Molecular Classification of Malignant Pleural Mesothelioma: Identification of a Poor Prognosis Subgroup Linked to the Epithelial-to-Mesenchymal Transition

Aurélien de Reyniès1, Marie-Claude Jaurand2,3, Annie Renier2,3, Gabrielle Couchy2,3, Ilir Hysi6, Nabila Elarouci1, Françoise Galateau-Salle8,9, Marie-Christine Copin7, Paul Hofman10, Aurélie Cazes4, Pascal Andujar11,12, Sandrine Imbeaud2,3, Fabien Petel1, Jean-Claude Pairon11,12, Françoise Le Pimpec-Barthes2,3,5, Jessica Zucman-Rossi2,3, and Didier Jean2,3

Abstract

Purpose: Despite research efforts to develop more effective diagnostic and therapeutic approaches, malignant pleural mesothelioma (MPM) prognosis remains poor. The assessment of tumor response to therapy can be improved by a deeper phenotypical classification of the tumor, with emphasis on its clinico-biological heterogeneity. The identification of molecular profiles is a powerful approach to better define MPM subclasses and targeted therapies.

Experimental Design: Molecular subclasses were defined by transcriptomic microarray on 38 primary MPM cultures. A three-gene predictor, identified by quantitative reverse transcription PCR, was used to classify an independent series of 108 frozen tumor samples. Gene mutations were determined in BAP1, CDKN2A, CDKN2B, NF2, and TP53. Epithelial-to-mesenchymal transition (EMT) markers were studied at the mRNA and protein levels.

Results: Unsupervised hierarchical clustering on transcriptomic data defined two robust MPM subgroups (C1 and C2), closely related to prognosis and partly to histologic subtypes. All sarcomatoid/desmoplastic MPM were included in the C2 subgroup. Epithelioid MPM were found in both subgroups, with a worse survival prognosis in the C2 subgroup. This classification and its association with histologic subtypes and survival were validated in our independent series using the three-gene predictor. Similar subgroups were found after classification of other MPM series from transcriptomic public datasets. C1 subgroup exhibited more frequent BAP1 alterations. Pathway analysis revealed that EMT was differentially regulated between MPM subgroups. C2 subgroup is characterized by a mesenchymal phenotype.

Conclusions: A robust classification of MPM that defines two subgroups of epithelioid MPM, characterized by different molecular profiles, gene alterations, and survival outcomes, was established. Clin Cancer Res; 20(5); 1323–34. ©2014 AACR.
Malignant pleural mesothelioma (MPM) is an extremely aggressive tumor with no known effective therapeutic option. Promising novel target therapies require better knowledge of molecular alterations and of the clino-biological heterogeneity of MPM tumors. Growing evidence suggests that a given cancer is a heterogeneous patho-physiological disease characterized by distinct molecular alterations, and MPM is no exception. Molecular classification with potential diagnostic, prognostic, and therapeutic interests is now available for several cancers but not for MPM. The only classification of MPM linked to survival outcome is based on histologic subtypes and does not reflect the biological heterogeneity, especially of epithelioid MPM, the most frequent subtype. In this study, a robust molecular classification consisting of two subgroups characterized by different molecular profiles, gene alterations, and survival outcomes is defined. This finding is a critical endpoint for translational research and may lead to the improved management of this heterogeneous tumor.

Translational Relevance
Malignant pleural mesothelioma (MPM) is an extremely aggressive tumor with no known effective therapeutic option. Promising novel target therapies require better knowledge of molecular alterations and of the clino-biological heterogeneity of MPM tumors. Growing evidence suggests that a given cancer is a heterogeneous patho-physiological disease characterized by distinct molecular alterations, and MPM is no exception. Molecular classification with potential diagnostic, prognostic, and therapeutic interests is now available for several cancers but not for MPM. The only classification of MPM linked to survival outcome is based on histologic subtypes and does not reflect the biological heterogeneity, especially of epithelioid MPM, the most frequent subtype. In this study, a robust molecular classification consisting of two subgroups characterized by different molecular profiles, gene alterations, and survival outcomes is defined. This finding is a critical endpoint for translational research and may lead to the improved management of this heterogeneous tumor.

abnormalities, gene mutations, epigenetic alterations, and gene expression (3–6). MPM is characterized by numerous chromosomal abnormalities involving alterations in both chromosome number and structure (7). Genetic alterations were identified in several tumor suppressor genes (TSG), but not in oncogenes. Mutations in BAP1 (BRCA1-associated protein-1), CDKN2A (cyclin-dependent kinase inhibitor 2A), CDKN2B (cyclin-dependent kinase inhibitor 2B), and NF2 (neurofibromin 2) have been reported in a high percentage of MPM, and TP53 (tumor protein p53) has been found mutated at a lower rate in comparison with other human cancers (6). High-throughput methylation analyses and large-scale microRNA (miRNA) studies also revealed heterogeneity at the epigenetic level, and transcriptomic studies underlined differential gene expression changes and signaling pathway deregulation in MPM (6). Furthermore, comparison between clinical trials indicated that this disease has variable prognoses (8) and different sensitivities to various anticancer drugs (9). The biological diversity of MPM is also suggested by the difficulty to define a single specific biomarker (10).

MPM is currently classified according to 3 main histologic subtypes: epithelioid, sarcomatoid, and biphasic. Although multiple patterns of each histologic subtype have been described, a molecular classification of MPM would make it possible to better describe their biological diversity (11). To date, only a few studies have proposed a classification based on molecular data and linked to the clinical characteristics of MPM. Classifications based on methylation profiles have defined MPM subgroups characterized by different clinical outcomes (12, 13). However, classifications based on transcriptomic data that make it possible to define gene expression and pathway profiles were not related to the prognosis. The two previous transcriptomic classifications only led to separate MPM cell lines and MPM tumor samples according to their histologic subtypes (14). Specific profiles of miRNA expression also defined MPM subgroups related to histologic subtypes (15, 16).

The aim of this study was to establish a transcriptomic classification of MPM using a large collection of MPM consisting of cells in culture and frozen tumor samples, in order to better characterize the biologic diversity of MPM and to develop a more in-depth understanding of pathway deregulation in molecular subgroups of MPM. These efforts are needed to identify new therapeutic approaches, in particular, biology-driven targeted therapies.

This is the first study that provides evidence that transcriptomic classification of MPM and especially of epithelioid MPM into 2 subgroups has prognostic value. Our molecular analysis made inroads into mesothelial carcinogenesis mechanisms by identifying distinct profiles of mutation and differential progression in the epithelial-to-mesenchymal transition (EMT) between the 2 molecular subgroups.

Materials and Methods

Patients
Human normal and tumor specimens were obtained with the informed consent of the patients. Tumor collection included 67 MPM in culture consisting of a discovery series (38 cases) and an extension series (29 cases), and 108 MPM frozen tumor samples. Exposure data of MPM patients to asbestos was collected by interviewer-administered questionnaires and was estimated by taking past occupational, domestic, and environmental exposure to asbestos into account. Epithelioid diagnosis was certified by histologic analysis using consensus guidelines established by the International Mesothelioma Interest Group (http://imig.org; ref. 17). Clinico-pathological and epidemiologic data of patients are reported in Table 1.

Mesothelioma cells in culture
Primary MPM cultures were established in our laboratory from surgical resection, pleural biopsies, or malignant pleural fluid of confirmed MPM cases, obtained from several French hospitals: Centre Hospitalier Régional Universitaire (CHRU) de Brest; Centre Hospitalier Universitaire (CHU) de Caen; Centre Hospitalier Intercommunal, Créteil; CHU de Marseille; Hôpital Européen Georges Pompidou (HEGP), Paris; CHU de Toulouse. These MPM in culture are primary lines and therefore cannot be authenticated, but all genetic and transcriptomic analyses were carried out with cultures used at low-passage number (<12 passages from its initial stocks). The first established MPM in culture were characterized by immunohistochemistry (18). Normal mesothelial cells were cultured from surgical resection of blebs from patients with spontaneous pneumothoraces. Cells were grown in RPMI 1640 medium supplemented with penicillin, streptomycin, Glutamax, and 10% fetal calf serum (Life Technologies; ref. 19).
and RNA, respectively. Ribogreen fluorescence assays (Life Technologies) for DNA concentrations were determined using Hoechst 33258 or lent 2100 Bioanalyzer (Agilent Technologies). Nucleic acid microarray analysis, electrophoregram profiles on an Agitometer (Thermo Fisher Scientific), agarose gels and, for the samples was assessed using a NanoDrop spectrophotometer and stainless steel beads (Qiagen). The quality of manufacturer’s protocol, after homogenization using the TRizol reagent (Life Technologies), according to the manufacturer’s instructions, including proteinase K (Qiagen) digestion and isopropanol precipitation. RNA was prepared using the RNA extraction procedure including proteinase K (Qiagen) digestion and isopropanol precipitation. RNA was prepared using the RNase-free DNase set (Qiagen) and purified by ExoSAP-IT (USB Corporation Cleveland) treatment. Primers used are listed in the Supplementary Table S3 for BAP1 and CDKN2B genes or were previously published for the 2 other genes (21, 22). The PCR products were analyzed and quantified using the Qiavelo genetic analyzer (Qiagen). Exon deletions were detected at this step and confirmed in a second independent PCR multiplex amplification. Fluorescent-based automated cycle sequencing was performed by the BigDye Terminator method (Life Technologies) using a multicapillary sequencer (ABI 3130 Genetic Analyzer). Mutations were detected using Sequencher software (Gene Codes Corp.) and verified by independent amplification and sequencing.

**Pangenomic expression profiling**

Expression profiles of primary cultures from 38 MPM and 3 normal pleural mesotheliums were obtained using Affymetrix HG-U133-plus-2.0 microarrays and deposited in the ArrayExpress repository (E-MTAB-1719). In addition, Affymetrix expression profiles of 2 public datasets (“Bott” and “Gordon” cohorts) of MPM samples from the GEO repository (http://www.ncbi.nlm.nih.gov/geo/): GSE29354 (20) and GSE2549 (14) were analyzed.

**Real-time quantitative reverse transcription PCR analysis**

Predefined TaqMan probe and primer sets for target genes were chosen from the Life Technologies database (http://www.appliedbiosystems.com) and factory loaded into 384-well microfluidic cards (Life Technologies). Single-stranded cDNA was prepared from 1.5 μg of total RNA from cells in culture and tumor samples using random primers and the High Capacity cDNA Reverse Transcription Kit (Life Technologies), according to the manufacturer’s instructions. Quantitative PCR was performed with microfluidic cards using the ABI PRISM 7900HT System (Life Technologies). Gene expression was normalized to internal control ribosomal 18S relative to the mean expression of the corresponding gene in three normal mesothelial cells in culture, according to the 2^−ΔΔCt method.

**Mutation analysis**

Mutations in BAP1, CDKN2A, CDKN2B, NF2, and TP53 genes in cells in culture were screened by Sanger sequencing. All coding exons for TSGs were amplified from genomic DNA by PCR using a Qiagen Multiplex Master Mix Kit (Qiagen) and purified by ExoSAP-IT (USB Corporation Cleveland) treatment. Primers used are listed in the Supplementary Table S3 for BAP1 and CDKN2B genes or were previously published for the 2 other genes (21, 22). The PCR products were analyzed and quantified using the Qiavelo genetic analyzer (Qiagen). Exon deletions were detected at this step and confirmed in a second independent PCR multiplex amplification. Fluorescent-based automated cycle sequencing was performed by the BigDye Terminator method (Life Technologies) using a multicapillary sequencer (ABI 3130 Genetic Analyzer). Mutations were detected using Sequencher software (Gene Codes Corp.) and verified by independent amplification and sequencing.

Genetic alterations in frozen tumor samples were determined by multiplex ligation-dependent probe amplification (MLPA). The MLPA was carried out following the manufacturer’s instructions using SALSA MLPA P417 and MEO24 Kits (MRc-Holland) for BAP1 and CDKN2A/CDKN2B loci, respectively. Products of the MLPA reaction were separated on the ABI 3130 Genetic Analyzer. GeneScan 500 LIZ (Life Technologies) was used as the internal size standard. Data analysis was performed using Coffalyser software (MRc-Holland). Four negative reference samples consisting of genomic DNA prepared from human whole blood (Promega) and normal liver tissues were included in each MLPA run. The reliability of the assay was determined

<table>
<thead>
<tr>
<th>Gender, n (%)</th>
<th>Cells in culture (n = 67)</th>
<th>Tumor samples (n = 108)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>53 (79)</td>
<td>79 (73)</td>
</tr>
<tr>
<td>Female</td>
<td>14 (21)</td>
<td>29 (27)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>Median ± SD</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td>65 ± 11</td>
<td>37–91</td>
</tr>
<tr>
<td>Histology, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epithelioid</td>
<td>48 (75)</td>
<td>80 (78)</td>
</tr>
<tr>
<td>Biphasic</td>
<td>9 (14)</td>
<td>13 (13)</td>
</tr>
<tr>
<td>Sarcomatoid</td>
<td>6 (9)</td>
<td>7 (7)</td>
</tr>
<tr>
<td>Desmoplastic</td>
<td>1 (2)</td>
<td>3 (3)</td>
</tr>
<tr>
<td>Lymphohistiocytoid</td>
<td></td>
<td>0 (0)</td>
</tr>
<tr>
<td>Asbestos exposure, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exposed</td>
<td>51 (80)</td>
<td>67 (71)</td>
</tr>
<tr>
<td>Nonexposed</td>
<td>13 (20)</td>
<td>27 (29)</td>
</tr>
<tr>
<td>Stage IMIG, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>nd</td>
<td>4 (5)</td>
</tr>
<tr>
<td>II</td>
<td>nd</td>
<td>13 (17)</td>
</tr>
<tr>
<td>III</td>
<td>nd</td>
<td>30 (39)</td>
</tr>
<tr>
<td>IV</td>
<td>nd</td>
<td>29 (38)</td>
</tr>
<tr>
<td>Survival (months)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>10.2</td>
<td>16.9</td>
</tr>
<tr>
<td>Range</td>
<td>0.3–118.8</td>
<td>0.1–69.6</td>
</tr>
</tbody>
</table>

Abbreviations: IMIG, International Mesothelioma Interest Group; nd, not determined.

**RNA and DNA extraction**

Genomic DNA was extracted from frozen MPM tumor samples and primary cell cultures using a standard extraction procedure including proteinase K (Qiagen) digestion and isopropanol precipitation. RNA was prepared using the TRizol reagent (Life Technologies), according to the manufacturer’s protocol, after homogenization using the TissueLyser and stainless steel beads (Qiagen). The quality of the samples was assessed using a NanoDrop spectrophotometer (Thermo Fisher Scientific), agarose gels and, for microarray analysis, electrophoregram profiles on an Agilent 2100 Bioanalyzer (Agilent Technologies). Nucleic acid concentrations were determined using Hoechst 33258 or Ribogreen fluorescence assays (Life Technologies) for DNA and RNA, respectively.
by testing MPM cells in culture well characterized by DNA sequencing. Peak height ratios lower than 0.7 were considered as indicative of deletion. All MLPA results were reproduced at least 2 times.

Western blots

Confluent MPM cells in culture were washed in PBS and cell extracts were prepared in radioimmunoprecipitation assay buffer (Millipore) containing complete protease inhibitor cocktail (Roche Diagnostics). Equal amounts of protein, as determined by the Bio-Rad DC protein assay (Bio-Rad), were separated by SDS-PAGE, transferred to polyvinylidene difluoride membrane (Millipore), and immunoblotted with anti-E-cadherin (CDH1; clone 24E10 1/1,000 dilution; Cell Signaling Technology) and anti-glyceraldehyde 3-phosphate dehydrogenase (GAPDH; clone 14C10 1/5,000 dilution; Cell Signaling Technology) antibodies. Bound antibodies were detected with HRP-conjugated secondary antibodies (Santa Cruz Biotechnology) and ECL reagents (GE Healthcare). The relative quantification of bands was determined using Quantity One Quantitation software (Bio-Rad).

Immunohistochemistry

Immunohistochemistry for E-cadherin was performed on 5-μm deparaffinized sections using an automated single-staining procedure (Benchmark Ultra; Ventana Medical Systems). Briefly, the mouse monoclonal anti-E-cadherin antibody (clone 4A2C7, 1/50 dilution) was added to the slides for 32 minutes after pretreatment with UltraC1 buffer (Ventana Medical Systems) for 36 minutes. The UltraView Universal DAB Detection Kit (Ventana Medical Systems) was then used according to the manufacturer’s instructions.

Statistical analyses

Affymetrix HG-U133 Plus 2.0 expression profiles of 38 primary MPM cultures (discovery cohort) and 3 normal mesothelial cells were normalized in batch using the RMA algorithm (23). Each of the 2 public datasets of MPM samples (GSE29354, GSE25499) was independently normalized in batch with the RMA algorithm. Consensus clustering analysis of the 38 MPM expression profiles from the discovery series and the public series was performed as already reported (24), yielding 2 robust clusters, C1 and C2. Briefly, 24 dendrograms were obtained by hierarchical clustering using 3 different linkages (ward, complete, average) and 8 lists of genes corresponding to the top 40% to 0.5% (8 different thresholds) most varying genes (assessed via a robust coefficient of variation). Each of the 24 dendrograms yielded a raw partition in 2 clusters; the 24 raw partitions in 2 clusters were then used to build a consensus dendrogram that itself was divided into 2 clusters, thus yielding a consensus partition in 2 clusters. Principal component analysis of these sample profiles was also performed. The first 3 principal components showed a clear separation between C1 and C2. To identify genes differentially expressed between C1 and C2 and to make it possible to discriminate C1 and C2 samples, moderate t tests (limma R package) and the area under the curve (AUC) criterion (PresenceAbsence R package) were used. Genes (n = 40) selected based on these criteria (AUC and t test), as well as their potential involvement in mesothelial carcinogenesis, were measured by quantitative reverse transcription PCR (qRT-PCR) on 67 MPM cultures (i.e., discovery + extension cohorts containing 29 additional MPM cultures) and 108 tumor samples (validation series). To assign each sample to either a C1 or C2 molecular subtype, a predictor was trained using these qRT-PCR measurements on the discovery cohort. This predictor was obtained via a forward stepwise procedure, optimizing the success rate criterion with 3 algorithms [DLDA, DQDA (R package sma), PAM (R package pamR; ref. 25). The predictor is defined by a majority vote across these 3 algorithms and is based on 3 genes [periplakin (PPL), uroplakin 3B (UPK3B), tissue factor pathway inhibitor (TFPI)]. It was applied to the remaining samples (29 MPM cultures + 108 tumor samples). Pathway analysis based on sample group comparisons was performed using a consensus of 4 methods [GSA (26): R package GSA; globaltest (27): R package globaltest; SAM- GS (28): original R code; Tukey (29): original R code]. Given 2 sample groups to be compared and a pathway of interest, each method will yield a P value: the lower the P value is, the more the genes from the pathway are differentially expressed between the 2 sample groups. To aggregate the results of the 4 methods, the list of pathways for each method is first sorted and the mean rank across the 4 methods is then calculated; the final order is based on this mean rank. Pathways and related genes were retrieved from 2 repositories: Gene Ontology (GO; www.geneontology.org) and KEGG ([www.genome.jp/kegg/pathway.html]). Moderate t tests and Mann–Whitney tests were used for differential expression analysis. Survival analysis was performed on overall survival, using Kaplan–Meier estimates and log-rank tests.

Results

Major clinico-pathological and epidemiologic characteristics of MPM in culture and tumor samples are summarized in Table 1. Histologic types were defined by the MESOPATH French National group, responsible for MPM diagnosis certification in France (17, 30). As expected, epithelioid was the most frequent histologic type (76%), and most of the patients were male (75%) and had been exposed to asbestos (75%). To further characterize MPM, mutation analyses were also carried out in genes involved in mesothelial carcinogenesis (CDKN2A, CDKN2B, BAP1, NF2) and in more common cancer-related genes (TP53).

Consensus clustering of the expression profiles of MPM in culture identifies 2 robust groups associated with histology and the BAP1 mutation

Molecular classification of MPM was constructed and validated according to a flowchart detailed in Supplementary Fig. S1. Pangenomic expression profiles of 38 MPM in culture (the discovery series) were analyzed by consensus clustering, which yielded 2 robust clusters, C1 and C2...
Fig. 1A and Supplementary Fig. S2). This classification was partly related to histologic types. Sarcomatoid and desmoplastic MPM were all found in C2, whereas epithelioid MPM were distributed in both C1 and C2. No significant difference was found between C1 and C2 for the TSGs screened for mutation. However, BAP1 and CDKN2A tend to be more
frequently mutated in C1 (Fisher exact test, $P = 0.09$ and 0.14, respectively; Fig. 1A). Of note, the analysis of the comparative genomic hybridization (CGH) array profiles of MPM in culture (7) revealed that the 3p21 chromosomal region containing the $BAP1$ locus was more frequently lost in C1 than in C2 (87% and 37%, respectively; Fisher exact test, $P < 0.01$; Supplementary Fig. S3).

To validate the molecular classification, we first used a supervised approach. We trained a C1/C2 predictor on the discovery series using qRT-PCR measurements of 40 genes found to be discriminant between C1 and C2 and obtained a predictor based on 3 genes—$PPL$, $UPK3B$, and $TFPI$—showing a 100% success rate in the discovery set (Supplementary Table S1 and Fig. S4). C1/C2 memberships were then predicted in the extension series of MPM in culture ($n = 29$) and in the validation series, including 108 MPM tumor samples (Supplementary Table S1). The same histologic subtype distribution was observed in MPM in culture and in tumor samples, with sarcomatoid MPM being exclusively found in the C2 subgroup, whereas epithelioid MPM were divided between C1 and C2 (Fisher exact test, $P < 0.04$; Fig. 1B). Biphasic MPM were also distributed between C1 and C2, in agreement with their bi-component feature. The more frequent mutation of $BAP1$ in C1 was confirmed (Fisher exact test, $P < 0.04$; Fig. 1B). Of note, these characteristics were independently observed both in MPM in culture and MPM tumor samples.

To further confirm the C1/C2 molecular subgroups, we used an unsupervised approach. We performed consensus clustering independently in 2 public MPM pangenomic expression profile series (refs. 14 and 20; Supplementary Fig. S5A–S5D). The consensus classifications derived from the Bott and Gordon series were then compared with the C1/C2 classification in terms of gene expression signatures and showed agreement between the classifications derived from our 3 MPM series (Supplementary Fig. S5E). The public series subgroups were referred to as "C1-like" and "C2-like." A similar histologic subtype distribution was found in these 2 independent datasets (Supplementary Fig. S6A).

MPM tumors with a C2 gene expression profile show a worse prognosis, independently of their histology

In the discovery cohort, C2-related patients showed a worse prognosis than those of C1 (overall survival log-rank test, $P = 0.02$), which is coherent with the known poor prognosis of sarcomatoid and biphasic MPM (Fig. 2A).

![Figure 2](https://example.com/figure2.png)

**Figure 2.** Overall survival differences between MPM molecular subgroups. Overall survival curves of patients related to C1 (blue) and C2 (red) molecular subgroups, in the discovery series (A and B) and the validation series (C and D), respectively, are calculated for all MPM (A and C) and epithelioid MPM (B and D). Differences in survival between molecular subgroups are assessed by log-rank tests. Kaplan–Meier survival analyses for all MPM and for epithelioid MPM showed that C1 and C2 subgroups have statistically significant different prognoses.
Interestingly, when only epithelioid MPM were taken into consideration, the overall survival rate of patients was also lower in C2 than in C1 (logrank test, \( P < 0.01 \); Fig. 2B). Both observations were confirmed in the validation series (logrank test, \( P < 0.01 \) and 0.05, respectively; Fig. 2C and D). Cox univariate and multivariate analyses for overall survival were performed on all MPM samples from the discovery series; the C1/C2 classification remained significant in the multivariate analysis (Table 2). In agreement, in the Gordon and Bott series, MPM patients with a C2-like expression profile showed a poorer prognosis (log-rank test, \( P < 0.01 \)), even when considering only epithelioid MPM (log-rank test, \( P = 0.03 \); Supplementary Fig. S6B).

### Pathway analysis reveals EMT in the poor prognosis C2 molecular subtype

To identify pathways specifically deregulated in C1 and C2 molecular subtypes, we performed an extensive pathway analysis in the discovery cohort, combining the results from 4 algorithms and screening both the Gene Ontology terms and KEGG pathways. The Gene Ontology terms, “cell migration” (GO:0016477) and “EMT” (GO:0001837), were found among the most deregulated in the pathway ranking; in agreement, most deregulated KEGG pathways included those related to cell adhesion, junctions, and extracellular matrices (Supplementary Table S2).

To continue with the comparison in the Bott and Gordon series, the same pathway analysis was performed independently in these series by comparing C2- and C1-like MPM, yielding results similar to those of the discovery cohort (Supplementary Table S2). By aggregating pathway analyses from the 3 series, the Gene Ontology terms, “cell migration” and “EMT” were ranked in positions 3 and 12, respectively, among more than 13,000 Gene Ontology terms. Similarly, pathways related to cell adhesion, junctions, and extracellular matrices were in the top ranks (<20) among the KEGG pathways. Interestingly, the TGF-β signaling pathway (KEGG hsa04350), known to induce EMT (31), was also at top ranked.

To further characterize the EMT in C1 and C2 subgroups, several sources of positive and negative EMT markers described in relevant publications were assembled (ref. 32–34; list given in Supplementary Table S2). In our discovery series, most of the positive EMT markers were found to be upregulated in C2 compared with C1, whereas most of the negative EMT markers were found to be downregulated in C2 compared with C1, supporting the hypothesis that the C2 subgroup is characterized by a mesenchymal phenotype (Fig. 3A). The same results were observed in the Bott and Gordon series when comparing C2- and C1-like MPM (Supplementary Fig. S7A and S7B).

The differential expression of EMT markers was validated using qRT-PCR on our collection of MPM in culture and in tumor samples (Fig. 3B). Genes upregulated during EMT, such as the adhesion molecules POSTN and VCAN or the transcription factors SNAI2, TCF4, and HMG2, had statistically significant higher expression in C2 than in C1 (Mann–Whitney test, \( P < 0.01 \)). In contrast, genes downregulated during EMT, such as CDH1 and CDH3, had statistically significant lower expression in C2 than in C1 (Mann–Whitney test, \( P < 0.03 \)). Consistent results between C2 and C1 were obtained when restricting the analysis to epithelioid MPM (Supplementary Fig. S7C and S7D).

Protein expression of CDH1, encoding E-cadherin, was studied by Western blot analysis in MPM in culture and by immunohistochemistry in MPM tumor samples. Figure 4A shows representative Western blots in 5 C1 epithelioid MPM and 3 C2 epithelioid MPM. Relative quantification...
of E-cadherin band intensity in 14 C1 epithelioid MPM and 11 C2 epithelioid MPM demonstrated higher expression in C1 MPM than in C2 MPM (Fig. 4B). Immunohistochemistry for E-cadherin performed on epithelioid and biphasic MPM of the C1 and C2 subgroups showed plasma membrane staining in cells of epithelioid morphology (Fig. 4C–F). This staining was stronger in C1 (Fig. 4C and E) than in C2 (Fig. 4D and F).

Discussion

MPM is a tumor with a terrible prognosis because of its resistance to conventional anticancer therapies (35). Recent reviews of targeted therapies in MPM (36, 37) concluded that none are likely to become standard yet. Of note, none of the studies referred to was biology driven. Indeed, efforts are still needed to provide better treatment for patients suffering from this cancer. In this study, as a first step toward the identification and validation of appropriate targeted therapies for MPM, our objective was to better characterize the biological diversity of MPM and to better understand the pathway deregulation in tumor molecular subgroups.

Most MPM transcriptomic studies reported in the literature have used tumor samples (6). In this study, our discovery cohort was based on MPM in culture, which has the considerable advantage of focusing on the tumor cells, without interference from the tumor microenvironment. To date, the largest transcriptomic studies by microarray on MPM in culture used 10 and 13 independent samples (38, 39). In our study, we used 67 different MPM in culture and 108 frozen tumor samples to validate the classification; this makes our study one of the largest in terms of MPM samples. Furthermore, previous transcriptomic studies mainly focused on the identification of MPM biomarkers to differentiate MPM from normal tissue or from other tumors. Only 2 studies classified MPM into subgroups and found correlations only with histologic types (14, 38). One of the most important points of our classification was the separation of epithelioid MPM into 2 molecular subtypes, C1 and C2, with different prognoses.

We identified a predictor based on 3 genes, PPL, UPK3B, and TFPI, which makes it possible to determine the molecular subtypes of MPM. To our knowledge, PPL, a member of the adhesion molecule plakin family, was not previously described in MPM. UPK3B encodes a plasmic transmembrane protein, originally considered to be specific to the urothelial lineage before also being reported as a marker of the mesothelial lineage (40). It was previously shown that UPK3B showed heterogeneous expression between epithelioid MPM cell lines (41). TFPI, a protease inhibitor that regulates blood coagulation and could be involved in the procoagulant response in the pathogenesis of pleural injury, was also shown to have a variable expression in MPM tumors (42, 43).

In the past, 3 groups proposed prognostic predictors based on gene expression profiles. A microarray-based predictor composed of 35 probe sets corresponding to 29 genes was developed to predict 1-year survival from operation (40). A 27-gene classifier made it possible to distinguish 2 MPM tumor groups with different survival rates (44). A 4-gene predictor was found to differentiate patients on the basis of postsurgical outcome in several independent retrospective MPM cohorts, and to predict overall survival in a prospective clinical trial of patients undergoing surgery for MPM (45–47). In contrast with these studies, our 3-gene predictor was not designed to discriminate clinical outcome but, instead, to discriminate 2 molecular subtypes obtained by unsupervised analysis. While being associated with different outcomes, these C1 and C2 subgroups are also related to specific molecular features that can be further used to define specific therapies. However, further validation of our 3-gene predictor is needed in prospective MPM cohorts.
Interestingly, the C1 MPM subgroup characterized by a better prognosis shows a higher overall mutation rate than the C2 MPM subtype in MPM in culture and a higher rate of \textit{BAP1} mutation in both MPM in culture and tumor samples. Germline and recurrent somatic mutations in the \textit{BAP1} gene were recently identified in MPM (20, 48). These results are in contrast with others found in different types of tumors. In clear-cell renal-cell carcinoma, \textit{BAP1}-mutated tumors were higher-grade tumors and were associated with worse outcomes than tumors mutated in another gene, \textit{PBRM1} (49). In uveal melanoma, \textit{BAP1} mutations were present in most of the highly metastatic tumors, but only in a few nonmetastatic tumors (50). No significant association of the \textit{BAP1} mutation with overall survival in MPM was observed in our study or in other studies (20, 51). However, we observed a trend suggesting that MPM with the \textit{BAP1} alteration would have a worse prognosis, especially in the C1 subgroup. More cases would need to be studied to determine the prognostic value of the \textit{BAP1} mutation in each molecular subgroup. The role of the \textit{BAP1} mutation in tumor aggressiveness could be tumor type specific. We did not find any link between genetic alterations in other TSGs (\textit{CDKN2A}, \textit{CDKN2B}, \textit{NF2}, and \textit{TP53}) and patient outcome. Previously, only \textit{CDKN2A} gene alterations were found to be associated with overall survival after surgery (40). This discrepancy could be because of the low number of MPM without \textit{CDKN2A} deletion (16 cases) in our collection of MPM in culture.

In our study, C1 and C2 differ in the expression of EMT markers. Higher expression of epithelial markers that are lost during EMT and lower expression of mesenchymal markers are found in C1 in comparison to C2. Among the 3 genes of our subtype predictor, \textit{PPL} and \textit{UP3KB} may also be considered as epithelial markers, based on their expression pattern in cell types such as keratinocytes and urothelial cells. Accordingly, they were underexpressed in C2. Our data show that differences in EMT marker mRNA expression and E-cadherin protein expression also distinguished epithelioid MPM in each subgroup. Regarding the molecular profile of the epithelioid MPM in the C2 subgroup, it may be suggested that these cells have undergone mesenchymal differentiation and acquired mesenchymal features, while...
still retaining some epithelial characteristics. Previous studies showed that epithelioid MPM had a higher level of expression of E-cadherin than the biphasic and sarcomatoid subtypes (52–54). However, the absence of E-cadherin expression in some subgroups of epithelioid MPM tumors despite their epithelial histologic features was previously observed by immunohistochemistry (52, 54). In these studies, E-cadherin staining varied between 40% and 60%, which is in agreement with our classification of epithelioid MPM. A previous study assessing the difference between epithelioid and nonepithelioid MPM also suggested that EMT is of prognostic significance (55). In this study, we showed the potential prognostic value of EMT in MPM and, more importantly, we identified a subgroup of epithelioid cells linked to an EMT process and characterized by a poor outcome. Interestingly, the TGF-β signaling pathway, a key regulator of EMT in tumor cells (56), is also one of the main pathways deregulated between C1 and C2. Specific TGF-β signaling inhibitors have been developed (57) and may be valuable drugs for targeted therapy of the C2 subgroup MPM.

miRNAs are emerging as key players in carcinogenesis and are aberrantly expressed in several tumors including MPM. Recent MPM miRNome studies comparing mainly MPM versus normal cell or tissue samples have shown deregulation of several miRNAs (Supplementary Table S4); expression of their targets is expected to be inversely deregulated. We performed an in silico analysis of the convergence between miRNA and mRNA deregulation by comparing the predicted targets of miRNAs deregulated in MPM to transcriptome data in different MPM series unrelated to the miRNome studies. The enrichment of the predicted miRNA targets in the gene expression deregulation signatures was assessed (Supplementary Table S4). Of the 68 miRNAs previously identified as deregulated between MPM and normal samples (28 downexpressed in MPM, 40 overexpressed in MPM), 13 (9 downexpressed in MPM) were found to have related predicted gene target expression significantly inversely deregulated in MPM versus normal samples. Interestingly, the same enrichment analysis, focusing on genes deregulated between the molecular subgroups C1 and C2, highlighted miRNAs previously involved in EMT such as miR-200 family members (miR-200b and miR-429), miR-203, miR-204, and miR-221 (58). The combined study of both gene and miRNA expression should lead to a better understanding of the carcinogenesis mechanisms in MPM and in subgroups of MPM.

In conclusion, our transcriptomic analysis has established a robust classification of MPM and defined 2 subgroups of epithelioid MPM with different outcomes. This has been validated by the definition of a qRTPCR predictor based on 3 genes that make it possible to classify MPM sample tumors in each molecular subgroup. The 2 MPM subgroups were differentiated partly by their mutation profile and strongly by specifically deregulated pathways such as EMT. Our classification provides knowledge about molecular features of the biological diversity of MPM and will thus help to define biology-driven therapies.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: M.-C. Jaurand, F. Le Pimpec-Barthes, J. Zucman-Rossi, D. Jean
Development of methodology: A. de Reyniès, A. Renier, G. Coucy
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): I. Hysi, F. Galateau-Salle, M.-C. Copin, P. Hofman, A. Cazes, P. Andujar, J.-C. Patron
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A. de Reyniès, M.-C. Jaurand, A. Renier, G. Coucy, I. Hysi, A. Cazes, S. Imbeaud, D. Jean
Writing, review, and/or revision of the manuscript: A. de Reyniès, M.-C. Jaurand, F. Galateau-Salle, M.-C. Copin, S. Imbeaud, J. Zucman-Rossi, D. Jean
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): N. Elarouci, F. Galateau-Salle, F. Petel, F. Le Pimpec-Barthes
Study supervision: M.-C. Jaurand, F. Le Pimpec-Barthes, J. Zucman-Rossi, D. Jean

Grant Support
This work was supported by the Ligue Nationale Contre le Cancer as part of the program, “Cartes d’Identité des Tumeurs” (CIT, http://www.ligue-cancer.net/article/273_le-programme-cartes-d-identite-des-tumeurs-4.html#), the Comité de l’Oise of the Ligue Contre le Cancer, INSERM, and the Chancellerie des Universités de Paris (Legs POIX).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 5, 2013; revised December 18, 2013; accepted December 23, 2013; published OnlineFirst January 17, 2014.

References
7. Jean D, Thomas E, Manie E, Renier A, de Reynies A, Lecomte C, et al. Syntetic relationships between genomic profiles of fiber-induced...


