

Title

Molecular classification of malignant pleural mesothelioma: Identification of a poor prognosis subgroup linked to the epithelial-to-mesenchymal transition

Authors and affiliations

1. Aurélien de Reyniès^{1 *}
2. Marie-Claude Jaurand^{2,3 *}
3. Annie Renier^{2,3}
4. Gabrielle Couchy^{2,3}
5. Ilir Hysi⁴
6. Nabila Elarouci¹
7. Françoise Galateau-Sallé^{5,6}
8. Marie-Christine Copin⁷
9. Paul Hofman⁸
10. Aurélie Cazes⁹
11. Pascal Andujar^{10,11}
12. Sandrine Imbeaud^{2,3}
13. Fabien Petel¹
14. Jean-Claude Paireon^{10,11}
15. Françoise Le Pimpec-Barthes^{2,3,12}
16. Jessica Zucman-Rossi^{2,3}
17. Didier Jean^{2,3}

* These two authors contributed equally to this work.

¹ Ligue Nationale Contre Le Cancer, Programme Cartes d'Identité des Tumeurs (CIT), 75013 Paris, France

² INSERM, UMR-674, IUH, 75010 Paris, France

³ Université Paris Descartes, Labex Immuno-Oncology, Sorbonne Paris Cité, Faculté de Médecine, 75005 Paris, France

⁴ CHRU Lille, Service de Chirurgie Cardiaque, 59037 Lille, France

⁵ CHU Caen, Service d'Anatomo-Pathologie, 14000 Caen, France

⁶ INSERM, U1086, 14000 Caen, France

⁷ CHRU Lille, Université de Lille II, Institut de Pathologie, Centre de Biologie-Pathologie, 59037 Lille, France

⁸ CHU Nice, Laboratoire de Pathologie Clinique et Expérimentale (LPCE) et Biobanque Humaine, Hôpital Pasteur, 06002 Nice, France

⁹ AP-HP, GHU Ouest, Hôpital Européen Georges Pompidou, Laboratoire d'Anatomie Pathologique, 75015 Paris, France

¹⁰ Centre Hospitalier Intercommunal de Créteil, Service de Pneumologie et de Pathologie Professionnelle, 94000 Créteil, France

¹¹ INSERM, U955, Equipe 4, 94000 Créteil, France

¹² AP-HP, GHU Ouest, Hôpital Européen Georges Pompidou, Service de Chirurgie Thoracique, 75015 Paris, France

Keywords

Mesothelioma - Tumor molecular classification - Transcriptome microarray - Genetic alteration - Asbestos

Running title

Molecular classification of mesothelioma

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://cit.ligue-cancer.net>), the Comité de l'Oise of the Ligue Contre le Cancer, INSERM, and the Chancellerie des Universités de Paris (Legs POIX).

Corresponding author

Didier JEAN
INSERM UMR-S 674
27 rue Juliette Dodu
75010 PARIS
FRANCE
Phone: +33 (0)1 53 72 51 89; Fax: +33 (0)1 53 72 51 92
E-mail address: didier.jean@inserm.fr

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Abbreviations

EMT: epithelial-to-mesenchymal transition; MPM: malignant pleural mesothelioma.

Authors' Contributions

Conception and design:

M.C. Jaurand, J. Zucman-Rossi, D. Jean.

Development of methodology:

A. de Reyniès, A. Renier, G. Couchy.

Acquisition of data (acquired patients clinical, histological and occupational annotations and tumor samples):

M.C. Jaurand, I. Hysi, F. Galateau-Sallé, M.C. Copin, P. Hofman, P. Andujar, J.C. Pairon, F. Le Pimpec-Barthes.

Analysis and interpretation of data (biological, biostatistical and computational analyses):

A. de Reyniès, M.C. Jaurand, A. Renier, G. Couchy, I. Hysi, A. Cazes, S. Imbeaud, D. Jean.

Writing, review and/or revision of the manuscript:

A. de Reyniès, M.C. Jaurand, S. Imbeaud, J. Zucman-Rossi, D. Jean.

Administrative, technical or material support (reporting or organizing data, constructing databases):

N. Elarouci, F. Petel.

Study supervision:

M.C. Jaurand, D. Jean.

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 clinico-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 the present 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.

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 RT-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 histological 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 histological 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.

Introduction

Malignant pleural mesothelioma (MPM) commonly affects the mesothelial surfaces of the pleural cavities. Its major risk factor is past exposure to asbestos fibers (1). Despite research efforts to develop more effective diagnostic and therapeutic approaches, MPM prognosis remains poor and patient median survival is very short, with a median of approximately 12 months. Conventional anti-cancer therapies (surgery, radiotherapy and chemotherapy) are ineffective, either alone or combined (2), underlining the strong need to develop new therapies and requiring better knowledge of molecular alterations in tumor cells. The current therapeutic strategy used is based on health status, imaging, histology and biomarkers, and can be improved through the use of new tools for a more in-depth phenotypic classification of the tumor that reflects the clinico-biological heterogeneity of MPM.

Large-scale molecular studies of MPM have demonstrated tumor heterogeneity at the level of chromosomal 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 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 anti-cancer

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 three main histological subtypes: epithelioid, sarcomatoid and biphasic. Although multiple patterns of each histological 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 histological subtypes (14). Specific profiles of miRNA expression also defined MPM subgroups related to histological subtypes (15, 16).

The aim of the present 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 biological 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 two 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-mesenchymal transition between the two 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 histological analysis using consensus guidelines established by the International Mesothelioma Interest Group (<http://imig.org>) (17). Clinico-pathological and epidemiologic data of patients are reported in **Table I**.

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 pneumothoraxes. Cells were grown in RPMI 1640 medium supplemented with penicillin, streptomycin, Glutamax and 10% fetal calf serum (Life Technologies, Courtaboeuf, France) (19).

Tumor samples

Frozen MPM tumor samples were retrieved from the French mesobank collection (biobanks of three French hospitals: CHU de Caen; CHRU de Lille; CHU de Nice) and the biobank collection of HEGP. The percentage of tumor cells in MPM samples was estimated by histological examination.

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, Courtaboeuf, France) 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, Illkirch, France), agarose gels and, for microarray analysis, electrophoregram profiles on an Agilent 2100 Bioanalyzer (Agilent Technologies, Massy, France). Nucleic acid concentrations were determined using Hoechst 33258 or Ribogreen fluorescence assays (Life Technologies) for DNA and RNA, respectively.

Pangenomic expression profiling

Expression profiles of primary cultures from 38 MPM and three 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 two 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 polymerase chain reaction (qRT-PCR) analysis

Pre-defined 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 $2^{-\Delta\Delta Ct}$ 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 tumor suppressor genes were amplified from genomic DNA by PCR using a Qiagen Multiplex Master Mix kit (Qiagen) and purified by ExoSAP-IT (USB Corporation Cleveland, OH, USA) treatment. Primers used are listed in the [Supplementary Table S3](#) for *BAP1* and *CDKN2B* genes or were previously published for the two other genes (21, 22). The PCR products were analyzed and quantified using the Qiaxcel 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 multi-capillary sequencer (ABI 3130 Genetic Analyzer). Mutations were detected using Sequencher software (Gene Codes Corp, Ann Arbor, MI, USA) 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 ME024 kits (MRC-Holland, Amsterdam, The Netherlands) 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, Charbonnières, France) and normal liver tissues were included in each MLPA run. The reliability of the assay was determined 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 two times.

Western blots

Confluent MPM cells in culture were washed in PBS and cell extracts were prepared in RIPA buffer (Millipore, St Quentin en Yvelines, France) containing complete protease inhibitor cocktail (Roche Diagnostics, Meylan, France). Equal amounts of protein, as determined by the Bio-Rad DC protein assay (Bio-Rad, Marnes la Coquette, France), were separated by SDS-PAGE, transferred to PVDF membrane (Millipore), and immunoblotted with anti-E-cadherin (*CDH1*; clone 24E10 1/1000 dilution; Cell Signaling Technology, Ozyme, Saint Quentin, France) and anti-glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*; clone 14C10 1/5000 dilution; Cell Signaling Technology) antibodies. Bound antibodies were detected with HRP-conjugated secondary antibodies (Santa Cruz Biotechnology, Heidelberg, Germany) and ECL reagents (GE Healthcare, Velizy, France). The relative quantification of bands was determined using Quantity One Quantitation software (Bio-Rad).

Immunohistochemistry

Immunohistochemistry for E-cadherin was performed on 5- μ m deparaffinized sections using an automated single-staining procedure (Benchmark Ultra, Ventana Medical Systems, Tucson, AR, USA). Briefly, the mouse monoclonal anti-E-cadherin antibody (clone 4A2C7, 1/50 dilution) was added to the slides for 32 minutes after pretreatment with UltraCC1 buffer (Ventana Medical Systems) for 36 min. 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 three normal mesothelial cells were normalized in batch using the RMA algorithm (23). Each of the two public datasets of MPM samples (GSE29354, GSE2549) 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 two robust clusters, C1 and C2. Briefly, 24 dendrograms were obtained by hierarchical clustering using three different linkages (Ward, Complete, Average) and eight lists of genes corresponding to the top 40% to 0.5% (eight different thresholds) most varying genes (assessed via a robust coefficient of variation). Each of the 24 dendrograms yielded a raw partition in two clusters; the 24 raw partitions in two clusters were then used to build a consensus dendrogram that itself was divided into two clusters, thus yielding a consensus partition in two clusters. Principal Component Analysis (PCA) of these sample profiles was also performed. The first three 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 AUC criterion (PresenceAbsence R package) were used. Genes (40) selected based on these criteria (AUC and t-test), as well as their potential involvement in mesothelial carcinogenesis, were measured by

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 three algorithms (DLDA, DQDA [R package sma], PAM [R package pamr (25)]. The predictor is defined by a majority vote across these three algorithms and is based on three genes (PPL [periplakin], UPK3B [uroplakin 3B], TFPI [tissue factor pathway inhibitor]). 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 four methods (GSA (26): R package GSA; globaltest (27): R package globaltest; SAM-GS (28): original R code; Tukey (29): original R code). Given two 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 two sample groups. To aggregate the results of the four methods, the list of pathways for each method is first sorted and the mean rank across the four methods is then calculated; the final order is based on this mean rank. Pathways and related genes were retrieved from two repositories: Gene Ontology (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 logrank tests.

Results

Major clinico-pathological and epidemiological characteristics of MPM in culture and tumor samples are summarized in [Table 1](#). Histological 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 histological 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 two 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 two robust clusters, C1 and C2 ([Fig. 1.A](#), [Supplementary Fig. S2](#)). This classification was partly related to histological 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 tumor suppressor genes screened for mutation. However, *BAP1* and *CDKN2A* tend to be more frequently mutated in C1 (Fisher exact test, $p=0.09$ and $p=0.14$, respectively; [Fig. 1.A](#)). Of note, the analysis of the 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 three genes - *PPL*, *UPK3B* and *TFPI* - showing a 100% success rate in the discovery set ([Supplementary Table S1](#), [Supplementary 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 histological 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. 1.B](#)). 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. 1.B](#)). 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 two public MPM pangenomic expression profile series (14, 20) ([Supplementary Fig. S5.A-D](#)). The consensus classifications derived from the Bott and Gordon series were then compared to the C1/C2 classification in terms of gene expression signatures and showed agreement between the classifications derived from our three MPM series ([Supplementary Fig. S5.E](#)). The public series subgroups were referred to as “C1-like” and “C2-like”. A similar histological subtype distribution was found in these two independent datasets ([Supplementary Fig. S6.A](#)).

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 logrank test, $p = 0.02$), which is coherent with the known poor prognosis of sarcomatoid and biphasic MPM ([Fig. 2.A](#)). 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. 2.B](#)). Both observations were confirmed in the validation series (logrank test, $p < 0.01$ and $p = 0.05$, respectively;

Fig. 2.C,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 (logrank test, $p < 0.01$), even when considering only epithelioid MPM (logrank test, $p = 0.03$; **Supplementary Fig. S6.B**).

Pathway analysis reveals epithelial-to-mesenchymal transition (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 four algorithms and screening both the Gene Ontology (GO) terms and KEGG pathways. The GO terms, 'cell migration' (GO:0016477) and 'epithelial-to-mesenchymal transition' (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-like and C1-like MPM, yielding results similar to those of the discovery cohort (**Supplementary Table S2**). By aggregating pathway analyses from the three series, the GO terms, 'cell migration' and 'epithelial-to-mesenchymal transition' were ranked in position 3 and 12, respectively, among more than 13,000 GO terms. Similarly, pathways related to cell adhesion, junctions and extracellular matrices were in the top ranks (< 20) among the KEGG pathways. Interestingly, the TGF-beta signaling pathway (KEGG hsa04350), known to induce EMT (31), was also 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 (32-34) (list given in **Supplementary Table S2**). In our discovery series, most of the positive EMT markers were found to be up-regulated in C2 compared to C1, while most of the negative EMT markers were found to be

down-regulated in C2 compared to C1, supporting the hypothesis that the C2 subgroup is characterized by a mesenchymal phenotype (Fig. 3.A). The same results were observed in the Bott and Gordon series when comparing C2-like and C1-like MPM (Supplementary Fig. S7.A-B).

The differential expression of EMT markers was validated using qRT-PCR on our collection of MPM in culture and in tumor samples (Fig. 3.B). Genes up-regulated during EMT, such as the adhesion molecules *POSTN* and *VCAN* or the transcription factors *SNAI2*, *TCF4* and *HMGA2*, had statistically significant higher expression in C2 than in C1 (Mann-Whitney test, $p < 0.01$). In contrast, genes down-regulated 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. S7.C-D).

Protein expression of *CDH1*, encoding E-cadherin, was studied by Western blot in MPM in culture and by immunohistochemistry in MPM tumor samples. Fig. 4.A shows representative Western blots in five C1 epithelioid MPM and three 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. 4.B). 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. 4.C-F). This staining was stronger in C1 (Fig. 4.C,E) than in C2 (Fig. 4.D,F).

Discussion

MPM is a tumor with a terrible prognosis due to its resistance to conventional anti-cancer 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 the present 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 two studies classified MPM into subgroups and found correlations only with histological types (14, 38). One of the most important points of our classification was the separation of epithelioid MPM into two molecular subtypes, C1 and C2, with different prognoses.

We identified a predictor based on three 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 pro-coagulant response in the pathogenesis of pleural injury, was also shown to have a variable expression in MPM tumors (42, 43).

In the past, three 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 two MPM tumor groups with different survival rates (44). A four-gene predictor was found to differentiate patients on the basis of post-surgical 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 three-gene predictor was not designed to discriminate clinical outcome but, instead, to discriminate two 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 three-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 *BAP1* mutation in both MPM in culture and tumor samples. Germline and recurrent somatic mutations in the *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, *BAP1*-mutated tumors were higher-grade tumors and were associated with worse outcomes than tumors mutated in another gene, *PBRM1* (49). In uveal melanoma, *BAP1* mutations were present in most of the highly metastatic tumors, but only in a few non-metastatic tumors (50). No significant association of the *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 *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 *BAP1* mutation in each molecular subgroup. The role of the *BAP1* mutation in tumor aggressiveness could be tumor type-specific. We did not find any link between genetic

alterations in other tumor suppressor genes (*CDKN2A*, *CDKN2B*, *NF2* and *TP53*) and patient outcome. Previously, only *CDKN2A* gene alterations were found to be associated with overall survival after surgery (40). This discrepancy could be due to the low number of MPM without *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 three genes of our subtype predictor, *PPL* and *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 under-expressed 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 histological 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 non-epithelioid MPM also suggested that EMT is of prognostic significance (55). In the present 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-beta 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-beta 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 vs. 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 down-expressed in MPM, 40 over-expressed in MPM), 13 (9 down-expressed in MPM, 4 over-expressed in MPM) were found to have related predicted gene target expression significantly inversely deregulated in MPM vs. 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 two subgroups of epithelioid MPM with different outcomes. This has been validated by the definition of a qRT-PCR predictor based on three genes that make it possible to classify MPM sample tumors in each molecular subgroup. The two 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.

References

1. Huang SX, Jaurand MC, Kamp DW, Whysner J, Hei TK. Role of mutagenicity in asbestos fiber-induced carcinogenicity and other diseases. *J Toxicol Environ Health B Crit Rev* 2011;14:179-245.
2. Scherpereel A, Astoul P, Baas P, Berghmans T, Clayson H, de Vuyst P, et al. Guidelines of the European Respiratory Society and the European Society of Thoracic Surgeons for the management of malignant pleural mesothelioma. *Eur Respir J* 2010;35:479-95.
3. Musti M, Kettunen E, Dragonieri S, Lindholm P, Cavone D, Serio G, et al. Cytogenetic and molecular genetic changes in malignant mesothelioma. *Cancer Genet Cytogenet* 2006;170:9-15.
4. Lindholm PM, Salmenkivi K, Vauhkonen H, Nicholson AG, Anttila S, Kinnula VL, et al. Gene copy number analysis in malignant pleural mesothelioma using oligonucleotide array CGH. *Cytogenet Genome Res* 2007;119:46-52.
5. Taniguchi T, Karnan S, Fukui T, Yokoyama T, Tagawa H, Yokoi K, et al. Genomic profiling of malignant pleural mesothelioma with array-based comparative genomic hybridization shows frequent non-random chromosomal alteration regions including JUN amplification on 1p32. *Cancer Sci* 2007;98:438-46.
6. Jean D, Daubriac J, Le Pimpec-Barthes F, Galateau-Salle F, Jaurand MC. Molecular changes in mesothelioma with an impact on prognosis and treatment. *Arch Pathol Lab Med* 2012;136:277-93.
7. Jean D, Thomas E, Manie E, Renier A, de Reynies A, Lecomte C, et al. Syntenic Relationships between Genomic Profiles of Fiber-Induced Murine and Human Malignant Mesothelioma. *Am J Pathol* 2011;178:881-94.
8. Kindler HL, Karrison TG, Gandara DR, Lu C, Krug LM, Stevenson JP, et al. Multicenter, double-blind, placebo-controlled, randomized phase II trial of gemcitabine/cisplatin plus bevacizumab or placebo in patients with malignant mesothelioma. *J Clin Oncol* 2012;30:2509-15.
9. Bertino P, Carbone M, Pass H. Chemotherapy of malignant pleural mesothelioma. *Expert Opin Pharmacother* 2009;10:99-107.
10. Betta PG, Magnani C, Bensi T, Trincerri NF, Orecchia S. Immunohistochemistry and molecular diagnostics of pleural malignant mesothelioma. *Arch Pathol Lab Med* 2012;136:253-61.
11. Husain AN, Colby T, Ordonez N, Krausz T, Attanoos R, Beasley MB, et al. Guidelines for Pathologic Diagnosis of Malignant Mesothelioma: 2012 Update of the Consensus Statement from the International Mesothelioma Interest Group. *Arch Pathol Lab Med* 2012;137:647-67.
12. Goto Y, Shinjo K, Kondo Y, Shen L, Toyota M, Suzuki H, et al. Epigenetic profiles distinguish malignant pleural mesothelioma from lung adenocarcinoma. *Cancer Res* 2009;69:9073-82.
13. Christensen BC, Houseman EA, Godleski JJ, Marsit CJ, Longacker JL, Roelofs CR, et al. Epigenetic profiles distinguish pleural mesothelioma from normal pleura and predict lung asbestos burden and clinical outcome. *Cancer Res* 2009;69:227-34.
14. Gordon GJ, Rockwell GN, Jensen RV, Rheinwald JG, Glickman JN, Aronson JP, et al. Identification of novel candidate oncogenes and tumor suppressors in malignant pleural mesothelioma using large-scale transcriptional profiling. *Am J Pathol* 2005;166:1827-40.
15. Guled M, Lahti L, Lindholm PM, Salmenkivi K, Bagwan I, Nicholson AG, et al. CDKN2A, NF2, and JUN are dysregulated among other genes by miRNAs in malignant mesothelioma -A miRNA microarray analysis. *Genes Chromosomes Cancer* 2009;48:615-23.
16. Xu Y, Zheng M, Merritt RE, Shrager JB, Wakelee H, Kratzke RA, et al. microRNA-1 Induces Growth Arrest and Apoptosis in Malignant Mesothelioma. *Chest* 2013.
17. Husain AN, Colby TV, Ordonez NG, Krausz T, Borczuk A, Cagle PT, et al. Guidelines for pathologic diagnosis of malignant mesothelioma: a consensus statement from the International Mesothelioma Interest Group. *Arch Pathol Lab Med* 2009;133:1317-31.

18. Zeng L, Fleury-Feith J, Monnet I, Boutin C, Bignon J, Jaurand MC. Immunocytochemical characterization of cell lines from human malignant mesothelioma: characterization of human mesothelioma cell lines by immunocytochemistry with a panel of monoclonal antibodies. *Hum Pathol* 1994;25:227-34.
19. Zeng L, Buard A, Monnet I, Boutin C, Fleury J, Saint-Etienne L, et al. In vitro effects of recombinant human interferon gamma on human mesothelioma cell lines. *Int J Cancer* 1993;55:515-20.
20. Bott M, Brevet M, Taylor BS, Shimizu S, Ito T, Wang L, et al. The nuclear deubiquitinase BAP1 is commonly inactivated by somatic mutations and 3p21.1 losses in malignant pleural mesothelioma. *Nat Genet* 2011;43:668-72.
21. Andujar P, Wang J, Descatha A, Galateau-Salle F, Abd-Alsamad I, Billon-Galland MA, et al. p16INK4A inactivation mechanisms in non-small-cell lung cancer patients occupationally exposed to asbestos. *Lung Cancer* 2010;67:23-30.
22. Andujar P, Pairon JC, Renier A, Descatha A, Hysi I, Abd-Alsamad I, et al. Differential mutation profiles and similar intronic TP53 polymorphisms in asbestos-related lung cancer and pleural mesothelioma. *Mutagenesis* 2013;28:323-31.
23. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003;31:e15.
24. Cairo S, Armengol C, De Reynies A, Wei Y, Thomas E, Renard CA, et al. Hepatic stem-like phenotype and interplay of Wnt/beta-catenin and Myc signaling in aggressive childhood liver cancer. *Cancer Cell* 2008;14:471-84.
25. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* 2002;99:6567-72.
26. Efron B, Tibshirani R. Empirical bayes methods and false discovery rates for microarrays. *Genet Epidemiol* 2002;23:70-86.
27. Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics* 2004;20:93-9.
28. Dinu I, Potter JD, Mueller T, Liu Q, Adewale AJ, Jhangri GS, et al. Improving gene set analysis of microarray data by SAM-GS. *BMC Bioinformatics* 2007;8:242.
29. Goeman JJ, Buhlmann P. Analyzing gene expression data in terms of gene sets: methodological issues. *Bioinformatics* 2007;23:980-7.
30. Levallet G, Vaisse-Lesteven M, Le Stang N, Ilg AG, Brochard P, Astoul P, et al. Plasma cell membrane localization of c-MET predicts longer survival in patients with malignant mesothelioma: a series of 157 cases from the MESOPATH Group. *J Thorac Oncol* 2012;7:599-606.
31. Zavadil J, Bottlinger EP. TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene* 2005;24:5764-74.
32. Thiery JP, Sleeman JP. Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol* 2006;7:131-42.
33. Zeisberg M, Neilson EG. Biomarkers for epithelial-mesenchymal transitions. *J Clin Invest* 2009;119:1429-37.
34. Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, Yang J, et al. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc Natl Acad Sci U S A* 2010;107:15449-54.
35. Dhalluin X, Scherpereel A. Treatment of malignant pleural mesothelioma: current status and future directions. *Monaldi Arch Chest Dis* 2010;73:79-85.
36. Greillier L, Marco S, Barlesi F. Targeted therapies in malignant pleural mesothelioma: a review of clinical studies. *Anticancer Drugs* 2011;22:199-205.
37. Jakobsen JN, Sorensen JB. Review on clinical trials of targeted treatments in malignant mesothelioma. *Cancer Chemother Pharmacol* 2011;68:1-15.

38. Hoang CD, D'Cunha J, Kratzke MG, Casmey CE, Frizelle SP, Maddaus MA, et al. Gene expression profiling identifies matriptase overexpression in malignant mesothelioma. *Chest* 2004;125:1843-52.
39. Gueugnon F, Leclercq S, Blanquart C, Sagan C, Cellierin L, Padiou M, et al. Identification of novel markers for the diagnosis of malignant pleural mesothelioma. *Am J Pathol* 2011;178:1033-42.
40. Lopez-Rios F, Chuai S, Flores R, Shimizu S, Ohno T, Wakahara K, et al. Global gene expression profiling of pleural mesotheliomas: overexpression of aurora kinases and P16/CDKN2A deletion as prognostic factors and critical evaluation of microarray-based prognostic prediction. *Cancer Res* 2006;66:2970-9.
41. Kanamori-Katayama M, Kaiho A, Ishizu Y, Okamura-Oho Y, Hino O, Abe M, et al. LRRN4 and UPK3B are markers of primary mesothelial cells. *PLoS One* 2011;6:e25391.
42. Idell S, Pueblitz S, Emri S, Gungen Y, Gray L, Kumar A, et al. Regulation of fibrin deposition by malignant mesothelioma. *Am J Pathol* 1995;147:1318-29.
43. Mutsaers SE, Prele CM, Brody AR, Idell S. Pathogenesis of pleural fibrosis. *Respirology* 2004;9:428-40.
44. Pass HI, Liu Z, Wali A, Bueno R, Land S, Lott D, et al. Gene expression profiles predict survival and progression of pleural mesothelioma. *Clin Cancer Res* 2004;10:849-59.
45. Gordon GJ, Jensen RV, Hsiao LL, Gullans SR, Blumenstock JE, Richards WG, et al. Using gene expression ratios to predict outcome among patients with mesothelioma. *J Natl Cancer Inst* 2003;95:598-605.
46. Gordon GJ, Rockwell GN, Godfrey PA, Jensen RV, Glickman JN, Yeap BY, et al. Validation of genomics-based prognostic tests in malignant pleural mesothelioma. *Clin Cancer Res* 2005;11:4406-14.
47. Gordon GJ, Dong L, Yeap BY, Richards WG, Glickman JN, Edenfield H, et al. Four-gene expression ratio test for survival in patients undergoing surgery for mesothelioma. *J Natl Cancer Inst* 2009;101:678-86.
48. Testa JR, Cheung M, Pei J, Below JE, Tan Y, Sementino E, et al. Germline BAP1 mutations predispose to malignant mesothelioma. *Nat Genet* 2011;43:1022-5.
49. Kapur P, Pena-Llopis S, Christie A, Zhrebker L, Pavia-Jimenez A, Rathmell WK, et al. Effects on survival of BAP1 and PBRM1 mutations in sporadic clear-cell renal-cell carcinoma: a retrospective analysis with independent validation. *Lancet Oncol* 2013;14:159-67.
50. Harbour JW, Onken MD, Roberson ED, Duan S, Cao L, Worley LA, et al. Frequent mutation of BAP1 in metastasizing uveal melanomas. *Science* 2010;330:1410-3.
51. Zauderer MG, Bott M, McMillan R, Sima CS, Rusch V, Krug LM, et al. Clinical Characteristics of Patients with Malignant Pleural Mesothelioma Harboring Somatic BAP1 Mutations. *J Thorac Oncol* 2013;8:1430-3.
52. Abutaily AS, Addis BJ, Roche WR. Immunohistochemistry in the distinction between malignant mesothelioma and pulmonary adenocarcinoma: a critical evaluation of new antibodies. *J Clin Pathol* 2002;55:662-8.
53. Fassina A, Cappellesso R, Guzzardo V, Dalla Via L, Piccolo S, Ventura L, et al. Epithelial-mesenchymal transition in malignant mesothelioma. *Mod Pathol* 2012;25:86-99.
54. Merikallio H, Paakko P, Salmenkivi K, Kinnula V, Harju T, Soini Y. Expression of snail, twist, and Zeb1 in malignant mesothelioma. *APMIS* 2013;121:1-10.
55. Schramm A, Opitz I, Thies S, Seifert B, Moch H, Weder W, et al. Prognostic significance of epithelial-mesenchymal transition in malignant pleural mesothelioma. *Eur J Cardiothorac Surg* 2010;37:566-72.
56. Wendt MK, Tian M, Schiemann WP. Deconstructing the mechanisms and consequences of TGF-beta-induced EMT during cancer progression. *Cell Tissue Res* 2012;347:85-101.
57. Akhurst RJ, Hata A. Targeting the TGFbeta signalling pathway in disease. *Nat Rev Drug Discov* 2012;11:790-811.

58. Lamouille S, Subramanyam D, Belloch R, Derynck R. Regulation of epithelial-mesenchymal and mesenchymal-epithelial transitions by microRNAs. *Curr Opin Cell Biol* 2013;25:200-7.

Tables

Table 1. Clinico-pathological characteristics and epidemiologic data of the series of MPM patients

	Cells in culture (n=67)	Tumor samples (n=108)
Gender (n [%])		
Male	53 [79]	79 [73]
Female	14 [21]	29 [27]
Age (years)		
Median \pm SD	65 \pm 11	67 \pm 10
Range	37-91	39-90
Histology (n [%])		
Epithelioid	48 [75]	80 [78]
Biphasic	9 [14]	13 [13]
Sarcomatoid	6 [9]	7 [7]
Desmoplastic	1 [2]	3 [3]
Lymphohistiocytoid	0 [0]	1 [1]
Asbestos exposure (n [%])		
Exposed	51 [80]	67 [71]
Non-exposed	13 [20]	27 [29]
Stage IMIG (n [%])		
I	nd	4 [5]
II	nd	13 [17]
III	nd	30 [39]
IV	nd	29 [38]
Survival (months)		
Median	10.2	16.9
Range	0.3-118.8	0.1-69.6

MPM: malignant pleural mesothelioma; *nd*: not determined.

Table 2. Univariate and multivariate analysis for overall survival of MPM patients

	Tumor samples (n=64)			
	Univariate HR (95% CI)	<i>P</i>	Multivariate ^a HR (95% CI)	<i>P</i>
Gender				
Female vs. male	1.51 (0.88-2.59)	0.18	-	-
Age				
< vs. > median	1.83 (1.10-3.02)	0.01	2.01 (1.17-3.44)	0.01
Histology				
Epithelioid vs. non-epithelioid	2.65 (1.22-5.77)	<0.01	2.20 (1.21-3.99)	0.01
Asbestos exposure				
Exposed vs. non-exposed	1.14 (0.63-2.09)	0.64	-	-
IMIG stage				
I/II vs. III/IV	1.60 (0.89-2.85)	0.14	-	-
Molecular subgroup				
C1 vs. C2	2.10 (1.21-3.64)	<0.01	2.54 (1.45-4.44)	<0.01

^a Multivariate analysis was performed on parameters with univariate *P* below 0.1. CI: confidence interval; HR: hazard ratio; MPM: malignant pleural mesothelioma.

Legends of figures

Figure 1. Unsupervised consensus classification of MPM in culture based on gene expression profiles

(A) A consensus dendrogram of MPM in culture (upper panel) derived from 24 raw partitions obtained using various experimental settings (see [Methods](#)) is shown. On the basis of this consensus dendrogram, a consensus partition of the samples is defined by a division into two clusters, C1 (blue) and C2 (red). In the lower panel, sample annotations are shown: histological subtypes (**E**: Epithelioid, **B**: Biphasic, **D**: Desmoplastic, **S**: Sarcomatoid), and mutation status (–: Wild type, **M**: Mutated) of *BAP1*, *CDKN2A*, *CDKN2B*, *TP53* and *NF2* genes. To assess the association between C1/C2 clusters and sample annotations, global Fisher exact tests were performed and related p-values are reported. (B) A three-gene qRT-PCR predictor was applied to predict C1/C2 subgroup membership in MPM in culture (discovery series + extension series, n=67) and in MPM tumor samples (validation series, n=108). Contingency tables are calculated separately for MPM in culture (left) and MPM tumor samples (right), rows corresponding to sample annotations (histological subtypes and genetic alterations of *BAP1*, *CDKN2A*, *CDKN2B*, *TP53* and *NF2* genes) and columns to C1/C2 predicted subgroups. P-values of Fisher exact tests are shown.

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-B**) and the validation series (**C-D**), respectively, are calculated for all MPM (**A,C**) and epithelioid MPM (**B,D**). Differences in survival between molecular subgroups are assessed by logrank tests. Kaplan-Meier survival analyses for all MPM and for epithelioid MPM showed that C1 and C2 subgroups have statistically significant different prognoses.

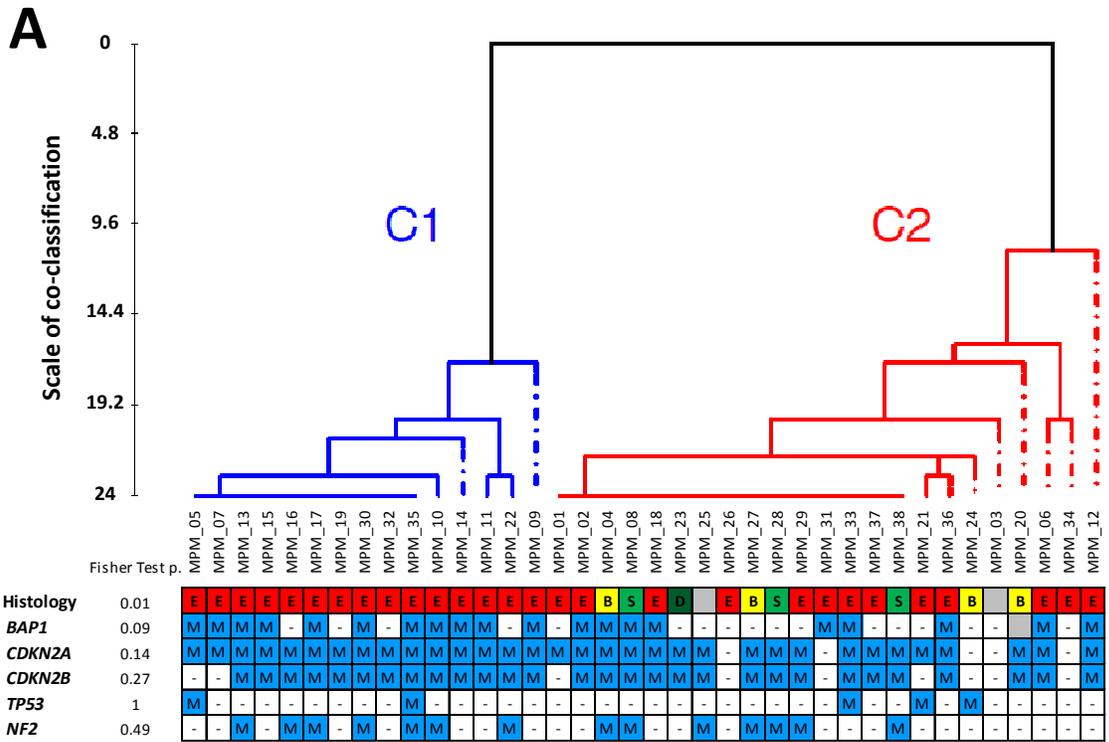
Figure 3. Differential mRNA expression of EMT markers in MPM molecular subgroups

(A) Genes reported as positive and negative EMT markers ([Supplementary Table S2](#)) were tested for differential expression between C1 and C2 samples from the discovery cohort using moderate t-tests. Each dot represents a marker (mesenchymal: red; epithelial: green); the y-axis gives the $-\log_{10}$ p-

value of the test; the x-axis gives the log₂ ratio of expression between C2 and C1 samples. **(B)** The expression of EMT positive markers (*POSTN*, *VCAN*, *SNAI2*, *TCF4*, *HMGA2*) and negative markers (*CDH1*, *CDH3*) was measured by qRT-PCR, and differential expression of these markers between C1 and C2 samples was tested using Mann-Whitney tests in MPM in culture (upper panel) and in MPM tumor samples (lower panel). For each tested marker, box plots of $-\Delta\Delta\text{Ct}$ normalized values are shown for C1 (blue) and C2 (red), respectively. MPM and p-values are indicated above. Data showed up-regulation of EMT positive marker expression and down-regulation of negative marker expression in C2 MPM compared to C1 MPM.

Figure 4. Differential protein expression of the EMT marker, E-cadherin, in MPM molecular subgroups

(A-B) Cell extracts prepared from 25 different epithelioid MPM in culture were submitted to SDS-PAGE and analyzed by Western blot using anti-E-cadherin antibody. GAPDH expression was also analyzed in controls. Representative Western blots are shown **(A)**. E-cadherin is detected as a protein of 120 kDa in all of the C1 MPM. Each band corresponding to E-cadherin and GAPDH was quantified, expressed using arbitrary units and normalized to the expression of MESO_07 cells. The ratio E-cadherin/GAPDH is shown on the graph for each MPM subgroup **(B)**. E-cadherin expression at the protein level is higher in C1 than in C2 MPM in culture. **(C-F)** E-cadherin immunostaining of paraffin-embedded tumors in representative cases of epithelioid **(C,D)** and biphasic **(E,F)** MPM is shown. Most of the cells with an epithelial morphology are stained in the C1 subgroup **(C,E)**, while only few cells with an epithelial morphology are stained in the C2 subgroup **(D)**. No staining was observed in sarcomatoid MPM cells **(E-F)**.



B

	MPM in culture			Tumor samples		
	Discovery and extension series (n=67)		P-value*	Validation series (n=108)		P-value*
	C1	C2		C1	C2	
Histology n (%)						
Epithelioid	19 (40)	29 (60)	0.04	55 (69)	25 (31)	<0.01
Sarcomatoid/Desmoplastic	0 (0)	7 (100)		0 (0)	10 (100)	
Biphasic	2 (22)	7 (78)		6 (46)	7 (54)	
Genetic alterations n (%)						
CDKN2A	19/20 (95)	28/43 (65)	0.01	16/27 (59)	8/16 (50)	0.75
CDKN2B	17/20 (85)	25/43 (58)	0.05	16/27 (59)	7/16 (44)	0.32
BAP1	16/20 (80)	14/41 (34)	<0.01	12/27 (44)	2/16 (13)	0.04
NF2	10/20 (50)	15/44 (34)	0.27	nd	nd	nd
TP53	2/20 (10)	4/42 (10)	1.00	nd	nd	nd

nd: not determined. * Fisher exact test

Figure 1

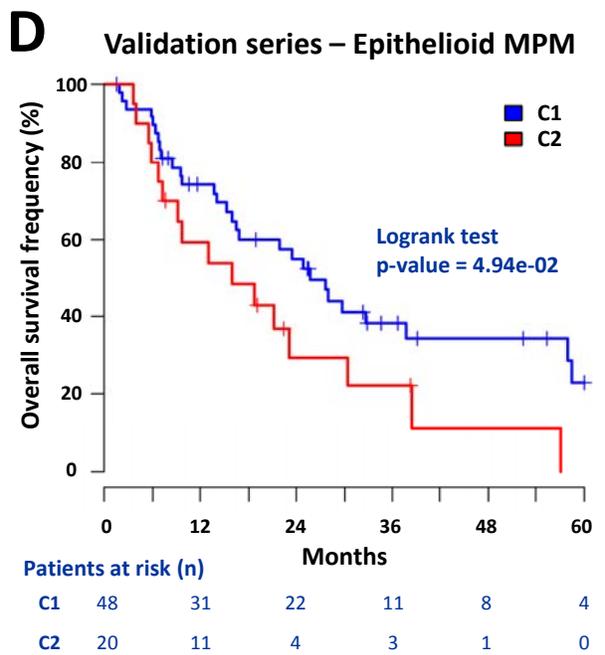
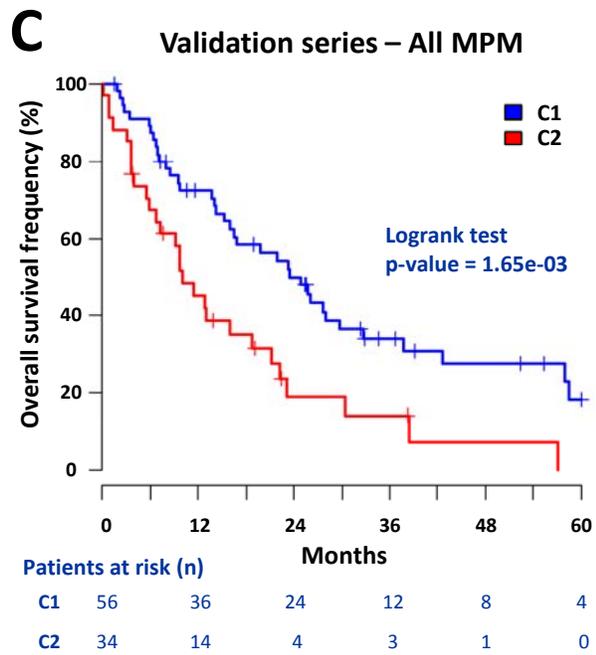
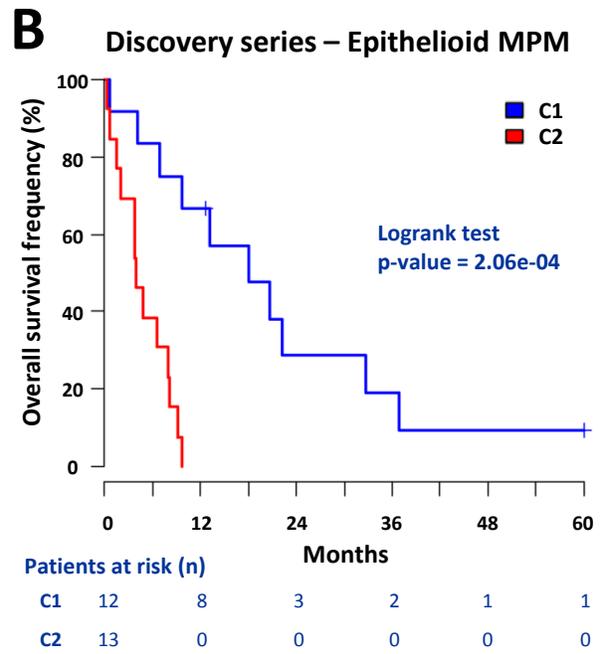
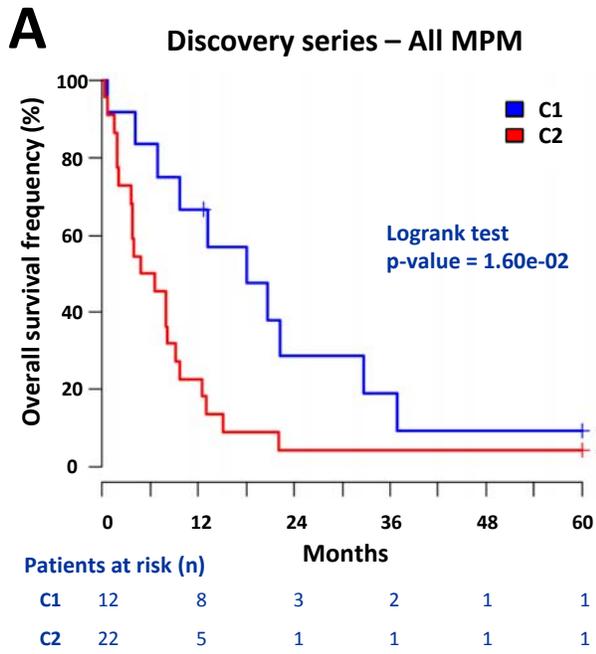
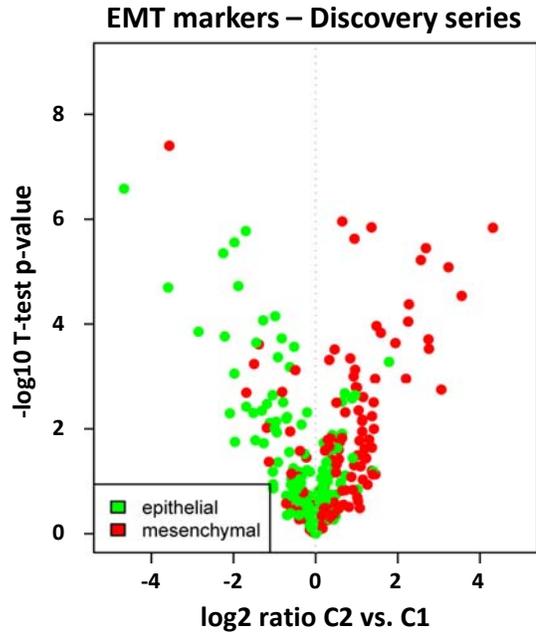
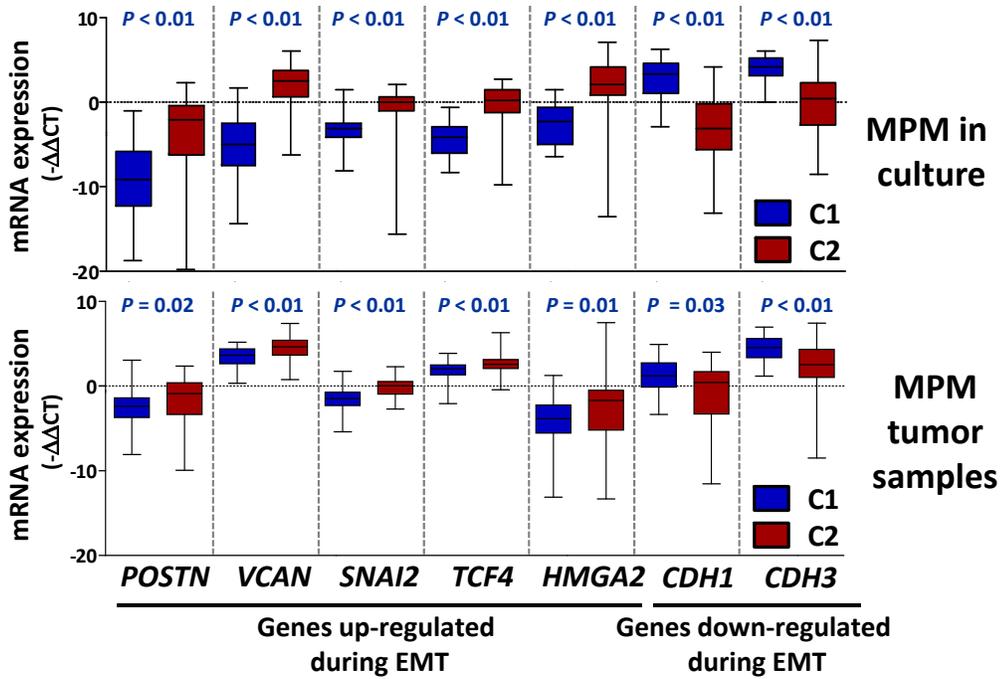


Figure 2

A**B****Figure 3**

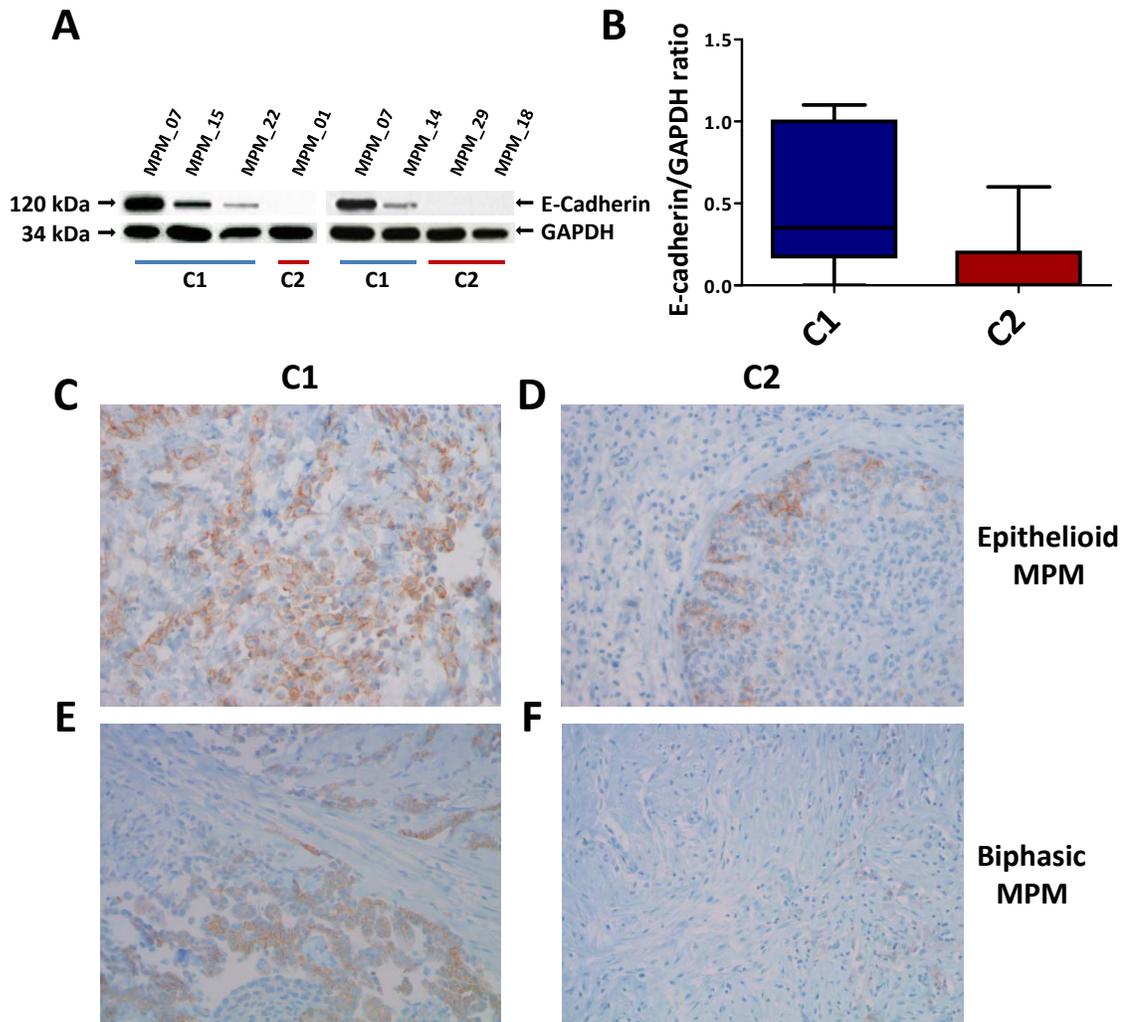


Figure 4