Whole-Exome Sequencing Reveals Frequent Genetic Alterations in BAP1, NF2, CDKN2A, and CUL1 in Malignant Pleural Mesothelioma

Guangwu Guo1,2, Juliann Chmielecki1,2, Chandra Goparaju3, Adriana Heguy4, Igor Dolgalev4, Michele Carbone5, Sara Seepo1, Matthew Meyerson1,2,6, and Harvey I. Pass5

Abstract

Malignant pleural mesothelioma (MPM) is an aggressive neoplasm associated with asbestos exposure. Although previous studies based on candidate gene approaches have identified important common somatic mutations in MPM, these studies have focused on small sets of genes and have provided limited view of the genetic alterations underlying this disease. Here, we performed whole-exome sequencing on DNA from 22 MPMs and matched blood samples, and identified 517 somatic mutations across 490 mutated genes. Integrative analysis of mutations and somatic copy-number alterations revealed frequent genetic alterations in BAP1, NF2, CDKN2A, and CUL1. Our study presents the first unbiased view of the genomic basis of MPM. Cancer Res; 1–6. ©2014 AACR.

Introduction

Malignant pleural mesothelioma (MPM) is a rare and highly aggressive tumor of serosal surfaces associated with exposure to asbestos, whose incidence is increasing worldwide (1). Effective treatment for MPM is difficult due to delays in diagnosis, nonuniformity of staging, and lack of effective medical therapy. Although several major advances in the management of patients with MPM have been made in the past few years, including the use of pemetrexed and cisplatinum, currently available treatments still appear to provide only modest benefit (1). Previous genetic analyses have identified several key genetic alterations, including inactivation of CDKN2A (2), NF2 (3), and BAP1 (4). In addition, frequent losses of chromosome arms 1p, 3p, 4q, 6q, 9p, 13q, 14q, and 22q and gains of chromosome arms 1q, 5p, 7p, 8q, and 17q have been reported in MPM (5). These studies provide broad insights into the biology of MPM and identification of new targets for therapy. However, no comprehensive analysis of MPM genomes using next-generation sequencing have yet been reported.

Advances in technologies for high-throughput sequencing of DNA have enabled comprehensive characterization of somatic mutations in different cancer types at an unprecedented resolution (6). In particular, exome sequencing represents a cost- and labor-efficient strategy to identify mutations in protein-coding exons in the human genome. In recent years, exome sequencing has been successfully applied to a variety of human cancers, enabling cancer genome discoveries with the potential to translate into advances in cancer diagnosis and treatment as well as basic cancer biology (7).

Here, we performed exome sequencing on genomic DNA from 22 MPM tissues and matched normal samples. We found frequent inactivation of BAP1, NF2, CUL1, and CDKN2A by somatic mutations and/or copy-number alterations. We also observed frequent alterations in several canonical cancer-related pathways. In summary, our data provide a first view of the comprehensive landscape of somatic genome alterations in MPM.

Materials and Methods

Patient samples collection and preparation

All 22 tumors and matched blood samples were obtained from individuals diagnosed with MPM at the New York University (NYU) General Thoracic Surgery Service (Supplementary Table S1). All patient samples were collected in the operating room, immediately snap-frozen in liquid nitrogen within 10 minutes of collection, and then stored at −80°C. Blood was withdrawn before the initiation of anesthesia for serum, plasma, and peripheral blood mononuclear cells (PBMC). The PBMCs, cryopreserved after processing, were used for normal genomic DNA matched to the specimens and were processed as per the manufacturer’s specifications (Becton Dickinson; BD CPT vacutainer tubes). Percentages of tumor within the tumors ranged from 70% to 90% based on histologic examination, and

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Corresponding Author: Harvey I. Pass, Department of Cardiothoracic Surgery, NYU Langone Medical Center, 530 First Avenue, 9V New York, NY 10016. Phone/ fax: 212-263-7417; E-mail: harvey.pass@nyumc.org

doi: 10.1158/0008-5472.CAN-14-1008

©2014 American Association for Cancer Research.

www.aacrjournals.org

Downloaded from cancerres.aacrjournals.org on April 13, 2017. © 2014 American Association for Cancer Research.
histologies of the specimens included biphasic (three tumors), desmoplastic (one tumor), and epithelial (18 tumors). Each subject was properly informed before recruitment for the study, according to the regulations of the NYU ethics review boards. Before whole-exome sequencing, germline SNP analysis was performed on tumor–blood pairs to confirm their identities; all tumors matched their normal pairs.

Massively parallel sequencing
For whole-exome sequencing, genomic DNAs from tumor–blood pairs were fragmented and subjected to whole-exome capture with the SureSelect Human All Exon 50-Mb Kit (Agilent Technologies) following standard protocols. Exome capture libraries were then sequenced on the Illumina HiSeq platform according to standard protocols (8).

Somatic mutation calling
Mutation analysis of exome sequencing data was performed as described previously (8). Briefly, basic read alignment and sequence quality control were done by using the Picard and Firehose pipelines at the Broad Institute.Mapped genomes were processed by the Broad Firehose pipeline to perform additional quality control, variant calling (MuTect algorithm; ref. 9), and mutation significance analysis (MutSig algorithm; ref. 10).

Validation of recurrently mutated genes in additional MPMs
The exonic regions of recurrently mutated genes PIK3C2B, RDX, and TAO1 were determined using the RefSeq transcript data retrieved from the UCSC Genome Browser, yielding 67 regions. Specific primers were designed using Primer3, to amplify the exonic regions with a buffer of at least 100 bps on both ends with a total length of 500 to 1,500 bps (Supplementary Table S2).

The regions of interest were amplified by PCR reactions using a touchdown PCR protocol with Kapa HiFi HotStart polymerase (Kapa Biosystems). The touchdown PCR method consisted of: 1 cycle of 95°C for 5 minutes; 3 cycles of 95°C for 30 seconds, 64°C for 15 seconds, 72°C for 30 seconds; 3 cycles of 95°C for 30 seconds, 62°C for 15 seconds, 72°C for 30 seconds; 3 cycles of 95°C for 30 seconds, 60°C for 15 seconds, 72°C for 30 seconds; 37 cycles of 95°C for 30 seconds, 58°C for 15 seconds, 72°C for 30 seconds; 1 cycle of 70°C for 5 minutes. All amplicons for each sample were pooled and used to generate libraries with unique indices using the Nextera DNA Sample Preparation Kit (Illumina), which generates smaller fragments randomly within each amplicon.

Sequencing of amplicons was carried out on an Illumina MiSeq instrument with 2 × 150 bps paired-end reads. The average coverage of each gene was approximately ×900. Somatic mutations in each gene were detected by the same methods as described in discovery stage.

Copy-number profiling
We analyzed copy number in the whole-exome sequencing by generating copy-number segments based on published methods (11), and then applied GISTIC2 (12) to identify focal regions of somatic copy-number alterations (SCNA) as described previously. A significant SCNA region was defined as having a Q value threshold of 0.25 from the permutation-derived null distribution. The output of focal GISTIC2 was used to define the key peaks of amplifications and deletions.

Pathway enrichment analysis
We performed pathway enrichment analysis using WebGestalt (http://bioinfo.vanderbilt.edu/webgestalt/) by examining the distribution of genes with somatic mutations or copy-number changes within the KEGG database. The significance of mutation enrichment was determined by a hypergeometric test and was adjusted for multiple testing with the Benjamini–Hochberg FDR.

Results
Landscape of somatic mutations in MPM
We captured and sequenced the exomes from paired tumor and blood DNA samples from 22 MPM cases (Supplementary Table S1). On average, 3.5 Gb of high-quality sequencing data were generated for each sample, resulting in coverage of targeted exome regions to a mean depth of 89× or greater (Table 1). More than 88.2% of the targeted regions were covered sufficiently for confident variant calling (≥14 reads in tumors and ≥8 reads in normal samples). By statistical comparison of paired tumor/blood sequence data (see Materials and Methods), mutation analysis identified a total of 517 somatic mutations (Supplementary Table S3), including 134 synonymous mutations, 295 missense mutations, 25 nonsense mutations, 25 splice-site changes, and 38 small coding insertions/deletions (indels). Comparison of somatic mutations from this study with the COSMIC database (13) yielded 53 mutations across 46 previously reported mutated genes. Of 295 missense mutations, 136 were predicted to have a high probability of being deleterious by PolyPhen (14).

We detected a mean of 23 somatic mutations per tumor (range 2–51) corresponding to an average of 0.79 mutations per megabase (range, 0.07–1.71), a considerably smaller number compared with other types of malignant tumors (Supplementary Table S4). Three of the 22 tumors, NYU589, NYU1396, and NYU269, harbored small numbers of somatic mutations with two, and three mutations, respectively; despite high estimated histologic purity, we were unable to determine sample purity by DNA analysis (15), and therefore are uncertain whether the low mutation rate represents low tumor purity or a biologic feature of these tumors. The mutation spectrum in the 22 MPMs was dominated by C:G→T:A transitions (Supplementary Fig. S1), as noted in several other adult cancers (16).

Significantly mutated genes in mesothelioma
Protein-altering mutations occurred across 490 genes, of which 97% (477/490) were mutated in only a single individual. To identify significantly mutated genes that were likely to be implicated in MPM tumorigenesis, we considered genes that were mutated in at least two tumors and were significantly more mutated than the background mutation rate (Q value < 0.1), as calculated by MutSig (10). We identified three significantly mutated genes (BAP1, NF2, and CUL1) meeting these two criteria (Fig. 1) and 10 additional genes that were recurrently but not significantly mutated (Supplementary Table S3).

BAP1 has the highest prevalence of protein-altering mutations identified in this MPM cohort (8/22 tumors = 36%), consistent with earlier MPM-sequencing analyses (4). BAP1 encodes a nucleolar ubiquitin carboxyterminal hydrolase (17). We found eight mutations in BAP1, seven of which (three nonsense mutations, three frame-shift indels, and 1 splice site change) were predicted to truncate the encoded protein (P < 0.001; Fig. 2A), consistent with
Figure 1. Significantly mutated genes in MPM. Significantly mutated genes (Q value < 0.1) identified by exome sequencing are sorted vertically by a Q value. Top bar graph, somatic mutation frequency for 22 MPMs. The middle matrix, BAP1, NF2, and CUL1 somatic mutation status with left-adjacent bar graph, indicating the total number of mutations and percentage of frequency for each gene. The boxplots below the matrix represent the distributions of allelic fractions observed in each tumor. Bottom plot, each tumor’s somatic mutation spectrum as a percentage of all mutations (right axis).
component of the SCF (SKP1–CUL1–F-box protein) E3 ubiquitin ligase complex and plays an important role in protein degradation and protein ubiquitination. All mutations in CUL1 were predicted to affect protein function by PolyPhen (14) and one missense mutation (E485K) was also identified in the COSMIC database.

The other recurrently mutated genes

In this study, 13 genes were recurrently mutated in two or more MPMs (Supplementary Table S3). Besides BAP1, NF2, and CUL1, four recurrently mutated genes (RDX, TP53, PIK3C2B, and TAOK1) harbored higher mutation rates than the background with a significance threshold at a moderately stringent level ($P < 0.01$; Supplementary Table S5). TP53 harbored somatic mutations in two cases, consistent with previously reported frequencies in MPM (4). To our knowledge, the three novel recurrently mutated genes (RDX, PIK3C2B, and TAOK1) have not been identified in MPM, although mutations in these genes have been reported in many cancer types in the COSMIC database (13). RDX, encoding a cytoskeletal protein that may be important in linking actin to the plasma membrane, is subject to two somatic mutations in this case series (S497R and Q48P). PIK3C2B encodes a protein that belongs to the PI3K family and is somatically mutated in two cases. TAOK1, encoding a serine/threonine protein kinase involved in various processes such as the MAPK signaling pathway and mitotic anaphase, was also mutated in two cases.

To further evaluate the mutation prevalence of RDX, PIK3C2B, and TAOK1 in MPM, we determined the coding sequencing of the three genes in an additional 10 tumor-normal pairs from patients with MPM and 9 mesothelioma cell lines by high-throughput deep sequencing. However, no somatic mutations in these three genes were found in the 10 tumor samples. All mutations detected in mesothelioma cell lines were known SNPs and none of them were present in the COSMIC database.

Copy-number alterations in MPM

We profiled the 22 MPMs for SCNAs using exome sequencing data using previously described methods (11) and found predominant loss of 22p and 22q harboring NF2 (Supplementary Fig. S2), a frequent genetic alteration in MPM (3). We also applied GISTIC2 (12) analysis to exome sequencing data to identify recurrent focal SCNAs, yielding six regions of focal deletion that encompass 505 genes (Supplementary Table S6). Among these regions, the most common focal deletion, altered in 10 tumors (45%), encompassed a region at 9p21 containing CDKN2A and CDKN2B (Supplementary Fig. S3), which is a widely reported genomic alteration in MPM (2). Consistent with our previous study (18), we found that MIR31 (microRNA 31), a microRNA gene located at 9p21 near CDKN2A, was frequently deleted in eight (36%) tumors (Supplementary Fig. S4).

Significantly altered pathways

To construct a comprehensive view of the common genetic alterations underlying the MPM, we performed integrative pathway analysis of the somatic mutations and focal SCNAs. We identified a number of pathways with recurrent alterations in MPM, including the cell cycle, MAPK, and Wnt signaling pathway (Supplementary Table S7). The cell cycle was the most frequently altered pathway, with alterations in 12 of the tumors (55%), including the alterations in CDKN2A/B, CUL1, and TP53. Genes within the MAPK and Wnt signaling pathways were altered in 11 and seven tumors, respectively.

Discussion

In this study, we analyzed somatic mutations in 22 MPMs compared with matched normal samples using whole-exome sequencing. We detected more than 500 somatic mutations and showed that 490 genes were somatically mutated in at least one of the tumors. Integrating mutation and SCNA data, we found that BAP1, NF2, and CDKN2A, known to be altered genes in MPM, were the three most frequently altered genes detected in our study (Fig. 3). We also found frequent genetic alterations in the cell cycle, MAPK signaling, and Wnt signaling pathway. The results are consistent with a previous study of MPM using both CGH arrays and Sanger sequencing (4).

The tumor-suppressor BAP1, one of several classes of deubiquitinating enzyme, may be involved in regulation of...
transcription, regulation of cell cycle and growth, and response to DNA damage and chromatin dynamics (17). Frequent mutations in BAP1 have been reported in many types of cancer, including uveal melanoma (19), renal cancer (20), and MPM (4). In the present study, nine tumors (41%) harbored somatic mutations or SCNAS in BAP1. We also found 11 genetic alterations in NF2, a component of the Hippo signaling pathway. A previous study reported 10 inactivating homozygous deletions or mutations of LATS2, another member of the Hippo pathway, among 45 MPM tumor tissues or cancer cell lines (21), although no mutations in LATS2 were found in our study. These studies suggest that the Hippo signaling pathway may be implicated in MPM tumorigenesis. However, no somatic mutations in genes of the Hippo pathway other than NF2 and LATS1 were found in our study (Supplementary Table S8).

We also found significant recurrence of mutations in CUL1, a gene in which mutations have not been previously reported in MPM, in two tumors. CUL1 encodes a core component of SCF E3 ubiquitin–protein ligase complexes that mediate the ubiquitination of proteins involved in cell-cycle progression, signal transduction, and transcription.

Finally, we used exome sequencing data to identify somatic CNAs and found six significant focal deletions. Among them, the deletion of 9p21 centered on CDKN2A was the most common event occurring in 10 of 22 MPMs. Moreover, seven of CDKN2A-deleted tumors also harbored genetic alterations in NF2 ($P = 0.19$). No somatic substitutions and coding indels were found in CDKN2A in our cohort.

Our study presents the first unbiased view of somatic mutations in MPM. However, one limitation of our study is that we have only analyzed a small number of tumors analyzed; thus, the sequencing of larger MPM cohorts will be necessary to determine a more comprehensive genomic landscape for MPM as well as a more precise prevalence frequency for the mutated genes detected in our analysis. In addition, further functional study for novel mutated genes such as CUL1 will be required to define their roles in MPM.

**Disclosure of Potential Conflicts of Interest**

M. Carbone has a pending patent application on Bap1 and provides consultation for mesothelioma expertise and diagnosis. M. Meyerson reports receiving a commercial research grant from Bayer, has ownership interest (including patents) in, and is a consultant/advisory board member for Foundation Medicine. No potential conflicts of interest were disclosed by the other authors.

**Authors’ Contributions**

Conception and design: J. Chmielecki, M. Carbone, M. Meyerson, H.I. Pass

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): A. Heguy, J. Chmielecki, M. Carbone, M. Meyerson

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): G. Guo, J. Chmielecki, A. Heguy, I. Dolgalev, M. Carbone, M. Meyerson

Writing, review, and/or revision of the manuscript: G. Guo, A. Heguy, M. Meyerson, H.I. Pass

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): C. Goparaju, S. Seepo, H.I. Pass

Study supervision: M. Meyerson, H.I. Pass

Other (project management): S. Seepo

**Acknowledgments**

The authors thank Cheng-Zhong Zhang and Andrew Cherniack for helpful discussions.

**Grant Support**

This work was supported by the Broad Institute through a grant from the National Human Genome Research Institute at the NIH (U54HG003067: Eric S. Lander, principal investigator), and by NCI EDRN 5U01CA111295-07 (H.I. Pass as principal investigator), the Stephen E. Banner Lung Cancer Foundation, and Belluck and Fox. J. Chmielecki was supported by an American Cancer Society AstraZeneca Postdoctoral Fellowship.

Received April 14, 2014; revised September 23, 2014; accepted October 28, 2014; published OnlineFirst December 8, 2014.

---

**References**


Whole-Exome Sequencing Reveals Frequent Genetic Alterations in \textit{BAP1, NF2, CDKN2A, and CUL1} in Malignant Pleural Mesothelioma

Guangwu Guo, Juliann Chmielecki, Chandra Goparaju, et al.

\textit{Cancer Res} Published OnlineFirst December 8, 2014.