High-density array-CGH with targeted NGS unmask multiple noncontiguous minute deletions on chromosome 3p21 in mesothelioma

Yoshie Yoshikawa, Mitsu Emi, Tomoko Hashimoto-Tamaoki, Masaki Ohmuraya, Ayuko Sato, Tohru Tsujimura, Seiki Hasegawa, Takashi Nakano, Masaki Nasu, Sandra Pastorino, Agata Szymiczek, Angela Bononi, Mika Tanji, Ian Pagano, Giovanni Gaudino, Andrea Napolitano, Chandra Goparaju, Harvey I. Pass, Haining Yang, and Michele Carbone

*Department of Genetics, Hyogo College of Medicine, Nishinomiya, Hyogo 663-8501, Japan; †Thoracic Oncology Program, University of Hawaii Cancer Center, Honolulu, HI 96813; ‡Division of Molecular Pathology, Department of Pathology, Hyogo College of Medicine, Nishinomiya, Hyogo 663-8501, Japan; §Department of Thoracic Surgery, Hyogo College of Medicine, Nishinomiya, Hyogo 663-8501, Japan; †Department of Internal Medicine, Hyogo College of Medicine, Nishinomiya, Hyogo 663-8501, Japan; and ‡Department of Cardiothoracic Surgery, New York University Langone Medical Center, New York, NY 10016

We used a custom-made comparative genomic hybridization array (aCGH; average probe interval 254 bp) to screen 33 malignant mesothelioma (MM) biopsies for somatic copy number loss throughout the 3p21 region (10.7 Mb) that harbors 251 genes, including BRCA1 (breast cancer 1)-associated protein 1 (BAP1), the most commonly mutated gene in MM. We identified frequent minute biallelic deletions (<3 kb) in 46 of 251 genes: four were cancer-associated genes: SETD2 (SET domain-containing protein 2) (7 of 33), BAP1 (8 of 33), PBRM1 (polybromo 1) (2 of 33), and SMARCC1 (switch/sucrose nonfermentable SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily c, member 1) (2 of 33). These four genes were further investigated by targeted next-generation sequencing (tNGS), which revealed sequence-level mutations causing biallelic inactivation. Combined high-density aCGH and tNGS revealed biallelic gene inactivation in SETD2 (9 of 33, 27%), BAP1 (16 of 33, 48%), PBRM1 (5 of 33, 15%), and SMARCC1 (2 of 33, 6%). The incidence of genetic alterations detected is much higher than reported in the literature because minute deletions are not detected by NGS or commercial aCGH. Many of these minute deletions were not contiguous, but rather alternated with segments showing oscillating copy number changes along the 3p21 region. In summary, we found that in MM: (i) multiple minute simultaneous biallelic deletions are frequent in chromosome 3p21, where they occur as distinct events involving multiple genes; (ii) in addition to BAP1, mutations of SETD2, PBRM1, and SMARCC1 are frequent in MM; and (iii) our results suggest that high-density aCGH combined with tNGS provides a more precise estimate of the frequency and types of genes inactivated in human cancer than approaches based exclusively on NGS strategy.

Significance

We found that gene mutations/deletions are frequent in mesothelioma and occur through a variety of DNA alterations. We identified genes implicated in malignant mesothelioma: SETD2, SMARCC1, PBRM1. Previous next-generation studies (NGS) underestimated the frequency of genetic alterations in malignant mesothelioma because NGS mainly identifies nucleotide level mutations. Our findings are of general relevance to the field of cancer research that relies almost exclusively on NGS to identify gene alterations in cancer biopsies, and uses this information to design specific molecular therapies. An integrated approach that includes a high-density comparative genomic hybridization array and NGS or targeted-NGS, as conducted here, may reveal additional genes that are inactivated by mechanisms other than point mutations. This information may inform us on how to design more effective molecular therapies.

Malignant mesotheliomas (MMs) are highly aggressive adult malignancies that arise from the mesothelial cells covering the pleural, peritoneal, and pericardial cavities. MMs are often caused by exposure to asbestos and other mineral fibers (1). Co-factors and genetic predisposition may also cause MM, especially upon asbestos exposure (2–4). Diagnosis is usually established at a late stage, when MMs are most resistant to medical and surgical therapies. Recent clinical trials targeting molecular pathways have failed to significantly improve MM prognosis (1). Promising biomarkers are being investigated for early detection (5, 6). It is hoped that a better understanding of the molecular pathogenesis of MM will lead to more specific and effective targeted therapies (7).

Using next-generation sequencing (NGS) for exome sequencing involves significant—and at times overlooked—challenges, including DNA quality and quantity, tumor heterogeneity, and the difficulty to detect a wide variety of complex genetic mutations (8). Conventional NGS studies revealed that the most commonly mutated genes in MMs are BRCAl-associated protein 1 (BAP1), CDKN2A (cyclin-dependent kinase inhibitor 2a), and NF2 (neurofibromin 2) (9–12). The reported frequency of BAP1 inactivation, the most commonly mutated gene in MM, using NGS or Sanger sequencing, was between 20% and 30% (9–15).

Overall, NGS studies revealed that each single MM biopsy has its own set of specific mutations, and that driver mutations—except for BAP1—are rare (reviewed in ref. 16). For example, Guo et al. (9) conducted NGS exome sequencing of the whole genome in 22 MM frozen biopsies: they found 490 mutated genes, of which 477 (97%) were mutated only in one biopsy, and found an average of 23 mutations per biopsy (range 2–51). Mutations in human cancers range from as low as one base substitution per exome (≤0.1 per megabase) in some pediatric malignancies, to thousands of mutations per exome (≥100 per megabase) in adult malignancies, such as lung cancer and melanoma (17). Thus, the low level of mutations detected by NGS in MM (9, 11, 18) was unexpected and highly unusual. However, past...
cytogenetic studies revealed that MMs have multiple numerical and structural chromosomal abnormalities (19), findings indicating that genetic alterations in MMs should be frequent, not rare. In support of this hypothesis, comparative genomic hybridization array (aCGH) and multiple ligation-dependent probe amplification studies revealed that several BAP1 deletions were missed by NGS and Sanger sequencing, because of their large size (20, 21).

Large DNA deletions and copy number (CN) changes are best detected by aCGH and SNP arrays. MM CN analyses using whole-genome aCGH and SNP arrays revealed the presence of occasional large genomic deletions (22–26); however, the presence of smaller deletions (i.e., minute deletions (<3 kb]) has not been specifically investigated. Deletions in that size range fall in a gray area: they are too large to be reliably detected by NGS and Sanger sequencing and too small to be reliably detected by commercial aCGH. We reasoned that such hypothetical minute deletions would not be detected by conventional whole-genome aCGH, as these arrays have only an average of 400,000–1,000,000 probes: because the human genome has over 3 billion base pairs, commercial aCGH provide only one probe every 3,000–8,000 bp (for example, the Agilent CGH array). Whole-exome NGS and targeted NGS (tNGS) sequencing are sensitive to detect minute deletions, but these techniques may produce false-positive results because they are not designed to measure DNA CN (8). Thus, it is not possible to draw definitive conclusions about CN changes relying exclusively on NGS technology, and the NGS results should always be validated by integrated genomic approaches (8).

Here, we hypothesized and tested whether, in addition to the relatively infrequent single-nucleotide mutations detected by NGS studies, larger DNA alterations, including minute deletions, might contribute to inactivate additional tumor-suppressor genes in MM. To screen for the presence of possible minute deletions, we designed a custom-made high-density oligonucleotide aCGH with probes designed at an average spacing of 8,000 bp (8,000 probes: because the human genome has over 3 billion base pairs, 8,000 bp would not be detected by conventional whole-genome aCGH, as these arrays have only an average of 400,000–1,000,000 probes: because the human genome has over 3 billion base pairs, commercial aCGH provide only one probe every 3,000–8,000 bp (for example, the Agilent CGH array). Whole-exome NGS and targeted NGS (tNGS) sequencing are sensitive to detect minute deletions, but these techniques may produce false-positive results because they are not designed to measure DNA CN (8). Thus, it is not possible to draw definitive conclusions about CN changes relying exclusively on NGS technology, and the NGS results should always be validated by integrated genomic approaches (8).

Using our custom-made high-density aCGH, we were able to identify deletions that would have been missed using conventional aCGH analysis (Figs. S1 and S2). In total, we identified biallelic minute deletions (<3 kb) or larger deletions involving some parts of the 3p21 region in 19 of 33 MMs (Fig. L4 and Table S3). These deletions were concentrated into two DNA regions: a telomeric region at 52.2–53.2 Mb (cluster 2) (Fig. L4). Representative deletions are shown in Fig. 1 B and C and in Fig. S2. Fig. 1B displays sample MM21, showing loss of heterozygosity (LOH) in two regions (43.8–45.4 Mb and 51.9–53.5 Mb). Within the centromeric region, we found biallelic deletions in three noncontiguous regions in gene cluster 2: (i) deletion of the BAP1 gene (BAP1 exons 1–16 and its adjacent PHF7 (PHD finger protein 7) gene; (ii) deletion of the NISCH (nischarin) gene (exons 3–15); and (iii) deletion of the PBRM1 gene (exons 22–30). Fig. 1C shows sample MM57 that contained LOH of the entire 3p21 region. We found biallelic deletions in three noncontiguous regions in telomeric cluster 1: (i) deletion of SETD2 (exon 1) and its adjacent FLJ39534 (gene exons 1–2); (ii) deletion of KLHL18 (kelch like family member 18) (exons 2–10); and (iii) deletion of SMARCC1 (exons 3–5). Moreover, we found

![Fig. 1. Segmental CN loss detected in 3p21 by high-density aCGH. (A) Frequency of biallelic deletion detected among 251 genes in 3p21 in 33 MMs. On the vertical-axis, genes are aligned along the genome location. The blue horizontal bars represent each of the 46 genes and the length of the bar represents the case frequency that showed biallelic deletions. (B and C) Representative profiles of CN change in MM21 and MM57, respectively. CN is shown as log2 ratio and the regions showing biallelic deletion are marked in red.](image-url)
with reduced transcript levels (Fig. S4). A strong positive correlation, suggesting that monoallelic gene loss was associated (Fig. S3); aCGH and mRNA data showed a significant positive correlation validated in a panel of MM cell lines and in a subset of MM samples (Fig. 2). Deregulated mRNA expression of these genes was confirmed. The most frequently mutated genes, comprising clusters 1 and 2, are shown in Fig. 2. Among the mutated genes, four genes previously associated with cancer contained frequent biallelic deletions; that is, SETD2 (biallelic deletion frequency: 7/33, 21%), SMARCC1 (2/33, 6%), BAP1 (8/33, 24%), and PBRM1 (3/33, 9%) (Figs. 1 and 2). These four genes were further investigated by tNGS, as described below. Among the remaining 42 genes that were found deleted (summarized in Table S3), the most frequently found were: SCAP (SREBP-steryl regulatory element binding transcription factor 1-cleavage activating protein) (5/33, 15%), PHF7 (5/33), NISCH (3/33), SEMAS3G (semaphorin 3G) (3/33), and FLJ39534 (3/33) genes. Several MMs (15/33, 45%) contained combined biallelic deletions of two or more genes in 3p21 detectable by aCGH (Table S3). Break points of these genome deletions were different among cases, except for a particular deletion in intron 1 of the SCAP gene, in which the deletion region (chr3: 47.49-47.5 Mb) coincided with the polymorphic CN-variable region in the human population (Database of Genomic Variants); thus, this was no longer investigated (representative samples shown in Fig. 1 and Fig. S2). Finally, the 3p21.3 region containing the tumor-suppressor gene cluster—which includes RBM5, TUSC2, HYAL1, and HYAL2—which is frequently deleted in several human tumors (27), did not show any biallelic deletions in MM.

Monoeallelic gene loss was also a frequent event: for example, SETD2 (monoeallelic deletion frequency: 7/33, 21%), SMARCC1 (14/33, 42%), BAP1 (10/33, 30%), and PBRM1 (12/33, 36%) (Fig. 2). Deregulated mRNA expression of these genes was confirmed in a panel of MM cell lines and in a subset of MM samples (Fig. S3); aCGH and mRNA data showed a significant positive correlation, suggesting that monoallelic gene loss was associated with reduced transcript levels (Fig. S4).

Validation of DNA CN Analysis of Selected Genes SETD2, BAP1, PBRM1, and SMARCC1 on 3p21 by tNGS Sequencing. SETD2, BAP1, PBRM1, and SMARCC1 were further investigated by tNGS. In these studies, the read depth was 150–200 for each target gene region, ensuring accuracy of the final sequence. The read depth was measured and compared by pair analysis between tumor DNA and its normal counterpart DNA from the same patient. All of the biallelic deletions detected by aCGH were consistently detected, and thus validated, by tNGS (Fig. 3 and Table S4). Moreover, tNGS analysis identified additional minute biallelic deletions that were not detected by aCGH, at frequencies of 7 of 33 MMNs (21%) for SETD2, 3 of 33 (9%) for BAP1, 6 of 33 (18%) for PBRM1, and 5 of 33 (15%) for SMARCC1 (Fig. 2 and Table S4). However, because these deletions were either hidden by nonnegligible noise of aCGH or very small, they could not be independently validated by high-density aCGH; therefore, we chose a prudent approach and they were not further considered.

Analysis of Sequence-Level Mutations in SETD2, PBRM1, BAP1, and SMARCC1. tNGS of these same tumors and normal DNAs detected somatic point mutations—which, as expected, were undetectable by aCGH—of these genes. SETD2 point mutations were found in three MMs (3 of 33, 9%); p.S1024fs (MM29), p.K1948X (NYU1250), and p.T1753fs (NYU1373), (Table 1). BAP1 point mutations were found in nine MMs (9 of 33, 27%); that is, p.S460X (MM35), p.W202X (MM45), p.V234- (MM29), c.2278-2283del/splice site del (NYU0866), p.I72fs (NYU0754), and p.E212X (NYU1306). Three of these nine MMs had two separate inactivating BAP1 somatic sequence-level mutations: p.S341fs and p.O280X (NYU0647), p.I499fs and p.V234fs (NYU0937), and p.Y646X and p.A648fs (NYU1353). BAP1 point mutations were identified in two MMs (2 of 33, 6%): p.V1582_1582del-stop lost (MM19) and p.Q235X (NYU0047). We did not detect any nucleotide mutation in the SMARCC1 gene. In addition, germ-line variants, each one in these four genes, were found; p.T1033A of SETD2 in MM case NYU0851, p.M1486f of PBRM1 and p.I1075H of SMARCC1 in MM35, and the splice site (NM_004656:c.438-2A>G) of BAP1 in NYU0754. Using tNGS, we were therefore able to identify both the germ-line and somatic mutation present in sample NYU0754, previously published as W-III-04 (14).

In total, combining all minute alterations and point mutations we confirmed biallelic inactivation of SETD2, BAP1, PBRM1, and SMARCC1 in at least 27% (9 of 33), 48% (16 of 33), 15% (5 of 33), and 6% (2 of 33) of 33 MMs, respectively (Fig. 2). Tissue-culture experiments revealed that silencing of SETD2, SMARCC1, or PBRM1 using siRNAs caused increased proliferation in a MM cell line wild-type for all of these three genes. These findings support the hypothesis that deletions of these genes are biologically relevant (Fig. S5).

We also investigated patient information for possible correlations between gene alterations and survival (patients with wild-type or monoeallelic deletions vs. patients with biallelic inactivations) using the Kaplan–Meier method stratified by stage. The results were not significant probably because of the small sample size.

![Fig. 2. Summary of genomic alterations in 33 malignant mesothelioma cases. Each column indicates genomic alterations in each case. Genes NBEAL2, SETD2, KLHL18, SCAP, and SMARCC1 comprise gene cluster 1; genes DNAH1, BAP1, PHF7, NISCH, PBRM1, and SFMBT1 comprise gene cluster 2. Biallelic deletion detected by both, high-density aCGH array and tNGS, are shown in dark blue. Monoallelic loss detected by both, high-density aCGH array and tNGS, are shown in light blue. Blue with white lines represents biallelic deletion detected by tNGS only. Bright pink: two separate somatic mutations of the same gene detected by tNGS at sequence level; light pink: one somatic mutation detected by tNGS at sequence level; dotted pink: one germ-line and one somatic mutation detected by tNGS at sequence level.](image-url)
of independent deletion events cannot be entirely ruled out. Using DNA segments, as in chromothripsis (28), although sequential rearrangements involving 46 genes on 3p21. Many of these deletions are likely due to somatic inactivation of tumor-suppressor genes. Our studies revealed multiple biallelic genome alterations of chromosome 3p21 in 33 MM using a high-density aCGH, followed by tNGS sequencing to validate the alterations detected and to investigate for nucleotide sequence mutations. We found multiple biallelic genome rearrangements involving 46 genes on 3p21. Many of these deletions were not contiguous, but rather they alternated along normal DNA segments, as in chromothripsis (28), although sequential independent deletion events cannot be entirely ruled out. Using high-density custom aCGH, we found frequent minute deletions of BAP1, SETD2, and PBRM1—the latter two genes had previously been found rarely mutated in MM (11, 29)—and of SMARCC1, not previously associated with MM. tNGS independently validated the deletions detected by high-density aCGH. In addition to minute biallelic deletions, tNGS revealed single-nucleotide inactivating mutations (i.e., truncating mutations) of SETD2, PBRM1, and BAP1 in several MMs. Although we detected CN alterations in SMARCC1, no nucleotide level mutations were found in this gene, confirming previous studies based on NGS (9–12) and underscoring the limitations of this technique when used alone. Together, the combination of high-density aCGH and tNGS detected a much higher percentage of genetic alterations in BAP1, SETD2, PBRM1, and SMARCC1 than reported in the MM literature, which is largely based on NGS sequencing (9–12). This is because NGS sequencing is a technique that is effective at identifying nucleotide mutations, but is not optimized for the identification of minute or larger chromosomal deletions (8). Moreover, minute genome changes are unlikely to be detected using commercial SNP arrays or whole-genome array CGH. Thus, an important implication of our study is that conventional NGS approaches are inefficient to identify reliably all of the different genetic alterations that occur in MM, and most likely in other cancers. For example, using a NGS approach, Bueno et al. (11) and Ugurluer et al. (12) reported inactivating nucleotide mutations of BAP1 and SETD2 in 20–30% and 8% of MMs, respectively. Instead, when combining high-density aCGH (to detect minute deletions of <3 kb) and tNGS (to detect nucleotide level mutations), we found that BAP1 and SETD2 were inactivated—either by minute biallelic deletion or LOH combined with point mutations—in 48% and 27% of MMs, respectively. Previous NGS studies underestimated the frequency of SETD2 deletions in MM because minute deletions, which are not reliably detected by NGS, were the most common deletions we found in this gene (i.e., seven of nine MM biopsies). Similarly, we detected biallelic deletions for PBRM1 in 9% of MMs and of SMARCC1 in 6% of MMs that were either underestimated or missed by previous NGS studies. Therefore, our data indicate that an integrated aCGH and tNGS approach is more accurate than NGS based-approaches to identify the different types of gene alterations present in human cancers.

Altersations of SETD2, PBRM1, and SMARCC1 have been linked to several human malignancies (30–39). Briefly, disruptions of the SETD2 gene have been frequently observed in clear cell renal cell carcinomas (ccRCC), gliomas, chronic lymphocytic leukemia, breast fibrop epithelial tumors, gastro-intestinal stromal tumors, and melanomas (30–37). PBRM1 is one of the most frequently mutated genes in ccRCC (31), and its alterations were also found in cholangiocarcinomas and liver cancers (38). Moreover, down-regulated PBRM1 expression was positively correlated with tumor stage and the overall survival in breast cancer patients (40). SMARCC1 has been directly shown to be a tumor-suppressor gene in colorectal and ovarian cancer models (41). Additionally, we found that silencing of these genes in a MM cell line results in increased proliferation, as expected for tumor-suppressor genes.

Because NGS and tNGS may overestimate the presence of biallelic deletions (8), only those deletions independently validated by both high-density aCGH and tNGS were included in the final results. Thus, it is possible that our results may underestimate the true frequency of biallelic deletions in BAP1, SETD2, PBRM1, and SMARCC1. Despite our conservative approach, we found a much higher frequency of gene alterations than the frequencies described in the published literature. Thus, our data suggest that we may need to reconsider the hypothesis that MMs are low mutation-frequency malignancies, as we and others had proposed based on NGS studies and Sanger sequencing (9–11).

The implications of our findings may be of general relevance to cancer researchers that currently focus almost exclusively on NGS studies to identify gene mutations in cancer biopsies and then use this information to design specific molecular therapies. A more comprehensive analysis of the whole set of genetic alterations that occur in cancer, as conducted here, may likely reveal additional mutations in genes that are inactivated by mechanisms other than point mutations, and this information may inform us on how to design more effective molecular therapies.

Our studies revealed multiple biallelic genome alterations of several genes in the 3p21 chromosomal region: 46 genes contained biallelic deletions in at least one MM biopsy. Many of these deletions were not contiguous, but rather alternated with segments showing oscillating CN changes along the 3p21 region. Our findings may be consistent with chromothripsis, a condition caused by catastrophic events, such as aborted apoptosis, which causes DNA fragmentation followed by chromosomal rearrangements and loss of some DNA sequences (28). Alternatively, these segmental losses on 3p21 may have occurred sequentially because of some inherited fragility of this chromosomal region in MM.

Whereas biallelic alterations of tumor-suppressor genes can be easily reconciled with the current models of carcinogenesis, monoallelic alterations might be biologically relevant only for
haplo-insufficient genes. The functional characterization of these monoallelic alterations should be the focus of future investigations and is beyond the scope of this current study.

The existing literature, however, supports the hypothesis that monoallelic losses of these genes are of biological relevance: (i) In mice, monoallelic losses of \( BAP1 \) alter asbestos-induced peritoneal inflammatory response (42) and induce ccRCC in the presence of a coexisting \( Vh \)l deficiency (43). (ii) In chronic lymphocytic leukemia, \( SETD2 \) alterations are very often monoallelic, and epigenetic inactivation of the wild-type allele is absent. Thus, this is a typical example of a haplo-insufficient tumor suppressor gene (36). (iii) In the ExAC database, which includes data from more than 60,000 genomes, truncating variants of \( SETD2 \), \( PBRM1 \), and \( SMARCC1 \) are not tolerated in the germ line—probability of loss-of-function intolerance score of 1.0—(44) implying that monoallelic losses of these genes, as observed in our MM samples, have relevant biological consequences. Based on our results and on these observations, we expect that context-specific and co-occurring events might result in haplo-insufficient behavior of at least some of the \( 3p21 \) genes that we identified.

\( BAP1 \) itself is part of the gene set undergoing biallelic inactivation in the \( 3p21 \) region and its loss-of-function might represent a positive feedback because \( BAP1 \) inactivation impairs DNA repair, possibly contributing to the catastrophic massive genomic rearrangement in this particular region of the chromosome. Thus, key events in MM carcinogenesis might include multiple CN losses caused by (i) fragmentation of \( 3p21 \) induced by oxidative stress caused by asbestos, the main carcinogen in MM (reviewed in ref. 45), and (ii) aberrant DNA repair because of \( BAP1 \) inactivation (reviewed in ref. 46). Similarly, alterations of \( SETD2 \), another epigenetic modifier involved in transcription elongation, RNA processing, and DNA repair, plays an important role in maintaining genomic integrity, as observed in ccRCC (35). Therefore, \( SETD2 \) alterations may also contribute to the genomic instability of the \( 3p21 \) region in MM. Loss of \( PBRM1 \) or \( SMARCC1 \), as members of the switching/sucrose nonfermenting (SWI/SNF) complex, may also lead to chromosomal instability because of their role in chromatin remodeling and sister chromatid cohesion (41, 47). In other words, each of these four genes that we found to be relatively frequently mutated in \( 3p21 \) in MM, is involved in the epigenetic modification of DNA, an unlikely coincidence.

An obvious question that remains to be investigated is whether our observations are specific to the \( 3p21 \) region or, more likely, extend to other genome regions in MM. Those DNA regions that include \( CDKN2A \) and \( NF2 \), which are frequently mutated in MM (9–12), should be studied to investigate the possible presence of similar clusters of gene rearrangements—as observed for \( 3p21 \)—in the corresponding chromosomal regions (9p21 and 22q21, respectively).

Materials and Methods

DNAs from MM Specimens. We studied DNAs from 9 MMs surgically resected at the Hyogo College of Medicine, and DNAs from 24 MMs surgically resected at the New York University. For the Japanese samples, DNA was extracted from matching blood samples and from primary tumor cells established in tissue culture, to enrich for tumor cells, from resected MM biopsies using DNeasy Blood & Tissue Kit (Qiagen) or QiAamp DNA Micro Kit (Qiagen 56304) according to the manufacturer’s instructions. For the United States samples, DNA was extracted from blood and from tumor cells obtained by laser-capture microdissection performed on MM biopsies with a MMI CellCut Plus (Molecular Machines & Industries) (21).

This study was approved by the Ethics Committee of Hyogo College of Medicine and by the Institutional Review Board of the New York University and performed in accordance with the Declaration of Helsinki (1995) of the World Medical Association (as revised in 2013 in Fortaleza, Brazil) (48). Written informed consent was received from all patients. Collection and use of patient information and samples were approved by the Institutional Review Board of the Hyogo College of Medicine and the New York University.

High-Density Custom-Made CGH Microarray Analysis. We designed a custom-made microarray containing 42,125 oligonucleotide probes targeting a 10.7-Mb genomic region of \( 3p21 \) (Chr 3: 43,700,000-54,400,000) of the National Center for Biotechnology Information Build 37, hg19) at an average spacing of 254 bp within the nonrepeat masked regions of the genome using the Agilent’s database (Agilent Technologies). For quality control and normalization, 10,148 probes covering all chromosomes were also contained in this custom array. MM samples and reference DNAs (each 250 ng) were labeled with Cy5 or Cy3, respectively, using the Agilent DNA labeling kit. Following the manufacturer’s recommendation, hybridization and washes were performed in stringent conditions. Then, the arrays were scanned at 5-μm resolution, using the Agilent microarray scanner and analyzed using Feature Extraction v10.7.3. Data analysis was performed using Agilent CytoGenomics 3.0.5.1;
The setting of data analysis is described in SI Materials and Methods. The CN data were plotted after moving average calculated using the linear algorithm, univariate average, using every probe log-ratio score within a 10 kb. The CN alteration of DOCK3 was not verified because this gene was the most frequent outlier for the CN ratio (CN ratio > log 0.4, < –log 0.4) by the selfexperiment using the same genomic DNA.

tNGS. NGS was performed on an Illumina MiSeq using paired-end 150-bp runs. Libraries were prepared from 100 ng of genomic DNA from MM tumors, paired normal and reference DNA, using a Haloplex Custom kit (Agilent Technologies) according to the manufacturer’s instructions. Illumina paired-end reads were each aligned to the human NCBI Build 37 reference sequence using bwa software (bio-bwa.sourceforge.net, v0.6.0). The aligned sequence files were sorted and merged using SAMTools (samtools.sourceforge.net, v0.1.18). GATK (https://software.broadinstitute.org/gatk/) was used for realignment, base quality score recalibration, single-nucleotide variant or indel (small insertions and deletions) variant calling, and variant quality recalibration. Snippy was used to categorize the effects of variants by impact. CN analysis was performed by StrandNGS 2.5, by estimating the contamination rate and ploidy in tumor cells. The estimated CN ratio is listed in Table S4.

ACKNOWLEDGMENTS. This work was supported in part by a Grant-in-Aid for Scientific Research Grants KAKENHI 24597015 and 15K08658 (to T.T.), 25460710 (to M.E.), and 26460689 (to Y.F.); a Grant-in-Aid for Researchers Grant from Hyogo College of Medicine, 2015 (to Y.F.); National Cancer Institute Grants NCIC-R01 CA198138 (to M.C.) and NCIC-R01 CA160715 (to H.Y.); the University of Hawai‘i Foundation, which received an unrestricted gift to support malignant mesothelioma research from Honeywell International Inc. (to M.C.); United-4-a-Cure (M.C. and H.Y.); and Bellocq and Fox (to H.I.P.).
Yoshikawa et al. 10.1073/pnas.1612074113

SI Materials and Methods

**Survival Statistics.** Kaplan–Meier survival estimates were performed using the LIFETEST procedure in the SAS 9.4 software (SAS Institute). We compared patients with at least one wild-type copy of each gene (wild/wild, wild/+, and wild/variant) to patients without any wild-type copy (+/−, −/−, and variant/variant), stratified by stage (1–2 vs. 3–4).

**Data Analysis of High-Density Custom-Made CGH Microarray.** Data analysis was performed using Agilent CytoGenomics 3.0.5.1. Normalization parameters were set on the GC correction (window size = 2 kb) and diploid peak centralization; the aberration detection was using algorithm ADM-2; and for the aberration filter, the minimum number of probes in regions was set to three and the minimum absolute average log-ratio of region was set to each array. The setting of this log ratio was important to judge whether the deletion in the region was monoallelic or biallelic, and the average log ratio per one allele was obtained using the normalized log-ratio distribution.

**Cell Line and In Vitro Experiments.** The PPM-Mill cell line was established from a surgically resected human MM specimen, and was characterized and provided by H.I.P. (New York University, New York). This cell line was selected for in vitro experiments because it is wild-type for PBRM1, SETD2, and SMARCC1. Of these PPM-Mill cells, 2 × 10⁶ cells were plated per well in a 96-well plate and grown in DMEM + 1% (vol/vol) FBS. PBRM1 (sc-76075), SETD2 (Dharmacon D-012448-01), SMARCC1 (Dharmacon D-010813-01), or scrambled control siRNAs (sc-37007, D-001810-10-05) were transfected using HiPerfect Transfection Reagent (Qiagen, cat. no.: 301705) according to the manufacturer’s conditions. To evaluate the level of silencing, total RNA was isolated from the cultured cells with TRizol reagent (Invitrogen) and quantified using microspectrophotometry (NanoDrop Technologies). Total RNA was reverse-transcribed using the high-capacity cDNA reverse-transcription kit (Invitrogen) according to the manufacturer’s protocols. Quantitative RT-PCR was performed with TaqMan probes and TaqMan universal PCR Master Mix (Invitrogen). The amplification was carried out with an ABI PRISM 7900 RT-PCR System. Triplicate assays were performed with RNA samples isolated from at least two independent experiments. Fold-changes in gene expression were calculated using the ΔΔCt method using β-actin as the normalization control. Results from cells transfected with siRNAs for one of the 3p21 genes were then normalized to those obtained with cells transfected with their respective scrambled control siRNA (siSCR). Growth of siPBRM1-, siSETD2-, siSMARCC1-, and siSCR-transfected cells was determined using the Alamar blue assay according to the manufacturer’s protocol (AlamarBlue, Thermo Fisher). The assay was performed in five replicates from at least two independent experiments.

**Gene Microarray Expression.** Gene-expression profiling was performed with Affymetrix Human Genome U133 Plus 2.0 arrays following the manufacturer’s instructions. Raw data (CEL files) were summarized using the “plier16” algorithm and quantile normalization was performed by GeneSpring GX11.0.2 (Agilent Technologies). Gene-expression levels were presented compared with the average of three reactive mesothelial cells as log-transformed.

![Fig. S1](image-url) High-density aCGH is superior to conventional aCGH: a representative MM sample. MM48 shows monoallelic deleted regions using conventional aCGH.
Fig. S2. (Continued)
Fig. S2. CN loss in 3p21 region, detected by high-density aCGH in representative MM samples. The regions judged to be biallelically deleted were marked in a red color and the length of these markers indicates the average log ratio of the consecutive probes in the deleted region. (A) MM14, MM48, SELF/SELF (technical control), which indicates that the same genomic DNA, HapMap Male (NA12891), was used for both Cy5 and Cy3 labeling in the high-density aCGH analysis. (B) MM35, MM45, NYU0093, NYU0937.
Fig. S3. Heap map showing relative log₂ ratio of mRNA expression in five MM cell lines and nine MM clinical samples. Both MM lines (H2052, H2452, H28, HMMME, MSTO-211H) and MM clinical samples (MM14, 19, 21, 26, 29, 35, 45, 48, 57) show deregulated mRNA levels of 3p21 genes.

Fig. S4. Correlation between high-density aCGH data and mRNA expression. Statistically significant correlation between high-density aCGH data and mRNA expression for SETD2, PBRM1, SMARCC1, and BAP1. For the x axis: 0.0 for biallelic deletion; 0.5 for monoallelic deletion; 1.0 for absence of CN variation. Monoallelic losses often presented intermediate levels of mRNA expression.
Silencing of SETD2, SMARCC1, or PBRM1 is associated to increased cell proliferation. (A) PBRM1, SETD2, and SMARCC1 genes were efficiently silenced in PPM-Mill cells at 72 h using siRNAs. (B) At the same time point, viability of silenced cells was significantly higher than control. Data were compared using Mann–Whitney test; *P < 0.05, **P < 0.01.

Other Supporting Information Files

Table S1 (DOCX)
Table S2 (DOCX)
Table S3 (DOCX)
Table S4 (DOCX)