BAP1 Loss Predicts Therapeutic Vulnerability in Malignant Peritoneal Mesothelioma

Raunak Shrestha[1,2,3,12], Noushin Nabavi[1,5,12], Yen-Yi Lin[1,3], Fan Mo[1,6,7], Shawn Anderson[1], Stanislav Volik[1], Hans H. Adomat[1], Dong Lin[1,5], Hui Xue[5], Xin Dong[5], Robert Shukin[1], Robert H. Bell[1], Brian McConeghy[1], Anne Haegert[1], Sonal Brahmbhatt[1], Estelle Li[1], Htoo Zarni Oo[1,3], Antonio Hurtado-Coll[1], Ladan Fazli[1], Joshua Zhou[1], Yarrow McConnell[4], Andrea McCart[8], Andrew Lowy[9], Gregg B. Morin[5], Mads Daugaard[1,3], S. Cenk Sahinalp[1,10], Faraz Hach[1,3], Stephane Le Bihan[1], Martin E. Gleave[1,3], Yuzhuo Wang[1,3,5], Andrew Churg[11,*], and Colin C. Collins[1,3,*]

[1]Vancouver Prostate Centre, 2660 Oak St, Vancouver, BC, V6H 3Z6, Canada

[2]Bioinformatics Training Program, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

[3]Department of Urologic Sciences, University of British Columbia, Vancouver, BC, V5Z 1M9, Canada

[4]Department of Surgery, University of British Columbia, Vancouver, BC, V5Z 1M9, Canada

[5]BC Cancer Research Centre, 675 W 10th Ave, Vancouver, BC, V5Z 1L3, Canada

[6]International Precision Medicine Research Centre, Zhejiang-California International Nanosystems Institute, Zhejiang University. Hangzhou, Zhejiang, 310058, China

[7]Neoantigen Therapeutics Inc., Hangzhou, Zhejiang, 310051, China

[8]Mount Sinai Hospital, 600 University Ave, Toronto, ON, M5G 1X5, Canada

[9]Moores Cancer Center, 3855 Health Sciences Dr, La Jolla, CA 92093, USA

[10]School of Informatics and Computing, Indiana University, Bloomington, IN, 47408, USA

[11]Department of Pathology, Vancouver General Hospital, Vancouver, BC, V5Z 1M9, Canada

[12]These authors contributed equally to this work

[*]Correspondence should be addressed to C.C.C. (ccollins@prostatecentre.com) and A.C. (achurg@mail.ubc.ca)

Running title: Molecular subtypes of malignant peritoneal mesothelioma

Abstract

Background: Malignant Peritoneal Mesothelioma (PeM) is a rare but frequently fatal cancer that originates from the peritoneal lining of the abdomen. Standard treatment of PeM is limited to cytoreductive surgery and/or chemotherapy, and no effective targeted therapies for PeM yet exist. In the search for novel therapeutic target candidates in PeM, we performed a comprehensive integrative multi-omics analysis of 19 treatment-naïve PeM tumors.

Results: The analysis identified PeM tumors with *BAP1* loss to form a distinct molecular subtype characterized by distinct expression patterns of genes involved in chromatin remodeling, DNA repair pathway, and immune checkpoint receptor activation. This PeM subtype could potentially benefit from immune checkpoint, PARP, or HDAC inhibition therapies.

Conclusions: Our findings uncover *BAP1* as a trackable prognostic and predictive biomarker, and refine PeM disease classification. This integrated molecular characterization provides a comprehensive foundation for developing PeM precision medicine.

Keywords: peritoneal mesothelioma; BAP1; genomics; tumor immunosurveillance; precision oncology

Background

Malignant mesothelioma is a rare but aggressive cancer that arises from internal membranes lining of the pleura and the peritoneum. While the majority of mesotheliomas are pleural in origin, peritoneal mesothelioma (PeM) accounts for approximately 10-20% of all mesothelioma cases. PeM emerges from mesothelial cells lining of the peritoneal/abdominal cavities. The incidence rate of PeM is estimated to be less than 0.5 per 100,000 with 400-800 cases reported annually in the United States of America alone [1]. Occupational asbestos exposure is a significant risk factor in the development of pleural mesothelioma (PM). However, epidemiological studies suggest that unlike PM, asbestos exposure plays a far smaller role in the etiology of PeM tumors [1].

Mesothelioma is typically diagnosed in the advanced stages of the disease. The best therapeutic approach to PeM is cytoreductive surgery (CRS) and hyperthermic intraperitoneal chemotherapy, sometimes followed by normothermic intraperitoneal chemotherapy; however, even with this regime a significant proportion of patients die of their disease [2]. Currently there are limited clinical trials for mesothelioma and they offer conventional cytotoxic chemotherapies to which PeM are mostly resistant. Actionable molecular targets for PeM critical for precision oncology remains to be defined.

Studies investigating genetic abnormalities of PeM [3–10] have revealed recurrent copy-number losses of cyclin dependent kinase inhibitor 2A (*CDKN2A*) on 9p21, neurofibromin 2 (*NF2*) on 22q and nuclear deubiquitinase *BRCA1* associated protein 1 (*BAP1*) on 3p21. In addition, these studies also reported recurrent mutations in *BAP1*, SET domain containing 2 (*SETD2*), and DEAD-box helicase 3, X-linked (*DDX3X*). However, genomic information alone is unlikely to successfully uncover candidate therapeutic targets if not analyzed in the context of transcriptomes and proteomes.

In this study, we performed an integrated analysis of the genome, transcriptome, and proteome of 19 PeM tumors – predominantly of epithelioid subtype.

Results

Patient cohort description

We assembled a cohort of 19 tumors from 18 patients (here we refer to this cohort as VPC-PeM) undergoing CRS at Vancouver General Hospital (Vancouver, Canada), Mount Sinai Hospital (Toronto, Canada), and Moores Cancer Centre (San Diego, California, USA) (**Additional file 2: Table S1**). We obtained 19 fresh-frozen primary treatment-naïve PeM tumors and adjacent benign tissues or whole blood from the 18 cancer patients. For one patient, MESO-18, two tumors from distinct sites were available. Immunohistochemical analyses using different biomarkers were evaluated by two independent pathologists (**Additional file 1: Figure S1-S4**). Both pathologist categorized all 19 tumors as epithelioid PeM with a content of higher than 75% tumor cellularity. To the best of our knowledge this is the largest cohort of PeM subjected to an integrative multi-omics analysis.

Landscape of somatic mutations in PeM tumors

To investigate the heterogeneity of somatic gene mutations in VPC-PeM, we performed high-coverage exome sequencing (Ion Proton Hi-Q) of 19 tumors and 16 matched normal samples (matched normal samples from remaining two patients were unavailable). We achieved a mean coverage of 180x for cancerous samples and 96x for non-cancerous samples (**Additional file 2: Table S3**), with at least 43-77% of targeted bases having a coverage of 100x. We identified 346 unique non-silent mutations (313 of which were not previously reported in the Catalogue of Somatic Mutations in Cancer (COSMIC) [11]) affecting 202 unique genes (**Additional file 1: Figure S5** and **Additional file 2: Table S3**). We observed an average of 0.015 protein-coding non-silent mutations per Mb per tumor sample. Patient MESO-18 had the highest mutation burden (0.04 mutations per Mb) and MESO-11 had the lowest mutation burden (0.001 mutations per Mb) (**Additional file 1: Figure S5**). The non-silent mutation burden in PeM is low compared to other adult cancers including many abdominal cancers (**Fig. 1a**), with the exception of prostate adenocarcinomas (PRAD), kidney chromophobe carcinomas (KICH), and testicular germ cell tumors (TGCT). Notably, the mutation burden in PeM was fairly similar to PM as well as pancreatic adenocarcinomas (PAAD). We also assessed the mutational process that contribute to alterations in tumors. Analysis of base-level transitions and transversions at mutated sites showed that C>T transitions were predominant

in PeM (**Fig. 1b** and **Additional file 1: Figure S6**). Using the software deconstructSigs [12], we found that mutational signature 1, 5, 12, and 6 were operative in PeM. Interestingly, signature 1 was often correlated with age at diagnosis, and signature 6 was associated with DNA mismatch repair and mostly found in microsatellite instable tumors [13].

The mutation landscape of PeM was found to be highly heterogeneous. The nuclear deubiquitinase *BAP1* was the most frequently mutated gene (5 out of 19 tumors) in PeM tumors. *BAP1* is a tumor-suppressor gene known to be involved in chromatin remodeling, DNA double-strand break repair, and regulation of transcription of many other genes [14]. Previous studies have also reported *BAP1* as the most frequently mutated gene in both PeM[3, 6] and PM [14, 15]. The *BAP1* missense mutation in MESO-18A/E resulted in a single amino-acid (AA) change in the ubiquitin carboxyl hydrolase domain keeping the rest of the amino acid chain intact (**Additional file 1: Figure S7**). In MESO-06 and MESO-09, a *BAP1* frameshift deletion resulted in a premature stop codon and chain termination. In MESO-09 approximately 91% of *BAP1* amino acid chains were intact, but in MESO-06 only 2% of *BAP1* amino acid chains were intact (**Additional file 1: Figure S7-S8**). We also observed a *BAP1* germline mutation in only one case (MESO-09). In three (15%) tumors, we identified a recurrent R396I mutation in *ZNF678* - a zinc finger protein containing zinc-coordinating DNA binding domains involved in transcriptional regulation. We compared the mutated genes in our VPC-PeM cohort with publically available datasets [3, 15, 16] of both PeM and PM (**Additional file 1: Figure S9**). *BAP1* was the only mutated gene common between the three PeM cohorts. Twenty-five genes including tumor suppressors *LATS1*, *TP53*, and chromatin modifiers *SETD2* were common between at least two PeM cohorts. Many mutated genes in VPC-PeM were also previously reported in PM. *BAP1* and *SETD2* were the two mutated genes found common between VPC-PeM and all four PM cohorts.

Next, we identified driver genes of PeM using our recently developed algorithm HIT'nDRIVE [17]. Briefly, HIT'nDRIVE measures the potential impact of genomic aberrations on changes in the global expression of other genes/proteins which are in close proximity in a gene/protein-interaction network. It then prioritizes those aberrations with the highest impact as cancer driver genes. HIT'nDRIVE prioritized 25 unique driver genes in 15 PeM tumors for which matched genome and transcriptome data were available (**Fig. 1c** and **Additional file 2:**

Table S4). Six genes (*BAP1*, *BZW2*, *ABCA7*, *TP53*, *ARID2*, and *FMN2*) were prioritized as drivers, harboring single nucleotide changes.

Copy number landscape in PeM tumors

To investigate the somatic copy number aberration (CNA) profiles of PeM, we derived copy-number calls from exome sequencing data using the software Nexus Copy Number Discovery Edition Version 8.0. The aggregate CNA profile of PeM tumors is shown in **Fig. 2a-b**. We observed a total of 1,281 CNA events across all samples (**Additional file 2: Table S5**). On an average, 10% of the protein-coding genome was altered per PeM tumor. MESO-14 had the highest CNA burden (42%) whereas MESO-11 had the least (0.01%). Interestingly, both mutation and CNA burden in PeM was strongly correlated (R = 0.74) (**Additional file 1: Figure S10**).

We also compared the CNA burden in protein-coding regions of the VPC-PeM cohort with different adult cancers from The Cancer Genome Atlas (TCGA) project. Similar to the mutation burden, VPC-PeM tumors were observed at the lower end of the pan-cancer CNA burden spectrum. Only UCEC, PRAD, and PAAD tumors had lower median CNA burden as compared to PeM tumors (**Fig. 2c**). CNA status and mRNA expression for around half of the genes were positively correlated ($R \geq 0.1$) and 16% of the genes had strong correlation ($R \geq 0.5$). To identify cancer genes, we compared aberrations in protein-coding genes with data from the Cancer Gene Census (CGC) [18]. Intriguingly, CNA status and mRNA expression for majority of CGC genes were positively correlated (**Additional file 1: Figure S10**).

To identify recurrent focal CNAs in PeM tumors, we used the GISTIC [19] algorithm which yielded 5 regions of focal deletions ($q < 0.05$) including in 3p21 and 22q13 which are characteristic of malignant mesotheliomas (**Fig. 2d** and **Additional file 2: Table S6**). Furthermore, GISTIC prioritized 8 regions of focal amplification ($q < 0.05$) which included genes such as *IGH*, *VEGFD*, *BRD9*, *FOXL1*, *EGFR*, and *PDGFA* (**Fig. 2d**). Copy-number status of these genes was also significantly correlated with their respective mRNA expression (**Additional file 1: Figure S10**). Chromosome 1 was the most aberrant region in PeM and chromosomes 13 and 18 were relatively unchanged except for MESO-14 (**Fig. 2b**).

Chromosome 3p21 was deleted in almost half of the tumors (8 of 19) in the cohort. Here, we call tumors with 3p21 (or *BAP1*) loss as *BAP1*^{del} and the rest of the tumors with 3p21 (or *BAP1*) copy-number intact as *BAP1*^{intact}. Interestingly, *BAP1* mRNA transcripts in *BAP1*^{del} tumors were expressed at lower levels as compared to those in *BAP1*^{intact} tumors (Wilcoxon signed-rank test *p*-value = $3x10^{-4}$) (**Fig. 2e**). Immunohistochemical (IHC) staining showed lack of BAP1 nuclear staining in the tumors with *BAP1* homozygous deletion (**Fig. 2f**). Tumors with *BAP1* heterozygous loss still displayed BAP1 nuclear staining (**Additional file 1: Figure S11**). We observed three *BAP1* mutated cases among *BAP1*^{intact} tumors. *BAP1* mRNA transcripts in these three tumors, were expressed at high levels. As mentioned in the previous section, the mutation analysis also predicted that despite mutation in *BAP1* in these three tumors, the entire *BAP1* amino-acid chain is still intact and may be functionally active (**Additional file 1: Figure S7**). Furthermore, we found DNA copy loss of 3p21 locus to include four concomitantly deleted cancer genes - *BAP1*, *SETD2*, *SMARCC1*, and *PBRM1*, consistent with [20]. Copy-number status of these four genes was significantly correlated with their corresponding mRNA expression (**Additional file 1: Figure S12**), suggesting that the allelic loss of these genes is associated with decreased transcript levels. These four genes are chromatin modifiers, and *PBRM1* and *SMARCC1* are part of SWI/SNF complex that regulates transcription of a number of genes. Furthermore, HIT'nDRIVE prioritized *BAP1*, *PBRM1*, and *SETD2* as driver genes of PeM (**Fig. 1c** and **Additional file 2: Table S4**).

CNA status of genes associated with a number of key cancer pathways was observed to be different between the PeM subtypes (i.e. *BAP1*^{del} and *BAP1*^{intact}). We observed many genes involved in chromatin remodeling, SWI/SNF complex and DNA repair pathway to be deleted in *BAP1*^{del} tumors as compared to *BAP1*^{intact} tumors (**Fig. 1c**). In contrast, we found copy-number gain of many genes in DNA repair pathways (*BRCA2*, *ATM*, *MGMT*, and *RAD51*) and the cell cycle (*MYC*, *CDK5*, *CCNB1*, and *CCND1*) in the *BAP1*^{intact} tumors. Furthermore, PeM tumors (both *BAP1*^{del} and *BAP1*^{intact}) harbored CNA events in carcinogenic pathways such as MAPK, PI3K, MTOR, Wnt, and Hippo pathways. Interestingly, *ESR1* copy number deletion is enriched in *BAP1*^{del} tumors while co-amplification of *EGFR* and *BRAF* were present in three *BAP1*^{intact} tumors. Notably, we identified copy-number loss of tumor suppressor *LATS1/2* and copy-number gain of *NF2* in one case, both of which has been previously associated with

mesotheliomas [15], in *BAP1*^{del} tumors. Notably, both LATS1/2 and NF2 are key regulators of the Hippo pathway [21].

Unsupervised consensus clustering of tumor samples based on copy-number segmentation mean values of the 3349 most variable genes identified four tumor sub-groups (**Fig. 2g** and **Additional file 1: Figure S13**). We observed that *BAP*1^{del} and *BAP1*^{intact} tumors were grouped into distinct clusters. This indicates that *BAP1*^{del} tumors have distinct copy-number profiles from those of *BAP1*^{intact} tumors. We identified 692 genes (p-value < 0.01, Kruskal-Wallis test) with significantly differential CNA genes segments between the clusters. These genes were mapped to eight distinct chromosome loci – 19p, 6q, 1q, and, 13q and were mostly gained in clusters 1 and 3, whereas Xq, 22q, and 7p loss were mostly in clusters 1 and 3 (**Additional file 1: Figure S13**).

Next, we compared the CNA profile of PeM with PM. We note that PeM and PM tumors displayed substantial overlap (**Additional file 1: Figure S14**). Both PeM and PM tumors harbored characteristic loss of chromosome segments 3p21, 9p21 and 22q. Copy number loss of *CDKN2A* (9p21) or *NF2* (22q) is observed in more than 40% of PM tumors [15]. However, in PeM, loss of *CDKN2A* and *NF2* was observed in only in a single case (MESO-14). We identified 1670 genes (p-value < 0.0005, Wilcoxon signed-rank test) with significant differential CNA segments between PeM and PM (**Additional file 2: Table S7**). Intriguingly, these 1670 genes mapped to eleven distinct chromosome loci 1q, 3p, 6p, 11p, 11q, 12p, 12q, and 17q that were mostly gained in PM and lost in PeM. Vice versa 22q, Xp, and Xq were mostly lost in PM and gained in PeM (**Additional file 1: Figure S14**).

Gene fusions in PeM

To identify the presence of gene fusions, we analyzed RNA-seq data in 15 PeM using deFuse algorithm [22]. Overall, 82 unique gene fusion events were identified using our filtering criteria (**Methods** and **Additional file 2: Table S8**), out of which we successfully validated 18 gene fusions using Sanger sequencing (**Additional file 1: Figure S15** and **Additional file 2: Table S9**). We observed more gene fusion events in *BAP1*^{del} tumors as compared to that in *BAP1*^{intact} tumors (**Fig. 3a-b**). Similar to mutation or copy-number profiles, the gene fusion landscape of PeM was highly heterogeneous. *MTG1-SCART1* was the most recurrent gene fusion observed in 7

cases. This was followed by *GKAP1-KIF27* and *KANSL1-ARL17B* (**Fig. 3c**) each of which was identified in 6 different cases. Three unique fusions were present in *PBRM1*, 2 in *KANSL1*, and 1 each in *BAP1* and *SETD2* all of which are involved in chromatin remodeling process (**Fig. 1c** and **3d-f**). Notably, *BAP1*, *SETD2*, *PBRM1*, and *KANSL1* were prioritized as a driver gene by HIT'nDRIVE on basis of gene-fusion. Fusions in these genes were mostly found in the $BAP1^{del}$ subtype.

The global transcriptome and proteome profile of PeM

To segregate transcriptional subtypes of PeM, we performed total RNA-seq (Illumina HiSeq 4000) and its quantification of 15 PeM tumor samples for which RNA were available (RNA for remaining four tumor samples did not pass the quality control checks). We first performed principal-component analyses and unsupervised consensus clustering of all PeM tumors to determine transcriptomic patterns using genes based on variance among tumor specimens. Consensus clustering revealed two distinct transcriptome sub-groups (**Fig. 4a**). We found $BAP1^{intact}$ and $BAP1^{del}$ have some distinct transcriptomic patterns; however, a few samples showed an overlapping pattern (**Additional file 1: Figure S16**).

We performed mass spectrometry (Fusion Orbitrap LC/MS/MS) with isobaric tagging for expressed peptide identification and its corresponding protein quantification using Proteome Discoverer for processing pipeline for 16 PeM tumors and 7 matched normal tissues (matched normal samples for the remaining tumors were not available). We identified 8242 unique proteins in 23 samples analyzed (we were surprised BAP1 protein was however not detected in our MS experiment, likely due to inherent technical limitations with these samples and/or processing. Quality control analysis of in solution Hela digests also have very low BAP1 with only a single peptide observed in occasional runs). First, we analyzed global matched mRNA-protein expression correlation. Although, ~58% (4715 of 8109) of proteins showed positive mRNA-protein correlation (Pearson correlation; $R \geq 0.1$), only 22.7% (1839) of the proteins were strongly correlated with their corresponding mRNA ($R \geq 0.5$) (**Additional file 1: Figure S17**). Expression of 2.4% (194) of proteins strongly negatively correlated with their corresponding mRNA ($R \leq -0.5$). To analyze the proteomic pattern across PeM tumors, we performed principal-component analyses and unsupervised consensus clustering following the same procedure as described above for the

transcriptome. Unlike in transcriptome profiles, the proteome profiles of *BAP1* PeM tumor sub-types did not group into distinct clusters (**Fig. 4b** and **Additional file 1: Figure S18**).

To identify differentially expressed genes (DEG) between *BAP1*^{intact} and *BAP1*^{del}, we performed Wilcoxon signed-rank test using mRNA and protein expression data independently. We identified 1520 and 466 DEG (with *p*-value < 0.05) using mRNA and protein expression data respectively (**Additional file 1: Figure S19** and **Additional file 2: Table S10-S11**). However, only 53 genes were found common between the two sets of DEG. As expected, *BAP1*, *PBRM1* and *SMARCA4*, *SMARCD3* were among the top-500 DEG. Many other important cancer-related genes were differentially expressed such as *CDK20*, *HIST1H4F*, *ERCC1*, *APOBEC3A*, *CDK11A*, *CSPG4*, *TGFB1*, *IL6*, *LAG3*, and *ATM*.

Transcriptional and post-transcriptional mechanisms regulate chromatin remodeling protein-complexes

Next, we aimed to study the extent to which changes in copy number profile affects its corresponding protein expression. For this, we calculated Pearson correlation between CNA-mRNA expression and CNA-protein expression. While, copy number profile of genes, on average, have good agreement with their corresponding mRNA expression, a number of detected proteins had poor correlation with their respective gene's copy number profile. Approximately 25% (1871 of 7462) of proteins were observed to have poor correlation with their genes copy number which we here define as "attenuated proteins" (**Methods**, **Fig. 4c,** and **Additional file 2: Table S12**). Among the attenuated proteins, we identified important chromatin remodeling proteins - PBRM1, SETD2, and SMARCC1. The attenuated proteins also included cancer genes such as NF2, EGFR, APC, PIK3CA, and MAP3K4. We observed that the attenuated proteins were significantly enriched with direct protein-protein interaction partners of the UBC (hypergeometric test *p*-value: 10^{-5}), BAP1 (10^{-3}), and PBRM1 (10^{-2}) in STRING v10 interaction network. Notably, geneset enrichment analysis revealed that attenuated proteins are more likely to form a part of a multimeric complex or bind to macromolecules (**Fig. 4d**). These results corroborate previous findings from studies analyzing breast, ovarian and colorectal cancer datasets [23]. These attenuated proteins were found to be involved in mRNA processing, DNA repair pathway, cell cycle regulation, the immune system, and in carbohydrate and lipid metabolism. Strikingly, we found that DEG between the PeM subtypes are significantly

associated with protein attenuation (Chi-Squared test *p*-value: 10^{-4} using mRNA expression DEG, 10^{-6} using protein expression DEG). These findings suggest that the effects of CNA are attenuated at the protein level via post-transcriptional modification.

To identify large protein complexes containing the attenuated proteins and that are variable (i.e. at least a protein subunit of the complex is differentially expressed) between PeM subtypes, we leveraged a manually curated set of core protein complexes from the CORUM database [24]. These included many protein complexes involved in DNA conformation modification, DNA repair, transcriptional regulation, post-translational modification including ubiquitination (**Additional file 2: Table S13-S14**). Using our data, we observed that the majority of the protein complexes were highly co-regulated at the protein level rather than at the mRNA level. Notably, we identified SWI/SNF (BAF and PBAF) and HDAC complex which were highly co-regulated (**Fig. 4e-g**). We found copy-number deletion in many subunits of SWI/SNF complex, mostly in the *BAP1*^{del} subtype (**Fig. 1c**). About one quarter of proteins in the BAF complex and half of proteins in PBAF were attenuated. PBRM1 was both attenuated at the protein level as well as differentially expressed between PeM subtypes. SMARCB1 and SMARCA4 were also differentially expressed between PeM subtypes in this complex (**Fig. 4h**). We further identified a number of HDAC complex components as highly co-regulated. The complex consisted of Histone deacetylase (HDAC1/2), which regulates expression of a number of genes through chromatin remodeling. About one-third of protein subunits in the complex were attenuated at the protein level. More importantly, HDAC1, CHD4, and ZMYM2 were differentially expressed between PeM subtypes in the protein complex, and different family members of HDAC protein family were highly expressed in the *BAP1*^{del} subtype (**Fig. 4i**). This indicates potential use of HDAC inhibitors to suppress the tumor growth in the *BAP1*^{del} subtype. We note that both SWI/SNF and HDAC complexes interact with BAP1. Expression pattern of many subunits of these complexes were either highly correlated or highly anti-correlated with BAP1 expression (**Fig. 4e-g**). Although mRNA transcripts are transcribed proportional to the changes in copy-number profile of the gene, the corresponding proteins are often stabilized when in complex, and free proteins in excess are usually ubiquitinated and targeted for proteosomal degradation to maintain stoichiometry [23].

$BAP1^{del}$ subtype is characterized by distinct expression patterns of genes involved in DNA repair pathway, and immune checkpoint receptor activation

To identify the pathways dysregulated by the DEG between the PeM subtypes, we performed hypergeometric test based geneset enrichment analysis (**Methods**) using the REACTOME pathway database. Intriguingly, we observed high concordance between pathways dysregulated by the two sets (mRNA and protein expression data) of top-500 DEG (**Fig. 5a-b**). The unsupervised clustering of pathways revealed two distinct clusters for $BAP1^{del}$ and $BAP1^{intact}$ tumors. This indicates that the enriched pathways, between the patient groups, are also differentially expressed. $BAP1^{del}$ patients demonstrated elevated levels of RNA and protein metabolism as compared to $BAP1^{intact}$ patients. Many genes involved in chromatin remodeling and DNA damage repair were differently expressed between the groups (**Additional file 1: Figure S19-S20**). Our data suggests that $BAP1^{del}$ tumors have repressed DNA damage response pathways. Most importantly, protein expression data revealed that PARP1 is highly expressed in $BAP1^{del}$ tumors as compared to $BAP1^{intact}$ tumors indicating potential inhibition of PARP1 for $BAP1^{del}$ tumors. Genes involved in cell-cycle and apoptotic pathways were observed to be highly expressed in $BAP1^{del}$ patients. Furthermore, glucose and fatty-acid metabolism pathways were repressed in $BAP1^{del}$ as compared to $BAP1^{intact}$. More interestingly, we observed a striking difference in immune-system associated pathways between the PeM subtypes. Whereas $BAP1^{del}$ patients demonstrated strong activity of cytokine signaling and the innate immune system; MHC-I/II antigen presentation system and Adaptive immune system were active in $BAP1^{intact}$ patients.

Prompted by this finding, we next analyzed whether PeM tumors were infiltrated with leukocytes. To assess the extent of leukocyte infiltration, we computed an expression (RNA-seq and protein) based score using the immune-cell and stromal markers proposed by [25]. We discovered that the immune marker gene score was strongly correlated with stromal marker gene score (**Methods** and **Fig. 5c-d**). Using CIBERSORT [26] software, we computationally estimated leukocytes representation in the bulk tumor transcriptome. We observed massive infiltration of T cells and CD4/8 cells in majority of the PeM tumors (**Fig. 5e**). A subset of PeM tumors had massive infiltration of B-cells in addition to T cells and CD4/8 cells. Interestingly, when we group the PeM tumors by their

BAP1 aberration status, there was a marked difference in the proportion of infiltrated plasma cells, natural killer (NK) cells, mast cells, B cells and CD8 cells between the groups. Whereas the proportions of plasma cells, NK cells and B cells were less in the *BAP1*^{del} tumors, there was more infiltration of mast cells and CD8 cells were in *BAP1*^{del} tumors as compared to *BAP1*^{intact} tumors. We performed TMA IHC staining of CD3 and CD8 antibody on PeM tumors. We observed that *BAP1*^{del} PeM tumors were positively stained for both CD3 and CD8 confirming infiltration of B and T cells in *BAP1*^{del} PeM tumors (**Fig. 5f** and **Additional file 1: Figure S21**). Combined, this strongly indicates that leukocytes from the tumor-microenvironment infiltrates the PeM tumor.

Finally, we surveyed the PeM tumors for expression of genes involved in immune checkpoint pathways. A number of immune checkpoint receptors were highly expressed in *BAP1*^{del} tumors relative to *BAP1*^{intact} tumors. These included *CD274* (*PD-L1*), *CD80*, *CTLA4*, *LAG3*, and *ICOS* (**Fig. 5g**) for which inhibitors are either clinically approved or are at varying stages of clinical trials. Gene expression of these immune checkpoint receptors were highly correlated with immune score (**Fig. 5h**). Moreover, a number of MHC genes, immuno-inhibitor genes as well as immuno-stimulator genes were differentially expressed between *BAP1*^{del} and *BAP1*^{intact} tumors (**Additional file 1: Figure S22**). Furthermore, we analyzed whether the immune checkpoint receptors were differentially expressed in tumors with and without 3p21 loss in PM tumors from TCGA. Unlike in PeM, we did not observe a significant difference in immune checkpoint receptor expression between the PM tumor groups (i.e. *BAP1*^{del} and *BAP1*^{intact}) (**Additional file 1: Figure S23**). These findings suggest that *BAP1*^{del} PeM tumors could potentially be targeted with immune-checkpoint inhibitors while PM tumors may less likely to respond.

Discussion

In this study, we present a comprehensive integrative multi-omics analysis of malignant peritoneal mesotheliomas. Even though this is a rare disease we managed to amass a cohort of 19 tumors. Prior studies of mesotheliomas, performed using a single omic platform, have established loss of function mutation or copy-number loss of *BAP1* as a key driver event in both PeM and PM. Our novel contribution to PeM is that we provide evidence from integrative multi-omics analyses that *BAP1* copy number loss (*BAP1*^{del}) forms a distinct molecular subtype of PeM. This subtype of PeM is characterized by distinct expression patterns of genes involved in chromatin remodeling, DNA repair pathway, and immune checkpoint activation. Our results suggest that *BAP1*^{del} tumors might be prioritized for immune checkpoint blockade therapies. Thus *BAP1* may be both a prognostic and predictive biomarker for PeM enabling better disease classification and patient treatment.

Structural alterations in PeM tumors were found to be highly heterogeneous, and occur at a lower rate as compared to most other adult solid cancers. The majority of SNVs and CNAs were typically unique to a patient. However, many of these alterations were non-randomly distributed to critical carcinogenic pathways. We observed many alterations in genes involved in chromatin remodeling, SWI/SNF complex, cell cycle and DNA repair pathway. SWI/SNF complex is an ATP-dependent chromatin remodeling complex known to harbor aberrations in almost one-fifth of all human cancers [27]. Our results show that SWI/SNF complex is differentially expressed between PeM subtypes which further regulates oncogenic and tumor suppressive pathways. Notably, we also identified another chromatin remodeling complex - HDAC complex which is differentially expressed between PeM subtypes. HDAC, known to be regulated by *BAP1*, is a potential therapeutic target for the *BAP1*^{del} PeM subtype. Recent in-vitro experiments demonstrated *BAP1* loss altered sensitivity of PM as well as uveal melanoma (UM) cells to HDAC inhibition [28, 29].

Loss of *BAP1* is known to alter chromatin architecture exposing the DNA to damage, and also impairing the DNA-repair machinery [30, 31]. Similar to BRCA1/2 deficient breast and ovarian cancers, *BAP1* deficient PeM tumors most likely depends on PARP1 for survival. This rationale can be utilized to test PARP inhibitors in *BAP1*^{del} PeM

subtype. The DNA repair defects thus drive genomic instability and dysregulate tumor microenvironment [32]. DNA repair deficiency leads to the increased secretion of cytokines, including interferons that promote tumor-antigen presentation, and trigger recruitment of both T and B lymphocytes to destroy tumor cells. As a response, tumor cells evade this immune-surveillance by increased expression of immune checkpoint receptors. The results presented here also indicate that PeM tumors are infiltrated with immune-cells from the tumor microenvironment. Moreover, the *BAP1*^{del} subtype displays elevated levels of immune checkpoint receptor expression which strongly suggests the use of immune checkpoint inhibitors to treat this subtype of PeM. However, in a small subset of PM tumors in TCGA dataset, the loss of *BAP1* did not elevate expression of immune checkpoint marker genes. This warrants further investigation on the characteristics of these groups of PM tumors. Furthermore, recently, *BAP1* loss has been shown to define a distinct molecular subtype of clear cell renal cell carcinoma (ccRCC) and UM [33–35]. These studies showed that, similar to *BAP1*^{del} PeM subtype, *BAP1*^{del} tumors from both ccRCC and UM also have dysregulated chromatin modifiers, impaired DNA repair pathway, and immune checkpoint receptor activation. More recent studies in ccRCC [36] and melanoma [37] demonstrated that inactivation of *PBRM1* (or PBAF complex) predicts response to immune checkpoint blocking therapies. Similarly, DNA repair defects have also been shown to be predictive of response to immune checkpoint blocking therapies [38–40]. This strongly indicates a pan-cancer mechanism of oncogenesis shared among tumors with *BAP1* copy-number loss.

The main challenge in mesothelioma treatment is that, all current efforts made towards testing new therapy options are limited to using therapies that have been proven successful in other cancer types, without a good knowledge of underlying molecular mechanisms of the disease. As a result of sheer desperation, some patients have been treated even though no targeted therapy for mesothelioma has been proven effective as yet. For example, a number of clinical trials exploring the use of immune checkpoint inhibitors (anti-PD1/PD-L1 or anti-CTLA4) in PM and/or PeM patients that progressed under chemotherapy, and are positive for immune checkpoint markers are currently under progress. The results of the first few clinical trials report either very low response rate or no benefit to the patients [41–44]. Notably, *BAP1* copy-number or mutation status were not assessed in these studies. We believe

that response rates for immune checkpoint blockade therapies in clinical trials for PeM will improve when patients are segregated by their *BAP1* copy-number status.

Conclusions

Our first-in-field multi-omics data underscore PeM tumors with *BAP1* loss as a distinct molecular subtype that could potentially be targeted with immune checkpoint, PARP, or HDAC inhibitors. This dataset can be used to inform targeted therapies for almost half of PeM cases for which no definite therapies yet exist. Our findings uncover BAP1 as a trackable prognostic and predictive biomarker, and refine PeM disease classification.

Methods

Clinical samples and pathology evaluation

Primary untreated PeM tumors and matched benign samples were obtained from cancer patients undergoing cytoreductive surgeries following protocols approved by the Clinical Research Ethics Board of the Vancouver General Hospital (Vancouver, BC, Canada), Mount Sinai Hospital (Toronto, ON, Canada), and Moores Cancer Centre (San Diego, CA, USA). This study was approved by the Institutional Review Board of the University of British Columbia and Vancouver Coastal Health (REB No. H15-00902 and V15-00902). All patients signed a formal consent form approved by the respective institutional ethics board. Histologic parameters and pathological scoring of tumors confirming PeM was established by three independent pathologists. Hematoxylin and eosin (H&E) and immunostained formalin-fixed paraffin-embedded (FFPE) slides were reviewed by at least two specialized pathologists to diagnose PeM and its subtype. H&E staining was used to determine the highest tumor cellularity (≥75%) from sections for sequencing. The surgical resections were snap frozen and processed at respective institutions. The tumors have a companion normal tissue specimen (either adjacent normal tissue or peripheral blood previously extracted for germline DNA control). Each tumor specimen was approximately 1cm^3 in size and weighed between 100-300 mg. Specimen were shipped overnight on dry ice that maintained an average temperature of less than -80°C. Upon receipt, the tissues were sectioned into 5 slices for DNA, RNA, and protein extraction as well as construction of tissue microarrays (TMA).

Construction of tissue microarrays (TMAs)

FFPE tissue blocks were retrieved from the archives of the Department of Pathology, Vancouver General Hospital (Vancouver, Canada). H&E stained slides from each block were reviewed by two pathologists to identify tumor areas. TMAs were constructed with 1 mm diameter tissue cores from representative tumor areas from FFPE blocks. Cores were transferred to a paraffin block using a semi-automated tissue array instrument (Pathology Devices TMArrayer, San Diego, CA). Duplicate tissue cores were taken from each specimen, resulting in a composite TMA block. Reactive mesothelial tissues from pleura were also included as benign controls. Following construction, 4µm thick sections were cut for H&E and immunohistochemical staining.

Immunohistochemistry and histopathology

Freshly cut TMA sections were analyzed for immunoexpression using Ventana Discovery Ultra autostainer (Ventana Medical Systems, Tucson, Arizona). In brief, tissue sections were incubated in Tris-EDTA buffer (CC1) at 37°C to retrieve antigenicity, followed by incubation with respective primary antibodies at room temperature or 37°C for 60-120 min. For primary antibodies, mouse monoclonal antibodies against CD8 (Leica, NCL-L-CD8-4B11, 1:100), CK5/Cytokeratin 5(Abcam, ab17130, 1:100), BAP1 (SantaCruz, clone C4, sc-28383, 1:50), rabbit monoclonal antibody against CD3 (Abcam, ab16669, 1:100), and rabbit polyclonal antibodies against CALB2/Calretinin (LifeSpan BioSciences, LS-B4220, 1:20 dilution) were used. Bound primary antibodies were incubated with Ventana Ultra HRP kit or Ventana universal secondary antibody and visualized using Ventana ChromoMap or DAB Map detection kit, respectively. All stained slides were digitalized with the SL801 autoloader and Leica SCN400 scanning system (Leica Microsystems; Concord, Ontario, Canada) at magnification equivalent to x20. The images were subsequently stored in the SlidePath digital imaging hub (DIH; Leica Microsystems) of the Vancouver Prostate Centre. Representative tissue cores were manually identified by two pathologists.

Whole exome sequencing

DNA was isolated from snap-frozen tumors with 0.2 mg/mL Proteinase K (Roche) in cell lysis solution using Wizard Genomic DNA Purification Kit (Promega Corporation, USA). Digestion was carried out overnight at 55°C before incubation with RNase solution at 37°C for 30 minutes and treatment with protein precipitation solution followed by isopropanol precipitation of the DNA. The amount of DNA was quantified on the NanoDrop 1000 Spectrophotometer and an additional quality check done by reviewing the 260/280 ratios. Quality check were done on the extracted DNA by running the samples on a 0.8% agarose/TBE gel with ethidium bromide.

For Ion AmpliSeq™ Exome Sequencing, 100ng of DNA based on Qubit® dsDNA HS Assay (Thermo Fisher Scientific) quantitation was used as input for Ion AmpliSeq™ Exome RDY Library Preparation. This is a polymerase chain reaction (PCR)-based sequencing approach using 294,000 primer pairs (amplicon size range 225-275 bp), and covers >97% of Consensus CDS (CCDS; Release 12), >19,000 coding genes and >198,000

coding exons. Libraries were prepared, quantified by quantitative PCR (qPCR) and sequenced according to the manufacturer's instructions (Thermo Fisher Scientific). Samples were sequenced on the Ion Proton System using the Ion PI™ Hi-Q™ Sequencing 200 Kit and Ion PI™ v3 chip. Two libraries were run per chip for a projected coverage of 40M reads per sample.

Somatic variant calling

Torrent Server (Thermo Fisher Scientific) was used for signal processing, base calling, read alignment, and generation of results files. Specifically, following sequencing, reads were mapped against the human reference genome hg19 using Torrent Mapping Alignment Program. The mean target coverage ranges from 78.62 to 226.44, thus sequencing depth ranges from 78 to 226X. Variants were identified by using Torrent Variant Caller plugin with the optimized parameters for AmpliSeq exome-sequencing recommended by Thermo Fisher. The variant call format (VCF) files from all sample were combined using GATK (3.2-2) [45] and all variants were annotated using ANNOVAR [46]. Only non-silent exonic variants including non-synonymous single nucleotide variations (SNVs), stop-codon gain SNVs, stop-codon loss SNVs, splice site SNVs and In-Dels in coding regions were kept if they were supported by more than 10 reads and had allele frequency higher than 10%. To obtain somatic variants, we filtered against dbSNP build 138 (non-flagged only) and the matched adjacent benign or blood samples sequenced in this study. Putative variants were manually scrutinized on the Binary Alignment Map (BAM) files through Integrative Genomics Viewer (IGV) version 2.3.25 [47].

Copy number aberration (CNA) analysis

Copy number changes were assessed using Nexus Copy Number Discovery Edition Version 9.0 (BioDiscovery, Inc., El Segundo, CA). Nexus NGS functionality (BAM ngCGH) with the FASST2 Segmentation algorithm was used to make copy number calls (a Circular Binary Segmentation/Hidden Markov Model approach). The significance threshold for segmentation was set at $5x10^{-6}$, also requiring a minimum of 3 probes per segment and a maximum probe spacing of 1000 between adjacent probes before breaking a segment. The log ratio thresholds for single copy gain and single copy loss were set at +0.2 and -0.2, respectively. The log ratio thresholds for gain

of 2 or more copies and for a homozygous loss were set at +0.6 and -1.0, respectively. Tumor sample BAM files were processed with corresponding normal tissue BAM files. Reference reads per CNA point (window size) was set at 8000. Probes were normalized to median. Relative copy number profiles from exome sequencing data were determined by normalizing tumor exome coverage to values from whole blood controls. The germline exome sequences were used to obtain allele-specific copy number profiles and generating segmented copy number profiles. The GISTIC module on Nexus identifies significantly amplified or deleted regions across the genome. The amplitude of each aberration is assigned a G-score as well as a frequency of occurrence for multiple samples. False Discovery Rate q-values for the aberrant regions have a threshold of 0.15. For each significant region, a "peak region" is identified, which is the part of the aberrant region with greatest amplitude and frequency of alteration. In addition, a "wide peak" is determined using a leave-one-out algorithm to allow for errors in the boundaries in a single sample. The "wide peak" boundaries are more robust for identifying the most likely gene targets in the region. Each significantly aberrant region is also tested to determine whether it results primarily from broad events (longer than half a chromosome arm), focal events, or significant levels of both. The GISTIC module reports the genomic locations and calculated q-values for the aberrant regions. It identifies the samples that exhibit each significant amplification or deletion, and it lists genes found in each "wide peak" region.

Whole transcriptome sequencing (RNA-seq)

Total RNA from 100µm sections of snap-frozen tissue was isolated using the mirVana Isolation Kit from Ambion (AM-1560). Strand specific RNA sequencing was performed on quality controlled high RIN value (>7) RNA samples (Bioanalyzer Agilent Technologies) before processing at the high throughput sequencing facility core at BGI Genomics Co., Ltd. (The Children's Hospital of Philadelphia, Pennsylvania, USA). In brief, 200ng of total DNAse treated RNA was first treated to remove the ribosomal RNA (rRNA) and then purified using the Agencourt RNA Clean XP Kit (Beckman Coulter) prior to analysis with the Agilent RNA 6000 Pico Chip to confirm rRNA removal. Next, the rRNA-depleted RNA was fragmented and converted to cDNA. Subsequent steps include end repair, addition of an `A' overhang at the 3' end, ligation of the indexing-specific adaptor, followed by purification with Agencourt Ampure XP beads. The strand specific RNA library prepared using TruSeq (Illumina Catalogue

No. RS-122-2201) was amplified and purified with Ampure XP beads. Size and yield of the barcoded libraries were assessed on the LabChip GX (Caliper), with an expected distribution around 260 base pairs. Concentration of each library was measured with real-time PCR. Pools of indexed library were then prepared for cluster generation and PE100 sequencing on Illumina HiSeq 4000.

RNA-seq quantification

Using splice-aware aligner STAR (2.3.1z) [48], RNA-seq reads (~200MB in size) were aligned onto the human genome reference (GRCh38) and exon-exon junctions, according to the known gene model annotation from the Ensembl release 80 (http://www.ensembl.org). Apart from protein coding genes, non-coding RNA types and pseudogenes are further annotated and classified. Based on the alignment and by using gene annotation (Ensembl release 80), gene expression profiles was calculated. Only reads unique to one gene and which corresponded exactly to one gene structure, were assigned to the corresponding genes by using the python tool HTSeq [49]. Normalization of read counts was conducted by R package DESeq [50], which was designed for gene expression analysis of RNA-seq data across different samples.

Identification of fusion transcripts and validation

We used the deFuse algorithm [22] to predict rearrangements in RNA sequence libraries. The deFuse fusion transcript prediction calls were further filtered using following criteria: a fusion gene candidate: (1) must be predicted to have arisen from genome rearrangement, rather than via a readthrough event; (2) must be predicted in no more than two sequence libraries; (3) must map unambiguously on both sides of the predicted breakpoints (that is, no multi-mapping reads); (4) must not map entirely to repetitive elements; (5) must be detected in >5 reads (either split or spanning) and (6) must have at least one of the fusion partner transcript expressed.

Prioritized putative gene fusions were verified by designing PCR primers around the predicted fusion sites. Specifically, reverse transcription PCR (RT-PCR) was used to amplify the predicted fusion gene junctions from the same starting RNA material (100ng) as was used for RNA-seq. Two primers (20-22 bp nucleotides) spanning the exon boundary of fused genes were designed using Primer3 (v. 0.4.0)[51] (**Additional file 2: Table S9**). PCR

was performed in 20µl reactions using Q5 buffer (NEB), 0.2mM dNTPs, 0.4 µM each primer, 0.12 units Q5 High-Fidelity DNA Polymerase (NEB) and 2 µl of the RT reaction. The PCR reaction was carried out with the following program: 95°C, 30 seconds, followed by 30 cycles of 95°C for 10 seconds, 57°C for 20 seconds and 72°C for 10 seconds. Resulting PCR products, ranging in size from 150bp to 250bp, were purified using AMPure beads (Agencourt) and sequenced using Sanger sequencing to verify fusion junctions.

Proteomics analysis using mass spectrometry

Fresh frozen samples dissected from tumor and adjacent normal were individually lysed in 50mM of HEPES pH 8.5, 1% SDS, and the chromatin content was degraded with benzonase. The tumor lysates were sonicated (Bioruptor Pico, Diagenode, New Jersey, USA), and disulfide bonds were reduced with DTT and capped with iodoacetamide. Proteins were cleaned up using the SP3 method [52, 53] (Single Pot, Solid Phase, Sample Prep), then digested overnight with trypsin in HEPES pH 8, peptide concentration determined by Nanodrop (Thermo) and adjusted to equal level. A pooled internal standard control was generated comprising of equal volumes of every sample (10μl of each of the 100μl total digests) and split into 3 equal aliquots. The labeling reactions were run as three TMT 10-plex panels (9+IS), then desalted and each panel divided into 48 fractions by reverse phase HPLC at pH 10 with an Agilent 1100 LC system. The 48 fractions were concatenated into 12 superfractions per panel by pooling every 4th fraction eluted resulting in a total 36 overall samples.

These samples were analyzed with an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Fisher Scientific) coupled to EasyNanoLC 1000 using a data-dependent method with synchronous precursor selection (SPS) MS3 scanning for TMT tags. A short description follows; more detailed overview is in [53]. Briefly, an in house packed reverse phase column run with a 2 hour low pH acetonitrile gradient (5-40% with 0.1% formic acid) was used to separate and introduce peptides into the MS. Survey scans covering m/z 350-1500 were acquired in profile mode at a resolution of 120,000 (at m/z 200) with S-Lens RF Level of 60%, a maximum fill time of 50 milliseconds, and automatic gain control (AGC) target of $4x10^5$. For MS2, monoisotopic precursor selection was enabled with triggering charge state limited to 2-5, threshold $5x10^3$ and 10 ppm dynamic exclusion for 60 seconds. Centroided MS2 scans were acquired in in the ion trap in Rapid mode after CID fragmentation with a maximum

fill time of 20 milliseconds and 1 m/z isolation quadrupole isolation window, collision energy of 30%, activation Q of 0.25, injection for all available parallelizable time turned ON, and an AGC target value of $1x10^4$. For MS3, fragment ions were isolated from a 400-1200 m/z precursor range, ion exclusion of 20 m/z low and 5 m/z high, isobaric tag loss exclusion for TMT, with a top 10 precursor selection. Acquisition was in profile mode with the Orbitrap after HCD fragmentation (NCE 60%) with a maximum fill time of 90 milliseconds, 50,000 m/z resolution, 120-750 m/z scan range, an AGC target value of $1x10^5$, and all available parallelizable ON. The total allowable cycle time was set to 4 seconds.

Peptide identification and protein quantification

Qualitative and quantitative proteomics analysis was done using ProteomeDiscoverer 2.1.1.21 (Thermo Fisher Scientific). To maintain consistency with transcriptome annotation, we used Ensembl GRCh38.87 human reference proteome sequence database for proteome annotation. Sequest HT 1.3 was used for peptide-spectral-matches (PSMs), with parameters specified as trypsin enzyme, two missed cleavages allowed, minimum peptide length of 6, precursor mass tolerance 10 ppm, and a fragment mass tolerance of 0.6 Da. We allowed up to 4 variable modifications per peptide from the following categories: acetylation at protein terminus, methionine oxidation, and TMT label at N-terminal residues and the side chains of lysine residues. In addition, carbamidomethylation of cysteine was set as a fixed modification. PSM results were filtered using q-value cut off of 0.05 to control for false discovery rate (FDR) determined by Percolator. Identified peptides from both high and medium-confidence level after FDR-filtering were included in the final stage to provide protein identification and quantification results. Reporter ions from MS3 scans were quantified with an integration tolerance of 20ppm with the most confident centroid. Proteins were further filtered to include only those found with minimum one peak in all samples. Proteome Discoverer processed data was exported for further statistical analysis.

Mutational signature analysis

We used deconstructSigs [12], a multiple regression approach to statistically quantify the contribution of mutational signature for each tumor. The mutational signature were obtained from the COSMIC mutational

signature database[54]. Both silent and non-silent somatic mutations were used together to obtain the mutational signatures. Only mutational signatures with a weight more than 0.06 were considered for analysis.

Prioritization of driver genes using HIT'nDRIVE

Non-silent somatic mutation calls, CNA gain or loss, and gene-fusion calls were collapsed in gene-patient alteration matrix with binary labels. Gene-expression values were used to derive expression-outlier gene-patient outlier matrix using Generalized Extreme Studentized Deviate (GESD) test. STRING ver10 [55] protein-interaction network was used to compute pairwise influence value between the nodes in the interaction network. We integrated these genome and transcriptome data using HIT'nDRIVE algorithm [17]. Following parameters were used: $\alpha = 0.9,\ \beta = 0.6,\ \gamma = 0.8$. We used IBM-CPLEX as the Integer Linear Programming (ILP) solver.

Consensus clustering

We used ConsensusClusterPlus [56] R-package to perform consensus clustering. We used the following parameters: maximum cluster number to evaluate: 10, number of subsamples: 10000, proportion of items to sample: 0.8, proportion of features to sample: 1, cluster algorithm: hierarchical, distance: pearson.

Protein attenuation analysis

For every gene/protein profiled for CNA (segment mean), RNA-seq (normalized $\log_2$ expression), and MS (normalized $\log_2$ expression), we performed the following analysis. For every gene/protein, the Pearson correlation coefficients were calculated for CNA-mRNA expression (R_{CNA_mRNA}) and CNA-protein expression ($R_{CNA_protein}$). The 75th percentile of the difference between the above two correlation coefficients i.e. $R_{diff} = R_{CNA_mRNA} - R_{CNA_protein}$ was found to be approximately 0.45. Therefore those proteins with $R_{diff} \geq 0.45$ were considered as attenuated proteins.

Pathway enrichment analysis

The selected set of differentially expressed genes were tested for enrichment against gene sets of pathways present in Molecular Signature Database (MSigDB) v6.0 [57]. A hypergeometric test based gene set enrichment analysis was used for this purpose (https://github.com/raunakms/GSEAFisher). A cut-off threshold of false discovery rate

(FDR) < 0.01 was used to obtain the significantly enriched pathways. Only pathways that are enriched with at least three differentially expressed genes were considered for further analysis. To calculate the pathway activity score, the expression dataset was transformed into standard normal distribution using 'inverse normal transformation' method. This step is necessary for fair comparison between the expression-values of different genes. For each sample, the pathway activity score is the mean expression level of the differentially expressed genes linked to the enriched pathway.

Stromal and immune score

We used two sets of 141 genes (one each for stromal and immune gene signatures) as described in[25]. We used 'inverse normal transformation' method to transform the distribution of expression data into the standard normal distribution. The stromal and immune scores were calculated, for each sample, using the summation of standard normal deviates of each gene in the given set.

Enumeration of tissue-resident immune cell types using mRNA expression profiles

CIBERSORT algorithm[26] was applied to the RNA-seq gene-expression data to estimate the proportions of 22 immune cell types (B cells naive, B cells memory, Plasma cells, T cells CD8, T cells CD4 naive, T cells CD4 memory resting, T cells CD4 memory activated, T cells follicular helper, T cells gamma delta, T cells regulatory (Tregs), NK cells resting, NK cells activated, Monocytes, Macrophages M0, Macrophages M1, Macrophages M2, Dendritic cells resting, Dendritic cells activated, Mast cells resting, Mast cells activated, Eosinophils and Neutrophils) using LM22 dataset provided by CIBERSORT platform. Genes not expressed in any of the PeM tumor samples were removed from the LM22 dataset. The analysis was performed using 1000 permutation. The 22 immune cell types were later aggregated into 11 distinct groups.

External datasets

TCGA datasets for 16 different cancer-types used in this study were downloaded from the National Cancer Institute-Genomic Data Commons (NCI-GDC; https://portal.gdc.cancer.gov/) on February 2017. For somatic mutation data, non-silent variant calls that were identified by at least three out of four different tools (MUSE,

MuTect2, SomaticSniper and VArScan2) were considered. CNA segmented data were further processed using Nexus Copy Number Discovery Edition Version 9.0 (BioDiscovery, Inc., El Segundo, CA) to identify aberrant regions in the genome. In case of the RNA-seq expression data, HTSeq-FPKM-UQ normalized data were used. AACR Project GENIE[16] dataset was downloaded from https://www.synapse.org/#!Synapse:syn7222066 on April 2017.

Additional files

Additional file 1: Figure S1-S23. **Figure S1.** Pathology of Peritoneal Mesothelioma. **Figure S2.** Pathology of Peritoneal Mesothelioma. **Figure S3.** TMA slides of PeM tumors IHC stained for CK5. **Figure S4.** TMA slides of PeM tumors IHC stained for CALB2. **Figure S5.** Summary of non-silent somatic mutation landscape of PeM. **Figure S6.** Distribution of SNV nucleotide substitutions. **Figure S7.** Effect of BAP1 somatic mutation on resulting amino-acid chain. **Figure S8.** Effect of BAP1 somatic mutation on resulting amino-acid chain. **Figure S9.** Common somatic mutated genes in mesothelioma patient cohorts. **Figure S10.** Correlation of copy-number with mutation and gene-expression profiles. **Figure S11.** TMA slides of PeM tumors IHC stained for BAP1. **Figure S12.** Co-deletion of four cancer-associated genes in chromosomal region 3p21. **Figure S13.** Consensus clustering of copy-number segment mean profiles of PeM. **Figure S14.** Comparison of SCNA profile of peritoneal and pleural mesothelioma. **Figure S15.** Gene Fusions in PeM. **Figure S16.** Consensus clustering of mRNA expression profiles of PeM. **Figure S17.** Correlation between mRNA and protein expression profiles. **Figure S18.** Consensus clustering of mRNA expression profiles of PeM. **Figure S19.** Significant differentially expressed genes/proteins between molecular subtypes of PeM. **Figure S20.** Differentially expressed DNA-repair genes between PeM subtypes. **Figure S21.** TMA slides of BAP1del PeM tumors IHC stained for CD3 and CD8. **Figure S22.** Differentially expressed immune genes between PeM subtypes. **Figure S23.** Immune cell infiltration in Pleural Mesothelioma.

Additional file 2: Table S1-S14. **Table S1.** Clinical information associated with the PeM patients. **Table S2.** Quality control statistics for WES data. **Table S3.** Somatic Mutation in PeM tumors. **Table S4.** Driver genes

prioritized by HIT'nDRIVE algorithm. **Table S5.** Copy number segmentation profiles of PeM tumors. **Table S6.** Significant copy number changed regions prioritized by GISTIC. **Table S7.** Significant differential copy number aberrated regions between PeM and PM. **Table S8.** Predicted gene fusion events in PeM tumors using deFuse algorithm. **Table S9.** List of gene fusion events validated using Sanger sequencing. **Table S10.** Differentially expressed transcripts between PeM molecular subtypes. **Table S11.** Differentially expressed proteins between PeM molecular subtypes. **Table S12.** Correlation between CNA-mRNA and CNA-Protein. **Table S13.** Differentially expressed CORUM complex measured using mRNA expression levels. **Table S14.** Differentially expressed CORUM complex measured using protein expression levels.

Abbreviations

AA: Amino-acid; BAP1: BRCA1 associated protein 1; BLCA: Urothelial Bladder Carcinoma; BRCA: Breast Invasive Carcinoma; ccRCC: Clear cell renal cell carcinoma; CDKN2A: cyclin dependent kinase inhibitor 2A; CGC: Cancer Gene Census; CNA: Copy number aberration; COAD: Colorectal carcinoma; COSMIC: Catalogue of Somatic Mutations in Cancer; DDX3X: DEAD-box helicase 3, X-linked; DEG: Differentially expressed genes; FFPE: Formalin-fixed paraffin-embedded; GBM: Glioblastoma Multiforme; IHC: Immunohistochemical; IGV: Integrative Genomics Viewer; KICH: Kidney Chromophobe; KIRC: Kidney Renal Clear Cell Carcinoma; KRIP: Kidney renal papillary cell carcinoma; LUAD: Lung adenocarcinoma; LUSC: Lung Squamous Cell Carcinoma; NF2: Neurofibromin 2 (NF2); OV: Ovarian cancer; PAAD: Pancreatic Adenocarcinoma; PCR: polymerase chain reaction; PeM: Peritoneal Mesothelioma; PM: Pleural Mesothelioma; PRAD: Prostate Adenocarcinoma; qPCR: Quantitative PCR; SETD2: SET domain containing 2; TGCT: Testicular Germ Cell Tumor; TMA: Tissue microarray; UCEC: Uterine Corpus Endometrial Carcinoma; UCS: Uterine Carcinosarcoma; UM: Uveal Melanoma.

Declarations

Acknowledgements

We acknowledge the contribution and support of Drs. Jessica McAlpine, Anna Tinker, and Jeff Simko as well as Emily Taylor (Mount Sinai Hospital), Donald Donaldson, Rose Schweigert, Matthew Sturgen, Sarah Padilla at Vancouver Coastal Health for supporting and coordinating the procurement of PeM tumors. We are thankful for the support of regulatory and ethics bodies: Margaret Luk, Wylo Kyle, Jacqueline Lee, Zahra Karim, Nandita Chowdhury, Jessica Gagliardi, Suzanne Richardson, and Sheila O'Donoghue at Vancouver General Hospital; Chantal Lackan at Mount Sinai Hospital; Ida Deichaite and Oudone Sisanachandeng at Moores Cancer Centre. Thanks to Shane Colborne and Dr. Christopher Hughes for help and suggestions regarding mass spectrometry experiments and data analyses. The authors thank all members of the Collins', Wang's, Hach's, and Sahinalp's labs for helpful suggestions. The results published here are in part based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/.

Funding

This study is funded by: BC Cancer Foundation (C.C.C. and Y.Z.W.) and Mitacs Fellowship (Y.Z.W. and C.C.C.). RS is supported by Mitacs Accelerate Awards, Prostate Cancer Foundation - BC Research Award and Canadian Institutes of Health Research (CIHR) Bioinformatics Training Program. NN is supported by Mitacs Accelerate Awards. The funding agencies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

The whole-exome and whole-transcriptome sequencing data from this study is available in the European Genome-phenome Archive (EGA; https://ega-archive.org/) under accession number EGAS00001002820. The proteome data from mass spectrometry is available in the PRIDE Archive (https://www.ebi.ac.uk/pride/archive/) under accession number PXD008873.

Authors' contributions

R.S., N.N., S.LB., Y.W. A.C., and C.C.C. conceived the study. R.S. and N.N. performed data analysis and wrote the manuscript. N.N. and S.LB. managed the project. Y.L., F.M., S.A., S.V., H.H.A, R.H.B., and J.Z. performed data analysis. N.N., Ro.S., B.M., A.H., S.B., H.H.A, and G.B.M. performed the specimen processing, quality control, sequencing and mass-spectrometry experiments, and validation experiments. E.L., HZ.O., A.H., and L.F. constructed TMAs, performed IHC experiments, and TMA scoring. A.C. and HZ.O. reviewed the tissue slides. N.N., D.L., H.X., and X.D., constructed patient derived mouse xenograft. Y.M., A.M., and A.L. contributed clinical specimens and clinical data. M.D., S.C.S., F.H., S.LB., M.E.G., Y.W., A.C., and C.C.C. supervised the project, contributed scientific insights, and edited the manuscript.

Ethics approval and consent to participate

This study was approved by the Institutional Review Board of the University of British Columbia and the Vancouver Coastal Health (REB Number. H1500902 and V15-00902). All samples and information were collected with written and signed informed consent from the participating patients.

Consent for publication

Not applicable.

Competing interests

Authors declare no competing interests.

References

1. Testa JR: *Asbestos and Mesothelioma*. Springer International Publishing; 2017. [*Current Cancer Research*]

2. Sugarbaker PH, Chang D: **Long-term regional chemotherapy for patients with epithelial malignant peritoneal mesothelioma results in improved survival**. *Eur J Surg Oncol* 2017, **43**:1228–1235.

3. Alakus H, Yost SE, Woo B, French R, Lin GY, Jepsen K, Frazer KA, Lowy AM, Harismendy O: **BAP1 mutation is a frequent somatic event in peritoneal malignant mesothelioma.** *J Transl Med* 2015, **13**:122.

4. Chirac P, Maillet D, Leprêtre F, Isaac S, Glehen O, Figeac M, Villeneuve L, Péron J, Gibson F, Galateau-Sallé F, Gilly FN, Brevet M: **Genomic copy number alterations in 33 malignant peritoneal mesothelioma analyzed by comparative genomic hybridization array**. *Hum Pathol* 2016, **55**:72–82.

5. Joseph NM, Chen Y, Nasr A, Yeh I, Talevich E, Onodera C, Bastian BC, Rabban JT, Garg K, Zaloudek C, Solomon DA: **Genomic profiling of malignant peritoneal mesothelioma reveals recurrent alterations in epigenetic regulatory genes BAP1, SETD2, and DDX3X.** *Mod Pathol* 2017, **30**:246–254.

6. Kato S, Tomson BN, Buys TPH, Elkin SK, Carter JL, Kurzrock R: **Genomic Landscape of Malignant Mesotheliomas**. *Mol Cancer Ther* 2016, **15**:2498–2507.

7. Sheffield BS, Tinker A V., Shen Y, Hwang H, Li-Chang HH, Pleasance E, Ch'ng C, Lum A, Lorette J, McConnell YJ, Sun S, Jones SJM, Gown AM, Huntsman DG, Schaeffer DF, Churg A, Yip S, Laskin J, Marra MA: **Personalized oncogenomics: Clinical experience with malignant peritoneal mesothelioma using whole genome sequencing**. *PLoS One* 2015, **10**:1–12.

8. Singhi AD, Krasinskas AM, Choudry HA, Bartlett DL, Pingpank JF, Zeh HJ, Luvison A, Fuhrer K, Bahary N, Seethala RR, Dacic S: **The prognostic significance of BAP1, NF2, and CDKN2A in malignant peritoneal mesothelioma.** *Mod Pathol* 2016, **29**:14–24.

9. Ugurluer G, Chang K, Gamez ME, Arnett AL, Jayakrishnan R, Miller RC, Sio TT: **Genome-based**

Mutational Analysis by Next Generation Sequencing in Patients with Malignant Pleural and Peritoneal Mesothelioma. *Anticancer Res* 2016, **36**:2331–8.

10. Leblay N, Leprêtre F, Le Stang N, Gautier-Stein A, Villeneuve L, Isaac S, Maillet D, Galateau-Sallé F, Villenet C, Sebda S, Goracci A, Byrnes G, McKay JD, Figeac M, Glehen O, Gilly FN, Foll M, Fernandez-Cuesta L, Brevet M: **BAP1 Is Altered by Copy Number Loss, Mutation, and/or Loss of Protein Expression in More Than 70% of Malignant Peritoneal Mesotheliomas**. *J Thorac Oncol* 2017, **12**:724–733.

11. Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, Cole CG, Ward S, Dawson E, Ponting L, Stefancsik R, Harsha B, YinKok C, Jia M, Jubb H, Sondka Z, Thompson S, De T, Campbell PJ: **COSMIC: Somatic cancer genetics at high-resolution**. *Nucleic Acids Res* 2017, **45**:D777–D783.

12. Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C: **deconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution**. *Genome Biol* 2016, **17**:31.

13. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio S a. JR, Behjati S, Biankin A V., Bignell GR, Bolli N, Borg A, Børresen-Dale A-L, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjörd JE, Foekens J a., Greaves M, Hosoda F, Hutter B, Ilicic T, Imbeaud S, Imielinski M, Imielinsk M, Jäger N, Jones DTW, Jones D, Knappskog S, et al.: **Signatures of mutational processes in human cancer.** *Nature* 2013, **500**:415–21.

14. Bott M, Brevet M, Taylor BS, Shimizu S, Ito T, Wang L, Creaney J, Lake R a, Zakowski MF, Reva B, Sander C, Delsite R, Powell S, Zhou Q, Shen R, Olshen A, Rusch V, Ladanyi M: **The nuclear deubiquitinase BAP1 is commonly inactivated by somatic mutations and 3p21.1 losses in malignant pleural mesothelioma.** *Nat Genet* 2011, **43**:668–672.

15. Bueno R, Stawiski EW, Goldstein LD, Durinck S, De Rienzo A, Modrusan Z, Gnad F, Nguyen TT, Jaiswal BS, Chirieac LR, Sciaranghella D, Dao N, Gustafson CE, Munir KJ, Hackney JA, Chaudhuri A, Gupta R,

Guillory J, Toy K, Ha C, Chen Y-J, Stinson J, Chaudhuri S, Zhang N, Wu TD, Sugarbaker DJ, de Sauvage FJ, Richards WG, Seshagiri S: **Comprehensive genomic analysis of malignant pleural mesothelioma identifies recurrent mutations, gene fusions and splicing alterations.** *Nat Genet* 2016, **48**:407–16.

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

17. Shrestha R, Hodzic E, Sauerwald T, Dao P, Wang K, Yeung J, Anderson S, Vandin F, Haffari G, Collins CC, Sahinalp SC: **HIT'nDRIVE: patient-specific multidriver gene prioritization for precision oncology.** *Genome Res* 2017, **27**:1573–1588.

18. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR: **A census of human cancer genes.** *Nat Rev Cancer* 2004, **4**:177–83.

19. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G: **GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers**. *Genome Biol* 2011, **12**:R41.

20. Yoshikawa Y, Emi M, Hashimoto-Tamaoki T, Ohmuraya M, Sato A, Tsujimura T, Hasegawa S, Nakano T, Nasu M, Pastorino S, Szymiczek A, Bononi A, Tanji M, Pagano I, Gaudino G, Napolitano A, Goparaju C, Pass HI, Yang H, Carbone M: **High-density array-CGH with targeted NGS unmask multiple noncontiguous minute deletions on chromosome 3p21 in mesothelioma.** *Proc Natl Acad Sci U S A* 2016, **113**:13432–13437.

21. Li W, Cooper J, Zhou L, Yang C, Erdjument-Bromage H, Zagzag D, Snuderl M, Ladanyi M, Hanemann CO, Zhou P, Karajannis MA, Giancotti FG: **Merlin/NF2 loss-driven tumorigenesis linked to CRL4(DCAF1)-mediated inhibition of the hippo pathway kinases Lats1 and 2 in the nucleus.** *Cancer Cell* 2014, **26**:48–60.

22. McPherson A, Hormozdiari F, Zayed A, Giuliany R, Ha G, Sun MGF, Griffith M, Heravi Moussavi A, Senz J, Melnyk N, Pacheco M, Marra MA, Hirst M, Nielsen TO, Sahinalp SC, Huntsman D, Shah SP: **deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data.** *PLoS Comput Biol* 2011, **7**:e1001138.

23. Gonçalves E, Fragoulis A, Garcia-Alonso L, Cramer T, Saez-Rodriguez J, Beltrao P: **Widespread Post-transcriptional Attenuation of Genomic Copy-Number Variation in Cancer**. *Cell Syst* 2017, **0**:1–13.

24. Ruepp A, Waegele B, Lechner M, Brauner B, Dunger-Kaltenbach I, Fobo G, Frishman G, Montrone C, Mewes HW: **CORUM: The comprehensive resource of mammalian protein complexes-2009**. *Nucleic Acids Res* 2009, **38**(SUPPL.1):497–501.

25. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine D a, Carter SL, Getz G, Stemke-Hale K, Mills GB, Verhaak RGW: **Inferring tumour purity and stromal and immune cell admixture from expression data.** *Nat Commun* 2013, **4**:2612.

26. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA: **Robust enumeration of cell subsets from tissue expression profiles.** *Nat Methods* 2015, **12**:453–7.

27. Kadoch C, Crabtree GR: **Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics**. *Sci Adv* 2015, **1**:e1500447–e1500447.

28. Landreville S, Agapova OA, Matatall KA, Kneass ZT, Onken MD, Lee RS, Bowcock AM, Harbour JW: **Histone deacetylase inhibitors induce growth arrest and differentiation in uveal melanoma**. *Clin Cancer Res* 2012, **18**:408–416.

29. Sacco JJ, Kenyani J, Butt Z, Carter R, Chew HY, Cheeseman LP, Darling S, Denny M, Urbé S, Clague MJ, Coulson JM: **Loss of the deubiquitylase BAP1 alters class I histone deacetylase expression and sensitivity of mesothelioma cells to HDAC inhibitors.** *Oncotarget* 2015, **6**:13757–71.

30. Yu H, Pak H, Hammond-Martel I, Ghram M, Rodrigue A, Daou S, Barbour H, Corbeil L, Hebert J, Drobetsky E, Masson JY, Di Noia JM, Affar EB: **Tumor suppressor and deubiquitinase BAP1 promotes DNA double-strand break repair**. *Proc Natl Acad Sci* 2014, **111**:285–290.

31. Ismail IH, Davidson R, Gagné J-P, Xu ZZ, Poirier GG, Hendzel MJ: **Germline mutations in BAP1 impair its function in DNA double-strand break repair.** *Cancer Res* 2014, **74**:4282–94.

32. Mouw KW, Goldberg MS, Konstantinopoulos PA, D'Andrea AD: **DNA Damage and Repair Biomarkers of Immunotherapy Response.** *Cancer Discov* 2017, **7**:675–693.

33. Chen F, Zhang Y, Şenbabaoğlu Y, Ciriello G, Yang L, Reznik E, Shuch B, Micevic G, De Velasco G, Shinbrot E, Noble MS, Lu Y, Covington KR, Xi L, Drummond JA, Muzny D, Kang H, Lee J, Tamboli P, Reuter V, Shelley CS, Kaipparettu BA, Bottaro DP, Godwin AK, Gibbs RA, Getz G, Kucherlapati R, Park PJ, Sander C, Henske EP, et al.: **Multilevel Genomics-Based Taxonomy of Renal Cell Carcinoma**. *Cell Rep* 2016, **14**:2476–2489.

34. Peña-Llopis S, Vega-Rubín-de-Celis S, Liao A, Leng N, Pavía-Jiménez A, Wang S, Yamasaki T, Zhrebker L, Sivanand S, Spence P, Kinch L, Hambuch T, Jain S, Lotan Y, Margulis V, Sagalowsky AI, Summerour PB, Kabbani W, Wong SWW, Grishin N, Laurent M, Xie X-J, Haudenschild CD, Ross MT, Bentley DR, Kapur P, Brugarolas J: **BAP1 loss defines a new class of renal cell carcinoma**. *Nat Genet* 2012, **44**:751–759.

35. Robertson AG, Shih J, Yau C, Gibb EA, Oba J, Mungall KL, Hess JM, Uzunangelov V, Walter V, Danilova L, Lichtenberg TM, Kucherlapati M, Kimes PK, Tang M, Penson A, Babur O, Akbani R, Bristow CA, Hoadley KA, Iype L, Chang MT, Abdel-Rahman MH, Akbani R, Ally A, Auman JT, Babur O, Balasundaram M, Balu S, Benz C, Beroukhim R, et al.: **Integrative Analysis Identifies Four Molecular and Clinical Subsets in Uveal Melanoma**. *Cancer Cell* 2017, **32**:204–220.

36. Miao D, Margolis CA, Gao W, Voss MH, Li W, Martini DJ, Norton C, Bossé D, Wankowicz SM, Cullen D, Horak C, Wind-Rotolo M, Tracy A, Giannakis M, Hodi FS, Drake CG, Ball MW, Allaf ME, Snyder A, Hellmann MD, Ho T, Motzer RJ, Signoretti S, Kaelin WG, Choueiri TK, Van Allen EM: **Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma.** *Science* 2018, **5951**(January):1–11.

37. Pan D, Kobayashi A, Jiang P, Ferrari de Andrade L, Tay RE, Luoma A, Tsoucas D, Qiu X, Lim K, Rao P, Long HW, Yuan G, Doench J, Brown M, Liu S, Wucherpfennig KW: **A major chromatin regulator determines resistance of tumor cells to T cell-mediated killing.** *Science* 2018, **1710**(January):1–12.

38. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, Biedrzycki B, Donehower RC, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Duffy SM, Goldberg RM, de la Chapelle A, Koshiji M, Bhaijee F, Huebner T, Hruban RH, Wood LD, Cuka N, Pardoll DM, Papadopoulos N, Kinzler KW, Zhou S, Cornish TC, et al.: **PD-1 Blockade in Tumors with Mismatch-Repair Deficiency**. *N Engl J Med* 2015, **372**:2509–2520.

39. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Lu S, Kemberling H, Wilt C, Luber BS, Wong F, Azad NS, Rucki AA, Laheru D, Donehower R, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Greten TF, Duffy AG, Ciombor KK, Eyring AD, Lam BH, Joe A, Kang SP, Holdhoff M, Danilova L, Cope L, Meyer C, et al.: **Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade.** *Science* 2017, **357**:409–413.

40. Germano G, Lamba S, Rospo G, Barault L, Magrì A, Maione F, Russo M, Crisafulli G, Bartolini A, Lerda G, Siravegna G, Mussolin B, Frapolli R, Montone M, Morano F, de Braud F, Amirouchene-Angelozzi N, Marsoni S, D'Incalci M, Orlandi A, Giraudo E, Sartore-Bianchi A, Siena S, Pietrantonio F, Di Nicolantonio F, Bardelli A: **Inactivation of DNA repair triggers neoantigen generation and impairs tumour growth**. *Nature* 2017.

41. Alley EW, Lopez J, Santoro A, Morosky A, Saraf S, Piperdi B, van Brummelen E: **Clinical safety and activity of pembrolizumab in patients with malignant pleural mesothelioma (KEYNOTE-028): preliminary results from a non-randomised, open-label, phase 1b trial**. *Lancet Oncol* 2017, **18**:623–630.

42. Calabrò L, Morra A, Fonsatti E, Cutaia O, Amato G, Giannarelli D, Di Giacomo AM, Danielli R, Altomonte M, Mutti L, Maio M: **Tremelimumab for patients with chemotherapy-resistant advanced malignant mesothelioma: An open-label, single-arm, phase 2 trial**. *Lancet Oncol* 2013, **14**:1104–1111.

43. Calabrò L, Morra A, Fonsatti E, Cutaia O, Fazio C, Annesi D, Lenoci M, Amato G, Danielli R, Altomonte M, Giannarelli D, Di Giacomo AM, Maio M: **Efficacy and safety of an intensified schedule of tremelimumab for chemotherapy-resistant malignant mesothelioma: An open-label, single-arm, phase 2 study**. *Lancet Respir Med* 2015, **3**:301–309.

44. Maio M, Scherpereel A, Calabrò L, Aerts J, Perez SC, Bearz A, Nackaerts K, Fennell DA, Kowalski D, Tsao AS, Taylor P, Grosso F, Antonia SJ, Nowak AK, Taboada M, Puglisi M, Stockman PK, Kindler HL: **Tremelimumab as second-line or third-line treatment in relapsed malignant mesothelioma (DETERMINE): a multicentre, international, randomised, double-blind, placebo-controlled phase 2b trial**. *Lancet Oncol* 2017:1–13.

45. DePristo MA, Banks E, Poplin R, Garimella K V, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ: **A framework for variation discovery and genotyping using next-generation DNA sequencing data**. *Nat Genet* 2011, **43**:491–498.

46. Wang K, Li M, Hakonarson H: **ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data**. *Nucleic Acids Res* 2010, **38**:1–7.

47. Thorvaldsdóttir H, Robinson JT, Mesirov JP: **Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration**. *Brief Bioinform* 2013, **14**:178–192.

48. Dobin A, Davis C a., Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: **STAR: Ultrafast universal RNA-seq aligner**. *Bioinformatics* 2013, **29**:15–21.

49. Anders S, Pyl PT, Huber W: **HTSeq-A Python framework to work with high-throughput sequencing data**. *Bioinformatics* 2015, **31**:166–169.

50. Anders S, Huber W: **Differential expression analysis for sequence count data.** *Genome Biol* 2010, **11**:R106.

51. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG: **Primer3--new capabilities and interfaces.** *Nucleic Acids Res* 2012, **40**:e115.

52. Hughes CS, Foehr S, Garfield DA, Furlong EE, Steinmetz LM, Krijgsveld J: **Ultrasensitive proteome analysis using paramagnetic bead technology**. *Mol Syst Biol* 2014, **10**:757–757.

53. Hughes CS, McConechy MK, Cochrane DR, Nazeran T, Karnezis AN, Huntsman DG, Morin GB: **Quantitative Profiling of Single Formalin Fixed Tumour Sections: proteomics for translational research**. *Sci Rep* 2016, **6**:34949.

54. Alexandrov LB, Jones PH, Wedge DC, Sale JE, Campbell PJ, Nik-Zainal S, Stratton MR: **Clock-like mutational processes in human somatic cells**. *Nat Genet* 2015, **47**:1402–1407.

55. Szklarczyk D, Franceschini a., Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth a., Santos a., Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C: **STRING v10: protein-protein interaction networks, integrated over the tree of life**. *Nucleic Acids Res* 2014, **43**:D447–D452.

56. Wilkerson MD, Hayes DN: **ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking**. *Bioinformatics* 2010, **26**:1572–1573.

57. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.** *Proc Natl Acad Sci U S A* 2005, **102**:15545–50.

Figure Legend

Fig. 1 Landscape of somatic mutations in PeM tumors. **a** Comparison of somatic mutation rate in protein-coding regions of PeM with different adult cancers obtained from TCGA. **b** Mutational signature present in PeM (top panel). Proportional contribution of different COSMIC mutational signature per tumor sample. **c** Somatic alterations identified in PeM tumors group by important cancer-pathways. LUSC: Lung Squamous Cell Carcinoma, LUAD: Lung adenocarcinoma, BLCA: Urothelial Bladder Carcinoma, COAD: Colorectal carcinoma, UCEC: Uterine Corpus Endometrial Carcinoma, OV: Ovarian cancer, KRIP: Kidney renal papillary cell carcinoma, KIRC: Kidney Renal Clear Cell Carcinoma, UCS: Uterine Carcinosarcoma, GBM: Glioblastoma Multiforme, BRCA: Breast Invasive Carcinoma, MESO-PM: Malignant Pleural Mesothelioma, MESO-PeM: Malignant Peritoneal Mesothelioma, PAAD: Pancreatic Adenocarcinoma, PRAD: Prostate Adenocarcinoma, KICH: Kidney Chromophobe, TGCT: Testicular Germ Cell Tumor.

Fig. 2 Landscape of copy number aberrations in PeM tumors. **a** Aggregate copy-number alterations by chromosome regions in PeM tumors. Important genes with copy-number changes are highlighted. b Sample-wise view of copy-number alterations in PeM tumors. **c** Comparison of copy-number burden (considering protein-coding regions only) in PeM with respect to other adult cancers obtained from TCGA. **d** Highly aberrant genomic regions in PeM prioritized by GISTIC. **e** mRNA expression pattern of *BAP1* across all PeM samples. The Wilcoxon signed-rank test *p*-value for *BAP1* mRNA expression compared between the PeM subtypes is indicated in the box. **f** Detection of *BAP1* nuclear protein expression in PeM tumors by immunohistochemistry (Photomicrographs magnification - 20x). **g** Unsupervised consensus clustering of tumor samples based on copy-number segmentation mean values of the 3349 most variable genes.

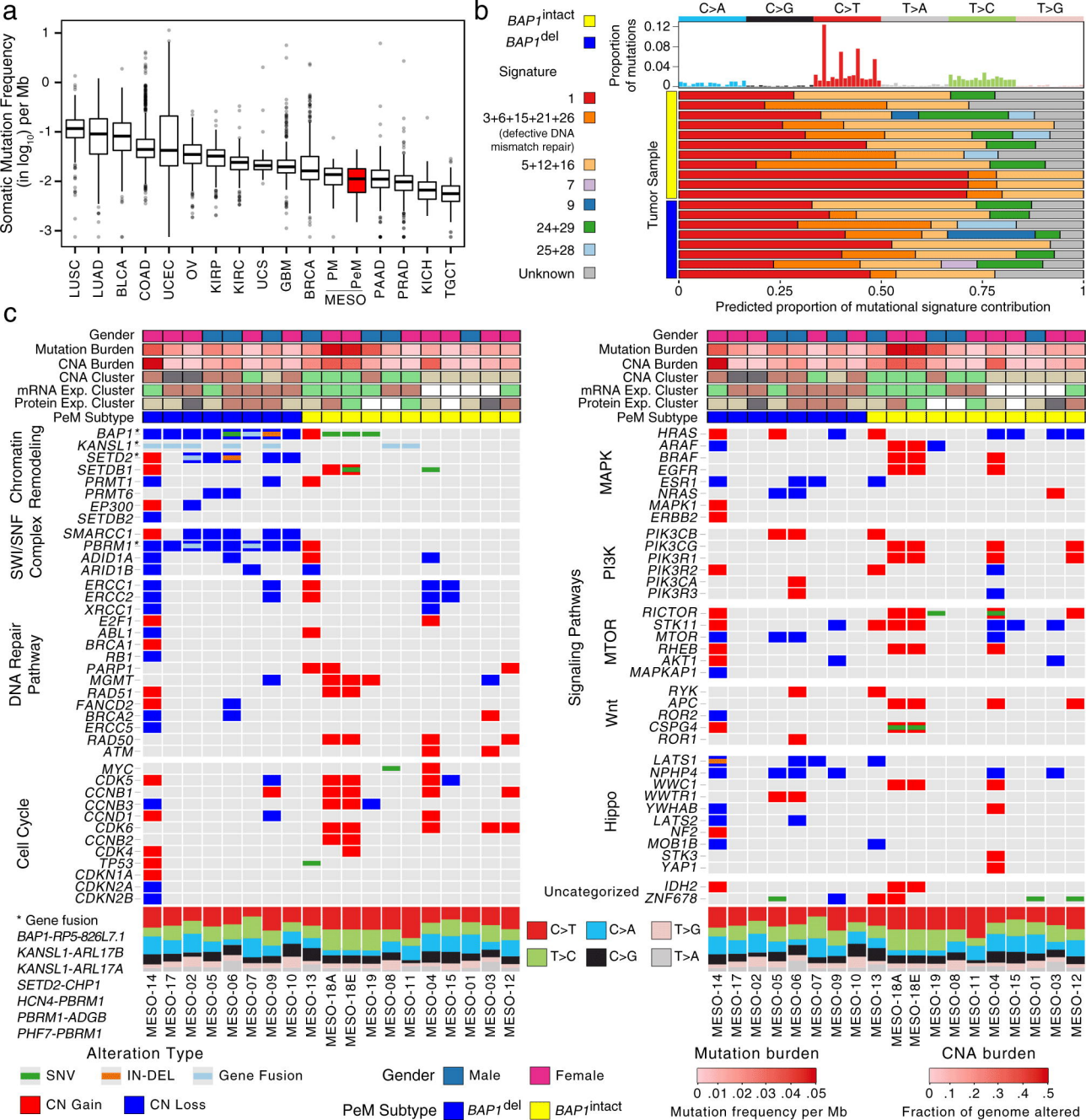
Fig. 3 Gene Fusions in PeM. **a-b** Circos plot showing the gene fusion events identified in PeM tumors. The lines with red and blue color indicates intra-chromosomal and inter-chromosomal gene fusion events respectively. **a** *BAP1*^{intact} subtype, **b** *BAP1*^{del} subtype. **c-f** Few selected gene fusion events identified in PeM tumors. The top and middle panel shows the chromosome and the transcripts involved in the gene fusion event. The bottom panel shows

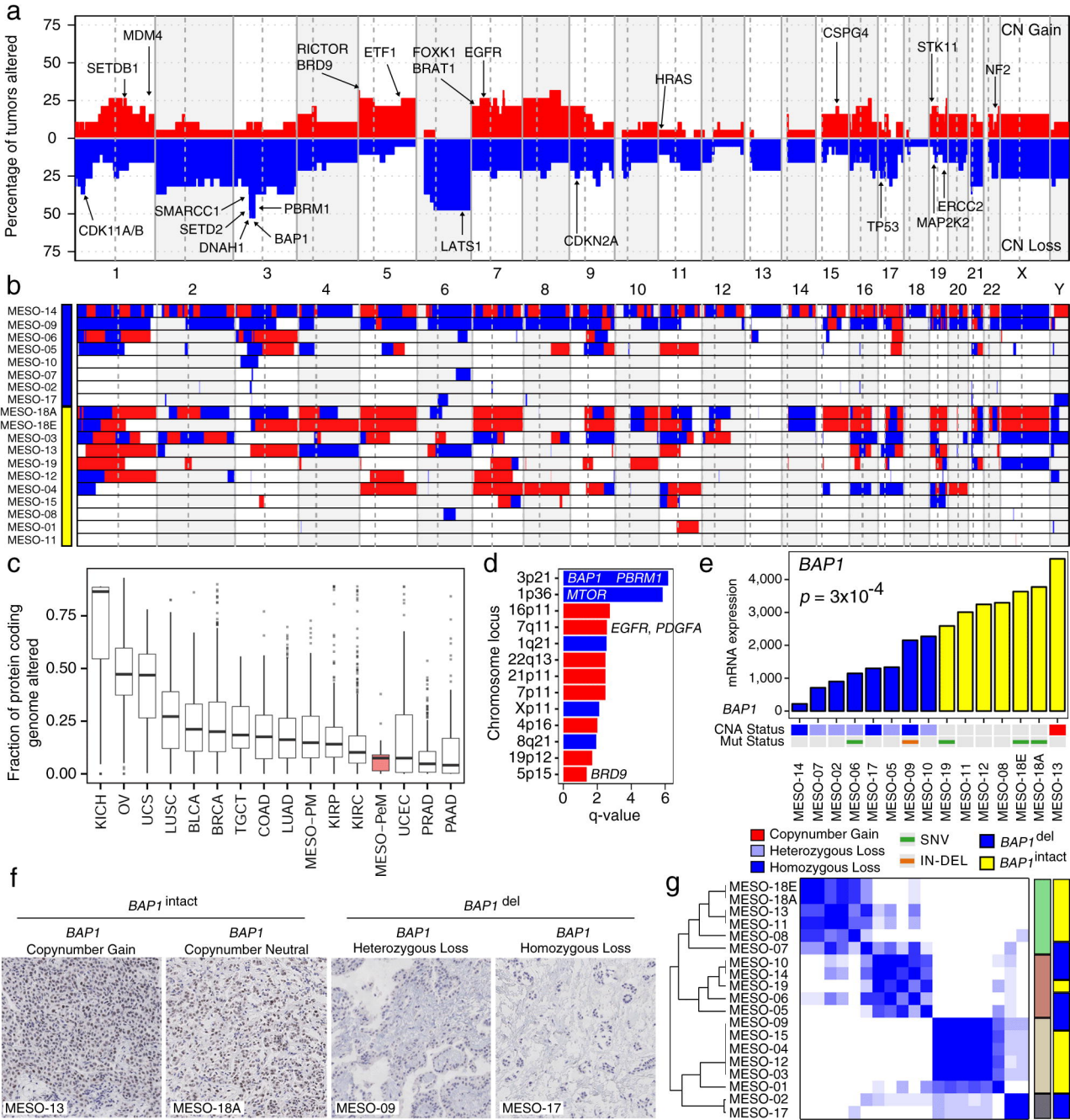
the RNA-seq read counts detected for the respective transcripts. **c** *KANSL1-ARL17B* fusion, **d** *PBRM1-ADGB* fusion, **e** *SETD2-CHP1* fusion, and **f** *PHF7-PBRM1* fusion. **g-j** The chromatogram showing the Sanger sequencing validation of the fusion-junction point.

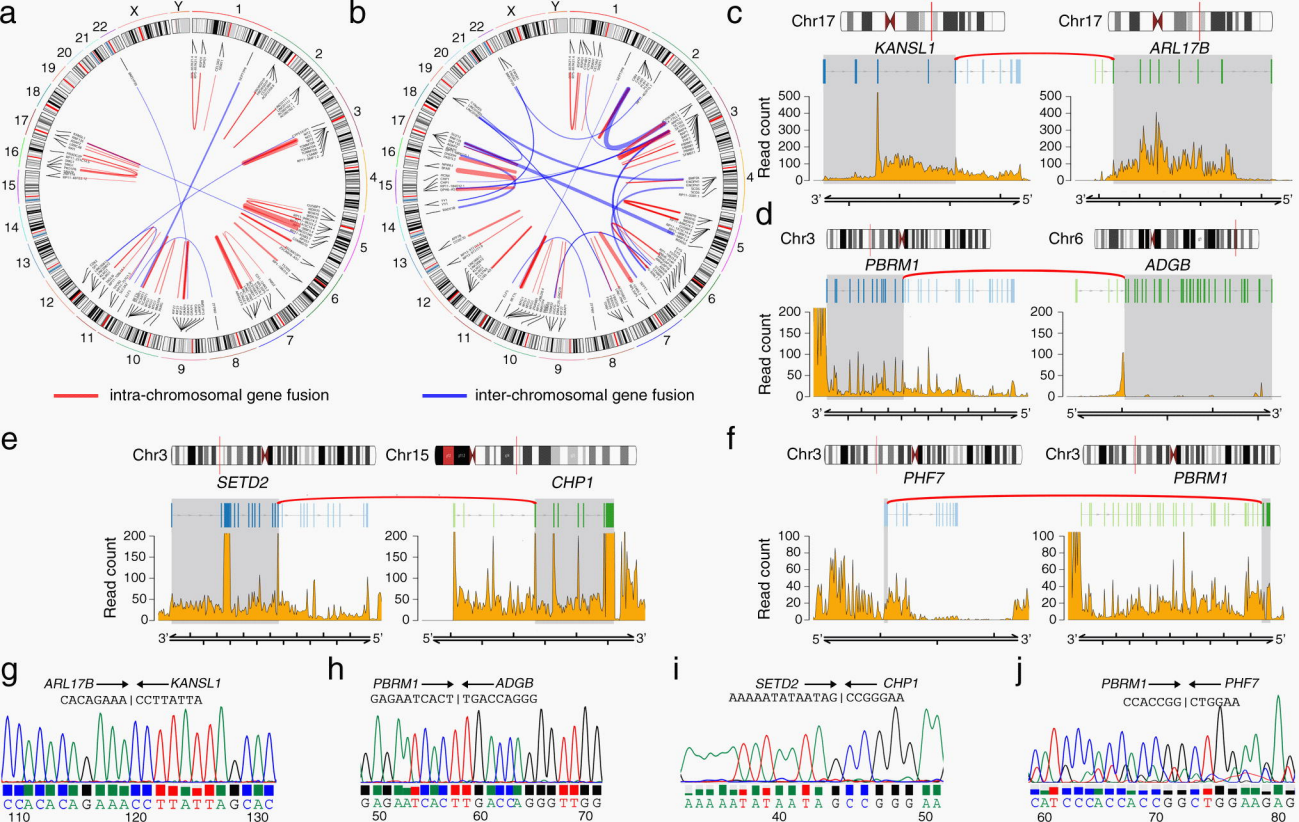
Fig. 4 Transcriptome and proteome profile of PeM. **a-b** Principal component analysis of PeM tumors using **a** transcriptome profiles and **b** proteome profiles. **c** Effects of CNA on transcriptome and proteome. In the scatterplot, each dot represents a gene/protein. The horizontal and vertical axes represent Pearson correlation coefficient between CNA-transcriptome and CNA-proteome respectively. Key cancer genes that undergo protein attenuation have been highlighted. **d** Geneset enrichment analysis of attenuated proteins against gene ontologies (left panel) and Reactome pathways (right panel). **e-g** CORUM core protein complexes regulated by PBRM1 and/or BAP1. The nodes represent individual protein subunit of the respective complex. The node color represents correlation of mRNA expression of respective gene with *BAP1*. The border color of the node indicates whether the respective protein is attenuated or not. The edge represents interaction between the protein subunits. The edge information were extracted from STRING v10 PPI network. The edge color (and edge thickness) represents correlation of protein expression between the respective interaction partners. **e** SWI/SNF complex B (PBAF), **f** SWI/SNF complex A (BAF), and **g** HDAC complex. **h-i** mRNA and protein expression level differences between PeM subtypes. **h** SWI/SNF complex and **i** HDAC complex. The expression levels are log2 transformed and mean normalized.

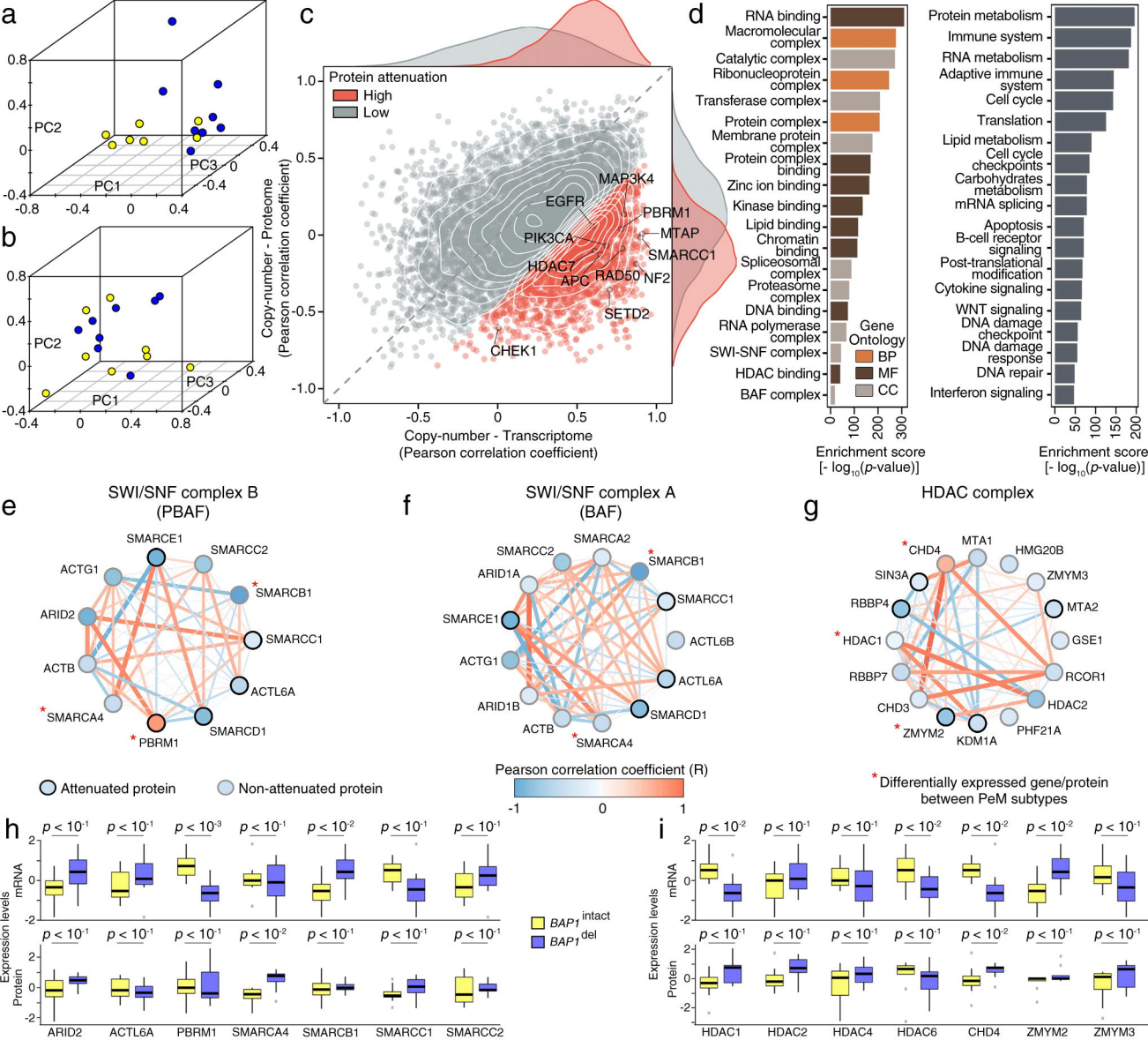
Fig. 5 Immune cell infiltration in PeM tumors. **a-b** Pathways enrichment of top-500 differentially expressed genes between PeM subtypes obtained using **a** mRNA expression and **b** protein expression. **c-d** Correlation between immune score and stromal score derived for each tumor sample using **c** mRNA expression and **d** protein expression. **e** Estimated relative mRNA fractions of leukocytes infiltrated in PeM tumors based on CIBERSORT analysis. **f** CD3 and CD8 immunohistochemistry showing immune cell infiltration on *BAP1*^{del} PeM tumor (Photomicrographs magnification - 20x). **g** mRNA expression differences in immune checkpoint receptors between PeM subtypes. The bar plot on the right represents negative log10 of Wilcoxon signed-rank test *p*-value

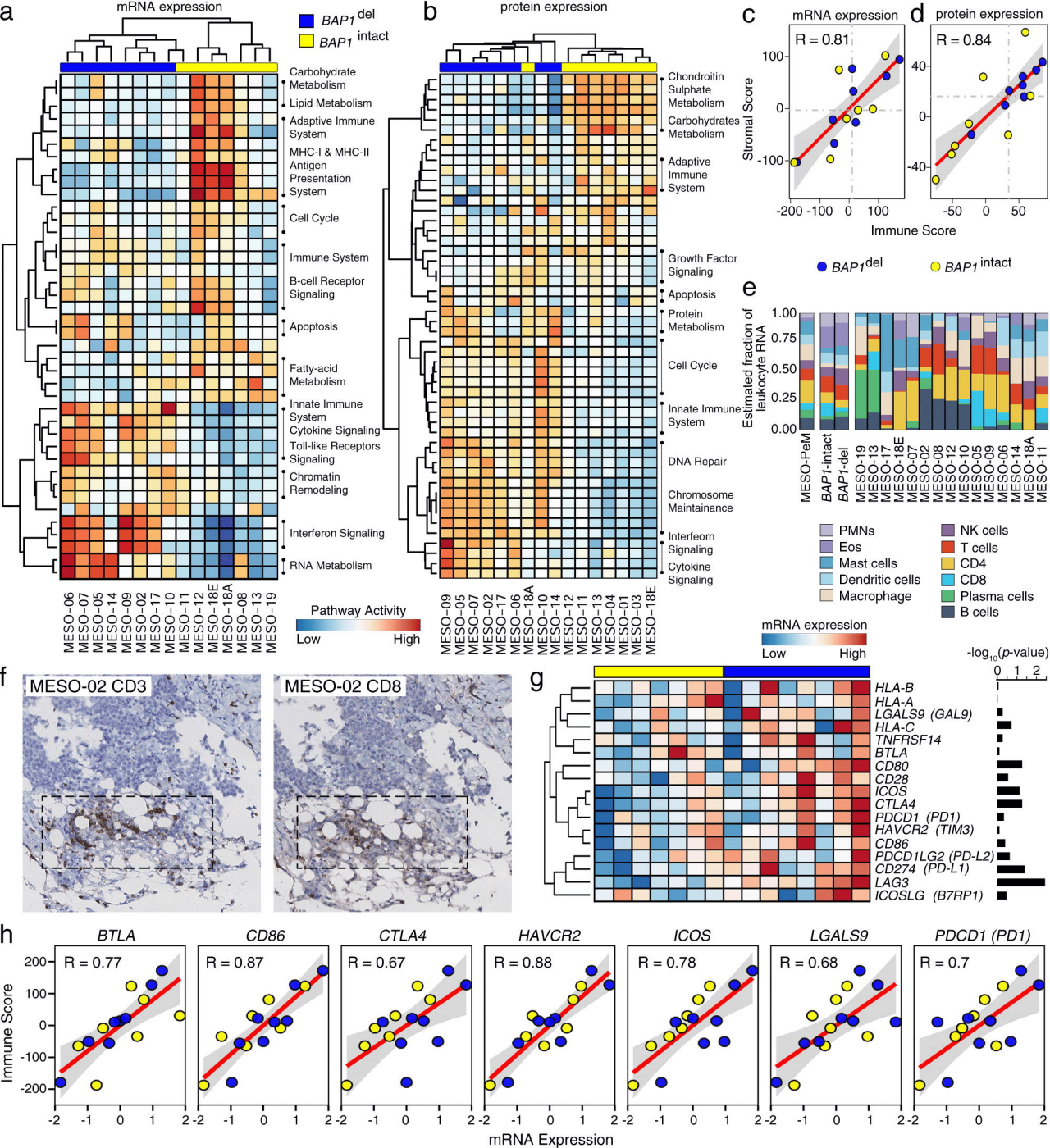
computed between PeM subtypes. **h** Correlation between immune score and mRNA expression of immune checkpoint receptors. The expression levels are log2 transformed and mean normalized


a
Somatic Mutation Frequency (in log10) per Mb
LUSC
LUAD
BLCA
COAD
UCEC
OV
KIRP
KIRC
UCS
GBM
BRCA
PM
PeM
MESO
PAAD
PRAD
KICH
TGCT
b
BAP1intact
BAP1del
Signature
1
3+6+15+21+26 (defective DNA mismatch repair)
5+12+16
7
9
24+29
25+28
Unknown
C>A
C>G
C>T
T>A
T>C
T>G
Proportion of mutations
Tumor Sample
Predicted proportion of mutational signature contribution
c
Gender
Mutation Burden
CNA Burden
CNA Cluster
mRNA Exp. Cluster
Protein Exp. Cluster
PeM Subtype
Chromatin Remodeling
BAP1*
KANSL1*
SETD2*
SETDB1
PRMT1
PRMT6
EP300
SETDB2
SWI/SNF Complex
SMARCC1
PBRM1*
ADID1A
ARID1B
DNA Repair Pathway
ERCC1
ERCC2
XRCC1
E2F1
ABL1
BRCA1
RB1
PARP1
MGMT
RAD51
FANCD2
BRCA2
ERCC5
RAD50
ATM
Cell Cycle
MYC
CDK5
CCNB1
CCNB3
CCND1
CDK6
CCNB2
CDK4
TP53
CDKN1A
CDKN2A
CDKN2B
Signaling Pathways
MAPK
HRAS
ARAF
BRAF
EGFR
ESR1
NRAS
MAPK1
ERBB2
PI3K
PIK3CB
PIK3CG
PIK3R1
PIK3R2
PIK3CA
PIK3R3
MTOR
RICTOR
STK11
MTOR
RHEB
AKT1
MAPKAP1
Wnt
RYK
APC
ROR2
CSPG4
ROR1
Hippo
LATS1
NPHP4
WWC1
WWTR1
YWHAB
LATS2
NF2
MOB1B
STK3
YAP1
Uncategorized
IDH2
ZNF678
* Gene fusion
BAP1-RP5-826L7.1
KANSL1-ARL17B
KANSL1-ARL17A
SETD2-CHP1
HCN4-PBRM1
PBRM1-ADGB
PHF7-PBRM1
MESO-14
MESO-17
MESO-02
MESO-05
MESO-06
MESO-07
MESO-09
MESO-10
MESO-13
MESO-18A
MESO-18E
MESO-19
MESO-08
MESO-11
MESO-04
MESO-15
MESO-01
MESO-03
MESO-12
C>T
C>A
T>G
T>C
C>G
T>A
Alteration Type
SNV
IN-DEL
Gene Fusion
CN Gain
CN Loss
Gender
Male
Female
PeM Subtype
BAP1del
BAP1intact
Mutation burden
0 .01 .02 .03 .04 .05
Mutation frequency per Mb
CNA burden
0 .1 .2 .3 .4 .5
Fraction of genome altered


a
Percentage of tumors altered
CN Gain
CN Loss
SETDB1
MDM4
RICTOR
BRD9
ETF1
FOXK1
BRAT1
EGFR
HRAS
CSPG4
STK11
NF2
CDK11A/B
SMARCC1
SETD2
DNAH1
PBRM1
BAP1
LATS1
CDKN2A
TP53
MAP2K2
ERCC2
b
MESO-14
MESO-09
MESO-06
MESO-05
MESO-10
MESO-07
MESO-02
MESO-17
MESO-18A
MESO-18E
MESO-03
MESO-13
MESO-19
MESO-12
MESO-04
MESO-15
MESO-08
MESO-01
MESO-11
c
Fraction of protein coding genome altered
KICH
OV
UCS
LUSC
BLCA
BRCA
TGCT
COAD
LUAD
MESO-PM
KIRP
KIRC
MESO-PeM
UCEC
PRAD
PAAD
d
Chromosome locus
BAP1 PBRM1
MTOR
EGFR, PDGFA
BRD9
q-value
e
BAP1
p = 3x10-4
mRNA expression
CNA Status
Mut Status
Copynumber Gain
Heterozygous Loss
Homozygous Loss
SNV
IN-DEL
BAP1 del
BAP1 intact
f
BAP1 intact
BAP1 del
BAP1 Copynumber Gain
BAP1 Copynumber Neutral
BAP1 Heterozygous Loss
BAP1 Homozygous Loss
MESO-13
MESO-18A
MESO-09
MESO-17
g


a
b
c
d
e
f
g
h
i
j
intra-chromosomal gene fusion
inter-chromosomal gene fusion
Chr17
KANSL1
ARL17B
Read count
Chr3
PBRM1
Chr6
ADGB
Chr15
SETD2
CHP1
PHF7
ARL17B → ← KANSL1
CACAGAAA | CCTTATTA
PBRM1 → ← ADGB
GAGAATCACT | TGACCAGGG
SETD2 → ← CHP1
AAAAATATAATAG | CCGGGAA
PBRM1 → ← PHF7
CCACCGG | CTGGAA


a
b
PC1
PC2
PC3
c
Protein attenuation
High
Low
Copy-number - Proteome
(Pearson correlation coefficient)
Copy-number - Transcriptome
(Pearson correlation coefficient)
MAP3K4
EGFR
PBRM1
MTAP
PIK3CA
SMARCC1
HDAC7
APC
RAD50
NF2
SETD2
CHEK1
d
RNA binding
Macromolecular complex
Catalytic complex
Ribonucleoprotein complex
Transferase complex
Protein complex
Membrane protein complex
Protein complex binding
Zinc ion binding
Kinase binding
Lipid binding
Chromatin binding
Spliceosomal complex
Proteasome complex
DNA binding
RNA polymerase complex
SWI-SNF complex
HDAC binding
BAF complex
Gene Ontology
BP
MF
CC
Enrichment score
[- log10(p-value)]
Protein metabolism
Immune system
RNA metabolism
Adaptive immune system
Cell cycle
Translation
Lipid metabolism
Cell cycle checkpoints
Carbohydrates metabolism
mRNA splicing
Apoptosis
B-cell receptor signaling
Post-translational modification
Cytokine signaling
WNT signaling
DNA damage checkpoint
DNA damage response
DNA repair
Interferon signaling
e
SWI/SNF complex B
(PBAF)
SMARCE1
SMARCC2
ACTG1
SMARCB1
ARID2
SMARCC1
ACTB
ACTL6A
SMARCA4
SMARCD1
PBRM1
f
SWI/SNF complex A
(BAF)
SMARCA2
SMARCC2
SMARCB1
ARID1A
SMARCC1
SMARCE1
ACTL6B
ACTG1
ACTL6A
ARID1B
SMARCD1
ACTB
SMARCA4
g
HDAC complex
CHD4
MTA1
HMG20B
SIN3A
ZMYM3
RBBP4
MTA2
HDAC1
GSE1
RBBP7
RCOR1
CHD3
HDAC2
ZMYM2
KDM1A
PHF21A
Attenuated protein
Non-attenuated protein
Pearson correlation coefficient (R)
-1
0
1
Differentially expressed gene/protein between PeM subtypes
h
i
Expression levels
mRNA
Protein
BAP1 intact
BAP1 del
ARID2
ACTL6A
PBRM1
SMARCA4
SMARCB1
SMARCC1
SMARCC2
HDAC1
HDAC2
HDAC4
HDAC6
CHD4
ZMYM2
ZMYM3


a
mRNA expression
BAP1 del
BAP1 intact
Carbohydrate Metabolism
Lipid Metabolism
Adaptive Immune System
MHC-I & MHC-II Antigen Presentation System
Cell Cycle
Immune System
B-cell Receptor Signaling
Apoptosis
Fatty-acid Metabolism
Innate Immune System Cytokine Signaling
Toll-like Receptors Signaling
Chromatin Remodeling
Interferon Signaling
RNA Metabolism
MESO-06 MESO-07 MESO-05 MESO-14 MESO-09 MESO-02 MESO-17 MESO-10 MESO-11 MESO-12 MESO-18E MESO-18A MESO-08 MESO-13 MESO-19
Pathway Activity
Low High
b
protein expression
Chondroitin Sulphate Metabolism
Carbohydrates Metabolism
Adaptive Immune System
Growth Factor Signaling
Apoptosis
Protein Metabolism
Cell Cycle
Innate Immune System
DNA Repair
Chromosome Maintainance
Interferon Signaling
Cytokine Signaling
MESO-09 MESO-05 MESO-07 MESO-02 MESO-17 MESO-06 MESO-18A MESO-10 MESO-14 MESO-12 MESO-11 MESO-13 MESO-04 MESO-01 MESO-03 MESO-18E
c
mRNA expression
R = 0.81
Stromal Score
d
protein expression
R = 0.84
Immune Score
BAP1 del
BAP1 intact
e
Estimated fraction of leukocyte RNA
MESO-PeM BAP1-intact BAP1-del MESO-19 MESO-13 MESO-17 MESO-18E MESO-07 MESO-02 MESO-08 MESO-12 MESO-10 MESO-05 MESO-09 MESO-06 MESO-14 MESO-18A MESO-11
PMNs
Eos
Mast cells
Dendritic cells
Macrophage
NK cells
T cells
CD4
CD8
Plasma cells
B cells
f
MESO-02 CD3
MESO-02 CD8
g
mRNA expression
Low High
-log10(p-value)
HLA-B
HLA-A
LGALS9 (GAL9)
HLA-C
TNFRSF14
BTLA
CD80
CD28
ICOS
CTLA4
PDCD1 (PD1)
HAVCR2 (TIM3)
CD86
PDCD1LG2 (PD-L2)
CD274 (PD-L1)
LAG3
ICOSLG (B7RP1)
h
BTLA R = 0.77
CD86 R = 0.87
CTLA4 R = 0.67
HAVCR2 R = 0.88
ICOS R = 0.78
LGALS9 R = 0.68
PDCD1 (PD1) R = 0.7
Immune Score
mRNA Expression