BAP1 loss defines a new class of renal cell carcinoma

Samuel Peña-Llopis1–3, Silvia Vega-Rubín-de-Celis1–3, Arnold Liao4, Nan Leng4, Andrea Pavia-Jiménez1–3, Shanshan Wang1–3, Toshinari Yamasaki1–3, Leah Zhrebker1–3, Sharanya Sivanand1–3, Patrick Spence1–3, Lisa Kinch5, Tina Hambuch5, Suneer Jain4, Yair Lotan5, Vitaly Margulis6, Arthur I Sagalowsky6, Pia Banerji Summerour3–7, Wareef Kabbani8, S W Wendy Wong9, Nick Grishin5, Marc Laurent4, Xian-Jin Xie3, Christian D Haudenschild4, Mark T Ross9, David R Bentley9, Payal Kapur8 & James Brugarolas1–3

Kidney cancer is estimated to have been diagnosed in over 60,000 individuals in the United States in 2011 (ref. 1). Most kidney tumors are RCC, and 70% are the clear cell type (ccRCC)2. Despite recent advances1, when metastatic, ccRCC remains largely incurable.

ccRCC is characterized by inactivation of the VHL gene (encoding the von Hippel-Lindau protein)3–6. VHL, which is on chromosome 3p25, is a two-hit tumor suppressor gene. One allele is typically inactivated through a point mutation (or indel), and the other is inactivated through a large deletion resulting in loss-of-heterozygosity (LOH)7,8. Also on chromosome 3p is PBRM1 (encoding Polybromo 1), which is frequently mutated in ccRCC9.

Other genes implicated in ccRCC development include SETD2 (ref. 10), KDM5C10 and KDM6A11, but the mutation frequency of each is estimated at <5% (refs. 10,11).

ccRCCs are classified into low- and high-grade tumors12, and nuclear grade is an important prognostic factor13,14. High-grade tumors have mammalian target of rapamycin (mTOR) complex 1 (mTORC1) activation15. mTORC1 is a critical regulator of cell growth and is negatively regulated by a complex formed by the tuberous sclerosis complex 1 (TSC1) and 2 (TSC2) proteins16. MTOR10,17 and TSC1 (ref. 18) are both mutated in sporadic ccRCC; however, mutations are infrequent19, and the genetic determinants of tumor grade remain largely unknown.

The molecular pathogenesis of renal cell carcinoma (RCC) is poorly understood. Whole-genome and exome sequencing followed by innovative tumorgraft analyses (to accurately determine mutant allele ratios) identified several putative two-hit tumor suppressor genes, including BAP1. The BAP1 protein, a nuclear deubiquitinase, is inactivated in 15% of clear cell RCCs. BAP1 cofractionates with and binds to HCF-1 in tumorgrafts. Mutations disrupting the HCF-1 binding motif impair BAP1-mediated suppression of cell proliferation but not deubiquitination of monoubiquitinated histone 2A lysine 119 (H2AK119ub1). BAP1 loss sensitizes RCC cells in vitro to genotoxic stress. Notably, mutations in BAP1 and PBRM1 anticorrelate in tumors (P = 3 × 10–5), and combined loss of BAP1 and PBRM1 in a few RCCs was associated with rhabdoid features (q = 0.0007). BAP1 and PBRM1 regulate seemingly different gene expression programs, and BAP1 loss was associated with high tumor grade (q = 0.0005).

Our results establish the foundation for an integrated pathological and molecular genetic classification of RCC, paving the way for subtype-specific treatments exploiting genetic vulnerabilities.

RESULTS
Identification of candidate two-hit tumor suppressor genes
We sequenced the genome of a sporadic, high-grade ccRCC and paired normal sample to >94% coverage and a mean depth of ≥35x (Supplementary Figs. 1 and 2). We found 6,571 somatically acquired single-nucleotide mutations or indels, including 59 in protein-coding regions (Supplementary Table 1). Every mutation evaluated was confirmed by Sanger sequencing (Table 1 and Supplementary Table 2). However, mutant allele ratios (MARs)—the fraction of mutant over mutant and wild-type alleles for each mutation—were low; few mutations reached a MAR of 0.5 (expected for heterozygous mutations), and no mutations reached 1 (expected for mutations accompanied by LOH) (Table 1 and Supplementary Table 2). For VHL, the MAR was 0.52 (Fig. 1a and Table 1), which suggested a heterozygous mutation. However, these results conflicted with DNA copy-number analyses showing that one copy of 3p was lost (Fig. 1b). We attributed the low MARs to tumor contamination by normal stroma. Contamination occurred despite careful sample selection (Supplementary Fig. 2c).

Tumor implantation in mice expands the neoplastic compartment, whereas human stroma is replaced by the host20. Therefore, tumorgrafts may be used to calculate MARs with accuracy. RCC tumors implanted orthotopically in mice preserve the characteristics of
human tumors. We performed Sanger sequencing of mutated genes in a tumorgraft derived from the index subject's tumor using human-specific primers. In comparison to tumor MAR (MAR<sub>T</sub>) values, tumorgraft MAR (MAR<sub>TG</sub>) values often increased to ~0.5, and, for several genes including VHL, they reached 1 (Table 1, Supplementary Fig. 3 and Supplementary Table 2).

To determine whether MAR<sub>TG</sub> values reflected those expected in the index subject's tumor, we asked whether a correlation existed between MAR<sub>TG</sub> and corresponding regional DNA copy numbers in the tumor (Fig. 1b, Table 1 and Supplementary Table 2). A correlation was found with MAR<sub>TG</sub> (P = 1.3 × 10<sup>-5</sup>) but not with MAR<sub>T</sub> (P = 0.054). These data suggest that MARs in tumors are more accurately determined by evaluating tumorgrafts. Consistent with the notion that tumorgrafts represent largely pure populations of human tumor cells, paired copy numbers (PCNs) and allele-specific copy numbers (ASCNs) in tumorgrafts more closely approached integer values (Fig. 1b, Table 1 and Supplementary Table 2).

To identify putative two-hit tumor suppressor genes, we searched for genes with MAR<sub>TG</sub> values of ~1. Some genes (STK40, UBE3B, HS6ST3, STK24, C1orf167, ADAMTS11 and C4orf43) were in regions of deletion (PCN<sub>TG</sub> values of ~1), whereas others (CRISPLD1, TMEM151A, TREH and CTND1) were in areas of copy-neutral LOH (PCN<sub>TG</sub> values of ~2); tumorgraft ASCN<sub>min</sub> values of ~0 and ASCN<sub>max</sub> values of ~2) (Fig. 1b and Table 1). Because mutations in regions of copy-neutral LOH could be either homozygous (for example, CRISPLD1) or heterozygous (for example, ZNF434) (Fig. 1b and Table 1), accurate MARs were essential to establish whether mutated genes were putative two-hit tumor suppressors.

Accurate MARs were also helpful in inferring whether, in areas of duplication (PCN<sub>TG</sub> of ~3), the allele amplified was mutated (for example, GFTP2; MAR<sub>TG</sub> = 0.65 (expected 0.66)) or wild type (for example, DIAPH1; MAR<sub>TG</sub> = 0.31 (expected 0.33)) (Table 1). In the case of GFTP2, the mutation may have preceded the duplication, whereas in the case of DIAPH1, the mutation is likely to have followed the duplication. Thus, analyses in tumorgrafts identified candidate tumor suppressor genes and shed light on the temporal sequence of mutation acquisition.

**Table 1** Integrated analysis of a subset of somatic mutations and DNA copy-number alterations in the index subject

| Gene   | Chr. | Position<sup>a</sup> | Nucleotide change | Illumina T | TG | PCN | Min | Max | PCN | Min | Max | Change |
|--------|------|----------------------|-------------------|------------|----|-----|-----|-----|-----|-----|-----|--------|
| C1orf167 | 1    | 11767238             | G>T               | 0.30       | 0.36 | 1.00 | 1.39 | 0.43 | 1.00 | 1.02 | 0.003 | 1.04 | Splice site |
| STK40  | 1    | 36593365             | C>A               | 0.30       | 0.32 | 1.00 | 1.39 | 0.43 | 1.00 | 1.00 | 0.003 | 1.04 | p.Met1313le |
| VHL    | 1    | 10166479             | C>G               | 0.37       | 0.52 | 1.00 | 1.39 | 0.43 | 1.07 | 0.98 | 0.003 | 1.05 | p.Leu158Val |
| DIAPH1 | 5    | 140885872            | C>T               | 0.26       | 0.20 | 0.31 | 2.53 | 0.97 | 1.55 | 3.05 | 0.97 | 2.00 | Arg1164Gln |
| GFTP2  | 5    | 179662025            | G>A               | 0.43       | 0.38 | 0.65 | 2.51 | 0.97 | 1.54 | 3.05 | 0.97 | 2.00 | Splice site |
| CRISPLD1 | 8    | 76088864             | G>A               | 0.57       | 0.56 | 1.00 | 2.02 | 0.41 | 1.63 | 1.98 | 0.003 | 2.00 | Vals200le |
| ADAMTS11 | 9    | 18767566             | del9              | 0.25       | 0.34 | 1.00 | 1.42 | 0.44 | 1.12 | 1.01 | 0.004 | 1.06 | p.Glu1114Lgin1116del |
| CTND1  | 11   | 57333402             | delG              | 0.36       | 0.38 | 1.00 | 1.74 | 0.41 | 1.16 | 1.95 | 0.003 | 1.96 | p.Val769Serfs*5 |
| TMEM151A | 11   | 65818643             | G>T               | 0.36       | 0.18 | 1.00 | 1.62 | 0.41 | 1.16 | 1.93 | 0.003 | 1.96 | Cys117Phe |
| TREH   | 11   | 118035289            | C>A               | 0.54       | 0.50 | 1.00 | 1.62 | 0.41 | 1.19 | 1.89 | 0.003 | 1.96 | Gly478Cys |
| UBE3B  | 12   | 108456960            | A>T               | 0.37       | 0.40 | 1.00 | 1.39 | 0.43 | 1.01 | 1.00 | 0.003 | 1.06 | p.Glu1066Tyr |
| HS6ST3 | 13   | 96283428             | A>T               | 0.38       | 0.38 | 1.00 | 1.39 | 0.43 | 1.06 | 0.98 | 0.004 | 1.04 | p.Try464Phe |
| STK24  | 13   | 97907504             | C>T               | 0.28       | 0.38 | 1.00 | 1.39 | 0.43 | 1.06 | 0.98 | 0.004 | 1.04 | Arg405Gln |
| C14orf43 | 14    | 73275194            | del50             | 0.32       | 0.35 | 1.00 | 1.39 | 0.44 | 1.07 | 0.99 | 0.004 | 1.05 | p.Gln408Glyfs*65 |
| ZNF434 | 16   | 3373160              | T>C               | 0.30       | 0.30 | 0.55 | 1.95 | 0.41 | 1.57 | 1.98 | 0.003 | 1.99 | p.Gln384Arg |

Mutation analyses of whole-genome sequences from a tumor-normal pair and the corresponding tumorgraft in the index subject. DNA copy numbers were inferred from segmented data at mutation sites. For heterozygous SNPs, min and max represent the ASCNs of the minor and major alleles, respectively. Bold copy numbers denote deletion (PCN < 1.5 or ASCN < 0.5) or amplification (PCN > 2.5 or ASCN > 1.5). Chr., chromosome. A complete list of mutations is provided in Supplementary Table 2.

*Annotated with NCBI36.1 and Ensembl build 54.

---

Evaluation of somatically mutated genes in a discovery cohort

Twenty-one genes mutated in the sequenced ccRCC tumor and not previously examined by the Sanger Institute were sequenced in a discovery set of 76 ccRCCs, and mutations were examined in the corresponding normal samples (Supplementary Table 3). As determined by VHL sequencing, which revealed somatically acquired mutations in 79% of tumors (Supplementary Data 1), sensitivity for mutation detection was excellent. Several putative two-hit tumor suppressor genes were mutated at higher than expected frequencies, including CRISPLD1, which was mutated in two additional tumors (q = 0.044), and TMEM151A, mutated in three additional tumors (q = 0.005). In addition, several other genes were recurrently mutated, including OCA2 and ND1 (also known as MT-ND1) (Supplementary Table 4). Germline mutations in OCA2 cause autosomal recessive oculocutaneous albinism type 2, and the two somatic mutations we identified (encoding p.Pro211Leu and p.Val443Leu alterations; Supplementary Table 4) are known disease-causing mutations. Two additional somatically acquired mutations were found in mitochondrial ND1 (Supplementary Table 4), a gene mutated in oncocytomas, a benign tumor type. The presence of these mutations in ccRCC suggests that oncocytomas could transform into malignant tumors. Transformation may result from VHL inactivation, which was observed in all the tumors with somatic ND1 mutations (Supplementary Data 1). VHL inactivation could change the morphological appearance of the tumor by affecting cellular metabolism and angiogenesis. In addition, three mutations were identified in TSC1, which we previously reported.

Exome sequencing identifies two-hit tumor suppressor gene BAP1

We performed exome sequencing of seven ccRCC primary tumors, including six of high grade, and corresponding normal samples. A metastasis from one affected individual was also sequenced. We performed exome sequencing of seven ccRCC primary tumors, including six of high grade, and corresponding normal samples. A metastasis from one affected individual was also sequenced. We found 345 somatically acquired mutations (Supplementary Table 5). In the tumor-metastasis pair, we observed 37 and 39 mutations, respectively, and 32 were shared.

To determine the accuracy of the mutations called, we performed Sanger sequencing. If concordance was >95%, Sanger sequencing of 82 mutations would predict >90% accuracy for the whole cohort.
Among 82 randomly selected mutations, 78 were confirmed with an accuracy of ≥95% (Supplementary Table 6 and Supplementary Data 2). For 5 tumors, there were tumorgrafts available, and sequencing analyses of mutated genes therein uncovered 16 potential two-hit tumor suppressor genes (Supplementary Table 6).

We focused on ten genes mutated in at least two tumors (Supplementary Table 7). All mutations were validated by Sanger sequencing. Whereas MAR<sub>T</sub>-analysis failed to identify any putative two-hit tumor suppressors, another gene in addition to VHL and PBRM1, BAP1, showed MAR<sub>TG</sub> values of ~1 (Supplementary Table 7).

BAP1 sequencing in the discovery set of 76 ccRCCs identified 11 nonsynonymous mutations, including 10 confirmed to be somatically acquired (Table 2). Examination of a validation ccRCC set (n = 92) with corresponding normal samples uncovered 11 additional nonsynonymous mutations, including 10 that were somatically acquired (Table 2). Two mutations in tumors without matching normal samples were truncating and likely deleterious. Altogether, the BAP1 mutation rate was 14% (24/176 tumors). BAP1 encodes a nuclear deubiquitinating (DUB) of the ubiquitin C-terminal hydrolase (UCH)-domain containing family family members Uch-L3 and Uch37 (Supplementary Table 7).

Development of a clinical assay for BAP1 detection
As most mutations were truncating, we developed in a laboratory with Clinical Laboratory Improvement Amendments (CLIA) certification an immunohistochemistry (IHC) test for the presence-absence of BAP1 protein. Genetically characterized ccRCC samples validated by protein blot were used as controls (Fig. 2c,d). Scoring was performed by a clinical pathologist who was blinded to the BAP1 genotype. IHC results were interpretable in 175 out of 176 tumors. Nuclear BAP1 was detected in 150 tumors, and 148 were wild type for BAP1 (Supplementary Fig. 4). The two discordant samples had missense mutations (encoding p.Gly13Val and p.Phe170Leu changes). Twenty-five samples were negative by IHC, and 22 of these had BAP1 mutations. Analysis of an IHC-negative sample that had wild-type BAP1 protein. Genetically characterized ccRCC samples validated by protein blot failed to reveal detectable BAP1 protein, suggesting that other mechanisms exist to inactivate BAP1. Overall, the positive and negative predictive values of the IHC test were ~100% and 98.6%, respectively.

Structural analyses of BAP1 missense mutations
To evaluate the effects of BAP1 missense mutations in a structural context, we generated a BAP1 protein model on the basis of the related family members Uch-L3 and Uch37 (Fig. 2b). Because ubiquitin binding orders a significant portion of the protein, the UCH domain...
Table 2 List of BAP1 mutations in ccRCCs and cell lines

| ID    | Coding sequence mutation | Protein                                      |
|-------|--------------------------|----------------------------------------------|
| 3575  | c.5_6dupAT               | p.Lys31Ilefs*33                              |
| 63    | c.21_32del12             | p.Glu77Asp,Leu8_Asp11del                    |
| T145  | c.38G>T                  | p.Gly13Val                                  |
| T211  | c.58G>T                  | p.Glu20*                                     |
| T16   | c.128T>G                 | p.Val43Gly                                   |
| T114  | c.193delT                | p.Leu65Trpfs*7                               |
| T166  | c.283G>C                 |                                             |
| T69   | c.335T>C                 | p.Glu297*                                    |
| T115  | c.430C>A                 | p.His144Asn                                  |
| 3397  | c.IVS438–1G>A            | Splice site                                   |
| T55   | c.458delC                | p.Pro153Leufs*34                             |
| T212  | c.510T>A                 | p.Phe170Leu*                                 |
| T184  | c.889G>T                 | p.Glu297*                                    |
| 209   | c.971delC                |                                             |
| 162   | c.1219delG               | p.Asp407Metfs*19                             |
| T149  | c.1256delA               |                                             |
| T26   | c.1271_1274delGGA        | p.Lys425Glnfs*4                              |
| 78    | c.1793delC               | p.Pro598Glnfs*19                             |
| 9575  | c.1981A>T                | p.Lys661*                                    |
| T163  | c.2028_2046del19         | p.Cys676Trps*18                              |
| T70   | c.2050C>T                | p.Gln68*                                    |
| T25   | c.2051delA               | p.Gln684Glnfs*16                             |
| 40    | c.2134C>T                | p.Gln712*                                    |
| 9145  | c.2188T>G                | p.Tyr33Asp*                                  |
| 769-P | c.97T>G                  |                                             |
| UMRC6 | c.430delC                | p.His144Metfs*94                             |

*Missense mutations not affecting protein levels.

of BAP1 was modeled after that of Uch-L3 bound to ubiquitin (Protein Data Bank (PDB) 1xd3). The interaction with the ULD domain was built by superimposing that of Uch37 (PDB 3hhr). Four alterations abrogated protein expression: three were predicted to destabilize the protein (p.Val43Gly and p.Leu112Pro removed side chains that contribute to the hydrophobic core, and p.Ala95Pro disrupted the backbone of a central α helix), and the fourth (p.His144Asn) disrupted the position of a flexible loop (Fig. 2b). Two alterations did not abrogate protein expression (p.Gly13Val and p.Phe170Leu). These amino-acid changes disrupted side chains implicated in either an intramolecular interaction with the ULD domain (Gly13) or ubiquitin binding (Phe170) and highlight the importance of these interactions for tumor suppressor function.

BAP1 suppresses RCC cell proliferation

Studies of the role of BAP1 in cell proliferation have given conflicting results25–27,30–33. To examine BAP1 in an appropriate context, ccRCC cell lines were sought in which natural selection had led to BAP1 inactivation. Among 12 ccRCC cell lines initially examined, only 769-P had a BAP1 mutation (Supplementary Table 8). The mutation (c.97T>G; p.Tyr33Asp) disrupted a residue binding ubiquitin and did not abrogate protein expression (Figs. 2b and 3a).

To determine the role of BAP1, 769-P cells were reconstituted with epitope-tagged wild-type BAP1 (or an empty vector control). BAP1 repressed cell proliferation without causing apoptosis (Fig. 3a and data not shown). However, BAP1 did not completely abrogate cell proliferation. To determine whether endogenous, mutant BAP1 acted in a dominant-negative fashion, we depleted endogenous BAP1 using small hairpin RNA (shRNA). However, depletion of mutant BAP1 did not increase the effects of ectopically expressed wild-type BAP1, indicating that mutant BAP1 did not function in a dominant-negative fashion (Fig. 3b).

BAP1 debiquitinates H2AK119ub1 in renal cancer cells

The protein encoded by the BAP1 ortholog in Drosophila melanogaster, Calypso, targets monoubiquitinated histone H2A (H2Aub1)34. An examination of H2AK119ub1 levels in 769-P cells reconstituted with wild-type BAP1 showed downregulation of basal H2Aub1 levels, indicating that mammalian BAP1 debiquitinates H2A in renal cancer cells (Fig. 3c).

HCF-1 binding is required for suppression of cell proliferation

BAP1 interacts with host cell factor–1 (HCF-1)31,33,35, which serves as a scaffold for several chromatin-remodeling complexes36. HCF-1 binds to multiple transcription factors, including several E2F37,38, and recruits histone-modifying enzymes, such as Set1/MLL1 histone methyltransferases39–41, LSD1 histone demethylase42, Sin3 histone deacetylase39 and MOF histone acetyltransferase43.

We asked whether BAP1 interacted with HCF-1 in 769-P cells. An interaction was confirmed by reciprocal immunoprecipitation experiments (Fig. 3d). Notably, HCF-1 immunoprecipitation depleted BAP1 from cell extracts to the same extent as BAP1 immunoprecipitation, suggesting that, as in other cell types35, the majority of BAP1 in renal cancer cells is bound to HCF-1 (Fig. 3d). BAP1 has been proposed to debiquitinicate H2A (refs. 31,33) and regulate HCF-1 levels31, but, consistent with other reports31, HCF-1 levels were similar in BAP1-deficient and -reconstituted 769-P cells (Fig. 3d).

We mutated sequences in BAP1 encoding the HCF-1 binding motif and evaluated this mutant (HBM) in cell proliferation assays. HBM suppressed HCF-1 binding and compromised the inhibitory effect of BAP1 on cell proliferation (Fig. 3e). However, the HBM mutant did not differ from wild-type BAP1 in its ability to debiquitiniate histone H2A (Fig. 3f). Thus, BAP1 binds to HCF-1, and binding to HCF-1 but not H2Aub1 debiquitinination is important for the inhibition of cell proliferation.

Next, we performed gel-filtration chromatography to further examine BAP1-containing complexes. Extracts from 769-P cells expressing either an empty vector or wild-type BAP1 were fractionated using a size-exclusion column and subjected to protein blotting. Most BAP1 was found in complexes of >1 MDa and eluted with HCF-1 (Supplementary Fig. 5).

BAP1 loss sensitizes RCC cells to radiation and PARP inhibitors

BAP1 is phosphorylated following DNA damage44,45, and we asked whether BAP1 loss affected the response to γ-irradiation. 769-P cells with or without wild-type BAP1 showed a similar pattern of foci of Rad51 and H2AX phosphorylated at Ser139 (γH2AX) (Supplementary Fig. 6a). However, BAP1-deficient cells were more sensitive to ionizing radiation (Supplementary Fig. 6b), and fewer colonies formed in clonogenic assays (Supplementary Fig. 6c). In addition, BAP1 loss sensitized cells to the PARP inhibitor olaparib (Supplementary Fig. 6d,e).

We examined four additional ccRCC cell lines (Supplementary Table 8). UMRC6 lacked BAP1 protein and had a frameshift mutation (c.430delC) (Supplementary Fig. 7a,b). As in 769-P cells, (i) cell proliferation was inhibited by wild-type BAP1 and substantially less so by an HBM mutant, (ii) the HBM mutant reduced H2Aub1 levels, (iii) BAP1 cofractionated with HCF-1, and (iv) restoration of BAP1 protected UMRC6 cells against genotoxic death (Supplementary Fig. 7).

BAP1 binds HCF-1 and elutes with HCF-1 in tumorgrafts

The usefulness of RCC cell lines is limited by the development of mutations and copy-number alterations as tumor cells adapt to...
growth in culture. 

Figure 2 BAP1 is a tumor suppressor in ccRCC. (a) Schematic of BAP1 with alterations. BRCA1, putative BRCA1-interacting domain; ULD, Uch37-like domain; NLS, nuclear localization signal; I, insertion; Δ, deletion; †, missense; *, nonsense; S, splice site; L, stop codon loss. (b) Structural model of the BAP1 UCH domain (purple) and ULD tail (green) bound to ubiquitin (cyan); structural elements that alter upon ubiquitin binding are colored salmon. Left, cartoon of BAP1 model: mutated residues abrogating protein expression are labeled in gray. Right, surface representation of BAP1 highlighting the positions of RCC alterations on interaction surfaces. Left inset, enlarged view of the DUB active site. Right inset, enlarged view of the Gly13 interaction with an aromatic residue in the ULD tail (dots indicate interaction radius). (c) Protein blots of extracts from tumors with defined BAP1 mutation and chromosome 3p status. 769-P cells transfected with either an empty vector (EV) or a vector expressing wild-type (WT) BAP1 were used as controls. PPIB (cyclophilin B) is shown as a loading control. Arrow indicates BAP1. (d) Representative IHC of tumors and tumorgrafts positive or negative for nuclear BAP1. Scale bar, 50 µm. Open arrows indicate tumor cells; simple arrows indicate endothelial cells and lymphocytes, which express BAP1 and serve as internal controls.

BAP1 loss is associated with high tumor grade

Deep-sequencing studies were largely focused on high-grade tumors. An analysis of all 176 tumors examined showed that BAP1 loss correlated with high Fuhrman nuclear grade ($q = 0.0005$) (Supplementary Data 1). Because nuclear grade is associated with mTORC1 activation, we tested whether a correlation existed between BAP1 loss and mTORC1 activity. As determined by the phosphorylation of both S6 and 4E-BP1, BAP1 loss was correlated with mTORC1 activation ($q = 3 \times 10^{-4}$ and 0.029, respectively) (Fig. 5a and Supplementary Data 1). This association did not seem to be direct, however, and similar levels of mTORC1 activation were observed in BAP1-deficient and wild-type BAP1–reconstituted cells (Supplementary Fig. 8).

BAP1 and PBRM1 mutations anticorrelate in ccRCC

To explore whether a relationship existed between loss of BAP1 and PBRM1, we first developed an IHC assay for PBRM1 (also known as BAF180) (Fig. 5a). Evaluation of the 176 tumors showed confident PBRM1 staining for 146 samples, and 53% were negative for PBRM1 (Supplementary Data 1). As PBRM1 was lost in ~50% of tumors, BAP1 loss should distribute equally between PBRM1-expressing and -deficient tumors. However, only 4 of 21 BAP1-deficient tumors were also deficient for PBRM1 (Supplementary Fig. 9a). These results suggest that PBRM1 and BAP1 loss are anticorrelated in tumors ($P = 7 \times 10^{-4}$).

To explore this further, we sequenced $PBRM1$ in the 176 ccRCCs. We identified 94 somatic mutations, including 6 missense mutations (Supplementary Data 1). Structural analyses of the effects of these mutations are shown in Supplementary Figure 10. We correlated sequencing data with the results from IHC: 90% of samples that were positive for PBRM1 by IHC had a mutation, and 90% of the samples that were positive had wild-type sequence ($P = 4 \times 10^{-23}$;
**Figure 3** HCF-1–dependent suppression of cell proliferation by BAP1. (a) Proliferation curves of 769-P cells stably transduced with an empty vector (EV) or with a vector encoding Flag-tagged BAP1. Inset, BAP1 protein blot. (b) Proliferation curves of 769-P cells stably expressing an shRNA targeting endogenous (mutant) BAP1 or a vector control (control) and transduced with vectors encoding wild-type BAP1 (WT), BAP1 p.Tyr33Asp (Y33D) or an empty vector. Protein blots are shown for cells transduced with shRNA vectors (A and B) (top) and subsequently transduced with the indicated expression vectors (bottom). (c) Protein blots of partially purified histone fractions from 769-P cells transduced with an empty vector or wild-type BAP1. (d) Protein blot of input (cytosolic (C) or nuclear (N) fractions) as well as of immunoprecipitates (from nuclear fractions) and corresponding flow-through from cells transduced with an empty vector (−) or Flag-tagged wild-type BAP1 (+). Both ectopically expressed wild-type and endogenous, mutant BAP1 bind HCF-1. (e) Proliferation curves of 769-P cells depleted of endogenous BAP1 by shRNA and stably transduced with an empty vector, wild-type BAP1 or an HCF-binding motif mutant (HBM). Protein blots are shown for input and BAP1 immunoprecipitates. (f) Protein blot of partially purified histone fractions or cell lysates from 769-P cells depleted of endogenous BAP1 and transduced with an empty vector, wild-type BAP1, an HBM mutant or BAP1 p.Tyr33Asp. Error bars, s.e.m. (n = 3). *P < 0.05; **P < 0.01.

**Supplementary Fig. 9b.** An analysis of BAP1 and PBRM1 mutations in tumors revealed that only 3 of 24 samples with BAP1 mutations had a somatically acquired PBRM1 mutation (Supplementary Fig. 9c). Thus, mutations in these two genes were found to anti-correlate in tumors (P = 3 × 10⁻³). For comparison, we evaluated the distribution of mutations in SETD2 and KDM5C with respect to PBRM1 mutations in ccRCCs from the Sanger Institute⁹,¹⁰. Among 348 ccRCCs genotyped for PBRM1, 15 mutations in SETD2 were observed, and these mutations were distributed equally between tumors with mutant and wild-type PBRM1 (in 8 and 7 tumors, respectively; Supplementary Data 3). KDM5C mutations were similarly distributed across tumors with mutant and wild-type PBRM1 (in five and four tumors, respectively; Supplementary Data 3).

Combining the IHC and mutation data, 5 out of 27 BAP1-deficient tumors were found to also be deficient in PBRM1. Assuming a binomial distribution of BAP1 loss, these data indicate that simultaneous inactivation of BAP1 and PBRM1 is negatively selected for in tumors (P = 0.0008). Notably, however, loss of BAP1 or PBRM1 was observed in 70% of ccRCCs (Fig. 5b).

To obtain further insight into the relationship between BAP1 and PBRM1, we performed gene expression analyses. We grouped tumors and tumorigrafts according to their BAP1 and PBRM1 status and evaluated differences with respect to wild-type tumors and tumorigrafts (Fig. 5c). Probe sets (probes) that we had previously determined, using tumorigrafts, to be driven by non-neoplastic cells²¹ were excluded from the analysis. We identified 1,451 probes that were deregulated in BAP1-deficient tumors relative to tumors that were wild type for both BAP1 and PBRM1 (q < 0.05) (Supplementary Data 4). A similar number of probes distinguished PBRM1-deficient tumors (Supplementary Data 4). These two data sets had 94 probes in common (Fig. 5d). However, the overlap expected to occur at random

**Figure 4** BAP1 binds to and elutes with HCF-1 in tumorigrafts. (a) Protein blots of input and BAP1 immunoprecipitates (BAP1-IP) from the indicated tumorigrafts. WT, wild type; M, mutant. (b) Protein blots of trichloroacetic acid (TCA)-precipitated gel-filtration fractions of tumorigrafts with either wild-type (TG143 and TG144) or mutant (TG26) BAP1. C, control lysate from 769-P cells. (c) Protein blots of partially purified histone fractions from tumorigrafts with the indicated BAP1 status.
was 67. Similarly, pathway analyses of the two expression signatures showed little overlap. These results suggest that BAP1 and PBRM1 do not function in the same pathway and that the tumorigenic advantage to mutating BAP1 and PBRM1 is context dependent.

Further supporting the notion that loss of BAP1 and PBRM1 in tumors is not equivalent, analyses of the 176 tumors showed that PBRM1 loss was not associated with high tumor grade \((q = 0.26)\) (Supplementary Data 4). In the 348 ccRCC tumors sequenced by the Sanger Institute\(^9^{10}\) (Supplementary Data 3), we found a non-significant correlation between PBRM1 loss and low tumor grade \((P = 0.074)\). Furthermore, when focusing the analyses of the 176 tumors on those that had exclusively lost PBRM1, a statistically significant correlation with low tumor grade was found \((q = 0.025)\).

Tumors with loss of BAP1 and PBRM1 have rhabdoid features

A few tumors had loss of both BAP1 and PBRM1 \((n = 5)\) (Supplementary Data 1). Although occurrence of mutations in tumors may not indicate their simultaneous presence in the same cell and there is substantial mutation heterogeneity in RCC\(^{17,47}\), in two tumors for which tumorigrafts were available, \(\Delta A R T\) values for both BAP1 and PBRM1 were \(-1\), and no wild-type alleles were detected (data not shown). These data suggest that the two mutations were indeed present in the same tumor cells and highlight another application of tumorigrafts.

Tumors deficient in both BAP1 and PBRM1 were uniformly of high grade and showed characteristic features: abundant acidophilic cytoplasm, eccentric nuclei and prominent macronucleoli (Fig. 5a). These features were consistent with rhabdoid morphology\(^{49}\), a form of dedifferentiation portending aggressive tumor behavior\(^{50}\). They were present in all tumors for which there was sufficient material for analysis (4/5), and, although not unique to tumors deficient in both BAP1 and PBRM1, the association was significant \((q = 0.0007)\); Supplementary Data 1).

DISCUSSION

These results implicate BAP1 as a tumor suppressor in ccRCC and establish the foundation for a molecular genetic classification of RCC. We show that 70% of ccRCCs lose either BAP1 or PBRM1, that tumors tend to segregate into BAP1- or PBRM1-deficient subtypes and that BAP1 loss but not PBRM1 loss is associated with high tumor grade.

BAP1 functions as a two-hit tumor suppressor in ccRCC, and, consistent with this, mutant BAP1 does not act in a dominant-negative fashion. Both copies of BAP1 are also lost in melanoma\(^{28,29,30}\) and mesothelioma\(^{36,51}\). Although the number of RCC samples with BAP1 mutations is small, it is notable that no second-hit point mutations or indels were observed. In contrast, both BAP1 alleles may be inactivated through a point mutation (or indel) in mesothelioma\(^{51}\). We speculate that the different modes of inactivation of the second BAP1 allele reflect tissue-specific tumor suppressor gene cooperativity. Indeed, in ccRCC, 3p loss may simultaneously inactivate several genes suppressing renal tumorigenesis, including, most importantly, VHL, which is rarely mutated in other tumor types. In metastatic uveal melanoma, whole chromosome 3 losses are frequent, and other melanoma metastasis suppressors may exist on 3q. Thus, the deletion architecture of tumors may reflect tissue-specific cooperativity of tumor suppressor genes.

We propose that subsequent to a VHL mutation, which likely represents an early event in tumorigenesis\(^17\), the loss of 3p leaves cells vulnerable to the loss of the remaining PBRM1 or BAP1 allele. The acquisition of a PBRM1 or BAP1 mutation may set the course for ccRCCs with different properties. PBRM1 and BAP1 likely affect different epigenetic programs, and BAP1 loss is associated with high tumor grade and mTORC1 activation. Notably, whereas mutations in SETD2, which is also at 3p, seem to distribute equally between PBRM1-deficient and wild-type tumors, this is not the case for BAP1 mutations. PBRM1 and BAP1 mutations anticorrelate in ccRCC. These data suggest that there is a genetic context to tumor suppressor function and that simultaneous loss of BAP1 and PBRM1 in most tumors is disadvantageous.

The clinical implications of BAP1 loss remain to be explored. As BAP1 loss was associated with high tumor grade and correlated with metastasis development in uveal melanoma\(^28\), BAP1 loss in ccRCC may be associated with poor prognosis. From a therapeutic standpoint, whereas RCC is considered radioresistant, BAP1-deficient tumors may be more sensitive. Evaluating the prognostic and therapeutic implications of BAP1 loss will be greatly facilitated.
by the development in a clinical laboratory of a highly sensitive and specific IHC assay.

Notably, BAP1 is mutated in the germline, where it predisposes to melanoma and mesothelioma.28,29,30,51. Given the role of BAP1 in sporadic ccRCC, germline BAP1 mutations may similarly predispose to RCC. In fact, a germline variant (c.121G>A; p.Gly41Ser) was identified in one individual who had two first-degree relatives and one second-degree relative with RCC; this individual had previously been screened for a germline VHL mutation, but no mutation was found. In addition, a recently reported pedigree had one individual with a germline BAP1 mutation who had RCC.51. Thus, BAP1 mutation in the germline may predispose to RCC, in which case, RCC development may also be initiated by loss of BAP1.

Multiple lines of evidence implicate HCF-1 in BAP1-mediated RCC tumor suppressor function. First, BAP1 binds to and cofractionates with HCF-1. Second, as determined in immunodepletion experiments, the majority of BAP1 is bound to HCF-1. HCF-1 is a very abundant protein, and this may explain why mutant BAP1 does not function as a dominant negative. Third, the growth inhibitory effect of BAP1 is compromised by a mutation that, although not disruptive to protein structure (as determined by retention of deubiquitinating activity), disrupts HCF-1 binding. Finally, the interaction with HCF-1 is unlikely to reflect an abnormal epigenetic state of tumor cell lines in culture, as BAP1 also binds to and cofractionates with HCF-1 in tumorgrafts. Notably, however, the HCF-1 binding motif in BAP1 is not conserved in the Drosophila Calypso protein.

The role of H2Aub1 in ccRCC requires further study. BAP1 binding to HCF-1 was required for the suppression of cell proliferation but was dispensable for H2Aub1 deubiquitination. Thus, these two functions of BAP1—HCF-1 binding and H2Aub1 deubiquitination—can be separated. We did not find a correlation between BAP1 inactivation and global H2Aub1 levels in tumors. Nevertheless, the levels of H2Aub1 were not uniform across tumors, and we cannot rule out the possibility that BAP1 may affect the levels of H2Aub1 at specific sites.

Our studies were greatly aided by the availability of tumorgrafts. Tumorgrafts were instrumental in determining MARs with accuracy and for the identification of putative two-hit tumor suppressor genes. They also made it possible to determine the co-occurrence of mutations in tumor cell lines, and, when mutations occurred in regions of amplification, they shed light on the temporal sequence of mutation acquisition. Finally, tumorgrafts provided a renewable source of tumor material, allowing us to evaluate the relevance of biochemical observations made in cell lines in culture.

While this manuscript was in preparation, a brief communication reported a list of 12 genes mutated in ccRCC, including TSC1, which we previously showed to be mutated in sporadic ccRCC, and BAP1. The mutation frequency reported for BAP1 was 8%, but a VHL mutation frequency of 27% suggests low sensitivity.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. SNP and gene expression microarray data have been deposited at the Gene Expression Omnibus (GEO) (GSE25540 and GSE36895, respectively). Whole-genome and exome sequences for individuals consenting to the deposit of their information are in the database of Genotypes and Phenotypes (dbGaP) (phs000491).

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We recognize the individuals who participated in the study and who donated samples. We thank O. Sepulveda, A. Husain and A. Yadlapalli for technical support, S. Cohenour and D. Sheppard for assistance with contracts and regulatory considerations, Y. Machida (Mayo Clinic) for providing plasmids, B. Grossman (MD Anderson Cancer Center) for providing the UMRC cells, C. Camacho and N. Tomimatsu for irradiating cells, and the staff of the UT Southwestern Tissue Resource. This work was supported by a fellowship of excellence from Generalitat Valenciana (BPOSTDOC06/040 to S.-P.L.) and by the following awards to J.B.: a grant from the Cancer Prevention and Research Institute of Texas (RP101075), a Clinical Scientist Development Award from the Doris Duke Charitable Foundation, an American Cancer Society Research Scholar grant (113579) and a grant from the US National Institutes of Health (RO1 CA129387). The tissue management shared resource is supported in part by the US National Cancer Institute (NCI) (1P09CA142543). J.B. is a Virginia Murchison Linthicum Scholar in Medical Research at UT Southwestern. The content herein is solely the responsibility of the authors and does not represent the official views of any of the granting agencies.

AUTHOR CONTRIBUTIONS

S.-P.L. processed, managed and extracted nucleic acids from tissues, evaluated and validated mutations, and performed bioinformatic analyses on the exome data, as well as copy-number, gene expression and statistical analyses. S.V.-R., M.C.-C., S.D., D.S., C.D.H. and J.B. supervised the whole-genome sequencing process, performed quality control measures and were responsible for the primary SNV analysis in the clinical laboratory. N.L. analyzed exome sequences under the supervision of C.D.H. A.P.-J. and P.S. helped with tissue processing and histology. S.W. helped with functional studies in UMRC6 cells. T.Y. assisted in mutation validation and mouse studies. L.Z. reviewed patient’s records. L.K. and N.G. performed in silico structural analyses for BAP1 and PBRM1. S.S. maintained the tumorgrafts and processed tissues. Y.L., V.M. and A.I.S. provided tissues and cell lines and assisted with the procurement of samples from the index subject. P.B.S. was the index subject’s genetic counselor. W.K. and P.K. evaluated the pathology slides, and P.K. was responsible for the IHC assays. X.-J.X. performed statistical analyses and revised statistics. S.W.W.W. performed the indel analysis. M.T.R. and D.R.B. supervised and managed the genome sequencing and annotation process. J.B. conceived the study, designed experiments, analyzed the data and wrote the manuscript, with input from S.-P.L. and other authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ng.2323. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Siegel, R., Ward, E., Brawley, O. & Jemal, A. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. Cancer J. Clin. 61, 212–236 (2011).
2. Baldeyrou, M.M. et al. Genetics and epigenetics of renal cell cancer. Biochim. Biophys. Acta 1785, 133–155 (2008).
3. Brugarolas, J. Renal-cell carcinoma—molecular pathways and therapies. N. Engl. J. Med. 356, 185–187 (2007).
4. Latif, F. et al. Identification of the von Hippel-Lindau disease tumor suppressor gene. Science 260, 1317–1320 (1993).
5. Gnarra, J.R. et al. Mutations of the VHL tumour suppressor gene in renal carcinoma. Nat. Genet. 7, 85–90 (1994).
6. Nickerson, M.L. et al. Improved identification of von Hippel-Lindau gene alterations in clear cell renal tumors. Clin. Cancer Res. 14, 4726–4734 (2008).
7. 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).
8. 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).
9. Varaia, I. et al. Exome sequencing identifies frequent mutation of the swirlsnf complex gene pbrm1 in renal carcinoma. Nature 469, 539–542 (2011).
10. Dalglish, G.L. et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature 463, 360–363 (2010).
11. van Haften, G. et al. Somatic mutations of the histone h3k27 demethylase gene utx in human cancer. Nat. Genet. 41, 521–523 (2009).
Allelotype of pancreatic adenocarcinoma using xenograft ATM and ATR substrate analysis reveals extensive protein Interplay between pVHL and mTORC1 pathways in clear-cell Single-cell exome sequencing reveals single-nucleotide mutation © 2012 Nature America, Inc. All rights reserved.

Leukemia proto-oncoprotein MLL forms a SET1-like histone (eds. Figlin, R.A., Rathmell, W.K. & Rini, B.I.) 27–47 (Springer

Prognostic relevance of the mTOR pathway in renal cell

Histone H2A deubiquitinase activity of the Polycomb

Machida, Y.J., Machida, Y., Vashisht, A.A., Wohlschlegel, J.A. & Dutta, A. The deubiquitinating enzyme BAP1 regulates cell growth via interaction with HCF-1. J. Biol. Chem. 284, 34179–34188 (2009).

Scheuermann, J.C. et al. Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature 465, 243–247 (2010).

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).

Knezevic, J., Piluso, D., Bilan, P. & Capone, J.P. Host cell factor–1 and E2F4 interact via multiple determinants in each protein. Mol. Cell. Biochem. 288, 79–90 (2006).

Tyagi, S., Chabes, A.L., Wysocka, J. & Her, W. E2F activation of S phase promoters via association with HCF-1 and the MLL family of histone H3K4 methyltransferases. Mol. Cell 27, 107–119 (2007).

Wysocka, J., Myers, M.P., Laherty, C.D., Eisenman, R.N. & Herr, W. Human Sn3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methylationase are tethered together selectively by the cell-proliferation factor HCF-1. Genes Dev. 17, 896–911 (2003).

Yokoyama, A. et al. Leskemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. Mol. Cell. Biol. 24, 5639–5649 (2004).

Narayanan, A., Ruyechan, W.T. & Kristie, T.M. The coactivator host cell factor–1 mediates Set1 and MLL1 H3K4 trimethylation at herpesvirus immediate early promoters for initiation of infection. Proc. Natl. Acad. Sci. USA 104, 10835–10840 (2007).

Liang, Y., Vogel, J.L., Narayanan, A., Peng, H. & Kristie, T.M. Inhibition of the histone demethylase LSD1 blocks α-herpesvirus lytic replication and reactivation from latency. Nat. Med. 15, 1312–1317 (2009).

Smith, E.R. et al. A human protein complex homologous to the Drosophila MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. Mol. Cell. Biol. 25, 9175–9188 (2005).

Stokes, M.P. et al. Profiling of UV-induced ATM/ATR signaling pathways. Proc. Natl. Acad. Sci. USA 104, 19855–19860 (2007).

Matsuoka, S. et al. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science 316, 1160–1166 (2007).

Kaelin, W. Jr. Molecular biology of clear cell renal carcinoma. in Renal Cell Carcinoma (eds. Figlin, R.A., Rathmell, W.K. & Rini, B.I.) 27–47 (Springer New York, 2012).

Xu, X. et al. Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell 148, 886–895 (2012).

Chapman-Fredricks, J.R. et al. Adult renal cell carcinoma with rhabdoid morphology represents a neoplastic dedifferentiation analogous to sarcomatoid carcinoma. Ann. Diagn. Pathol. 15, 333–337 (2011).

Gökden, N. et al. Renal cell carcinoma with rhabdoid features. Am. J. Surg. Pathol. 24, 1329–1338 (2000).

Abdel-Rahman, M.H. et al. Germline BAP1 mutation predisposes to uveal melanoma, lung adenocarcinoma, meningioma, and other cancers. J. Med. Genet. 48, 856–859 (2011).

Álvarez, A. et al. Germline BAP1 mutations predispose to malignant mesothelioma. Nat. Genet. 43, 1022–1025 (2011).

Wysocka, J., Reilly, P.T. & Herr, W. Loss of HCF-1 expression is associated with gene expression alterations and cell cycle abnormalities. Mol. Cell. Biol. 21, 3820–3829 (2001).

Grujic, G. et al. Frequent mutations of genes encoding ubiquitin-mediated proteolysis pathway components in clear cell renal cell carcinoma. Nat. Genet. 44, 17–19 (2012).
Tumor samples were labeled with a number or a number preceded by a T if those samples were also used for tumorgraft generation. Tumorgrafts were labeled with the same number as tumors with the prefix TG and followed by a cohort c number (when applicable) referring to the tumor passage (for example, c0 for a primary tumorgraft).

Tumor stage was determined on the basis of tumor, node, metastasis (TNM) classification from the American Joint Committee on Cancer. Samples were annotated according to the edition corresponding to the date of surgery. Per the seventh edition, all tumors with lymph node metastases were referred to as pNI.

### Tissue selection

cRCs and adjacent normal kidney samples were frozen fresh in liquid nitrogen and stored at −80 °C. Tumor content and quality was inferred by a pathologist from perpendicular sections immediately flanking 1–3 mm thick fragments that were oriented using pathology dyes (Supplementary Fig. 2c). For whole-genome sequencing a sample was selected with >90% tumor content in both sections. For the Discovery Set, 76 cRCs samples with ≥80% tumor cellularity were selected among 431 fresh-frozen tumor samples from 133 patients. Seven tumor samples and a metastasis with ≥80% tumor cellularity were selected for exome sequencing among 16 patients with tumorgrafts growing in mice. For the Validation Set, 92 cRCs samples with ≥70% tumor cellularity were selected among 533 fresh-frozen tumor samples from 165 RCC patients. Genomic DNA and RNA were simultaneously extracted from each tissue (detailed in the Supplementary Note). Reference DNA, extracted from either adjacent normal kidney or peripheral blood mononuclear cells (PBMCs), was available for 71/76 tumors in the Discovery Set and 82/92 tumors of the Validation Set.

### Whole-genome sequencing of paired-end libraries from tumor and matched normal genomes

Tumor and PBMC samples were processed in a CLIA-certified and College of American Pathologists (CAP)-accredited laboratory. The preparation of short-insert (212–263 bp) Illumina paired-end sequencing libraries, flow cells and clusters has been described previously. Paired-end sequence reads of 100 bases were generated using the Genome Analyzer IIx. Image analysis, base calling and Phred quality scoring were performed using the Illumina analysis pipeline (RTA, v1.5). Sequence reads were filtered out from clusters whose proximity to others resulted in mixed sequence data.

### Whole-genome somatic substitution and indel detection

Single-nucleotide variants (SNVs) from the reference sequence (human NCB136.1) were determined separately for the tumor and normal genomes using CASAVA, v1.6. Prediction of a homozygous SNV required a minimum allele score of 10 (equivalent to at least three high-quality (Q33) base calls). Additionally, for heterozygous calls, the second allele was required to have a score of at least 6 (equivalent to two Q40 base calls), and the ratio of the two allele scores had to be ≥3, so that allele ratios did not deviate from the expected 1:1 ratio for heterozygous calls. Indels relative to human NCB136.1 were predicted using GROUPER. SNVs and indels in tumors were only considered as candidate somatic events if the read depth at the equivalent site in the normal genome build was at least 10. SNVs and indels observed in both genomes were subtracted from the tumor calls. Previously known SNPs (those in dbSNP130) were also removed. Indels in the tumor overlapping a contig of assembled shadow reads in the normal genome were removed. The impact of somatic changes on protein-coding and non-coding genes was annotated using Ensembl version 54.

### Exome capture and sequencing

Exome capture was performed by Illumina FastTrack using Illumina Truseq exome-target enrichment. Details are provided in the Supplementary Note.

### Exome mutation detection and validation

The sequences of tumor, metastasis and normal samples were compared to NCBI reference sequence, and SNVs and indels were determined independently using CASAVA, v1.8.0.a4, without any filtering. Mutations predicted in tumors that were also present in the corresponding normal samples were eliminated (Supplementary Table 5). Synchronous mutations were removed, and the resulting mutations were inspected visually using the Integrative Genomics Viewer (IGV; see URLs) to confirm their presence in tumor but not normal sequence reads (Supplementary Table 5). All genes with recurrent mutations were validated by Sanger sequencing (Supplementary Table 7). For the remaining genes, a mutation calling accuracy of >95% for 82 mutations would signify >90% accuracy for the whole cohort, according to a cumulative hypergeometric distribution. Sanger sequencing of 82 randomly selected mutations (proportional to the number of mutations in 7 tumors and 1 metastasis) showed 4 false positives (<5%), all of which had been scored on the basis of 2 mutant reads (Supplementary Table 6). Among the remaining genes, those with just two mutant reads (three) were inferred to represent false positives (Supplementary Data 2).

### Mutation analyses and mutant allele ratios

SNVs and indels in chromotags were scored with Mutation Surveyor, v3.30 and v3.98 (Softgenetics), using an overlapping factor of 0.2 and a dropping factor of 0.1. Reference sequences were obtained from NCBI. Only bidirectionally observed somatic mutations are reported. Mutations within 7 nt upstream or downstream of an exon boundary were considered to be splice-site mutations. A somatically acquired PBRM1 variant outside this range (15 nt upstream of exon 8) was not included in the analyses (sample 78).

MARs refer to the fraction of mutant allele for a particular mutation over the sum of mutant and wild-type alleles (MAR of 1, only mutant allele detected; MAR of 0.5, mutant and wild-type alleles detected at similar frequencies). MARs were calculated by measuring the nucleotide intensities of chromotags using ImageJ (see URLs) (whole-genome data) or with the Mutation Quantifier function of Mutation Surveyor, v3.98 (exome data). For indels, MARs were calculated by taking the average of the measurements of at least five nucleotides. MARs were scored as <0.10 if they accounted for <10% of all alleles but were clearly present in tumorgrafts. For Illumina tracings, MARs were calculated on the basis of the number of mutant over total reads.

### Copy-number analyses

Genomic DNA was hybridized to Affymetrix SNP Arrays 6.0 at the Genome Science Resource (Vanderbilt University) using standard procedures. Several tumor and tumorgraft SNP arrays were previously evaluated for other purposes and have been reported elsewhere. Copy-number analyses were performed using the Partek Genomics Suite. Copy-number analyses were performed on the basis of the number of mutations in 7 tumors and 1 metastasis (sample 78).

### Establishment and maintenance of tumorgrafts

Tumorgraft studies were approved by the UT Southwestern Institutional Animal Care and Use Committee (IACUC). Fresh tumor fragments (~2 mm in diameter) were implanted in the kidney of NOD/SCID mice as described elsewhere.

### Gene expression analyses

RNA samples were labeled with biotin and hybridized to Affymetrix Human Genome U133 Plus 2.0 arrays by the UT Southwestern Microarray Core. Gene expression arrays on 13 of 29 tumors and tumorgrafts were previously evaluated to identify a tumor-specific signature and have been previously reported. CEL intensity files were analyzed as described elsewhere. Probe sets with nonspecific hybridization
were removed (8,696, 16%). Further, 2,443 probe sets representing signal attributed to a stromal and/or immune signature were removed. Tumors and tumorgrafts with mutations in either BAP1 or PBRM1 (but not both) were compared to tumors and tumorgrafts that were wild type for both BAP1 and PBRM1 using t tests and a Benjamini-Hochberg false discovery rate (FDR) correction. Probe sets with FDR q < 0.05 were analyzed with Ingenuity Pathways Analysis (IPA).

**Statistics.** To determine whether a correlation (inverse correlation) existed between regional DNA copy numbers and MAR (or MAR), a two-tailed Spearman correlation test was used (data not normally distributed according to a Shapiro-Wilk test). Correlations were compared as previously described. The P values for the identification of 2 or 3 additional gene mutations among the 76 individuals in the discovery set were calculated using a binomial distribution, assuming, on the basis of data from the index subject, that the probability of identifying a nonsynonymous mutation in a gene was 0.0022 (47 mutations among the 21,099 protein-coding genes annotated in the GRCh37.p6 assembly). All Fuhrman grade 3 and 4 samples were reviewed for the presence of rhabdoid features. For the Sanger Institute data set, the highest tumor grade was used for each tumor. Throughout, a Fisher’s exact test was used to determine whether there were nonrandom associations between two binary variables. A Benjamini-Hochberg FDR correction of P values (q values) was calculated to account for multiple comparisons. SPSS Statistics 17.0 and SAS 9.0 were used to analyze data.

Primer sequences and antibody information are provided in Supplementary Tables 9 and 10, respectively. Further details and descriptions of other experimental methods and materials are available in the Supplementary Note.

54. Bentley, D.R. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59 (2008).
55. Olshen, A.B., Venkatraman, E.S., Lucito, R. & Wigler, M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* **5**, 557–572 (2004).
56. Peña-Llopis, S. *et al.* Regulation of TFEB and V-ATPases by mTORC1. *EMBO J.* **30**, 3242–3258 (2011).
57. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J. R. Stat. Soc., B* **57**, 289–300 (1995).
58. Steiger, J.H. Tests for comparing elements of a correlation matrix. *Psychol. Bull.* **87**, 245–251 (1980).
Erratum: BAP1 loss defines a new class of renal cell carcinoma

Samuel Peña-Llopis, Silvia Vega-Rubín-de-Celis, Arnold Liao, Nan Leng, Andrea Pavia-Jiménez, Shanshan Wang, Toshinari Yamasaki, Leah Zhrebker, Sharanya Sivanand, Patrick Spence, Lisa Kinch, Tina Hambuch, Suneer Jain, Yair Lotan, Vitaly Margulis, Arthur I Sagalowsky, Pia Banerji Summerour, Wareef Kabbani, SW Wendy Wong, Nick Grishin, Marc Laurent, Xian-Jin Xie, Christian D Haudenschild, Mark T Ross, David R Bentley, Payal Kapur & James Brugarolas

Nat. Genet. 44, 751–759 (2012); published online 10 June 2012; corrected after print 21 June 2012

In the version of this article initially published, the P value given in the abstract for the anticorrelation between BAP1 and PBRM1 mutations was given incorrectly as $9 \times 10^{-6}$ instead of $3 \times 10^{-5}$. The definition of mutant allele ratios (MARs) in the text on p. 1 of the PDF has been corrected. The titles and legends of Figure 1 and Table 1 incorrectly stated that data were presented from multiple tumors and tumorgrafts; these have been corrected to reflect that data came from one index subject. The title of Figure 2 originally referred to mutated residues; this has now been corrected to state that alterations in BAP1 are shown. On p. 7 of the PDF the text has been corrected to state that “a few tumors had loss of both BAP1 and PBRM,” whereas the text originally incorrectly cited somatic mutations in both genes. The errors have been corrected in the HTML and PDF versions of the article.