nature genetics

The nuclear deubiquitinase BAP1 is commonly inactivated by somatic mutations and 3p21.1 losses in malignant pleural mesothelioma

Matthew Bott[1,2], Marie Brevet[1], Barry S Taylor[3], Shigeki Shimizu[1], Tatsuo Ito[1], Lu Wang[1], Jenette Creaney[4], Richard A Lake[4], Maureen F Zakowski[1], Boris Reva[3], Chris Sander[3], Robert Delsite[5], Simon Powell[5], Qin Zhou[6], Ronglai Shen[6], Adam Olshen[6], Valerie Rusch[2] & Marc Ladanyi[1,7]

Malignant pleural mesotheliomas (MPMs) often show *CDKN2A* and *NF2* inactivation, but other highly recurrent mutations have not been described. To identify additional driver genes, we used an integrated genomic analysis of 53 MPM tumor samples to guide a focused sequencing effort that uncovered somatic inactivating mutations in *BAP1* in 23% of MPMs. The BAP1 nuclear deubiquitinase is known to target histones (together with ASXL1 as a Polycomb repressor subunit) and the HCF1 transcriptional co-factor, and we show that BAP1 knockdown in MPM cell lines affects E2F and Polycomb target genes. These findings implicate transcriptional deregulation in the pathogenesis of MPM.

MPM, an aggressive neoplasm of the serosal lining of the chest etiologically linked to asbestos, is diagnosed in approximately 2,000 to 3,000 individuals annually in the United States, most of whom die within 2 years of diagnosis. Studies of MPM have identified only two frequent genetic alterations, the most common being homozygous 9p21 deletions centered on *CDKN2A*, seen in up to 72% of tumors and over 80% of MPM cell lines[1]. Other recent studies have implicated *NF2* loss through monosomy 22 or 22q deletions as another common event[2]. Aside from these two tumor suppressors, other driver genes in MPM remain obscure. Here we used an integrated genomics approach to define recurrent genomic copy number alterations (CNAs) in MPM and to select candidates for sequencing.

We profiled CNAs in 53 primary MPM tumor samples on Agilent 244K CGH arrays. Unsupervised clustering of the copy number data identified two major clusters, with most MPMs (41/53; 77%) showing a predominance of losses, whereas in a minority of MPMs (12/53; 23%), genomic gains predominated (**Supplementary Fig. 1**). Based on the analysis of minimal common regions of CNA, the three most common deletions were at 9p21, 22q and 3p21. Although the known or presumed drivers for the former two are *CDKN2A* and *NF2*, respectively, the gene or genes whose loss drives selection for 3p21 deletions have not been identified despite prior efforts[3]. To refine the definition of candidate regions, we also analyzed the data using the RAE algorithm, a method that assigns statistical significance to CNAs in a given set of cancer samples (**Fig. 1**). Genes within significant CNAs as defined by the RAE analysis were reviewed for biological plausibility as oncogenic drivers. We also integrated Affymetrix expression array data on these cases[4] as an additional tool for gene selection, particularly in broader areas of gain or loss, where identifying a single or a few driver genes by analysis of CNAs alone has traditionally proven difficult. Using these approaches, we compiled a list of 25 potential driver genes for sequencing in our 53 MPM tumors. Among the mutations thus identified (**Supplementary Table 1**), the gene with the highest rate of non-synonymous mutations was *BAP1* (in 12/53 MPMs; 23%) (**Supplementary Table 2**), located at the epicenter of 3p21.1 losses seen in 30% of MPMs (**Fig. 2**). We detected only one silent somatic mutation in *BAP1*, representing a very high ratio of non-synonymous to synonymous somatic mutations. *BAP1* (encoding BRCA1-associated protein 1) encodes a ubiquitin COOH-terminal hydrolase originally identified through its interaction with BRCA1 (ref. 5). Sequencing also confirmed frequent inactivating mutations in *NF2* (11/53; 21%) and identified previously undescribed missense mutations in *LATS2* (2/53; 3.8%) and *LATS1* (2/53; 3.8%). Notably, other common tumor suppressors (*RB1*, *PTEN*, *RASSF1* and *TP53*) showed few or no mutations, consistent with previous studies in MPM. We also identified a *SMARCB1* nonsense mutation in a tumor with normal *NF2* and *BAP1*.

In all, 42% of MPM tumors harbored either *BAP1* loss, mutation or both. *BAP1* losses were confirmed by fluorescent *in situ* hybridization (FISH) (**Fig. 2**) and were also found in 6 of 25 MPM cell lines (24%; **Supplementary Fig. 2**). We confirmed the frequency of

[1]Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. [2]Department of Surgery, Memorial Sloan-Kettering Cancer Center (MSKCC), New York, New York, USA. [3]Computational Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. [4]National Centre for Asbestos Disease Research, School of Medicine and Pharmacology, University of Western Australia, Sir Charles Gairdner Hospital, Nedlands, Australia. [5]Department of Radiation Oncology, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. [6]Epidemiology-Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. [7]Human Oncology & Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. Correspondence should be addressed to M.L. (ladanyim@mskcc.org).

Received 3 November 2010; accepted 16 May 2011; published online 5 June 2011; doi:10.1038/ng.855

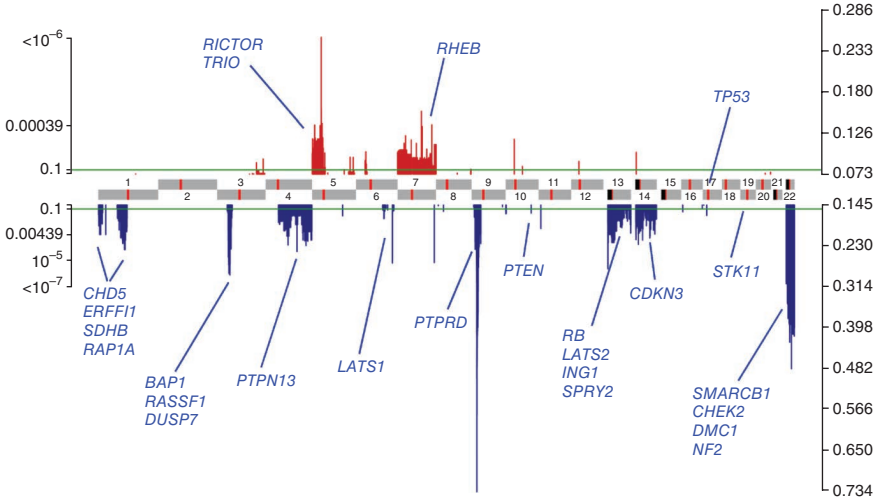

Figure 1 RAE analysis of genomic gains and losses in 53 MPM tumors. The effective frequency of gain (red) or loss (blue) and corresponding statistical significance (false discovery rate (FDR), *q* values) of aberrations are indicated (*y* axis, right and left, respectively), as is their chromosomal position (center, gray; centromeres in red). Green lines adjacent to the axis indicate the cutoff for statistical significance (*q* value = 0.1, FDR = 10%). Genes selected for sequencing are identified near their relevant peaks (*CDKN2A*, the target of the 9p21 deletion, was not sequenced in the current study).

BAP1 mutations in an independent collection of 68 MPM tumors (12 mutations, 18%) (**Supplementary Table 3**) and in MPM cell lines (8/25 mutated, 32%) (**Supplementary Table 4**). Overall, the 32 *BAP1* mutations identified included 6 nonsense mutations, 5 missense mutations, 13 frameshifting indels and 8 mutations at or near splice sites (**Fig. 3**). Of the 24 mutations in MPM tumors, we confirmed the mutations to be somatic in 10 of 11 cases with matched normal tissue available. The majority of the truncating mutations are predicted to result in loss of the nuclear localization signal and/or the C-terminal protein binding domain (**Fig. 3**). Immunohistochemistry (IHC) for BAP1 on paraffin-embedded sections of 47 of 53 MPM tumors from the original discovery set showed that *BAP1* loss and/or mutation was associated with an absence of BAP1 protein expression by IHC (Fisher's exact test, $P = 0.002$) (**Fig. 2**). Likewise, tumors with *BAP1* loss and tumors with *BAP1* loss and/or mutation expressed significantly less *BAP1* mRNA (*t*-test, $P < 0.0001$ for both). Finally, 25% of tumors showed no BAP1 staining by IHC despite apparently normal *BAP1* status by array CGH and sequencing and no significant difference in *BAP1* transcript levels (in the Affymetrix array data), raising the possibility of post-translational deregulation of BAP1 in an additional subset of cases.

We next assessed the functional importance of *BAP1* missense and splice site mutations. Of the 32 *BAP1* mutations, 7 (in 8 samples) occurred within 4 bp of a splice site. For six samples with available material containing five different splice site mutations, we confirmed that four of the five mutations resulted either in partially retained introns or missing exons (**Supplementary Table 5**), leading to abnormal stop codons just downstream. To assess whether *BAP1* missense mutations (resulting in p.Ser63Cys, p.Ala95Asp and p.Cys91Trp) were functionally abnormal, we purified FLAG-tagged mutant proteins by immunoprecipitation and subjected them to an *in vitro* enzymatic assay (Online Methods) in 293T cells and in two BAP1-null MPM cell lines (H28, Meso37). Two of the three mutants showed decreased levels of ubiquitin cleavage activity compared with wild-type BAP1 (**Supplementary Fig. 3** and **Supplementary Table 6**). Notably, the remaining mutant (p.Ser63Cys) showed an increase in ubiquitin cleavage compared to wild-type BAP1. Although the latter finding may represent an *in vitro* assay-dependent abnormality, it is conceivable that the ubiquitination status of targets must remain at a crucial equilibrium and that either over- or under-ubiquitination leads to instability, a concept already proposed in the regulation of other ubiquitin-modulated systems[6,7].

Next, we examined the relationship of *BAP1* mutation to other genetic alterations (**Fig. 4**) and clinicopathologic features. There were no significant associations between *CDKN2A* loss and loss or mutation of *BAP1* or *NF2*, supporting these as three independent genetic events in MPM. Regarding clinical and pathologic variables, there was a significant correlation of *BAP1* mutation with age (mean age of 66.7 years in mutant *BAP1* compared to 58.6 years in

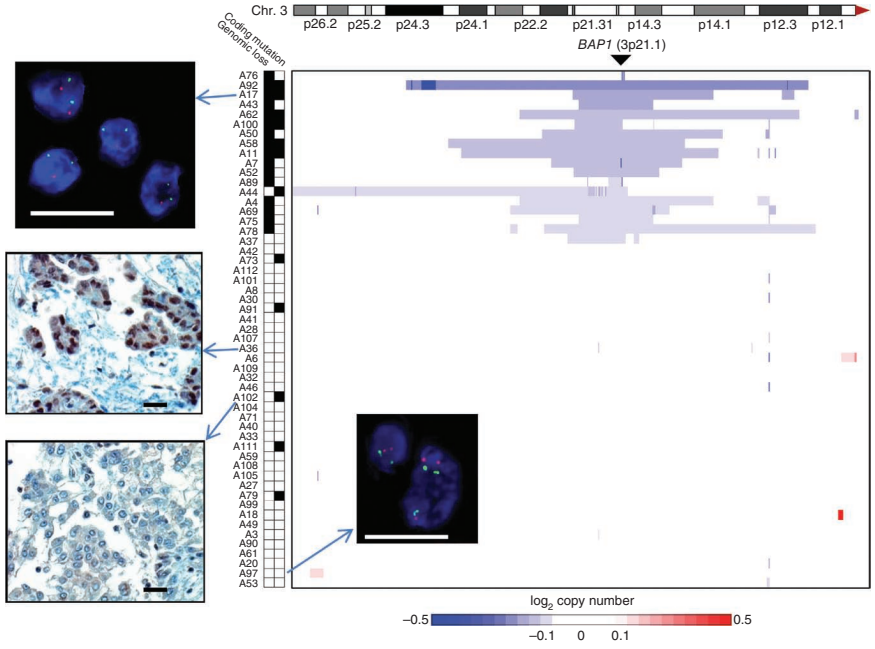

Figure 2 Heat map of chromosome 3p in MPM tumors. Genomic loss is indicated in blue. A subset of cases harbor more focal alterations centered on *BAP1*. Overall, 16 out of the 53 tumors (30%) have at least single copy genomic loss of the *BAP1* locus. In the insets (scale bars, 50 μm), BAP1 immunohistochemistry shows loss of nuclear staining for BAP1 protein in tumors with a *BAP1* mutation (for example, tumor A102 with a 1-bp insertion in exon 15) compared to those with wild-type *BAP1* (for example, tumor A36). Also shown is the *BAP1* FISH analysis (red signals, *BAP1* probe; green signals, chromosome 3 centromere) confirming single copy loss in tumor A17 and normal diploid *BAP1* status in tumor A97, validating array CGH results.

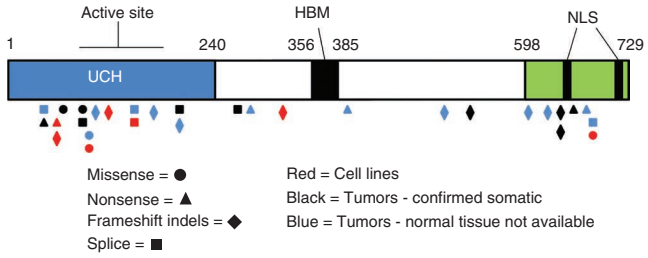

Figure 3 Distribution of *BAP1* mutations relative to functional domains. Shown are the N-terminal ubiquitin hydrolase domain (blue), the HCF1-binding domain (HBM) and the C-terminal protein interaction domain (green) containing two nuclear localization signals (black boxes).

wild-type *BAP1*; *t*-test, $P = 0.03$), but there was no significant correlation with other variables such as sex, overall survival, histologic subtype or asbestos exposure.

Because BAP1 was originally identified as a BRCA1-associated protein[5], we examined the possibility that defects in BAP1 might impair the BRCA1-dependent DNA damage response. We exposed cell lines with wild-type *BAP1* (Meso9, Meso10) and mutated or deleted *BAP1* (Meso37, H28, H226 and H2452) to 10 Gy of ionizing radiation. There were no consistent differences in RAD51 or BRCA1 complex formation between BAP1 wild-type and BAP1-deficient cell lines (**Supplementary Fig. 4**), suggesting no essential role for BAP1 in the formation of these DNA repair complexes. Additionally, because functional defects in the BRCA1-mediated DNA repair pathway confer synthetic lethality to PARP inhibition[8], we performed *in vitro* studies using the PARP inhibitor MK4827 (gift of Merck) but observed no differential sensitivity to PARP inhibition in BAP1 mutant MPM lines compared to BAP1 wild-type MPM cells (data not shown).

We also examined the effects of BAP1 knockdown on cell proliferation in MPM lines expressing wild-type BAP1. We performed transient small interfering RNA (siRNA) knockdown of BAP1 in three MPM lines containing wild-type BAP1 (HMeso, MSTO211H and H2373). All three lines showed decreased proliferation upon BAP1 knockdown (**Supplementary Fig. 5**). Conversely, the reintroduction of wild-type BAP1 in BAP1-null MPM lines (Meso37, H28) resulted in a modest increase in proliferation (**Supplementary Fig. 6**). Although paradoxical for a gene whose genetic alterations support a tumor suppressor function, these data are consistent with previous studies of non-MPM cell lines in which BAP1 knockdown caused proliferation defects with an accumulation of cells in S phase[9–11]. Indeed, flow cytometry analysis of BAP1 knockdown in the MSTO211H line showed an accumulation of cells in S phase, and conversely, reintroduction of BAP1 into H28 MPM cells (BAP1 null) resulted in a reduction of cells in S phase (**Supplementary Fig. 7**).

As BAP1 has been reported to bind and deubiquitinate HCF1 in other cell lines[11,12], we confirmed that BAP1 and HCF1 also co-immunoprecipitate in MPM lines (**Supplementary Fig. 8**). Studies have shown that the BAP1-HCF1 interaction may be important for HCF1-mediated growth effects. HCF1 acts in part through modulation of transcription at E2F-responsive promoters[13]. We therefore derived expression signatures of BAP1 knockdown in MPM lines using Affymetrix U133A 2.0 arrays and first looked specifically at downstream targets of E2F, including Cyclin A2, E2F1, p107 (RBL1) and CDC25C. We found that all of these effectors were downregulated in MPM cells following BAP1 knockdown (**Supplementary Table 7**), consistent with the possibility that BAP1 loss results in post-translational inactivation of HCF1 and thus downregulation of downstream E2F-responsive genes important for cell cycle progression. These findings were also validated by quantitative RT-PCR after BAP1 knockdown using a different siRNA (**Supplementary Table 7**). In 31 MPM tumors with wild-type BAP1, we sequenced the *HCF1* kelch domain, which is required for BAP1 binding[11,12], hypothesizing that *HCF1* mutations that disrupt the BAP1-HCF1 interaction might substitute for BAP1 mutations in MPM tumorigenesis, but we found no mutations.

Recently, it has also been shown that *BAP1* is the human counterpart of the *Drosophila* PcG gene *calypso* and that the Calypso-ASX complex corresponds functionally to the human BAP1-ASXL1 complex, representing a Polycomb repressive deubiquitinase (PR-DUB) complex[14]. We confirmed that BAP1 and ASXL1 co-immunoprecipitate; although the domains responsible for this interaction have not been mapped, several BAP1 point mutants tested did not seem to affect it (**Supplementary Fig. 9**). Notably, *ASXL1* is known to be frequently mutated in hematologic malignancies[15], usually within exon 12. We sequenced exon 12 of *ASXL1* in 31 MPM tumors with known wild-type *BAP1* and found no somatic mutations. A previous study suggested that normal Polycomb target gene silencing may require a balance between histone ubiquitination by PRC1 and dRAF and histone deubiquitination by PR-DUB[14]. Based on this link between BAP1-ASXL1 and Polycomb gene expression, we examined the relationship of the BAP1 knockdown expression signature in MPM to three published Polycomb target gene sets (Online Methods). We performed BAP1 knockdown in three MPM cell lines with wild-type BAP1 (HMeso, H2373, MSTO211H), pooled the gene expression results and compiled a list of significantly altered genes (up or down greater than twofold and $P < 0.05$). Comparison of this gene list to three published lists of PRC target genes[16–18] showed highly significant enrichment of all three in the BAP1 signature ($P < 10^{-19}$ for all three) (**Fig. 5**), supporting a role for BAP1 in the regulation of Polycomb target gene expression in MPM cells. Ingenuity pathway analysis of the 77 genes that appear in the BAP1 profiles and at least two Polycomb signatures (**Supplementary Table 8**) showed highly significant overrepresentation of genes involved in DNA replication, recombination and repair, including, among others, several cyclins, cyclin-dependent kinase inhibitors, *CDC* and *MCM* genes and *E2F1*.

Prior to this study, relatively few human cancer samples had been found to harbor truncating or missense mutations of *BAP1*, notably the lung cancer cell lines NCI-H1466, NCI-H226 and NCI-H1963 (ref. 5), as well as 2 of 188 cases in the TSP lung adenocarcinoma sequencing project[19]. Another study published after our initial submission has reported frequent mutational inactivation of *BAP1* in uveal melanomas[20]. Our data expand the scope of *BAP1* as a common target of mutational inactivation and a principal driver of 3p21 losses in MPM.

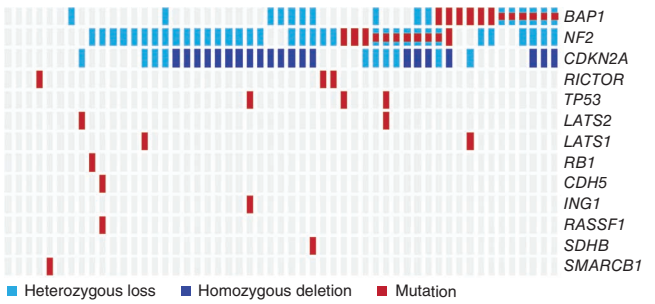

Figure 4 Integrated gene map showing *BAP1* losses and mutations in relation to other identified genomic events in all 53 samples. Three samples had no alterations in the genes listed.

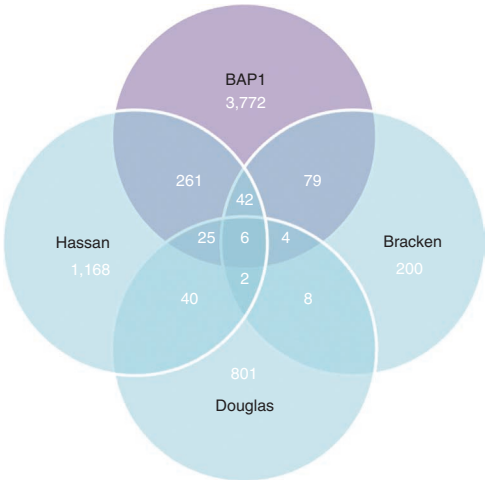

Figure 5 Relationship between the BAP1 knockdown expression signature in MPM and three published Polycomb target gene sets. The Bracken signature was derived by siRNA depletion of four PRC complex proteins (EZH2, EED, SUZ12 and BMI1) in embryonic fibroblasts. The Hassan signature reflects PRC targets identified by previous ChIP-on-chip experiments using SUZ12 in human embryonic stem cells. We generated the Douglas list using siRNA targeting BMI1 in A4573 Ewing Sarcoma cells. The significance of enrichment (overrepresentation) was assessed using the hypergeometric distribution. We noted significant overlap between genes in the BAP1 list and genes in all three PRC sets, with each pairwise overlap with the BAP1 signature showing $P < 10^{-19}$. The 77 genes that appear in the BAP1 profile and at least two other Polycomb signatures are listed in **Supplementary Table 8**.

Several groups have shown a role for BAP1 in cell cycle progression and its function as a tumor suppressor[9,10,21,22]. Re-expression of wild-type BAP1 in cells expressing mutant BAP1 protein causes growth suppression *in vitro* and *in vivo*[5,22]. This BAP1-mediated growth suppression requires an intact ubiquitin hydrolase domain and nuclear localization but does not require wild-type BRCA1 (ref. 22). It may seem paradoxical that BAP1 knockdown also appears to slow growth and downregulate E2F-responsive growth-related genes. It is possible that, rather than promoting tumorigenesis by accelerating growth kinetics, BAP1 contributes to a delayed but more permissive G1/S checkpoint. Thus, cells may grow more slowly but uncontrollably. This scenario may be consistent with our clinical correlation of older patient age for BAP1 mutant MPMs. Notably, *ASXL1* mutations are also associated with proliferation defects and older age at diagnosis in acute myeloid leukemias[15,23].

To better understand the cellular roles of BAP1, several groups[11,12,24] have examined the proteins bound and possibly de-ubiquitinated by BAP1. Proteins identified in more than one study include HCF1, ASXL1, ASXL2, FOXK1, FOXK2, ANKRD17, HAT1, UBE2O and AOF1. Notably, none of these recent studies identified BRCA1 as a significant binding partner. As the BAP1 nuclear deubiquitinase appears to be involved in maintaining an appropriate level of regulatory ubiquitination of target histones (together with ASXL1 as the Polycomb repressor PR-DUB subunit), the HCF1 transcriptional co-factor and possibly other transcriptional proteins, our finding of common BAP1 inactivation in MPM implicates transcriptional deregulation in the pathogenesis of this highly lethal cancer. In this light, the clinical and cell line responses to histone deacetylase inhibitors in a subset of MPMs may warrant further investigation in relation to BAP1 status[25,26]. Finally, our findings support the growing recognition of ubiquitin ligases and ubiquitin hydrolases as cancer genes and may also be relevant to emerging insights into oncogenic disruption of the histone code.

URLs. The Agilent 244K array CGH array data, the Affymetrix 250K SNP array data and the Affymetrix U133A and U133 plus 2.0 array data are available at http://cbio.mskcc.org/Public/Ladanyi_lab_mesothelioma_datasets/. Database of Genomic Variants, http://projects.tcag.ca/variation/; Integrative Genomics Viewer, http://www.broadinstitute.org/igv/; Matlab software, http://www.broadinstitute.org/mpr/publications/projects/NMF/nmf.m; Mutation Assessor algorithm, http://mutationassessor.org/.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Accession codes. For *BAP1*, the mutation nomenclature is based on RefSeqs NM_004656.2 and NP_004647.1. For mutations in other genes, the RefSeq used is listed in **Supplementary Table 1**. The copy number and expression microarray data on the tumor samples and the copy number data on the cell lines have been deposited in the GEO database under accession number GSE29211. They are also available at http://cbio.mskcc.org/Public/Ladanyi_lab_mesothelioma_datasets/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank A. Holloway for help in combining MPM cell line microarray datasets, R. Levine and O. Abdel-Wahab for helpful discussions and Merck for providing PARP inhibitor MK4827. We thank A. Heguy and the personnel of the Beene Translational Oncology Core Facility (MSKCC), A. Lash and Y. Liang of the Bioinfomatics Core Facility (MSKCC), A. Viale and the personnel of the Genomics Core Facility (MSKCC) and M. Asher in the Pathology Core Facility (MSKCC). The project was supported by a generous donation from an anonymous private donor. M. Brevet was supported in part by La Fondation de France. B.S.T. is the David H. Koch Fellow in Cancer Genomics (MSKCC). R.A.L. is supported by the National Health and Medical Research Council (NHMRC) of Australia.

AUTHOR CONTRIBUTIONS

M.L. designed and oversaw the study. V.R. and M.F.Z. oversaw the tumor sample procurement and histopathologic review, respectively. S.S. performed tumor sample selection and analyte processing for tumor samples and cell lines. J.C. and R.A.L. contributed microarray data and analytes from additional cell lines. M. Bott and M.L. reviewed microarray data and selected genes for sequencing. M. Bott obtained and analyzed additional sequencing and genotyping data. M. Bott and T.I. performed functional validation experiments. M. Brevet analyzed immunohistochemistry data. L.W. performed and analyzed FISH studies. R.D. performed functional assays for DNA repair foci, and R.D. and S.P. interpreted the results. B.S.T., B.R., C.S., Q.Z., R.S. and A.O. performed statistical and bioinformatics analyses. M. Bott and M.L. drafted the manuscript. All authors contributed to critical review of the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Illei, P.B., Rusch, V.W., Zakowski, M.F. & Ladanyi, M. Homozygous deletion of *CDKN2A* and codeletion of the methylthioadenosine phosphorylase gene in the majority of pleural mesotheliomas. *Clin. Cancer Res.* **9**, 2108–2113 (2003).
2. Thurneysen, C. *et al.* Functional inactivation of NF2/merlin in human mesothelioma. *Lung Cancer* **64**, 140–147 (2009).
3. Lu, Y.Y., Jhanwar, S.C., Cheng, J.Q. & Testa, J.R. Deletion mapping of the short arm of chromosome 3 in human malignant mesothelioma. *Genes Chromosom. Cancer* **9**, 76–80 (1994).
4. López-Ríos, F. *et al.* Global gene expression profiling of pleural mesotheliomas: overexpression of aurora kinases and p16/*CDKN2A* deletion as prognostic factors and critical evaluation of microarray-based prognostic prediction. *Cancer Res.* **66**, 2970–2979 (2006).

5. Jensen, D.E. *et al.* BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression. *Oncogene* **16**, 1097–1112 (1998).
6. Oestergaard, V.H. *et al.* Deubiquitination of FANCD2 is required for DNA crosslink repair. *Mol. Cell* **28**, 798–809 (2007).
7. Kim, J.M. *et al.* Inactivation of murine Usp1 results in genomic instability and a Fanconi anemia phenotype. *Dev. Cell* **16**, 314–320 (2009).
8. Farmer, H. *et al.* Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005).
9. Kittler, R. *et al.* Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nat. Cell Biol.* **9**, 1401–1412 (2007).
10. Schlabach, M.R. *et al.* Cancer proliferation gene discovery through functional genomics. *Science* **319**, 620–624 (2008).
11. 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).
12. Misaghi, S. *et al.* Association of C-terminal ubiquitin hydrolase BRCA1-associated protein 1 with cell cycle regulator host cell factor 1. *Mol. Cell. Biol.* **29**, 2181–2192 (2009).
13. Tyagi, S., Chabes, A.L., Wysocka, J. & Herr, 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).
14. Scheuermann, J.C. *et al.* Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. *Nature* **465**, 243–247 (2010).
15. Chou, W.C. *et al.* Distinct clinical and biological features of *de novo* acute myeloid leukemia with additional sex combs-like 1 (*ASXL1*) mutations. *Blood* **116**, 4086–4094 (2010).
16. Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H. & Helin, K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev.* **20**, 1123–1136 (2006).
17. Hassan, K.A., Chen, G., Kalemkerian, G.P., Wicha, M.S. & Beer, D.G. An embryonic stem cell-like signature identifies poorly differentiated lung adenocarcinoma but not squamous cell carcinoma. *Clin. Cancer Res.* **15**, 6386–6390 (2009).
18. Douglas, D. *et al.* BMI-1 promotes Ewing sarcoma tumorigenicity independent of CDKN2A repression. *Cancer Res.* **68**, 6507–6515 (2008).
19. Ding, L. *et al.* Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **455**, 1069–1075 (2008).
20. Harbour, J.W. *et al.* Frequent mutation of *BAP1* in metastasizing uveal melanomas. *Science* **330**, 1410–1413 (2010).
21. Mallery, D.L., Vandenberg, C.J. & Hiom, K. Activation of the E3 ligase function of the BRCA1/BARD1 complex by polyubiquitin chains. *EMBO J.* **21**, 6755–6762 (2002).
22. Ventii, K.H. *et al.* BRCA1-associated protein-1 is a tumor suppressor that requires deubiquitinating activity and nuclear localization. *Cancer Res.* **68**, 6953–6962 (2008).
23. Fisher, C.L. *et al.* Loss-of-function Additional sex combs like 1 mutations disrupt hematopoiesis but do not cause severe myelodysplasia or leukemia. *Blood* **115**, 38–46 (2010).
24. Sowa, M.E., Bennett, E.J., Gygi, S.P. & Harper, J.W. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**, 389–403 (2009).
25. Krug, L.M. *et al.* Potential role of histone deacetylase inhibitors in mesothelioma: clinical experience with suberoylanilide hydroxamic acid. *Clin. Lung Cancer* **7**, 257–261 (2006).
26. Crisanti, M.C. *et al.* The HDAC inhibitor panobinostat (LBH589) inhibits mesothelioma and lung cancer cells *in vitro* and *in vivo* with particular efficacy for small cell lung cancer. *Mol. Cancer Ther.* **8**, 2221–2231 (2009).

ONLINE METHODS

MPM samples and microarray hybridization. Tumor samples were collected and processed according to MSKCC institutional review board approved tissue collection protocols. Genomic DNA and total RNA were purified from MPM tumors and cell lines using DNeasy and RNeasy column purification kits (QIAGEN), respectively. All MPM tumor tissues were pleural primaries, except for one sample that was a tongue metastasis (tumor A76). We performed genomic copy number profiling on Agilent 244K CGH arrays using these purified DNA specimens. All samples were confirmed to have a robust array CGH profile and a p16 (*CDKN2A*) status concordant with that previously obtained by fluorescent *in situ* hybridization (FISH)[4]. The initial 53 MPM tumor samples were previously studied by gene expression profiling on Affymetrix U133A arrays[4]. DNAs from cell lines were hybridized to Affymetrix 250K SNP arrays according to the manufacturer's instructions. RNA from both tumors and cell lines was hybridized to Affymetrix U133A and U133 plus 2.0 arrays, respectively. Gene expression levels were estimated using the robust multichip average method[27]. Basic clinical and pathologic data on the set of 53 subjects with MPMs are provided as **Supplementary Data**.

Data analysis. Array CGH data were analyzed using both a traditional minimal common region (MCR)-type technique[28] as well as a RAE analysis (see below). To identify the MCRs, the normalized $\log_2$ tumor to normal ratios for each sample were segmented using circular binary segmentation (CBS) with $\alpha = 0.01$ and sd.undo = 1 (ref. 29). A 'gain' was a segment more than one median absolute deviation from the median segmented value, where the median absolute deviation was estimated from the deviation from the segmented means, and a 'loss' was similarly defined but in the opposite direction. Other segments were called 'normal'. Before constructing MCRs, the calls were smoothed and an adjustment was made for possible germ-line CNVs. To smooth the calls, normal segments less than 1 Mb flanked on both sides by gains or losses were changed to gains or losses, respectively, and altered segments less than 0.5 Mb surrounded on both sides by normal segments were changed to normal. To adjust for CNVs, altered regions less than 3 Mb that overlapped CNVs found in the Database of Genomic Variants (see URLs) and flanked on both sides by normal regions were changed to normal. 'Common regions' were consecutive markers for which the proportion of samples gained or lost exceeded 10%. 'Minimal common regions' were local maximums within the common regions of size greater than 1 Mb. Gene expression data were integrated with genomic copy number data to identify the subset of genes with copy number alterations that had the appropriate changes in their expression patterns, particularly in broad areas of gain or loss where identifying driver genes by copy number analysis alone was difficult. Differences in gene expression between altered and non-altered regions were identified using the Limma method[30].

RAE analysis. Raw data were obtained and normalized as previously described[31]. Probe-level data were segmented with circular binary segmentation ($\alpha = 0.01$, sd.undo = 1) and analyzed with RAE, both as previously described[29,32]. Regions of statistically significant copy number alteration (CNA) were determined after multiple hypothesis correction with the Benjamini and Hochberg false discovery rate (FDR) procedure (FDR < 10%, q value < 0.1). The deletion status of *BAP1* was assessed in each tumor based on the overlapping regions of the unified breakpoint profile inferred from segmentation spanning its RefSeq transcript coordinates (hg18). Samples were considered to harbor either heterozygous loss ($D_0 \geq 0.9$ and $D_1 < 0.9$) or homozygous deletion ($D_1 \geq 0.9$). Copy number segmentation was visualized in the Integrative Genomics Viewer (see URLs).

Clustering. The 53 copy number samples were clustered using non-negative matrix factorization as adapted for copy number[33,34]. See the URLs for the Matlab software for the analysis that was provided by the original authors. The stability of the clustering was used to choose the number of clusters.

Sequencing. Sanger sequencing of selected genes was carried out by a commercial vendor on a high-throughput platform (Beckman Coulter Genomics). Candidate mutations were confirmed by repeat PCR and resequencing. Sequencing of normal DNA from matched specimens was carried out in cases where such tissue was available. Sequencing primers are listed in **Supplementary Table 9**.

Immunohistochemistry. IHC was performed on formalin-fixed paraffin-embedded (FFPE) sections using BAP1 antibody (Santa Cruz, sc-28383) and a Ventana immunostainer Discovery XT. After pretreatment (CC1 solution), primary antibody was applied (dilution 1:50) for 30 min. The detection KIT-OMNIMAP (Ventana) was used for detection following the manufacturer's recommendations. Slides were scored based on nuclear staining intensity as follows: the percentage of positive tumor cells was multiplied by a factor of 1 to 3 (0, no staining; 1, faint staining; 2, moderate staining; and 3, strong staining), giving a final score between 0 and 300. Triplicate tumor tissue cores from the TMA were scored independently, and a final composite result (average of the three numbers) was given for each case. Cases were scored negative if the result was 0. The relationship between *BAP1* mRNA expression, BAP1 IHC expression and age was examined using an ANOVA test (GraphPad software).

Fluorescent *in situ* hybridization. The *BAP1* locus-specific FISH probe was made from BAC clone RP11- 630O10. BAC DNA was extracted using BACMAXTM DNA purification kit (Epicentre Biotechnologies) and labeled with SpectrumOrange-dUTP (red) using the nick-translation kit (Vysis/Abbott Molecular). Chromosome 3 centromere probe (CEP3, Cytocell) was labeled with green fluorescence; nuclei were counterstained with DAPI (blue). Four micron (4 μm) FFPE sections cut from tissue microarray blocks containing subject tumor specimens were pretreated by deparaffinizing in xylene and dehydrating in ethanol. Dual-color FISH was performed according to the protocol for FFPE sections from Vysis/Abbott Molecular with a few modifications. FISH analysis and signal capture were performed on a fluorescence microscope (Zeiss) coupled with the ISIS FISH Imaging System (MetaSystems). Fifty cells from each tumor specimen were analyzed.

RT-PCR and sequencing of *BAP1* mRNA transcripts. RT-PCR was performed on total RNA from samples harboring potential splice site mutations using the OneStep RT-PCR Kit (QIAGEN), and the primers used are listed in **Supplementary Table 10**. Products were confirmed on a 2% agarose gel and excised. Products were purified using Qiaquick PCR purification kit (QIAGEN) and subjected to Sanger sequencing.

Plasmids. Full-length *BAP1* complementary DNA was amplified by RT-PCR using Pfx DNA Polymerase (Invitrogen) from pENTR(tm)221-BAP1(Invitrogen) (for primers, see **Supplementary Table 11**) and subcloned into the BamH1 and Sal1 sites of the mammalian expression vector p3X-FLAG-CMV-24 (for 3xFlag tag) (Invitrogen). Site-directed mutagenesis of p3XFLAG-BAP1 was performed by QuikChange II (Stratagene) using the primers listed in **Supplementary Table 11**, and all introduced mutations were confirmed by sequencing.

Tissue culture, transfections and siRNA knockdown. All MPM cell lines were grown in Roswell Park Memorial Institute (RPMI) medium supplemented with 10% FCS and penicillin/streptomycin. HEK293T cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% FCS and penicillin/streptomycin. For knockdown expression studies, MPM cell lines expressing wild-type BAP1 (HMeso, MSTO211H and H2373) were treated with 10 nM of either BAP1 siRNA (s15822, si15820; Ambion) or negative control siRNA (AllStars, QIAGEN) using HiPerfect transfection reagent (QIAGEN). After 96 h, total RNA from each cell line was extracted, labeled and hybridized to Affymetrix U133A 2.0 arrays. Gene expression changes were identified using the Limma test[30]. For gene-specific validation, quantitative RT-PCR was performed using TaqMan gene expression assays (Applied Biosystems).

For knockdown proliferation studies, 3×10^4 cells of wild-type BAP1 cell lines (HMeso, Meso211H and H2373) were plated in 6-well plates on day 0. On day 1, transfection was performed with 10 nM of the same siRNA reagents as above. Cells were grown for 96 h, trypsinized and counted. For transfection growth studies, cells were transfected using Gene Jammer (Stratagene) and 125 ng of *BAP1* plasmids. Cell were grown for 4 days, trypsinized and counted. All proliferation experiments were performed in triplicate and the results were averaged.

Immunoprecipitations. Cells were washed once with ice-cold phosphate-buffered saline before scraping them off at 4 °C with 1 ml of phosphate-buffered

saline. Cells were resuspended in radioimmune precipitation assay buffer (RIPA) buffer (50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1% Nonidet P-40, 1 mM EDTA, 2.5 mM EGTA, 10% glycerol and the phosphatase inhibitors P-5726 and P-2850 (Sigma)). Cell extracts were then sonicated and centrifuged. Supernatants were used as crude extracts for immunoprecipitations. Nonspecific binding was reduced by pre-incubation of extracts with protein G-Sepharose (P-4691; Sigma) for 30 min. Pellets were discarded, and extracts were incubated with immune sera or controls for 2–4 h. Flag immunoprecipitations were performed with 5 µl of anti-FLAG Beads (M2; Sigma). BAP1 and HCF1 immunoprecipitations were performed with 5 µl of anti-BAP1 (C4; Santa Cruz) or anti-HCF1 (A301-399A, Bethyl Labs) and Protein A beads (16-157, Millipore).

Ubiquitin-AMC assay. The procedure was done as described[12]. Wild-type and mutant BAP1-Flag proteins were immunopurified with anti-Flag beads (30 µl) and eluted in four washes (50 µl each) of NP-40 buffer containing 2 mM DTT, 5% glycerol and 0.5 mg of three-Flag peptide/ml. To assess the UCH activity, the fluorogenic substrate 7-amido-4-methylcoumarin derivatized ubiquitin (ubiquitin- AMC; Boston Biochem) was diluted to a final concentration of 340 nM in 190 µl of the assay buffer (50 mM HEPES pH 7.5, 0.5 mM EDTA and 1 mM DTT). Samples were incubated for 2 h at room temperature (25 °C), and the levels of hydrolyzed AMC were measured by excitation at 380 nm and emission at 460 nm. UCH-L3 (400 µg; Boston Biochem) was used as a positive control, and 10 µl of BAP1-Flag eluate was used in each assay. All samples were tested in triplicate.

Predicted functional impact of somatic mutations. To predict the likely functional impact of missense mutations, we used the computation method Mutation Assessor (see URLs), which predicts the effect of a given mutated residue on protein function using evolutionary information from protein-family sequence alignments and residue placement in known or homology-deduced three-dimensional protein and complex structures[35].

Statistical analysis of *BAP1* loss or mutation status versus clinicopathologic variables. Overall survival was computed from the date of surgery and compared to each factor (*BAP1* loss or mutation status, BAP1 IHC expression, sex, histologic subtype, stage of the disease and asbestos exposure). Survival curves were estimated by the Kaplan-Meier method with a log-rank test (MedCalc Software).

Immunofluorescence microscopy of damage-induced DNA repair foci. Cells were seeded in 8-well chamber slides (BD Biosciences) at 3–5 × 10^4 cells per well and cultured overnight at 37 °C, 5% CO_2. Cells were either sham treated or exposed to 10 Gy of ionizing radiation using a Shepherd Mark-1 137-Cesium irradiator. Following 5 h incubation, cells were fixed and permeabilized in 2% paraformaldehyde, 0.1% Triton X-100 in PBS for 30 min at room temperature. Cells were further permeabilized with 0.5% Triton X-100 in PBS for 10 min at room temperature and rinsed, and non-specific antibody binding was blocked by treatment with 10% Bovine Growth Serum (Hyclone Laboratories) in PBS for 1 h at room temperature. Antibodies were diluted in PBS containing 1% bovine serum albumin, 0.1% Triton X-100, and subsequent rinses were performed with PBS, 0.1% Triton X-100. Cells were rinsed and incubated in rabbit anti-Rad51 IgG (Ab-1, Calbiochem; PC130, 1:600 dilution) and mouse anti-BRCA1 IgG (D9, Santa Cruz Biotechnologies; sc-6954, 1 µg/ml) overnight at 4 °C. Cells were rinsed three times for 10 min and incubated in Alexa Fluor 488-conjugated chicken anti-rabbit IgG (Molecular Probes/Invitrogen; 1.7 µg/ml) and Alexa Fluor 594-conjugated goat anti-mouse IgG (Molecular Probes/Invitrogen; 1.7 µg/ml) for 1 h at room temperature. Cells were rinsed three times with PBS, 0.1% Triton X-100, and once with PBS. Coverslips were mounted on slides with Vecta-Shield mounting medium containing DAPI (Vector Laboratories) and sealed. Images were acquired using a Zeiss LSM510 confocal microscope and analyzed using ZEN 2007 Image Analysis Software (Carl Zeiss Microimaging, Inc.).

Cytotoxicity assays and PARP inhibition. Cell lines with wild-type BAP1 (MSTO211H, Meso9, HMeso) or mutant BAP1 (H226, Meso37, H2452) were plated in 12-well culture dishes. After 24 h, cells were treated with either DMSO or various concentrations of PARP inhibitor, MK4827 (gift of Merck) (range 40 nM to 10 uM). Control and drug-treated media were replaced every 4–5 days. At 14 days, cell viability was assessed using a WST-1 assay (Roche Diagnostics).

Overrepresentation analysis. The goal of the overrepresentation analysis was to assess the significance of enrichment of the recently published Polycomb signatures[16–18] in the BAP1 knockdown list. Hypergeometric *P* values were calculated for overrepresentation of each Polycomb signature in the BAP1 knockdown list.

27. Irizarry, R.A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003).
28. Aguirre, A.J. *et al.* High-resolution characterization of the pancreatic adenocarcinoma genome. *Proc. Natl. Acad. Sci. USA* **101**, 9067–9072 (2004).
29. Venkatraman, E.S. & Olshen, A.B. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* **23**, 657–663 (2007).
30. Smyth, G.K. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article3 (2004).
31. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061–1068 (2008).
32. Taylor, B.S. *et al.* Functional copy-number alterations in cancer. *PLoS ONE* **3**, e3179 (2008).
33. Brunet, J.P., Tamayo, P., Golub, T.R. & Mesirov, J.P. Metagenes and molecular pattern discovery using matrix factorization. *Proc. Natl. Acad. Sci. USA* **101**, 4164–4169 (2004).
34. Carrasco, D.R. *et al.* High-resolution genomic profiles define distinct clinico-pathogenetic subgroups of multiple myeloma patients. *Cancer Cell* **9**, 313–325 (2006).
35. Reva, B., Antipin, Y. & Sander, C. Determinants of protein function revealed by combinatorial entropy optimization. *Genome Biol.* **8**, R232 (2007).

doi:10.1038/ng.855