Genetic alterations of malignant pleural mesothelioma: association to tumor heterogeneity and overall survival Author names

Lisa Quetel, Clément Meiller, Jean-Baptiste Assié, Yuna Blum, Sandrine Imbeaud, François Montagne, Robin Tranchant, Julien de Wolf, Stefano Caruso, Marie-Christine Copin, et al.

DR. ROBIN TRANCHANT (Orcid ID : 0000-0001-5789-0153)
DR. DIDIER JEAN (Orcid ID : 0000-0001-5823-7404)

Received Date : 23-Aug-2019
Revised Date : 13-Dec-2019
Accepted Date : 14-Feb-2020
Article type : Research Article

Genetic alterations of malignant pleural mesothelioma: association to tumor heterogeneity and overall survival

Author names

Lisa Quetel[1], Clément Meiller[1], Jean-Baptiste Assié[1,*], Yuna Blum[2,*], Sandrine Imbeaud[1,*], François Montagne[1,*,†], Robin Tranchant[1,‡], Julien de Wolf[1,¥], Stefano Caruso[1], Marie-Christine Copin[3,4], Véronique Hofman[5,6], Laure Gibault[7,8], Cécile Badoual[7,8], Ecaterina Pintilie[9], Paul Hofman[5,6], Isabelle Monnet[10], Arnaud Scherpereel[4,11,12], Françoise Le Pimpec-Barthes[1,7,13], Jessica Zucman-Rossi[1,7], Marie-Claude Jaurand[1], Didier Jean[1]

Addresses

[1] Centre de Recherche des Cordeliers, Inserm, Sorbonne Université, Université de Paris, Functional Genomics of Solid Tumors, F-75006 Paris, France.

[2] Programme Cartes d'Identité des Tumeurs (CIT), Ligue Nationale Contre Le Cancer, Paris, France

[3] Institut de Pathologie, Centre de Biologie-Pathologie, CHRU de Lille, F-59037 Lille, France.

[4] Université de Lille, F-59045 Lille, France.

[5] Laboratoire de Pathologie Clinique et Expérimentale (LPCE) et Biobanque (BB-0033-00025), CHRU de Nice, F-06003 Nice, France.

[6] FHU OncoAge, Université Côte d'Azur, F-06108 Nice, France.

[7] Assistance Publique-Hôpitaux de Paris, Hôpital Européen Georges Pompidou, F-75015, Paris, France.

[8] Service d'Anatomopathologie et cytologie, Hôpital Européen Georges Pompidou, F-75015, Paris, France.

[9] Service de Chirurgie Thoracique, Hôpital Calmette - CHRU de Lille, F-59000, Lille, France.

[10] Service de Pneumologie et Pathologie Professionnelle, Centre Hospitalier Intercommunal de Créteil, F-94010 Créteil, France.

[11] Service de Pneumologie et d'Oncologie Thoracique, Hôpital Calmette - CHRU de Lille, F-59000, Lille, France.

[12] Réseau National Expert pour le Mésothéliome Pleural Malin (MESOCLIN), F-59000, Lille, France.

[13] Service de Chirurgie Thoracique, Hôpital Européen Georges Pompidou, F-75015, Paris, France.

[*] These authors contributed equally to this work.

[†] Present address: Service de Chirurgie Générale et Thoracique, CHU de Rouen, F-76000, France.

[‡] Present address: Laboratoire de Biochimie (LBC), ESPCI Paris, PSL Research University, CNRS UMR8231 Chimie Biologie Innovation, F-75005, Paris, France.

[¥] Present address: Service de Chirurgie Thoracique et Transplantation Pulmonaire, Hôpital Foch, Suresnes, F-92150, France.

Corresponding author

Didier JEAN

Centre de Recherche des Cordeliers - INSERM UMR-1138

Equipe 28 « Génomique fonctionnelle des tumeurs solides »

27, rue Juliette Dodu

F-75010 Paris

France

Phone: +33 (0)1 44 27 80 95

E-mail address: **didier.jean@inserm.fr**

Running title

Mesothelioma mutations linked to heterogeneity and prognosis

Keywords

Thoracic cancer, gene mutations, tumor heterogeneity, tumor molecular classification, prognosis.

Abbreviations

MPM: Malignant pleural mesothelioma; **MME**: epithelioid MPM; **MMB**: biphasic MPM; **MMS**: sarcomatoid MPM; **MMD**: desmoplastic MPM; **TSG**: Tumor suppressor gene; **TERT_prom**: TERT promoter.

Abstract

Development of precision medicine for malignant pleural mesothelioma (MPM) requires a deep knowledge of tumor heterogeneity. Histologic and molecular classifications and histo-molecular gradients have been proposed to describe heterogeneity, but a deeper understanding of gene mutations in the context of MPM heterogeneity is required and the associations between mutations and clinical data need to be refined.

We characterized genetic alterations on one of the largest MPM series (266 tumor samples), well-annotated with histologic, molecular and clinical data of patients. Targeted next generation sequencing was performed focusing on the major MPM mutated genes and the *TERT* promoter. Molecular heterogeneity was characterized using predictors allowing classification of each tumor into the previously described molecular subtypes and the determination of the proportion of epithelioïd-like and sarcomatoïd-like components (E/S.scores).

The mutation frequencies are consistent with literature data, but this study emphasized that *TERT* promoter, not considered by previous large sequencing studies, was the third locus most affected by mutations in MPM. Mutations in *TERT* promoter, *NF2* and *LATS2* were more frequent in non-epithelioid MPM and positively associated to the S.score. *BAP1*, *NF2*, *TERT* promoter, *TP53* and *SETD2* mutations were enriched in some molecular subtypes. *NF2* mutation rate was higher in asbestos unexposed patient. *TERT* promoter, *NF2* and *TP53* mutations were associated with a poorer overall survival.

Our findings lead to a better characterization of MPM heterogeneity by identifying new significant associations between mutational status and histologic and molecular heterogeneity. Strikingly, we highlight the strong association between new mutations and overall survival.

Molecular Oncology (2020) © 2020 The Authors. Published by FEBS Press and John Wiley & Sons Ltd.

1 Introduction

Malignant pleural mesothelioma (MPM) is a rare, severe and rarely curable tumor arising in the pleura. MPM development is associated with occupational asbestos exposure that is the main etiological factor and remains a major public health concern even in countries that have banned asbestos. The evolution of our knowledge about tumor pathology taught us that MPM, such as other tumor types, presents specific molecular specificities for each patient. The lack of effective curative treatment for this cancer highlights the need to improve our knowledge of molecular alterations in the context of MPM heterogeneity with the aim to further design adapted therapeutic strategies and to implement precision medicine for this cancer.

The heterogeneity of MPM between patients was described at the clinical, histologic and molecular levels. Histology defines three major types: epithelioid (MME), sarcomatoid (MMS) and biphasic (MMB). However, this classification in three types partially reflects the tumor heterogeneity at both the molecular and clinical levels (Jean et al., 2012), and different histologic subtypes are also described (Husain et al., 2013). Recent researches based on large-scale genomic studies have identified molecular subtypes that go beyond the histologic classification (Bueno et al., 2016; de Reynies et al., 2014; Hmeljak et al., 2018). A first classification in two transcriptomic subtypes (C1 and C2) allowed separating MME, the most frequent histologic type, according to patient outcome, the MME with a better prognosis being classified in C1 and the MME with a worse prognosis in C2 (de Reynies et al., 2014). Classifications in four transcriptomic subtypes were also proposed and were associated to histology and prognosis (Blum et al., 2019; Bueno et al., 2016). The study based on The Cancer Genome Atlas (TCGA) data also identified four novel prognostic subgroups of MPM (iClusters 1 to 4) using an integrative multi-omics classification (Hmeljak et al., 2018). More recently, we described a new way to take into account MPM heterogeneity based on a deconvolution approach, which allows to define epithelioid-like and sarcomatoid-like entities and to determine their proportions, the E.score and the S.score, respectively, in a given tumor sample. These E/S.scores are highly associated with the prognosis and may have an impact on personalized therapeutic strategies in MPM, particularly targeted therapies and immunotherapies (Blum et al., 2019). A recent publication has also shown that MPM heterogeneity is well described by a continuum (Alcala et al., 2019).

A further step requires the integration of somatic mutations in the histologic and molecular profiles of tumors will permit to take into account genetic alterations of key genes, extend our biological information on the tumors and improve the transfer to clinical practices (Horlings et al., 2015; Huntsman and Ladanyi, 2018). The genetic landscape of MPM, recently specified by Next Generation Sequencing

studies (NGS) (Bueno et al., 2016; Hmeljak et al., 2018), confirmed the complexity and the heterogeneity of MPM mutation profiles between patients already suggested by previous sequencing studies (Andujar et al., 2016). However, the links between the mutation profile and, on one hand, the clinical characteristic of patients and, on the other hand, MPM heterogeneity are not clearly established. Some associations between specific gene mutations and clinical data or histologic and molecular subtypes have been already described, such as the association of *TP53* mutations with survival (Bueno et al., 2016), *TERT* promoter (*TERT*_prom) mutations with histology (Tallet et al., 2014) or *BAP1* mutations with the C1 and the iCluster 1 subtypes, both enriched in MME tumor (de Reynies et al., 2014; Hmeljak et al., 2018). However, these associations deserved to be validated and new associations explored in independent large MPM series.

In order to refine the association between mutations and clinical data and to better define the MPM heterogeneity at the genetic level, we made a deep characterization of the mutations in 21 genes of interest selected based on literature data including the major MPM mutated genes (*BAP1*, *NF2*, *TP53*, *SETD2*, *LATS2* …) and the *TERT*_prom using a collection of 266 MPM tumor samples with extended clinical annotations. To characterize the association between the mutations and the molecular heterogeneity described in MPM, tumor samples of the same cohort were classified into 2 and 4 subtypes, and the proportions of molecular components E/S.scores were determined.

2 Materials and Methods

2.1 MPM frozen tumors samples

The Inserm series was exclusively composed of frozen MPM tumor samples collected from 266 patients from biobanks of French hospitals (CHRU of Lille and Nice, Hôpital Européen Georges Pompidou of Paris and Centre Hospitalier Intercommunal of Créteil) linked to the French Mesobank network and certified by Mesopath as MPM (Galateau-Salle et al., 2014). This tumor collection (Inserm series) included biopsies or surgery resections of patients diagnosed between 2001 and 2017. The experiments were undertaken with the understanding and written consent of each subject. The study methodologies were conformed to the standards set by the Declaration of Helsinki and approved by a local medical ethics committee (CPP Ile-de-France II). The collected samples were registered in a database (DC-2016-2771) validated by the French research ministry. Samples were annotated with detailed clinico-pathological and epidemiologic information obtained from pathology reports (**Table 1**; **Table S1**). Overall survival was calculated based on the initial date of diagnosis. The percentage of tumor cells in MPM samples was estimated by histologic examination by the pathology department of each hospital. For 21 patients, it was possible to obtain several tumor samples from diagnostic biopsies, surgery resections (Extended

Pleurectomy/Decortication or Extra-pleural pneumonectomy) and tumor samples of recurrence. MPM primary cell lines were established in our laboratory from 12 tumor samples.

2.2 MPM primary cell cultures

MPM primary cell lines were established in our laboratory from 12 tumor samples included in the Inserm series and cultured based on a previous established protocol (Zeng et al., 1993). Briefly, fresh Malignant pleural mesothelioma (MPM) tumor samples were reduced into pieces of less than 0.5 mm3 with a scalpel and transferred to a 24-well tissue-culture plate (TPP, Dutscher, Issy les Moulineaux, France) for adhesion in culture medium containing RPMI 1640, GlutaMAX™, HEPES buffer supplemented with 10% fetal bovine serum and 1% Penicillin-Streptomycin (Gibco, Thermo Fisher Scientific, Villebon sur Yvette, France). Cultures were examined with a phase-contrast microscope to detect monolayer growth. When the cells were confluent, a trypsin-EDTA mixture (Trypsine/EDTA 0.05% in PBS, Gibco, Thermo Fisher Scientific) was used to detach the cells that were then amplified first in a 25 cm2 and then in a 75 cm2 flask (TPP). Cells were sub-cultured approximately every 2 weeks, depending on the cell line. MPM primary cell cultures were frozen in complete RPMI medium described before supplemented with 10% dimethyl sulfoxide (DMSO). When necessary, the cells are thawed rapidly at 37°C, washed with 15 mL of complete RPMI medium, centrifuged and cultured in flasks. The medium is changed after 24 hours and then every 3 days. The cultures were used between passages 6 and 10.

2.3 DNA and RNA extraction

Genomic DNA and total RNA were extracted using a standard isopropanol precipitation procedure and Trizol (Thermo Fisher Scientific), respectively or using the AllPrep DNA/RNA/miRNA Universal kit (Qiagen, Courtaboeuf, France) according to the manufacturer's protocol.

2.4 Primer design for targeted sequencing

The design of the primers was performed on the genome assembly GRCh37 (hg19). The primer pairs specific for each gene were designed to amplify regions of up to 300bp covering all the exon sequences at least twice with overlapping amplicons and thus at least 150bp of adjacent intron sequences on either side using the Primer3 program (Untergasser et al., 2012). The primers were also tested by PCR *in silico* (UCSC) and SNPcheck3 (https://secure.ngrl.org.uk/SNPCheck/snpcheck.htm) to contain no SNPs with a frequency greater than 5% in the general population. Universal adapter sequences (one for sense primers and another for antisense primers) were added to each primer used for library preparation and sequencing.

2.5 Gene sequencing

Genomic DNA was quantitated using Hoechst dyes and a microplate reader. Gene sequencing was performed on a MiSeq® System (Illumina, Evry, France). Targeted sequencing focused on 21 genes and the *TERT*_prom (**Table S2**) with an expected coverage of 400X. Only tumor samples with tumor content at least of 10% after histologic examination were used. Two types of libraries were generated by PCR, one covering the entire target regions ("whole library"), the other specific to GC-rich regions difficult to amplify ("GC-rich library"). The protocol of sequencing was established in our laboratory and slightly modified compared to the published protocol (Calderaro et al., 2017). For each library (whole and GC-rich), 300 ng of DNA were air dried at room temperature. Enrichment was performed on a Fluidigm Biomark® instrument using ROX dye, Fluidigm loading reagent, and either TaqMan® preamp master mix for whole library or Qiagen Multiplex PCR kit and DMSO for GC-rich libraries. The primers pairs (1715 and 316 for the whole and the GC-rich libraries, respectively) were multiplexed at 36 primer pairs per well for the whole library and 7 for the GC-rich library. Amplification programs were: 95°C 5', 98°C 2', and during 16 cycles: 98°C 15'', 60°C 4' for whole library, and 94°C 15' then 3 times 2 cycles: 94°C 15'', 62°C (61°C et 60°C for the next rounds) 30'', 72°C 2' then 2 times 3 cycles: 94°C 15'', 59°C (58°C for the next round) 30'', 72°C 2' then 2 times 4 cycles: 94°C 15'', 57°C 30'' (56°C for the next round), 72°C 2', then 72°C 5' for GC-rich library. After purification with Agencourt AMPure® XP magnetic beads (Beckman-Coulter, Villepinte, France), products were PCR amplified with adaptors containing unique index sequences and sequencing adaptors (P5/P7 sequences) on a GeneAmp® PCR 9700 system (Applied Biosystems, Thermo Fisher Scientific). Amplification programs were: 98°C 30'', during 7 cycles: 98°C 10'', 60°C 20'', 72°C 60'' then 72°C 5' for whole library, and 94°C 15', during 12 cycles: 94°C 15'', 60°C 30'', 72°C 2' then 72°C 5' for GC-rich library. A second sizing procedure was performed and PCR products were quantified (relative quantification) on an Applied Biosystems® 7900HT Fast Real-Time PCR System with Syber Green® and probes targeting P5/P7 sequences. After equimolar pooling of samples, DNA was concentrated using Agencourt AMPure XP magnetic beads. Libraries were then quantified (absolute quantification) using KAPA® Library Quantification Kits (Roche, Rosny-sous-Bois, France) following manufacturer's instructions. Libraries are then loaded on flow cells and processed on the MiSeq® System (Illumina) according to manufacturer's instructions. Sequencing was paired-end in all cases.

2.6 Sequencing data analysis

The downstream data analyses were performed on the Illumina FASTQ files generated by the Illumina MiSeq Reporter software (version 2.5.1). The primer sequences were removed using

fastx_trimmer function from the fastx Toolkit (v0.0.14). The reads were aligned on the genome assembly GRCh37 (hg19) using BWA version 0.7.5a and bam files were generated using samtools v1.3 (Li and Durbin, 2009; Li et al., 2009). GATK local realignment around known indels and GATK base quality score recalibration was used to recalibrate the reads in BAM files according the GATK best practices (version 3.5). After the alignment step, coverage statistics were generated using GATK DepthOfCoverage algorithm restricted to the targeted coding sequences. These statistics were used to evaluate the quality of each sample sequencing. Both snv and indel variants were called using Unified Genotyper with default arguments, except that no downsampling was done. Finally, functional effects were predicted using the Oncotator annotation algorithm, along with the ensembl Variant Effect Predictor (VEP) algorithm using RefSeq sequence database downloaded from NCBI on Jan 26, 2015 and also Annovar annotation (Ramos et al., 2015; Wang et al., 2010).

2.7 Variant classification

To ensure only high-confidence mutation, the resulting vcf files were filtered according to the criteria in the following order: (i) *Low quality*: QUAL < 100, Allelic Depth of Alternative variant < 6, Genotype Quality < 99, Variant Frequency < 0.1; (ii) *Frequent polymorphism*: referenced in databases (ExAC version 0.3, 1000 genomes phase 1, version 3 and gnomAD) over 0.1% (1000 Genomes Project Consortium et al., 2012; Lek et al., 2016); (iii) *Rare polymorphism*: referenced in databases under 0.1%; (iv) *Artefacts*: detected in a series of 71 non-tumoral frozen tissue samples from hepatocellular carcinoma (HCC) patients; (v) *Variant silent*: "Variant_Classification" = Silent; (vi) *Variant in non-coding region*: "Variant_Classification" = Intron, IGR, 3'UTR, 5'UTR, 5'Flank, 3'Flank; (vii) *Variant with structural consequence*: "Variant_Classification" = Nonsense, Missense, Splice site, Inframe deletion, Inframe insertion, Frameshift deletion and Frameshift insertion.

For missense substitutions, we used Polyphen2 HDIV, SIFT, Mutation Taster and CADD to obtain the prediction of the functional impact on the protein and only substitutions with a CADD score superior to 20 were conserved.

Variants with structural consequence were tagged (i) M1: damaging variants (nonsense, splice site, in-frame deletion and insertion, frameshift deletion and insertion, de novo start in-frame); M2: missense substitutions predicted as damaging by the three tools; M3: missense substitutions predicted as damaging by two of the three tools. Finally, all the damaging variants and *TERT*_prom hotspot mutation sites were confirmed using the Integrative Genomics Viewer (IGV) software (Broad Institute).

2.8 Quantitative Real Time PCR (qRT-PCR) analysis

Total RNA (1.5 µg) was reverse transcribed in a final volume of 50 µl using the High Capacity cDNA Reverse Transcription kit (Thermo Fisher Scientific). qRT-PCR reactions were performed using TaqMan probes and the high throughput BioMark HD system (Fluidigm, Les Ulis, France) following manufacturer's instructions. Pre-amplifications of 6 ng cDNA were performed using PreAmp Master Mix (Fluidigm) with a primers mix combining each primer used in the present study except the *18S* probe due to its very high gene expression level. Expression data (Ct values) were acquired using the Fluidigm Real Time PCR Analysis software. The mean of 5 housekeeping genes (*18S*, *ACTB*, *CLTC*, *GAPDH*, *TBP*) was used for the normalization of expression data (ΔCt).

2.9 Classification in molecular subtypes

To assign each sample to the molecular subtypes of the classification in two clusters (C1 and C2) (de Reynies et al., 2014) or four clusters (C1A, C1B, C2A and C2B) (Blum et al., 2019), a 9-gene predictor was developed by the "Cartes d'identité des tumeurs" (CIT) program founded by the French "Ligue Contre le Cancer". This predictor was defined using the qRT-PCR measurements, obtained on a Fluidigm BioMark HD system based on a selection of 69 genes differentially expressed between these subtypes (moderate t.test for the comparison between the 2 main subtypes C1 and C2 and between the subtypes intra C1 and intra C2, with fdr<0.05, absolute Fold Change >1.5 and AUC >0.8) (**Table S3**). Gene selection was computed by the *varSelRF* function of the R-package *varSelRF* on the Fluidigm dataset restricted to the 63 samples previously assigned to each subtype of both classification systems (C1/C2 and C1A/C1B/C2A/C2B) (Blum et al., 2019; Diaz-Uriarte, 2007). In brief, *varSelRF* minimizes the out-of-bag error, by successively eliminating the least important variables from random forests. The number of trees "ntree" was set to 10000 and "ntreeIterat" to 10000, default parameters were used otherwise. This procedure resulted in the following selection: *ADAM19*, *ETS1* and *PDCD1LG2* genes for C1 and C2; *CLDN1*, *DSC3* and *SLC24A3* for C1A and C1B; *CHL1*, *ECM2*, *PTPN13* for C2A and C2B. The predictor was trained using the restricted dataset (63 samples). The subtype prediction was defined by a majority vote across 3 algorithms [DLDA, DQDA (R package sma), PAM (R package pamr) and was applied to the remaining samples of the inserm series. Only tumor samples with tumor content of at least 30% after histologic examination were predicted and only predictions with a synthetized score higher than 60 for a particular subtype were taken into account to prevent misleading classification.

2.10 Estimation of E.score and S.score

E.score and S.score were estimated using the Wisp R package (https://cit-bioinfo.github.io/WISP/), also developed by CIT program, on qRT-PCR data with a signature of 55 genes as detailed elsewhere (Blum

et al., 2019). Only samples with a cumulated E.score and S.score higher than 50% were taken into account to ensure sufficient tumor content for correct estimation and scores were rescaled after removing the non-tumoral component for association analysis with genetic mutations.

2.11 Data and statistical analysis

Mutation frequencies and types were retrieved from the release v87 of COSMIC database (https://cancer.sanger.ac.uk/cosmic) (Tate et al., 2019). Data analysis was performed to separate *TERT*_prom and *TERT* core gene mutations from COSMIC database. Dataset of three others series were used: TCGA (Mesothelioma-TCGA, PanCancer Atlas), GENIE (GENIE Cohort v4.1-public) (AACR Project GENIE Consortium, 2017) and Bueno series (Bueno et al., 2016). For TCGA and GENIE series, mutation and clinical data were retrieved on November, 11, 2018 from cBioPortal for Cancer Genomics (www.cbioportal.org) (Cerami et al., 2012; Gao et al., 2013). TCGA iCluster subtypes and, mutation and clinical data of Bueno series were retrieved from published manuscript (Bueno et al., 2016; de Reynies et al., 2014; Hmeljak et al., 2018). Copy number alterations or fusion, listed in these series, were not taken into account for comparison with the Inserm series data. Clusters comparison of the different molecular classifications was performed by correlating the centroids of their corresponding meta-profiles as described (Blum et al., 2019).

Statistical tests were performed using GraphPad Prism version 6.07 software except Fisher's exact tests for contingency tables and univariate and multivariate Cox regression analysis, for which R statistical software was used. Lolliplots were drawn using mafTools (Mayakonda et al., 2018).

3 Results

3.1 Genetic alterations in MPM

Mutations in 21 key altered genes in MPM and in the *TERT*_prom were determined by targeted sequencing (see **Table S2** for gene selection based on literature data). We identified 200 variants with structural consequences (**Table S1**; **Table S4**). Damaging variants were found in 52.3% (139/266) of the MPM tumors (153/266 i.e. 57.5% taking into account mutations in *TERT*_prom), consistent with the percentage of mutated tumors in TCGA series (48.8%), when restricted to the same set of genes with putative driver mutation. Mutations were found in 19 genes and frequencies are shown in **Fig. 1a**. The top most frequently mutated genes were *BAP1*, *NF2* and *TERT*_prom with a mutation frequency of 24.5%, 19.2% and 12.0%, respectively. The mutation frequencies were similar to those reported in the COSMIC database or other series, except for *TP53* genes (**Fig. 1A**; **Fig. S1**).

The distribution of mutations in MPM are shown in the heat map for the 6 most frequently mutated genes with more than 3% of mutation and *TERT*_prom (**Fig. 1B**). Up to four of these key genes were mutated in a given tumor. A significant co-occurrence of mutations in *TERT*_prom and *NF2* was found (p=0.0004) *i.e. NF2* mutations were more frequent in *TERT*_prom mutant MPM than in wild-type MPM, 45.2% and 15.8%, respectively. Mutually exclusive mutations were found between *BAP1* and *TERT*_prom (p=0.013). Mutations in *BAP1* and *TP53* seem also to be mutually exclusive (not significant) with only 1 tumor case with both mutated genes. Interestingly, mutations in genes of the SWI/SNF family (*ARID1A*, *ARID2* and *SMARCA4*) and genes related to histone methylation (*KMT2D*, *SETD2*) were also mutually exclusive. This was not the case for genes belonging to Hippo signaling pathway (*NF2* and *LATS2*), with 2 MPM showing mutations in both genes, in agreement with our previous observation in MPM primary cell lines (Tranchant et al., 2017).

We mapped the variants on schematic representations of the protein for the 6 most frequently mutated genes (**Fig. 1C**; **Fig. S2**). The proportion of mutation types in Inserm series was also compared to those of COSMIC data (**Fig. S3**). The mutation profiles have typical tumor suppressor gene (TSG) profile. They are enriched in truncating mutations dispersed randomly all along the gene with specificities according to the genes. *BAP1* and *SETD2* showed a high proportion of frameshift deletions. *NF2* showed a majority of nonsense mutations, but very few missense substitutions contrary to the other genes such as *BAP1* or *TP53* that show a substantial proportion of missense substitutions in their functional domains. Our data were consistent with COSMIC in MPM, but were significantly different from COSMIC pan-cancer data for *BAP1* and *NF2* (p=0.008 and p=0.004, respectively), suggesting specific mutation types for both genes in MPM.

Three different mutations were identified in the promoter of *TERT*, one in the core promoter (9.7%), that was also the most frequently found in cancers, and two in the 5'-untranslated region (UTR) of *TERT* mRNA (1.9% and 0.4%) (**Fig. 1D**). We verified that these mutations were associated to overexpression of *TERT* mRNA as previously described (Tallet et al., 2014). As expected, MPM with *TERT*_prom mutation including mutations in the *TERT* 5'UTR showed a higher expression of *TERT* mRNA than wild type MPM (p=0.0015) (**Fig. S4A**).

3.2 Associations between mutation profile and MPM heterogeneity

MPM heterogeneity was characterized at the histologic and molecular level in Inserm series. First, we focused on the link between the mutation profile, considering the 6 most frequently mutated genes and *TERT*_prom, and the main three MPM histologic types *i.e.* MME, MMB and MMS. Mutations in *TERT_prom*, *NF2* and *LATS2* were significantly less frequent in MME than in non_MME samples (MMB and

MMS) (**Fig. 2A**; **Fig. S5A**). Associations between *NF2* and *LATS2* mutation profiles and histologic types were confirmed by analyzing three other MPM series: Bueno, TCGA and GENIE series (**Fig. S5B-F**).

Recently, we described a new method to take into account MPM heterogeneity based on a deconvolution approach on gene expression, which allows to define epithelioid-like and sarcomatoid-like components and to determine their proportions, the E.score and the S.score, respectively, in a given tumor sample (Blum et al., 2019). These scores were determined in Inserm series (**Table S1**) and their associations with the mutational profile were investigated. Significant positive associations were found between the S.score and *TERT_prom*, *NF2* and *LATS2* mutations, the same mutations associated with the non_MME histologic subtypes (**Fig. 2B**). The E.score and the S.score were also estimated for the MPM samples of Bueno and TCGA series (Blum et al., 2019). Analysis of *NF2* and *LATS2* mutation data showed also significant statistic associations in these series **Fig. S6**).

In a previous publication, we defined a molecular classification of MPM in two subtypes C1 and C2 (de Reynies et al., 2014). More recently, we also defined intra-subtypes (C1A, C1B, C2A and C2B) by subdividing C1 and C2 into two groups (Blum et al., 2019). We compared these intra-subtypes with subtypes already published (Bueno et al., 2016; Hmeljak et al., 2018) and those defined by unsupervised clustering from other public series (Lopez and Gordon series) (Blum et al., 2019) by correlating the centroids of their corresponding meta-profiles (**Fig. S7A**). The correlation matrix showed two main groups of highly correlated clusters present in all datasets, which contained the C1A subtype associated with Epithelioid and iCluster 1 subtypes and the C2B subtype associated with Sarcomatoid and iCluster 4 subtypes from Bueno and TCGA series, respectively. Tumor samples of Inserm series were predicted in 2- and 4- subtypes molecular classification using a predictor based on 9 genes : *ADAM19, ETS1* and *PDCD1LG2* for C1 and C2; *CLDN1*, *DSC3* and *SLC24A3* for C1A and C1B; *CHL1*, *ECM2*, *PTPN13* for C2A and C2B subtype prediction (**Fig. S8**). We showed a significant enrichment of *BAP1* mutations in the C1 subtype, and of *TERT*_prom in the C2 subtype (**Fig. 2C**). Distribution of mutations between intra-subtypes (C1A, C1B, C2A and C2B) are shown in **Fig. S7B**. Comparison of mutation profiles between C1A and C2B subtypes highlighted increase frequency of *BAP1* and *SETD2* mutations in C1A subtype and, *TERT*_prom, *NF2*, *TP53*, *ARID2* and *LATS2* mutations in C2B subtypes (**Fig. 2D**). These associations were only significant for *BAP1*, *TP53* and *TERT*_prom in Inserm series, but pan series analysis showed that these associations were all significant except for *LATS2* and *ARID2* (**Fig. S7C-I**).

3.3 Associations between mutation profile and clinical and epidemiological data

We did not find a significant association between mutation profile and age, gender or tobacco consumption. *NF2* mutations were significantly associated to asbestos exposure status (p=0.036) and

were more frequent in non-exposed patients (31%) than in exposed patient (17%) (**Fig. S9A**). *TERT*_prom and *NF2* mutations were significantly associated to the tumor stage (p=0.025 and p=0.007, respectively) and showed significant higher mutation rate in patients with stage IV tumors (20% and 28%, respectively) than in patients with stage I/III tumors (9% and 13%, respectively) (**Fig. S9B**). Strong significant associations were observed with overall survival. Overall survival frequency was lower in patients with MPM mutated for *TERT*_prom, *TP53* and *NF2* compared to patients with MPM wild-type (**Fig. 3a**). These associations are also found when considering only MME samples and non_MME samples except for *TP53* in non_MME samples (**Fig. S10**). Multivariate analysis considering age at diagnostic, tumor stage, histology, S.score based on a threshold of 0.22, which was shown to be the more predictive for prognosis (Blum et al., 2019), and mutation status in these three genes showed that the mutation status is predictive of prognosis only for *TP53* and not for *TERT*_prom and *NF2* (**Fig. S11**). However, multivariate analysis considering MPM samples with at least one mutation in *TP53*, *NF2* or *TERT*_prom highlighted the strong prognosis value of all three genes considered together (**Fig. 3B**).

3.4 Mutations profile between samples from the same patient

Our MPM collection includes several different tumor samples including diagnostic biopsies, surgery resections and tumor samples of recurrence from the same patient: (i) Eleven pairs of tumor samples corresponding to diagnostic biopsies and surgery resections, with neoadjuvant therapy, collected with a spacing of 2.9 to 5.4 months. (ii) Six pairs of samples corresponding to diagnostic biopsies and surgery resections, without neoadjuvant therapy, collected with spacing of 1.1 to 2.4 months. (iii) Four pairs of tumor samples corresponding to primary tumors and recurrence tumors, with a spacing of 18.1 to 161.5 months (**Table S5**). Mutations were identified in 12 sample pairs and there was no difference between mutational status between both samples in all the cases. We also established 12 MPM primary cell lines from tumor samples included in the Inserm series. Among 6 sample pairs with characterized mutations, mutations were identical between tumor samples and cell lines (**Fig. 4**).

4 Discussion

In previous studies, we defined a transcriptomic molecular classification of MPM and investigated inter- and intra-tumor heterogeneity using a deconvolution approach (Blum et al., 2019; de Reynies et al., 2014). The present study characterizes the genetic alterations in the most frequently mutated genes in MPM in a large series of 266 frozen tumors well annotated for clinical and histologic and molecular heterogeneity.

Inserm series was one of the largest series with the one of Bueno *et al*. (Bueno et al., 2016) used so far to screen mutations in key altered genes in MPM, and the largest for *TERT*_prom mutations that are not evaluated by exome sequencing. In the Inserm series, *TERT*_prom is the third locus most affected by mutations in MPM and deserved to be considered in the genetic landscape of MPM given the importance of telomerase upregulation in cancer (Pestana et al., 2017). The present study allows also drawing up an accurate assessment of the frequency of gene mutations in a homogeneous series of MPM. We highlight gene such as *ARID2*, only mentioned as mutated in a previous study on MPM cell lines (Yoshikawa et al., 2015). At the opposite, we show that genes that have been suggested as frequently mutated in small series of MPM are in fact only rarely mutated, such as *CUL1* (Guo et al., 2015). Even if the large majority of the mutations identified in this study are likely somatic mutations, we cannot exclude that some of the mutations correspond to germline mutations especially in *BAP1* gene, the gene with the highest frequency of germline mutations in MPM (Panou et al., 2018; Pastorino et al., 2018). Among the 65 patients showing *BAP1* mutation of Inserm series (**Table S1**), none had previous cutaneous or uveal melanoma, renal cell carcinoma, basal cell carcinoma, meningioma or cholangiocarcinoma, which were recently described in a study on 181 families carrying *BAP1* germline variants as the core tumor spectrum for the *BAP*1 tumor predisposition syndrome (Walpole et al., 2018).

Mutation frequencies are consistent between Inserm series and other series. The only significant discrepancy was for *TP53* mutation frequencies suggesting a variability between MPM series for this gene (**Fig. S1**). Of note, this study did not take into account large exon deletions as the targeted sequencing does not allow detecting accurately large deletions even based on the sequencing depth of coverage. This does not have an impact on *TP53* or *TERT*_prom alterations frequencies, but those of *CDKN2A*, *BAP1* and *NF2* are most likely under-evaluated as it is the case in most NGS studies due to the contamination of tumor samples by normal cells. For *NF2*, we verified that taking into account large deletions did not change the association between *NF2* mutation status and, histologic and molecular subtypes or gradients (**Fig. S12**).

Our detailed mutation analysis also give a precise overview of the mutation types profile observed in the main TSG mutated in MPM and in the *TERT*_prom. *BAP1* and *SETD2* are enriched in deletion consistent with their chromosome localizations in 3p21 region, which harbors multiple noncontiguous minute deletions in MPM (Yoshikawa et al., 2016). *NF2* shows few missense mutations but several mutations leading to the production of truncated protein forms, consistent with its role as a multifunctional protein interacting with several partners through different parts of the protein (Sato and Sekido, 2018). As for *BAP1* and *NF2*, *TERT*_prom mutation types present characteristics specific to MPM. The three mutations in *TERT*_prom were previously described in other cancers and generate *de novo* ETS

binding sites (Huang et al., 2015). However, while the C228T mutation is the most frequent in MPM as in other tumors, the C250T, the second most common mutation in tumors, is not found in MPM. The two others less frequent mutation sites A161C and C1581 in MPM are mostly found in the bladder transitional carcinoma in COSMIC database. We also validated the correlation between promoter mutation and *TERT* overexpression in MPM, observed previously (Tallet et al., 2014). Interestingly, literature data suggest that *BAP1* could downregulate *TERT* expression (Linne et al., 2017). As we found mutually exclusive mutations between both alterations, we compared *TERT* expression based on the *BAP1* mutation status, but did not find any significant association in MPM (**Fig. S4b**).

Most of the mutated genes in MPM are TSG and untargetable gene directly. However, we identified mutations in genes known as being targetable genes (**Table S3**). For example, the mutation in the oncogene *KRAS* Q61H has been referenced in COSMIC database (COSM555) and is a hotspot known to be oncogenic and found in several other malignancies such as large intestine and lung carcinomas. Pre-clinical and preliminary clinical data suggest that cancers with *KRAS*-mutant may be sensitive to MEK or ERK inhibitors (Sullivan et al., 2018). The *NRAS* Q61K mutation is also an oncogenic hotspot (COSM580) and there is promising clinical data in patients with oncogenic *NRAS*-mutant melanoma treated with the MEK1/2-inhibitor, binimetinib (Dummer et al., 2017). Unfortunately, this concerns only a very small subset of MPM patients. For most patients, genes or pathways deregulated as a consequence of TSG inactivation should be targeted, as it was suggested for hippo signal pathway linked to *NF2* and *LATS2* inactivation (Felley-Bosco and Stahel, 2014; Sato and Sekido, 2018).

The key input of this study is to precise the genetic landscape taking into account MPM heterogeneity. We identified significant increase of the mutation rate for *LATS2, NF2* and *TERT*_prom in non_MME allowing us to confirm in a larger MPM series the association with *TERT*_prom previously observed (Tallet et al., 2014) and to demonstrate at the statistical level the previously suggested association with *NF2* (Sato and Sekido, 2018). Mutations in *LATS2*, *NF2* and *TERT*_prom were also positively associated to the S.score, bringing new insights on the intra-tumor heterogeneity and strengthening the link between these mutations and the sarcomatoid cell type. *NF2* and *TP53* mutations were previously found to be associated to the S.score, based on TCGA series data (Blum et al., 2019). Here we confirmed association of *NF2* mutations in this larger Inserm series, possibly due to the variability of *TP53* mutations between MPM series mentioned above. Another important point concerns the occurrence of gene mutations according to the molecular classifications in subtypes. We confirmed the previously described association of *BAP1* mutations to the C1 subtype (de Reynies et al., 2014) and identified a significant enrichment of *TERT*_prom mutations in the C2 subtype. Further classifications in four subtypes were also proposed in several studies (Blum et al., 2019; Bueno et al., 2016; Hmeljak et al.,

2018). Here, we focused on the extreme subtypes, i.e. the C1A/Epithelioid/iCluster 1 and the C2B/Sarcomatoid/iCluster 4 subtypes, as they are detected in all MPM series. Hmeljak *et al*. reported a strong significant association of *BAP1* mutation with iCluster 1 and an enrichment of *LATS2* mutation in iCluster 4 (Hmeljak et al., 2018). We also observed the same associations in Inserm series. Moreover, we highlighted new significant associations between *TERT_prom* and *TP53* mutations, and C2B subtype. Pan series analysis confirmed all these associations and revealed a significant association of *NF2* mutations with C2B subtype. Interestingly, *TERT*_prom and *NF2* mutations are associated with histologic and molecular classifications, and molecular gradients, but not *TP53* and *BAP1* mutations (**Fig. 5**). These results highlight the complexity of MPM heterogeneity and suggest that classification in subtypes even if related to histologic types take into account another degree of heterogeneity. *BAP1* and *TP53* mutated tumors may form specific subtypes inside epithelioid and sarcomatoid enriched tumors, respectively. In a previous study, we demonstrate the impact of epigenetic mechanism in the establishment of epithelioid and sarcomatoid-related cell entities (Blum et al., 2019). Altogether, our new results also highlight the contribution of different genetic related mechanism and support different ways for mesothelial cell neoplastic transformation.

Previous studies reported association between loss of specific chromosome regions and asbestos exposure (Borczuk et al., 2016; Jean et al., 2011). However, to our knowledge, our study is the first to identify a link between gene mutations and asbestos status. In our series, *NF2* mutation was the most frequent alteration in asbestos non-exposed patients with a third of patients carrying *NF2* mutation. Only one study reported a MPM patient with constitutional *NF2* mutation missense mutation (Baser et al., 2005). Furthermore, in two recent studies screening germline cancer susceptibility mutations in large cohort of MPM patients, *NF2* was not identified as a cancer susceptibility gene (Panou et al., 2018; Pastorino et al., 2018), suggesting that *NF2* mutations observed in our series in unexposed patient are likely somatic. This high frequency of mutations supports *NF2* as a key driver of asbestos independent mesothelial carcinogenesis that was previously suggested in mice models. Development of peritoneal mesothelioma was observed in genetically engineered mice heterozygous in *Nf2* without asbestos exposure (Giovannini et al., 2000). Conditional mouse model leading to both *Ink4a/Arf* and *Nf2* inactivation was shown to develop malignant thoracic mesothelioma at a high incidence without asbestos exposure, mostly of sarcomatoid type (Jongsma et al., 2008). Interestingly, *LATS2* is the second gene showing the mutations more frequent in non-exposed patients (7%) than in exposed patient (2%). Both *NF2* and *LATS2* belong to the hippo signal pathway known to be crucial for asbestos driven carcinogenesis and, based on our data, also for asbestos independent mesothelial carcinogenesis.

One of the major strengths of our study is to demonstrate the strong link between the mutation status of *TERT*_prom, *NF2* and *TP53*, and overall survival. Accordingly, *TERT*_prom and *NF2* mutations were significantly more frequent in MPM with an advanced stage. The prognosis interest based on the mutational status was already reported for *TP53* by Bueno *et al*. (Bueno et al., 2016), but not for other genes. Of note, in agreement with the *NF2* mutations prognosis value, an immunohistochemistry study reported that low merlin expression is an indicator for poor prognosis in MPM patients (Meerang et al., 2016). *TERT*_prom mutation has been associated with worse prognosis in some cancers including meningioma but not in MPM (Lu et al., 2019). Multivariate analysis confirm the prognosis value of the S.score and highlight the prognosis value of the three genes together that could be an alternative for evaluating the prognosis in clinic.

The rapid evolution of MPM is challenging for targeted therapy. The comparison of mutation profile of tumor samples collected at different time points from a same patient did not show any difference. One tumor sample pair (T004LE and T288LE), corresponding to primary versus recurrence tumors, showed two *BAP1* mutations (K337fs and N157fs) present in both samples. It is impressive to find the same mutation since the two samples were collected within a 13-years interval of time. We can not completely exclude a germline mutation in *BAP1*, but to our knowledge two co-occurring *BAP1* germline mutations were not previously identified in a same family. Obviously, this stability over time of the mutational profile should be confirmed in larger series. As MPM cell lines are useful for studying mesothelial carcinogenesis and for identifying new therapies by testing anti-cancer drugs, we also characterized the mutations in primary cell lines established from sequenced tumor samples. Our results show that mutations in primary cell lines are representative from the mutations present in the tumor of the patient.

5 Conclusion

Overall, the present study provide a comprehensive overview of the genetic landscape of MPM taking into account the histologic and molecular heterogeneities. This better understanding of heterogeneity at the genetic level should facilitate the implementation of strategies to develop precision medicine for MPM, which is crucial for this incurable cancer. Our findings also highlight the strong prognostic value of genetic alterations relevant for clinical application.

Acknowledgements

This work was supported by Inserm, the Ligue Contre le Cancer (Ile de France committee and national program Cartes d'Identité des Tumeurs (CIT)), the Fondation ARC pour la recherche sur le cancer, Hadassah France, and the Chancellerie des Universités de Paris (Legs POIX). LQ, JBA, FM and JdW were supported by grants from Cancéropôle Région Île-de-France, from Fondation pour la Recherche Médicale (FRM), from the Institut thématique multi-organismes (ITMO) Cancer (Plan Cancer 2014-2019) and from the Fondation ARC, respectively. JdW and FM were also supported by grant of the Société Française de Chirurgie Thoracique et Cardio-Vasculaire (SFCTCV). The Inserm research unit is supported by the Labex OncoImmunology (investissement d'avenir), Coup d'Elan de la Fondation Bettencourt-Shueller, the SIRIC CARPEM, the Institut thématique multi-organismes (ITMO) Cancer (Plan Cancer 2014-2019) and Cancéropôle Île-de-France.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Author contributions

CM, JZR, MCJ and DJ are responsible for the study concept and design. VH, PH, MCC, EP, AS, IM, LG, CB, FLPB provided patient materials. VH, PH, FM, JdW and JBA collected and organized clinical data. VH, MCC, LG, CB performed histologic analysis of tumor samples. LQ, CM, RT, JdW, FM, JBA contributed to sample preparation and quality control. LQ, CM performed the acquisition of genetic data. SC, YB, CM and SI developed the bioinformatics tools and the pipeline analysis. YB, CM, LQ and DJ performed the analysis and interpretation of data. CM, LQ, MCJ and DJ were major contributors in writing the manuscript. FLPB, JZR, MCJ and DJ are responsible for the study supervision. All authors read and approved the final manuscript.

References

1000 Genomes Project Consortium, Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., McVean, G.A., 2012. An integrated map of genetic variation from 1,092 human genomes. Nature. 491, 56-65.

AACR Project GENIE Consortium, 2017. AACR Project GENIE: Powering Precision Medicine through an International Consortium. Cancer Discov. 7, 818-831.

Alcala, N., Mangiante, L., Le-Stang, N., Gustafson, C.E., Boyault, S., Damiola, F., Alcala, K., Brevet, M., Thivolet-Bejui, F., Blanc-Fournier, C., et al., 2019. Redefining malignant pleural mesothelioma types as a continuum uncovers immune-vascular interactions. EBioMedicine. 48, 191-202.

Andujar, P., Lacourt, A., Brochard, P., Pairon, J.C., Jaurand, M.C., Jean, D., 2016. Five years update on relationships between malignant pleural mesothelioma and exposure to asbestos and other elongated mineral particles. J Toxicol Environ Health B Crit Rev. 19, 151-172.

Baser, M.E., Rai, H., Wallace, A.J., Evans, D.G., 2005. Neurofibromatosis 2 (NF2) and malignant mesothelioma in a man with a constitutional NF2 missense mutation. Fam Cancer. 4, 321-322.

Blum, Y., Meiller, C., Quetel, L., Elarouci, N., Ayadi, M., Tashtanbaeva, D., Armenoult, L., Montagne, F., Tranchant, R., Renier, A., et al., 2019. Dissecting heterogeneity in malignant pleural mesothelioma through histo-molecular gradients for clinical applications. Nat Commun. 10, 1333.

Borczuk, A.C., Pei, J., Taub, R.N., Levy, B., Nahum, O., Chen, J., Chen, K., Testa, J.R., 2016. Genome-wide analysis of abdominal and pleural malignant mesothelioma with DNA arrays reveals both common and distinct regions of copy number alteration. Cancer Biol Ther. 17, 328-335.

Bueno, R., Stawiski, E.W., Goldstein, L.D., Durinck, S., De Rienzo, A., Modrusan, Z., Gnad, F., Nguyen, T.T., Jaiswal, B.S., Chirieac, L.R., et al., 2016. Comprehensive genomic analysis of malignant pleural mesothelioma identifies recurrent mutations, gene fusions and splicing alterations. Nat Genet. 48, 407-416.

Calderaro, J., Couchy, G., Imbeaud, S., Amaddeo, G., Letouze, E., Blanc, J.F., Laurent, C., Hajji, Y., Azoulay, D., Bioulac-Sage, P., et al., 2017. Histological subtypes of hepatocellular carcinoma are related to gene mutations and molecular tumour classification. J Hepatol. 67, 727-738.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al., 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2, 401-404.

de Reynies, A., Jaurand, M.C., Renier, A., Couchy, G., Hysi, I., Elarouci, N., Galateau-Salle, F., Copin, M.C., Hofman, P., Cazes, A., et al., 2014. Molecular classification of malignant pleural mesothelioma: identification of a poor prognosis subgroup linked to the epithelial-to-mesenchymal transition. Clin Cancer Res. 20, 1323-1334.

Diaz-Uriarte, R., 2007. GeneSrF and varSelRF: a web-based tool and R package for gene selection and classification using random forest. BMC Bioinformatics. 8, 328.

Dummer, R., Schadendorf, D., Ascierto, P.A., Arance, A., Dutriaux, C., Di Giacomo, A.M., Rutkowski, P., Del Vecchio, M., Gutzmer, R., Mandala, M., et al., 2017. Binimetinib versus dacarbazine in patients with advanced NRAS-mutant melanoma (NEMO): a multicentre, open-label, randomised, phase 3 trial. Lancet Oncol. 18, 435-445.

Felley-Bosco, E., Stahel, R., 2014. Hippo/YAP pathway for targeted therapy. Transl Lung Cancer Res. 3, 75-83.

Galateau-Salle, F., Gilg Soit Ilg, A., Le Stang, N., Brochard, P., Pairon, J.C., Astoul, P., Frenay, C., Blaizot, G., Chamming's, S., Ducamp, S., et al., 2014. [The French mesothelioma network from 1998 to 2013]. Annales de pathologie. 34, 51-63.

Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al., 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 6, pl1.

Giovannini, M., Robanus-Maandag, E., van der Valk, M., Niwa-Kawakita, M., Abramowski, V., Goutebroze, L., Woodruff, J.M., Berns, A., Thomas, G., 2000. Conditional biallelic Nf2 mutation in the mouse promotes manifestations of human neurofibromatosis type 2. Genes Dev. 14, 1617-1630.

Guo, G., Chmielecki, J., Goparaju, C., Heguy, A., Dolgalev, I., Carbone, M., Seepo, S., Meyerson, M., Pass, H.I., 2015. Whole-exome sequencing reveals frequent genetic alterations in BAP1, NF2, CDKN2A, and CUL1 in malignant pleural mesothelioma. Cancer Res. 75, 264-269.

Hmeljak, J., Sanchez-Vega, F., Hoadley, K.A., Shih, J., Stewart, C., Heiman, D., Tarpey, P., Danilova, L., Drill, E., Gibb, E.A., et al., 2018. Integrative Molecular Characterization of Malignant Pleural Mesothelioma. Cancer Discov. 8, 1548-1565.

Horlings, H.M., Shah, S.P., Huntsman, D.G., 2015. Using Somatic Mutations to Guide Treatment Decisions: Context Matters. JAMA Oncol. 1, 275-276.

Huang, D.S., Wang, Z., He, X.J., Diplas, B.H., Yang, R., Killela, P.J., Meng, Q., Ye, Z.Y., Wang, W., Jiang, X.T., et al., 2015. Recurrent TERT promoter mutations identified in a large-scale study of multiple tumour types are associated with increased TERT expression and telomerase activation. Eur J Cancer. 51, 969-976.

Huntsman, D.G., Ladanyi, M., 2018. The molecular pathology of cancer: from pan-genomics to post-genomics. J Pathol. 244, 509-511.

Husain, A.N., Colby, T., Ordonez, N., Krausz, T., Attanoos, R., Beasley, M.B., Borczuk, A.C., Butnor, K., Cagle, P.T., Chirieac, L.R., et al., 2013. Guidelines for pathologic diagnosis of malignant mesothelioma:

2012 update of the consensus statement from the International Mesothelioma Interest Group. Arch Pathol Lab Med. 137, 647-667.

Jean, D., Daubriac, J., Le Pimpec-Barthes, F., Galateau-Salle, F., Jaurand, M.C., 2012. Molecular changes in mesothelioma with an impact on prognosis and treatment. Arch Pathol Lab Med. 136, 277-293.

Jean, D., Thomas, E., Manie, E., Renier, A., de Reynies, A., Lecomte, C., Andujar, P., Fleury-Feith, J., Galateau-Salle, F., Giovannini, M., et al., 2011. Syntenic Relationships between Genomic Profiles of Fiber-Induced Murine and Human Malignant Mesothelioma. Am J Pathol. 178, 881-894.

Jongsma, J., van Montfort, E., Vooijs, M., Zevenhoven, J., Krimpenfort, P., van der Valk, M., van de Vijver, M., Berns, A., 2008. A conditional mouse model for malignant mesothelioma. Cancer Cell. 13, 261-271.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al., 2016. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 536, 285-291.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Genome Project Data Processing, S., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25, 2078-2079.

Linne, H., Yasaei, H., Marriott, A., Harvey, A., Mokbel, K., Newbold, R., Roberts, T., 2017. Functional role of SETD2, BAP1, PARP-3 and PBRM1 candidate genes on the regulation of hTERT gene expression. Oncotarget. 8, 61890-61900.

Lu, V.M., Goyal, A., Lee, A., Jentoft, M., Quinones-Hinojosa, A., Chaichana, K.L., 2019. The prognostic significance of TERT promoter mutations in meningioma: a systematic review and meta-analysis. J Neurooncol. 142, 1-10.

Mayakonda, A., Lin, D.C., Assenov, Y., Plass, C., Koeffler, H.P., 2018. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 28, 1747-1756.

Meerang, M., Berard, K., Friess, M., Bitanihirwe, B.K., Soltermann, A., Vrugt, B., Felley-Bosco, E., Bueno, R., Richards, W.G., Seifert, B., et al., 2016. Low Merlin expression and high Survivin labeling index are indicators for poor prognosis in patients with malignant pleural mesothelioma. Mol Oncol. 10, 1255-1265.

Panou, V., Gadiraju, M., Wolin, A., Weipert, C.M., Skarda, E., Husain, A.N., Patel, J.D., Rose, B., Zhang, S.R., Weatherly, M., et al., 2018. Frequency of Germline Mutations in Cancer Susceptibility Genes in Malignant Mesothelioma. J Clin Oncol. 36, 2863-2871.

Pastorino, S., Yoshikawa, Y., Pass, H.I., Emi, M., Nasu, M., Pagano, I., Takinishi, Y., Yamamoto, R., Minaai, M., Hashimoto-Tamaoki, T., et al., 2018. A Subset of Mesotheliomas With Improved Survival Occurring in Carriers of BAP1 and Other Germline Mutations. J Clin Oncol. JCO2018790352.

Pestana, A., Vinagre, J., Sobrinho-Simoes, M., Soares, P., 2017. TERT biology and function in cancer: beyond immortalisation. J Mol Endocrinol. 58, R129-R146.

Ramos, A.H., Lichtenstein, L., Gupta, M., Lawrence, M.S., Pugh, T.J., Saksena, G., Meyerson, M., Getz, G., 2015. Oncotator: cancer variant annotation tool. Hum Mutat. 36, E2423-2429.

Sato, T., Sekido, Y., 2018. NF2/Merlin Inactivation and Potential Therapeutic Targets in Mesothelioma. Int J Mol Sci. 19, pii: E988.

Sullivan, R.J., Infante, J.R., Janku, F., Wong, D.J.L., Sosman, J.A., Keedy, V., Patel, M.R., Shapiro, G.I., Mier, J.W., Tolcher, A.W., et al., 2018. First-in-Class ERK1/2 Inhibitor Ulixertinib (BVD-523) in Patients with MAPK Mutant Advanced Solid Tumors: Results of a Phase I Dose-Escalation and Expansion Study. Cancer Discov. 8, 184-195.

Tallet, A., Nault, J.C., Renier, A., Hysi, I., Galateau-Salle, F., Cazes, A., Copin, M.C., Hofman, P., Andujar, P., Le Pimpec-Barthes, F., et al., 2014. Overexpression and promoter mutation of the TERT gene in malignant pleural mesothelioma. Oncogene. 33, 3748-3752.

Tate, J.G., Bamford, S., Jubb, H.C., Sondka, Z., Beare, D.M., Bindal, N., Boutselakis, H., Cole, C.G., Creatore, C., Dawson, E., et al., 2019. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res. 47, D941-D947.

Tranchant, R., Quetel, L., Tallet, A., Meiller, C., Renier, A., de Koning, L., de Reynies, A., Le Pimpec-Barthes, F., Zucman-Rossi, J., Jaurand, M.C., et al., 2017. Co-occurring mutations of tumor suppressor genes, LATS2 and NF2, in malignant pleural mesothelioma. Clin Cancer Res. 23, 3191-3202.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3--new capabilities and interfaces. Nucleic Acids Res. 40, e115.

Walpole, S., Pritchard, A.L., Cebulla, C.M., Pilarski, R., Stautberg, M., Davidorf, F.H., de la Fouchardiere, A., Cabaret, O., Golmard, L., Stoppa-Lyonnet, D., et al., 2018. Comprehensive Study of the Clinical Phenotype of Germline BAP1 Variant-Carrying Families Worldwide. J Natl Cancer Inst. 110, 1328-1341.

Wang, K., Li, M., Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38, e164.

Yoshikawa, Y., Emi, M., Hashimoto-Tamaoki, T., Ohmuraya, M., Sato, A., Tsujimura, T., Hasegawa, S., Nakano, T., Nasu, M., Pastorino, S., et al., 2016. High-density array-CGH with targeted NGS unmask multiple noncontiguous minute deletions on chromosome 3p21 in mesothelioma. Proc Natl Acad Sci U S A. 113, 13432-13437.

Yoshikawa, Y., Sato, A., Tsujimura, T., Otsuki, T., Fukuoka, K., Hasegawa, S., Nakano, T., Hashimoto-Tamaoki, T., 2015. Biallelic germline and somatic mutations in malignant mesothelioma: multiple mutations in transcription regulators including mSWI/SNF genes. Int J Cancer. 136, 560-571.

Zeng, L., Buard, A., Monnet, I., Boutin, C., Fleury, J., Saint-Etienne, L., Brochard, P., Bignon, J., Jaurand, M.C., 1993. In vitro effects of recombinant human interferon gamma on human mesothelioma cell lines. Int J Cancer. 55, 515-520.

Tables

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

	Patients (n=266)
Gender (n [%])	
Male	203 [76]
Female	63 [24]
Age (years)	
Median ± SD	69.0 ± 10.9
Range	20-91
Histology (n [%])	
Epithelioid	201 [78]
Biphasic	30 [12]
Sarcomatoid	21 [8]
Desmoplasic	5 [2]
Lymphohistiocytoid	2 [1]
Asbestos exposure (n [%])	
Exposed	186 [81]
Non-exposed	45 [19]
Tobacco consumption (n [%])	
Smoker	142 [55]
Non-smoker	116 [45]
Stage IMIG (n [%])	
I	5 [2]
II	32 [14]
III	99 [45]
IV	86 [39]
Surgical treatment (n [%])	
EP	70 [26]
PD	36 [14]

AR	8 [3]
None	152 [57]
Chemotherapy treatment (n [%])	
Yes	189 [77]
No	56 [23]
Survival status (n [%])	
Deceased patients	204 [82]
Alive patients	46 [18]
Survival (months)	
Median	19.8
Range	0.1-178.3

EP: Extrapleural pneumonectomy; PD: Pleurectomy with decortication; AR: Atypical resection

Legends of Figures

Figure 1

Genetic alterations in MPM. (A) Mutation frequencies in the Inserm and COSMIC series. P-values were determined by the Fisher's exact test (*: $p<0.05$). (B) Distribution of mutations in MPM. MPM tumor samples with at least one mutation in the *TERT* promoter or the 6 genes most frequently mutated (142 cases) are shown. Number of mutated genes in each sample is indicated by a blue gradient color at the top. Histogram on the right corresponds to -log10(p-value) of the Fisher's exact test comparing association between *TERT* promoter mutations and other mutations. Lateral bars in magenta and green colors represent significant mutually exclusive and associated mutations, respectively. The black dashed line corresponds to a p-value threshold of 0.05. WT: wild type; M: mutated; M1: nonsense substitutions, in-frame or frameshift indels and splice sites; M2: missense substitutions damaging; M3: missense substitutions probably damaging. (C) Schematic representation of *BAP1* and *NF2* (Merlin) proteins with mutations mapped (Inserm series). Point mutations are represented as lollipops. Legends of the protein domains and the mutation types are indicated at the bottom and at the top, respectively. (D) Schematic representation of the *TERT* promoter annotated with the localizations of the *TERT* transcription (TSS) and translation (ATG) start sites and the hotspot mutation sites as blue lollipops. Nucleotide numbering indicates the position on chromosome 5 in the GRCh37 assembly. Numbers of mutation at each site are indicated in arrow boxes. On the right, the histogram and the pie chart show the percentage of mutation and the proportions of mutation at each site, respectively.

Figure 2

Associations between mutation profile and heterogeneity at the histologic and molecular levels. (A) Associations between mutation profile and histologic types. MMS and MMD were classified together. (B) Heat map of mutation profile in tumor samples along the E.score and S.score (n=231). Distribution of mutations are shown only for genes, which are characterized by a significant association with the E.score or the S.score. Histogram on the right corresponds to -log10(p-value) of the Student's t test comparing for a specific gene the E.score or the S.score between MPM with or without any alterations. The black dashed line corresponds to a p-value threshold of 0.05. (C) Associations between mutation profile and transcriptomic subtypes C1 and C2. (D) Associations between mutation profile and transcriptomic subtypes C1A and C2B. P-values were determined by the Fisher's exact tests (*: $p<0.05$; **: $p<0.01$; ***: $p<0.001$) (A, C and D). MME: epithelioid MPM; MMB: biphasic MPM; MMS: sarcomatoid MPM; MMD: desmoplastic MPM; WT: wild type; M: mutated; M1: nonsense substitutions, in-frame or frameshift indels and splice sites; M2: missense substitutions damaging; M3: missense substitutions probably damaging.

Figure 3

Associations between mutation profile and overall survival. (A) Kaplan-Meier plots of overall survival in patients with wild-type (blue curve) or mutated (red curve) *NF2*, *TERT* promoter and *TP53*. P-values were determined by the Log-rank tests. (B) Univariate and multivariate Cox regression analysis of overall survival in MPM patients. Forest plots show hazard ratios (HR) and 95% confidence interval (CI) for overall survival according to age at diagnostic, tumor stage, histology, S.score based on a threshold of 0.22 and mutation status. For histology, MMB, MMS and MMD were classified as non_MME. For mutation status, samples were discriminated for the presence or the absence of at least one mutation in one of the genes *TP53* or *NF2*, or in the *TERT* promoter (*3-genes* mutation). P-values of the Wald test for all variables are indicated at the right of each forest plot. MME: epithelioid MPM; MMB: biphasic MPM; MMS: sarcomatoid MPM; MMD: desmoplastic MPM; WT: wild type; M: mutated.

Figure 4

Mutation profile of tumor samples from the same patient. Heat map shows the genetic alterations identified in tumor samples collected from the same patient. Frozen tumor samples were collected either from diagnostic biopsy or surgery resection, with or without neo-adjuvant chemotherapy, and from primary and recurrence tumors. Cell lines were also established from MPM and compared to frozen

tumor samples. Legends are indicated at the bottom. n: number of tumor sample pairs; WT: wild type; M: mutated.

Figure 5

Schematic representation of the link between the genetic landscape and tumor heterogeneity in MPM. Solid lines with arrows indicate significant associations between mutated genes and histologic or molecular classifications, histo-molecular gradients or prognosis. Dotted lines with arrows or dashes indicate significant association or exclusion between mutated genes, respectively.

Supporting Information

Supplementary Table S1. Clinical and molecular annotations of MPM samples (Inserm series).

Supplementary Table S2. Genes of the targeted sequencing.

Supplementary Table S3. Deregulated genes between molecular subtypes (Table S3A) and associations to the different subtypes (Tables S3B-D).

Supplementary Table S4. Variants with structural consequences in genes and variants in *TERT* promoter identified by targeted sequencing.

Supplementary Table S5. Tumor samples from the same patient.

Supplementary Figures. Figure S1. Mutation frequencies in four MPM series. **Figure S2.** Genetic alterations in MPM. **Figure S3.** Proportions of different mutation types in MPM. **Figure S4.** Gene expression of *TERT* gene in MPM. **Figure S5.** Associations between mutation profile and histologic types. **Figure S6.** Associations between mutation profile and molecular gradients. **Figure S7.** Associations between mutation profile and molecular subtypes. **Figure S8.** Gene expression of the 9-gene predictor. **Figure S9.** Associations between mutation profile and asbestos exposure status and tumor stage. **Figure S10.** Associations between mutation profile and overall survival. **Figure S11.** Prognostic value of *NF2*, *TERT* promoter and *TP53* mutations. **Figure S12.** Associations between *NF2* mutation status including large deep deletions and, histologic and molecular subtypes or gradients.

Figure 1

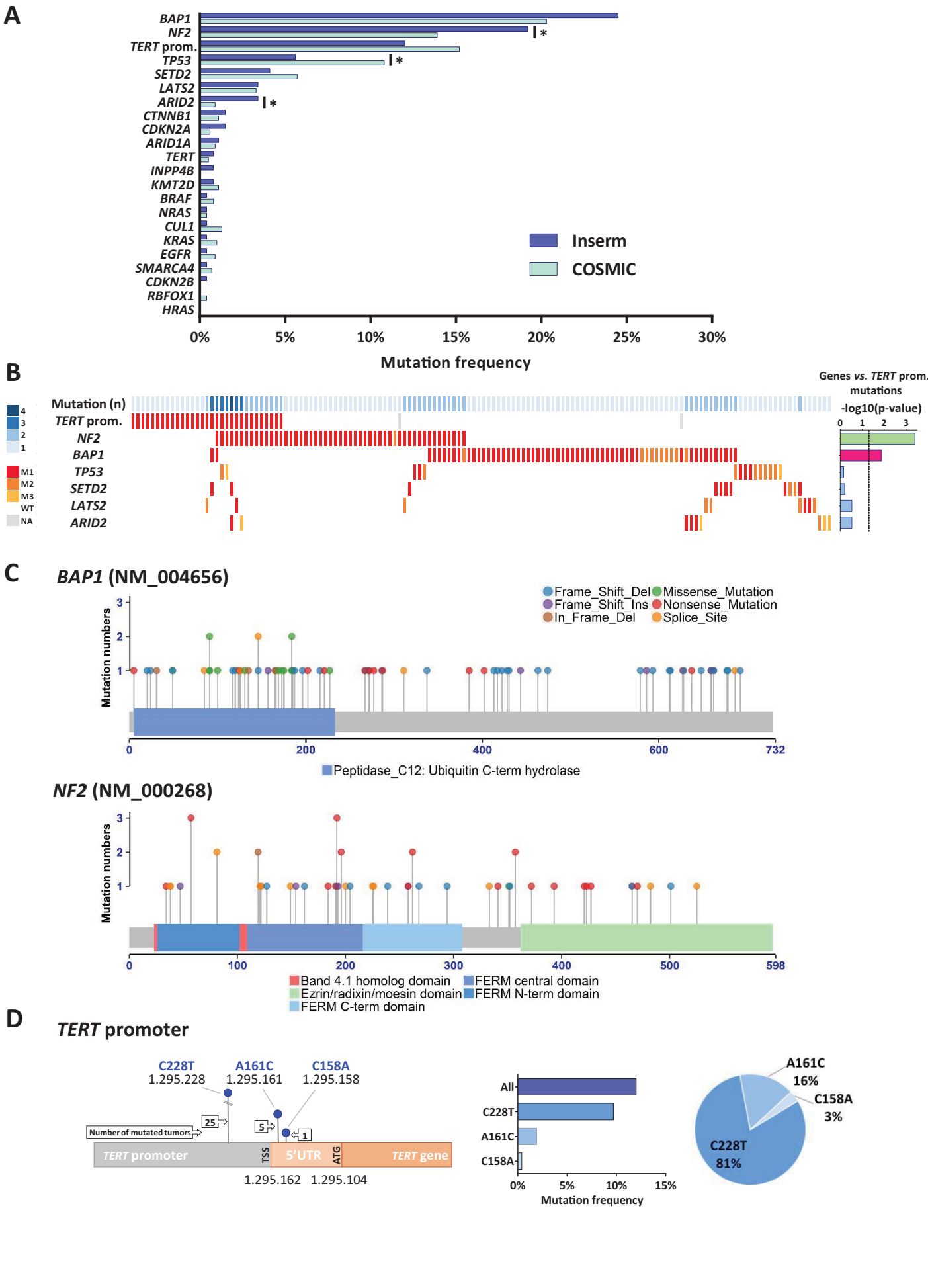

A
BAP1
NF2
TERT prom.
TP53
SETD2
LATS2
ARID2
CTNNB1
CDKN2A
ARID1A
TERT
INPP4B
KMT2D
BRAF
NRAS
CUL1
KRAS
EGFR
SMARCA4
CDKN2B
RBFOX1
HRAS
Inserm
COSMIC
0%
5%
10%
15%
20%
25%
30%
Mutation frequency
B
Mutation (n)
TERT prom.
NF2
BAP1
TP53
SETD2
LATS2
ARID2
M1
M2
M3
WT
NA
Genes vs. TERT prom. mutations
-log10(p-value)
C
BAP1 (NM_004656)
Frame_Shift_Del
Missense_Mutation
Frame_Shift_Ins
Nonsense_Mutation
In_Frame_Del
Splice_Site
Mutation numbers
Peptidase_C12: Ubiquitin C-term hydrolase
NF2 (NM_000268)
Band 4.1 homolog domain
FERM central domain
Ezrin/radixin/moesin domain
FERM N-term domain
FERM C-term domain
D
TERT promoter
C228T
1.295.228
A161C
1.295.161
C158A
1.295.158
Number of mutated tumors
TERT promoter
TSS
5'UTR
ATG
TERT gene
1.295.162
1.295.104
All
C228T
A161C
C158A
Mutation frequency
A161C
16%
C158A
3%
C228T
81%

Figure 2

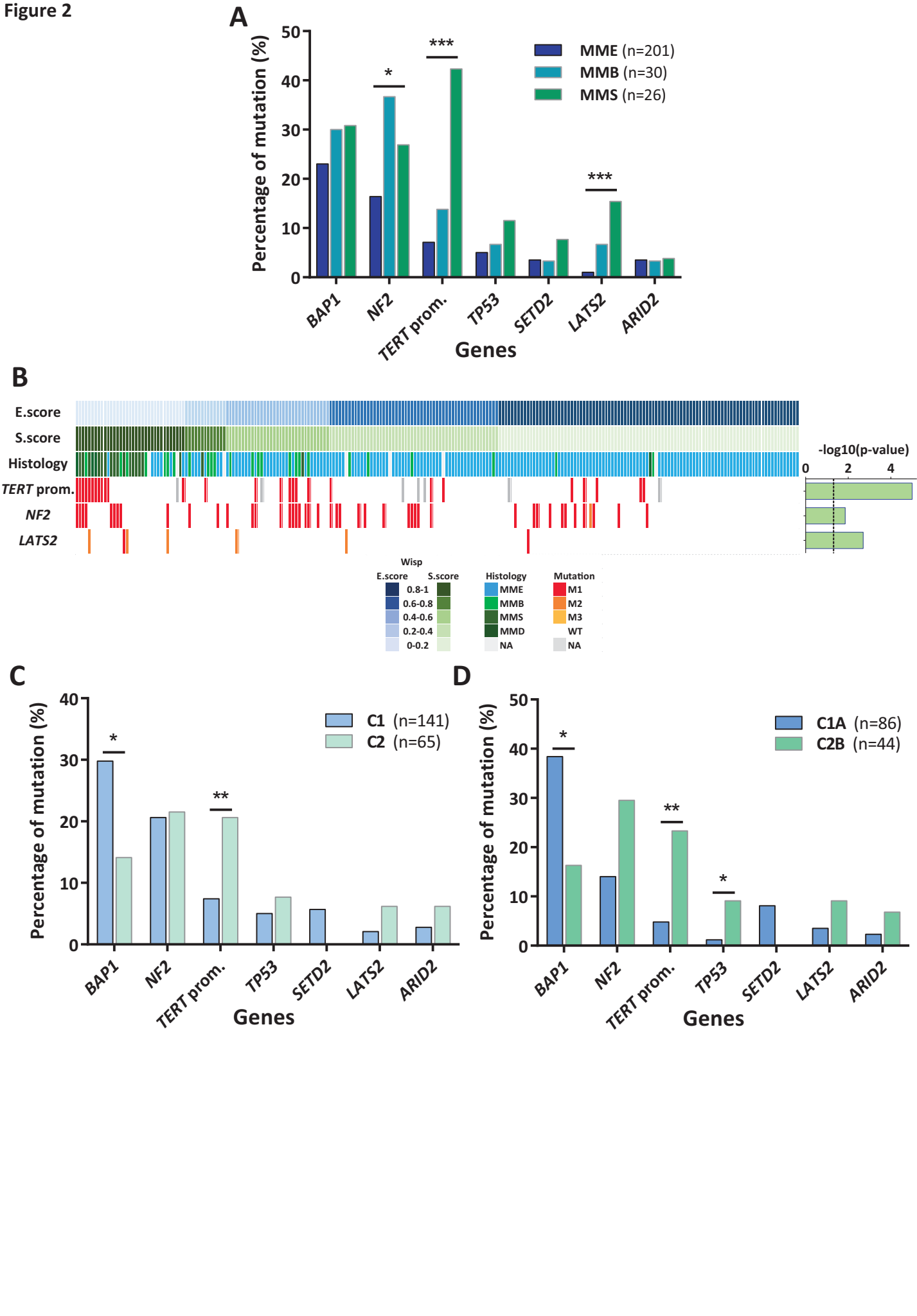

A
Percentage of mutation (%)
50
40
30
20
10
0
*
***
***
MME (n=201)
MMB (n=30)
MMS (n=26)
BAP1
NF2
TERT prom.
TP53
SETD2
LATS2
ARID2
Genes
B
E.score
S.score
Histology
TERT prom.
NF2
LATS2
-log10(p-value)
0
2
4
Wisp
E.score
0.8-1
0.6-0.8
0.4-0.6
0.2-0.4
0-0.2
S.score
Histology
MME
MMB
MMS
MMD
NA
Mutation
M1
M2
M3
WT
NA
C
Percentage of mutation (%)
40
30
20
10
0
*
**
C1 (n=141)
C2 (n=65)
BAP1
NF2
TERT prom.
TP53
SETD2
LATS2
ARID2
Genes
D
Percentage of mutation (%)
50
40
30
20
10
0
*
**
*
C1A (n=86)
C2B (n=44)
BAP1
NF2
TERT prom.
TP53
SETD2
LATS2
ARID2
Genes

Figure 3

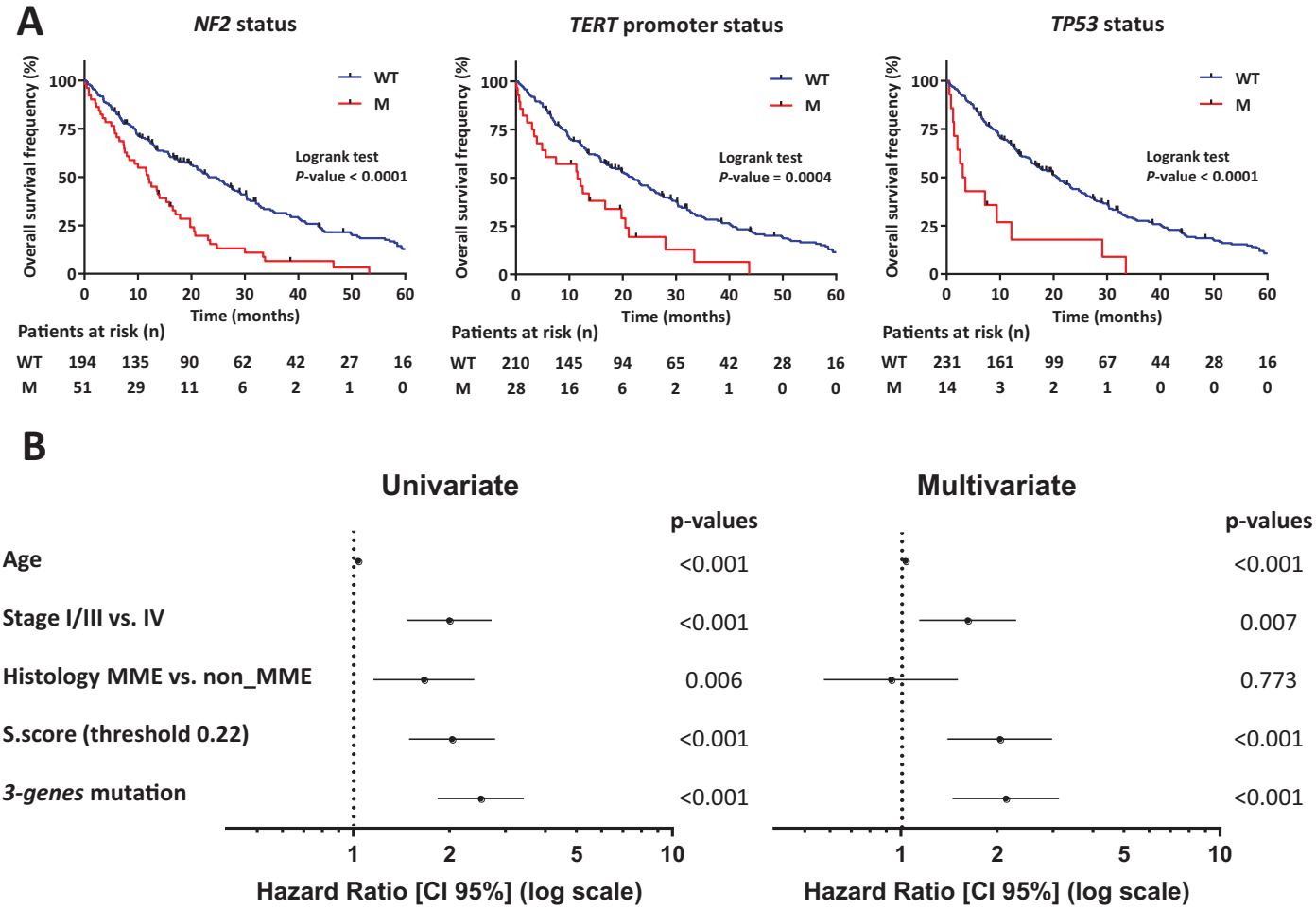

A
NF2 status
TERT promoter status
TP53 status
Overall survival frequency (%)
WT
M
Logrank test
P-value < 0.0001
P-value = 0.0004
P-value < 0.0001
Time (months)
Patients at risk (n)
WT 194 135 90 62 42 27 16
M 51 29 11 6 2 1 0
WT 210 145 94 65 42 28 16
M 28 16 6 2 1 0 0
WT 231 161 99 67 44 28 16
M 14 3 2 1 0 0 0
B
Univariate
Multivariate
p-values
Age
Stage I/III vs. IV
Histology MME vs. non_MME
S.score (threshold 0.22)
3-genes mutation
<0.001
<0.001
0.006
<0.001
<0.001
0.007
0.773
<0.001
<0.001
Hazard Ratio [CI 95%] (log scale)

Figure 4

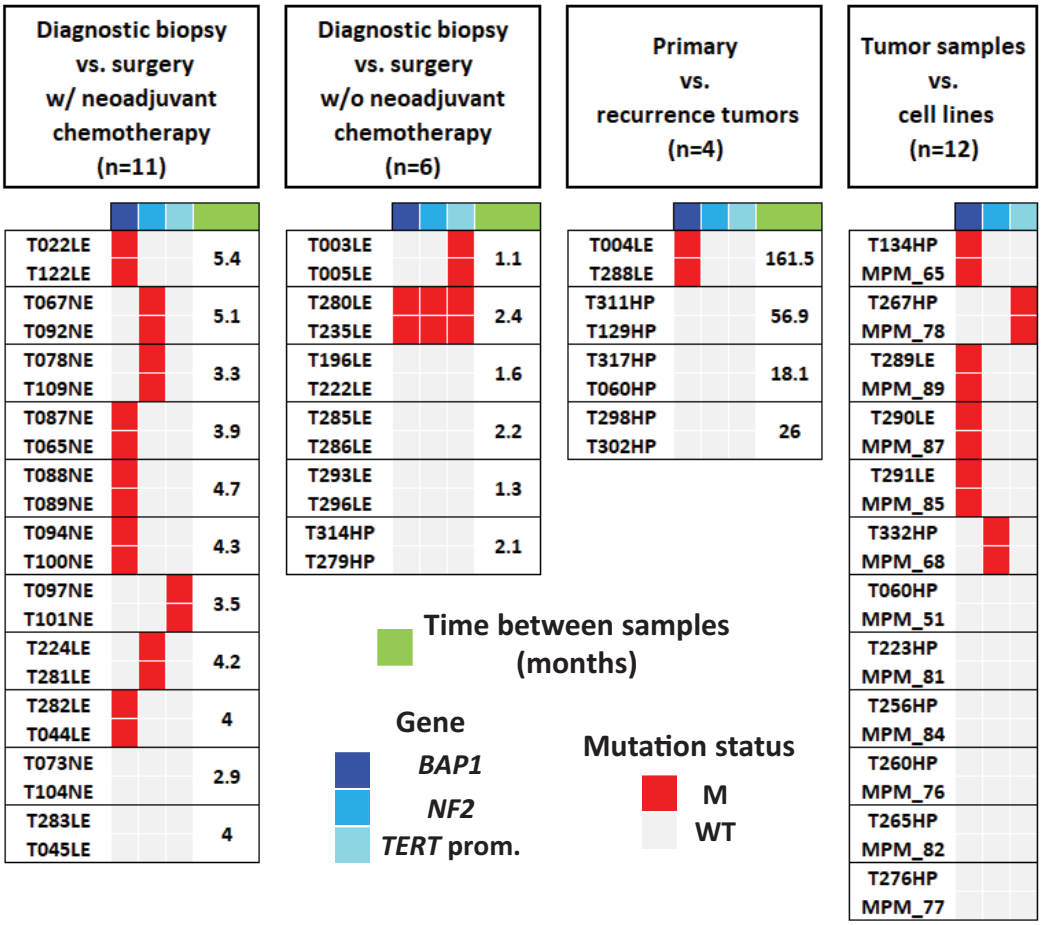
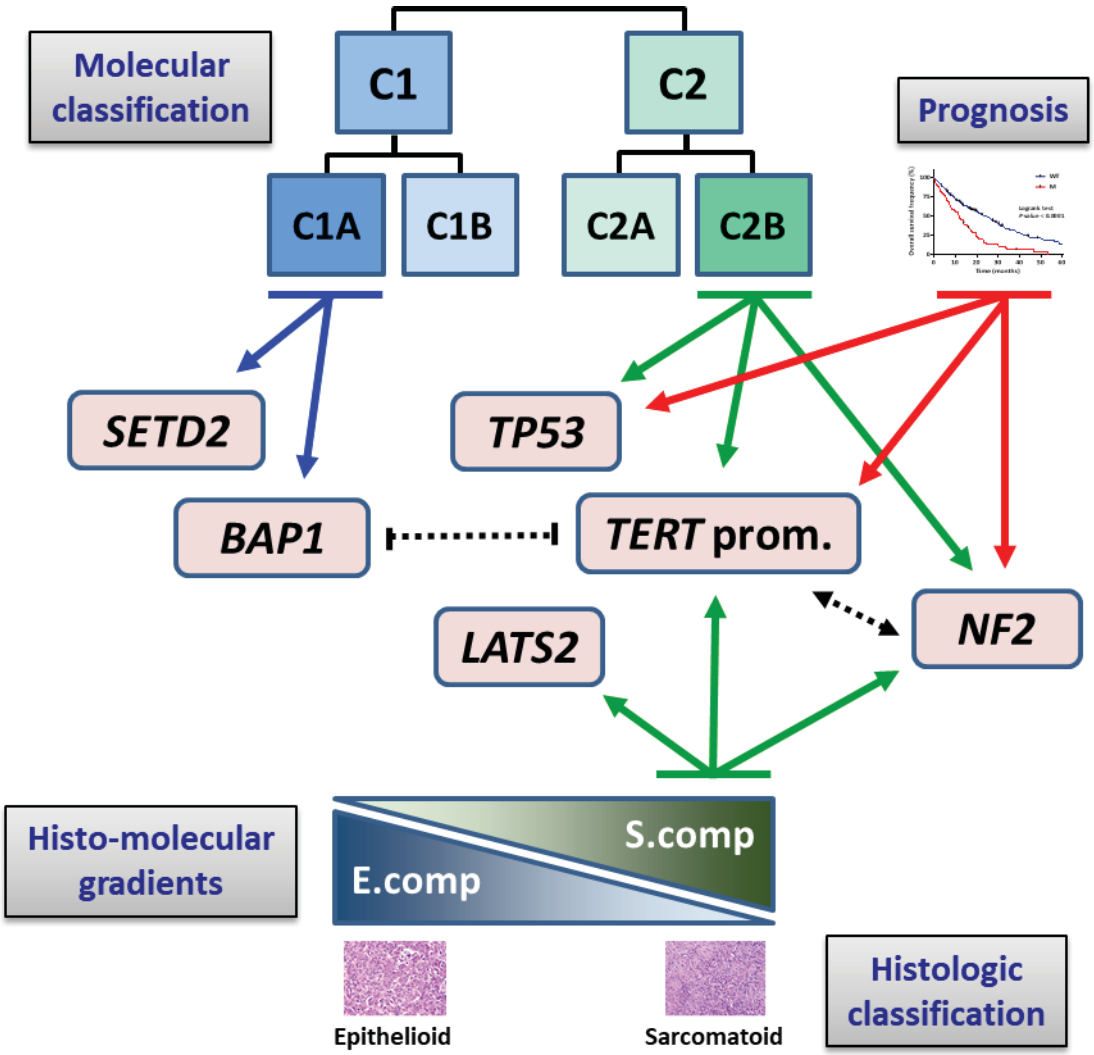

Diagnostic biopsy vs. surgery w/ neoadjuvant chemotherapy (n=11)
T022LE T122LE 5.4
T067NE T092NE 5.1
T078NE T109NE 3.3
T087NE T065NE 3.9
T088NE T089NE 4.7
T094NE T100NE 4.3
T097NE T101NE 3.5
T224LE T281LE 4.2
T282LE T044LE 4
T073NE T104NE 2.9
T283LE T045LE 4
Diagnostic biopsy vs. surgery w/o neoadjuvant chemotherapy (n=6)
T003LE T005LE 1.1
T280LE T235LE 2.4
T196LE T222LE 1.6
T285LE T286LE 2.2
T293LE T296LE 1.3
T314HP T279HP 2.1
Primary vs. recurrence tumors (n=4)
T004LE T288LE 161.5
T311HP T129HP 56.9
T317HP T060HP 18.1
T298HP T302HP 26
Tumor samples vs. cell lines (n=12)
T134HP MPM_65
T267HP MPM_78
T289LE MPM_89
T290LE MPM_87
T291LE MPM_85
T332HP MPM_68
T060HP MPM_51
T223HP MPM_81
T256HP MPM_84
T260HP MPM_76
T265HP MPM_82
T276HP MPM_77
Time between samples (months)
Gene
BAP1
NF2
TERT prom.
Mutation status
M
WT

Figure 5


Molecular classification
C1
C2
C1A
C1B
C2A
C2B
Prognosis
SETD2
TP53
BAP1
TERT prom.
LATS2
NF2
Histo-molecular gradients
S.comp
E.comp
Epithelioid
Sarcomatoid
Histologic classification