2	wild type	
RPMI-SE	wild type	
SW156	wild type	
UMRC3	wild type	
UMRC5	wild type	
UMRC6	c.430delC	Fs
769-P	c.97T>G	p.Tyr33Asp
786-O	wild type	

Fs, frameshift.

Supplementary Table 9. Sequencing primers for mutation validation

Gene	Sequence	Reference
ABP1	5'- GGGAAGCCCGTGCCGTCATC-3'	
	5'- GCCCAGTGCCCAGCATCTG-3'	
ADAMTSL1	5'- CGACCTCGGAGGAGGACCCG-3'	
	5'- GTCCCCGGTGGAGGAGGTCC-3'	
C1orf167	5'- TGGTGCTGGGCTCTGTGGGT-3'	
	5'- CGCTTGCTGAGTGGCTGGCT-3'	
C14orf43	5'- CCCTTCTGCAGGACTCAGC-3'	Sjöblom *et al. Science* **314**, 268-274 (2006)
	5'- GTGCTCTGGATCACTCCGC-3'	
COL12A1	5'- GGAGTGTCAGGTTGCAGGGGC-3'	
	5'- TCACGTGGCTGTTAAGTGGCTGA-3'	
CPN1	5'- GCTGTTTCTATCCCTGATCCT-3'	Dalgliesh *et al. Nature* **463**, 360-363 (2010)
	5'- ATCCCCACCTTAGTTTTAAAGAA-3'	
CRISPLD1	5'- TGAATGATAAACGTTGGCTTCTC-3'	Sjöblom *et al. Science* **314**, 268-274 (2006)
	5'- GGAACAATCAGGAACGTAAAGG-3'	
CTNND1	5'- TGTGGACAGCACCTACACTTG-3'	Dalgliesh *et al. Nature* **463**, 360-363 (2010)
	5'- ATCCCAAATAGCGACACAAGA-3'	
DIAPH1	5'- ACAATGCCCATGTGTTCCTT-3'	Sjöblom *et al. Science* **314**, 268-274 (2006)
	5'- AAATCATTTGCCAGGAGGTG-3'	
ERBB4	5'- AAGCCAATTCTTTAGAATATGATATGG-3'	Sjöblom *et al. Science* **314**, 268-274 (2006)
	5'- TGCCTTAGAGTGTTCCTCAATG-3'	
ESPL1	5'- CTGGGAAAGAGGCAAAGAGA-3'	Dalgliesh *et al. Nature* **463**, 360-363 (2010)
	5'- GGCAACGAAACCTTTACATAGAA-3'	
FAM117B	5'- TGGGTTGTCTGTTCTATGTCTAGGCC-3'	
	5'- ACAGCCGACCAGAAGCGCAC-3'	
FRMD1	5'- CTGCCACCTTCCAGCGACCACA-3'	
	5'- GCCCTGACCTTACCCACTGT-3'	
GFPT2	5'- AAAGAACAGAAAGGAACTCTGGGGT-3'	
	5'- ATGCTATTTCTCTGCTCTAGGGTCGCC-3'	
GTF2A1L	5'- AACCGTGCTACAGCAACCCG-3'	
	5'- GTTACTGGCCACTATCTGTCCCCT-3'	
HS6ST3	5'- TACAACCTGGCTAACAATCGC-3'	Sjöblom *et al. Science* **314**, 268-274 (2006)
	5'- GAATGGCTCAGCTTCTCCAAG-3'	
MYH8	5'- GTGGAAAGGTCTAATGCAGCC-3'	Sjöblom *et al. Science* **314**, 268-274 (2006)
	5'- TCTGTGGTTGAACAGACAGGAG-3'	
NME5	5'- GAACCTTTTTAATGATTGTATAGTA-3'	
	5'- GGGTTTCACCATGTTAGTCAGGCTGG-3'	
PDGFB	5'- GTTGAAGGGCGTGAGAAAGAG-3'	Sjöblom *et al. Science* **314**, 268-274 (2006)
	5'- GGAAGCCTGGTCAGGTATGAG-3'	
PLB1	5'- ACCAGAGCAAGGCTAGGCCA-3'	
	5'- ACCGTCCCCTTTCTCCTGGGC-3'	
PMS2	5'- CGCAACCCCAAACACAAAG-3'	Dalgliesh *et al. Nature* **463**, 360-363 (2010)
	5'- GCCCTAAACTTCCTGTAATTCTGTTC-3'	
PTPRK	5'- AGAACCCCCTACATGCAAAT-3'	Dalgliesh *et al. Nature* **463**, 360-363 (2010)
	5'- GGTGTGTTAATAGTCAACCCCTT-3'	
SCN1A	5'- TCCTAGATGGAAGGCACATTAGCA-3'	
	5'- TGGGAGCCCTGATCCAGTCTGT-3'	
SPHKAP	5'- TGTGATGTGATGTTTTGAGAATG-3'	Dalgliesh *et al. Nature* **463**, 360-363 (2010)
	5'- TCATTCCACCTAATCCTCTGC-3'	
STK24	5'- TGAACAGCATGCTTCCCAG-3'	Sjöblom *et al. Science* **314**, 268-274 (2006)
	5'- GGAACAGACAGGCCAAGAATC-3'	
UBE3B	5'- GTGTCTGGGGCTTGACCTC-3'	Dalgliesh *et al. Nature* **463**, 360-363 (2010)
	5'- CCACAAGGAGTCTCATGGCT-3'	
WDFY4	5'- TGGGACAAGTGGGTGAAGGGGA-3'	
	5'- AACCAGGGCCTCTGGGCAGG-3'	
ZNF434	5'- TCGGTGGGCACTGAAGTGGGA-3'	
	5'- CCCCACAAGCCAGCGCTCAA-3'	

Supplementary Table 10. Antibodies

Antibody	Vendor	Type	IHC conditions
BAP1	Santa Cruz Biotechnology Inc., CA	mouse monoclonal, clone C-4	1:50, 60 min
FLAG	Sigma-Aldrich, MO	mouse monoclonal, clone M2	
H2Aub1	EMD Millipore Corp., MA	mouse monoclonal, clone E6C5	
H2A	EMD Millipore Corp., MA	rabbit polyclonal	
H2B	EMD Millipore Corp., MA	rabbit polyclonal	
γH2A.X (Ser139)	EMD Millipore Corp., MA	mouse monoclonal, clone JBW301	
HCF-1	Bethyl Laboratories Inc., TX	rabbit polyclonal	
Menin	Bethyl Laboratories Inc., TX	rabbit polyclonal	
PBRM1	Bethyl Laboratories Inc., TX	rabbit polyclonal	1:250, 32 min
PPIB	Abcam, MA	rabbit polyclonal	
Ⓟ-S6 (Ser240/244)	Cell Signaling Technology Inc., MA	rabbit polyclonal	
Ⓟ-S6 (Ser235/236)	Cell Signaling Technology Inc., MA	rabbit monoclonal	1:200, 60 min
Ⓟ-S6K (Thr389)	Cell Signaling Technology Inc., MA	rabbit polyclonal	
Ⓟ-4E-BP1 (Thr37/46)	Cell Signaling Technology Inc., MA	rabbit monoclonal, clone 236B4	1:200, 60 min
Rad51	Santa Cruz Biotechnology Inc., CA	rabbit polyclonal H-92	
Tubulin	Sigma-Aldrich, MO	mouse monoclonal, clone B-5-1-2	
4E-BP1	Cell Signaling Technology Inc., MA	rabbit polyclonal	