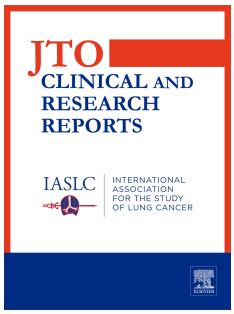
Journal Pre-proof



Viral mimicry response is associated with clinical outcome in pleural mesothelioma

Suna Sun, Weihong Qi, Hubert Rehrauer, Manuel Ronner, Ananya Hariharan, Martin Wipplinger, Clément Meiller, Rolf Stahel, Martin Früh, Ferdinando Cerciello, Jean-François Fonteneau, Didier Jean, Emanuela Felley-Bosco

PII: S2666-3643(22)00154-0

DOI: https://doi.org/10.1016/j.jtocrr.2022.100430

Reference: JTOCRR 100430

To appear in: *JTO Clinical and Research Reports*

Received Date: 5 July 2022

Revised Date: 18 October 2022

Accepted Date: 24 October 2022

Please cite this article as: Sun S, Qi W, Rehrauer H, Ronner M, Hariharan A, Wipplinger M, Meiller C, Stahel R, Früh M, Cerciello F, Fonteneau J-F, Jean D, Felley-Bosco E, Viral mimicry response is associated with clinical outcome in pleural mesothelioma, *JTO Clinical and Research Reports* (2022), doi: https://doi.org/10.1016/j.jtocrr.2022.100430.

Viral mimicry response is associated with clinical outcome in pleural mesothelioma

Suna Sun[1], Weihong Qi[2], Hubert Rehrauer[2], Manuel Ronner[1], Ananya Hariharan[1], Martin Wipplinger[1], Clément Meiller[3], Rolf Stahel[4,5], Martin Früh[5,6,7], Ferdinando Cerciello[7], Jean-François Fonteneau[8], Didier Jean[3], Emanuela Felley-Bosco[1]

1. Laboratory of Molecular Oncology, Department of Thoracic Surgery, University Hospital Zurich
2. Functional Genomics Center, ETH Zurich, University of Zurich
3. Centre de Recherche des Cordeliers, Inserm, Sorbonne Université, Université Paris Cité, Functional Genomics of Solid Tumors, F-75006 Paris, France
4. Coordinating Center, European Thoracic Oncology Platform (ETOP), Bern, Switzerland,
5. Swiss Group for Clinical Cancer Research (SAKK)
6. Department of Medical Oncology/Hematology, Cantonal Hospital of St. Gallen, St. Gallen, Switzerland;
7. Department of Medical Oncology, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland;
8. Nantes Université, Inserm UMR 1307, CNRS UMR 6075, Université d'Angers, CRCI2NA, F-44000 Nantes, France

Corresponding author:

Emanuela Felley-Bosco, Laboratory of Molecular Oncology, Department of Thoracic Surgery, University Hospital Zurich, Sternwartstrasse 14, 8091 Zürich, Switzerland

Email: emanuela.felley-bosco@usz.ch

Disclosure of conflict of interest: Dr. Früh reports grants from BMS and Astra Zeneca, other from AstraZeneca, Merck Sharp & Dohme; Roche, Bristol-Myers Squibb; Boehringer Ingelheim, Pfizer, Takeda, outside the submitted work

Abstract

Introduction: The aim of this study was to investigate endogenous retrovirus expression (ERV) expression and type-I-interferon (IFN) activation in human pleural mesothelioma (PM) and their association with clinical outcome.

Methods: ERVs' expression was determined from PM cohorts and mesothelial precursors RNA-seq data. ERV's expression was confirmed by qPCR. Methylation of genomic DNA was assessed by quantitative methylation specific PCR. DNA demethylation was induced in cells by demethylating agent 5-Aza-2'-deoxycytidine (5-Aza-CdR) treatment. To block the type-I IFN signaling, cells were treated with Ruxolitinib or MAVS silencing. IFN stimulated genes (ISGs) expression was determined by qPCR and Western Blot. Circulating ERVs were detected by qPCR.

Results: Long-terminal-repeats (LTR) represent the most abundant transposable elements upregulated in PM. Within LTR, *ERVmap_1248* and *LTR7Y*, which are specifically enriched in PM, were further analyzed. 5-Aza-CdR treatment increased the levels of *ERVmap_1248* expression and induced *ERVmap_1248* promoter demethylation in mesothelial cells. In addition, *ERVmap_1248* promoter was more demethylated in mesothelioma tissue compared to non-tumor tissue. 5-Aza-CdR treatment of mesothelial cells also increased levels of ISGs. Basal ISGs expression was higher in mesothelioma cells compared to mesothelial cells and it was significantly decreased by Ruxolitinib treatment or MAVS silencing. Furthermore, ISGs expression was higher in tumor tissue with high expression levels of *ERVmap_1248*. High expression of *ERVmap_1248* was associated with longer overall survival and *BAP1* mutations. *ERVmap_1248* and *LTR7Y* can be detected in PM plasma.

Conclusions: We provide clues for patient stratification especially for immunotherapy where best clinical responses are associated with an activated basal immune response.

Keywords: pleural mesothelioma, endogenous retroviruses, type-I interferon

Abbreviations:

5-Aza-CdR: 5-Aza-2'-deoxycytidine; BAP1: BRCA-associated protein 1; CDKN2A/B: Cyclin-dependent kinase inhibitor 2A/B; DNMTi: DNA methylthransferase inhibitor; dsRNA: double-stranded RNA; ERV: endogenous retrovirus; IFN: interferon, IRF3: interferon regulatory factor 3; ISG: interferon-stimulated genes; ISG15: interferon-stimulated gene 15; KZFP: KRAB-Zinc Finger Proteins; LINE: long-interspersed nuclear elements; LTR: long-terminal-repeat; MAVS: mitochondrial antiviral signaling protein; MDA5: melanoma differentiation-associated protein 5; NF2: neurofibromatosis type 2; PM: Pleural mesothelioma; qMSP: quantitative methylation specific PCR; RBM8A: RNA Binding Motif 8A; RIG-I: retinoic acid

inducible gene-I; RNA-seq: RNA-sequencing; SINE: short-interspersed nuclear elements; STAT1: signal transducer and activator of transcription 1; SVA: SINE-VNTR-Alu, TE: transposable elements

Introduction

Pleural mesothelioma (PM) is a rapidly fatal disease arising from the monolayer tissue lining the walls of pleural cavity and the internal organs housed inside [1]. Major drivers include *Cyclin-dependent kinase inhibitor 2A/B* (*CDKN2A/B*) and the more PM specific *BRCA-associated protein1* (*BAP1*), and *neurofibromatosis type 2* (*NF2*) mutations [1]. Recent data suggests that subclonal *NF2* mutations may occur later in mesothelioma development [2, 3]. Traditionally, the major histologic types of mesothelioma have been the main histologic indicators of prognosis. Indeed, patients with sarcomatoid and biphasic tumors have significantly worse overall survival compared to patients with epithelioid tumors [4]. Recent studies based on multi omics approaches [5-8] have refined the classification into four groups or into gradients based on molecular profiles.

PM is mostly associated with previous exposure to asbestos fibers [1] and we have recently shown that exposure to asbestos in mice leads to increased expression levels of endogenous retrovirus (ERV) sequences [9].

ERVs are integrated retroviral elements that cover 8% of the human genome [10]. They are part of the so-called transposable elements (TE), which include retrotransposons using RNA as an intermediate that is reverse transcribed into DNA and integrated in the genome, and DNA transposons directly excising themselves from one location before reinsertion (reviewed in[11]). ERV and long-interspersed nuclear elements (LINE) are autonomous retroelements encoding required proteins for retrotransposition, while short-interspersed nuclear elements (SINE) and SINE-VNTR-Alu (SVA) elements require the machinery from autonomous retrotransposons.

Many ERV sequences are expressed during embryo development and are subsequently epigenetically silenced [12]. However, certain ERV sequences are actively transcribed and are elevated in cancer [13]. Most ERVs in the human genome are non-autonomous long-terminal-repeat (LTR) elements that are either solitary (solo)-LTR or LTR flanking a small segment of internal ERV sequences and are short in length. They are likely to serve as genomic regulators and affect the transcription in *cis* [14]. On the other hand, autonomous LTRs are composed of LTRs that flank potential protein coding sequences and are near full-length proviral sequences, which could encode disease-associated antigens or functional RNA that regulated gene expression in *trans* [14]. Besides the effects as transcription regulators, the expression of ERV has been recently explored for its property as inducers of viral mimicry response, especially in immunotherapy context, and we and others have observed that expression of interferon-induced genes is associated with the clinical outcome in mesothelioma patients [9, 15].

Studying ERV expression has not been regularly implemented in high throughput studies because repetitive elements are not frequently investigated and the analysis of the few RNA-seq data in mesothelioma has mostly focused on the investigations of known genes [5, 7, 16]. Cancer-specific LTR-retroelements are mostly cancer-type specific [17] and mesothelioma is one of the cancer types with the highest number of expressed cancer-specific LTR-retroelement (eighth out of 31 cancer types)[17], however for the time being ERV expression in mesothelioma has not been thoroughly explored.

In this study, we extended previous work on ERV expression in human cancers [17] and mesothelioma experimental animal models [9], and we reveal the expression of ERVs in human mesothelioma that can be detected in the blood and are associated with type-I interferon (IFN) signaling as well as better overall survival.

Materials and Methods

ERV analysis

Mesothelioma RNA-seq reads included in the analysis were: the TCGA- Mesothelioma cohort (n = 87) downloaded from the NCBI database of Genotypes and Phenotypes (dbGaP) in 2019, under phs000178.v10.p8; the Pleural Mesothelioma cohort from the Bueno study (n=211) downloaded from the European Genome-phenome Archive (EGA) in 2020, under EGAS00001001563 (EGAD00001001915 and EGAD00001001916); the genetically characterized pleural mesothelioma primary cultures (FunGeST, n=64) provided by Didier Jean's team in 2022 for which RNA-Seq was performed as described in [2]; the human embryonic stem cell derived mesothelium (n=10) downloaded from the NCBI Gene Expression Omnibus (GEO) in 2020, under GSE113090 (GSM3096389- GSM3096398). The choice of using human embryonic stem cell derived mesothelium as control was dictated by the fact that RNA-Seq data on normal mesothelial cells or pleura are not available and ERV expression is cancer specific. RNA-seq reads were pre-processed using fastp (0.20.0). For analysis of transposable elements (TE) expression, TEtranscripts [18] was used to obtain TE counts, followed by differential expression analysis using DESeq2. TE loci were considered to be significantly differentially expressed when the adjusted p-values were <0.01, and where the log2 of the fold change was >1 for upregulated loci and < -1 for downregulated loci. Full length ERV sequences were downloaded from ERVmap [19]. ERV expression was quantified using featureCounts in the Bioconductor package Rsubreads, where reads mapped uniquely to ERVs were counted. Mutational status of *BAP1*, *NF2* and *CDKN2A* was extracted from TCGA [7] dataset.

PM patients and healthy donors

Tumor tissue was collected from 155 PM patients and non-tumor tissue from 6 non-PM patients between 1999-2015 [20](Supplementary Table 1A). The study was approved by the Ethical Committee Zürich (KEK-

ZH-2012-0094 and BASEC-No. 2020-02566), and patients either signed informed consent or waiver of consent was granted by the Ethical Committee (BASEC-No. 2020-02566). The study methodologies were conformed to the standards set by the Declaration of Helsinki. Tissue samples were processed immediately for total RNA extraction or frozen as previously described [20]. In order to assess that extracted RNA corresponded to tumor tissue containing at least 50% tumor content, we evaluated the PM score, which is based on *MSLN*, *CALB2* and *PDPN* expression levels as previously described [21].

Plasma was collected from 76 PM patients between 2005-2012 and 42 healthy donors between 2002-2017(Supplementary Table 1B). PM patients were enrolled in the trial SAKK 17/04 (NCT00334594)[22]. The study was approved by the Ethics Committee of the Zurich University Hospital (KEK-StV-Nr. 24/05), and patients either signed informed consent or waiver of consent was granted by the Ethical Committee (KEK-StV-Nr. 24/05). The study methodologies were conformed to the standards set by the Declaration of Helsinki.

Statistical analysis

The figures represent the mean values from at least three independent experiments. Paired and unpaired t-test, Mann–Whitney, Chi-square test, Gehan-Breslow-Wilcoxon test or Pearson correlation analysis were used and have been specified when used. Error bars indicate the standard error of the mean. Statistical analysis was performed using Prism 8 (Graphpad 8.0.0).

All other methods

Detailed description of the methods used is available in Supplementary Methods.

Results

LTR represent the most abundant TE upregulated or downregulated in PM patient datasets

We employed the RNA-sequencing (RNA-seq) data from TCGA[7], Bueno's [5] tumor samples cohorts, or mesothelioma primary cultures cell lines of FunGeST series [6, 23] as well as human embryonic stem cell derived mesothelium [24] as non-tumor control, to determine which TE subfamilies are differentially expressed in PM tumors or primary cultures. We observed that LTR represent the most abundant TE upregulated or downregulated in PM patient datasets compared to mesothelial precursors. 86%, 57% and 80% of LTR represent more than two-fold upregulated TE in TCGA, Bueno and FunGeST datasets, respectively, 38%, 56% and 51% of LTR represent more than two-fold downregulated TE, respectively (**Fig. 1A and** Supplementary Table 3).

ERV1 family has the highest number of upregulated or downregulated loci in LTR (**Fig. 1B**). Within ERV1 family, *LTR48B*, *LTR7Y*, and *LTR6*, were commonly upregulated in mesothelioma samples compared to mesothelial precursors (**Fig. 1C,** left panel). *LTR7Y* is an ERV specifically upregulated in blastocyst stage of human embryo [25]. It is noteworthy that some *LTR7Y* upregulated in blastocyst stage are located near genes such as *WNT16* and *FAM3C*, which are enriched in PM translatome [26] or *NCF2* which is correlated to the so-called S-score, which defines the sarcomatoid component proportion of PM by transcriptome analysis [8] (**Supplementary Fig. 1**). ERV1 family constituted 43% of the commonly downregulated TE in mesothelioma samples compared to mesothelial precursors (**Fig. 1C,** right panel).

Next we quantified locus-specific ERV expression comparing RNA-seq and reference ERV sequences using ERVmap [19], which uses stringent filtering criteria for RNA-seq reads that map to ERV loci and a 3220 ERV reference database of full-length proviral sequences [27].

We were particularly interested to full-length proviral sequences since it has recently been suggested that viral mimicry primed cancers include elevated baseline expression of retrotransposons that form double-stranded RNA (dsRNA) and elevated levels of retrotransposons-derived antigenic peptide that form tumor-associated antigens [28]. We stratified the ERVmaps by their total counts and specific enrichment in tumor. Two representative ERVs: *ERVmap_1248* (hg38, chr3:177,657,043-177,671,140) and *ERVmap_1064* (hg38, chr3:112,413,019-112,423,381), both belonging to ERVH group of the ERV1 family, were identified based on their enrichment in tumors. Another ERV sequence, *ERVmap_k48* (also called *HERV-K15*[29]), where counts almost do not differ between PM cohorts and precursors, was also selected as control (**Fig. 2A**). Coverage of *ERVmap_1248* locus was also enriched in FunGeST series, however to a lower extent **Supplementary Fig. 2A**), likely due to the polyA RNA-seq protocol used which has been shown to fail to detect several classes of repeat RNA [30]. Of note, the patients expressing levels of *ERVmap_1248* above the average were enriched for *BAP1* mutations but not *NF2* or *CDKN2*A mutations (**Supplementary Fig. 2B**). We confirmed that the expression of *ERVmap_1248* and *ERVmap_1064* but not of the control *ERVmap_k48* was lower in normal compared to tumor tissue (**Fig. 2B**). The difference was maintained when females were excluded (**Supplementary Fig. 2C**). We next compared the chromosomal location of the two ERVmaps enriched in mesothelioma to the chromosomal location of ERVs specifically expressed in mesothelioma extracted from a pan-cancer analysis [17]. Of note, 28 out of 357 (8%) of those ERVs are located near genes relevant for mesothelioma such as *MSLN*, *MET*, *UPKB1*, *UPK3B*, *LINC00578*, some interferon-stimulated genes (*RSAD2*, *IFI44*, *OAS2*), or DNA damage related genes (*CHEK2*, *MSH2*) (**Fig. 2C**). *MSLN* locus on chromosome 16 is especially enriched in ERV expression in mesothelioma (11 ERVs). The ERVmaps enriched in mesothelioma that we identified were located in ERV expression locations previously described [17] (**Fig. 2C**).

ERVmap_1248 can be induced by 5-Aza-2'-deoxycytidine treatment in mesothelial cells

Among the selected ERVs, the most abundant one, *ERVmap_1248*, and the control ERV, *ERVmap_k48*, were further analyzed. Basal expression of *ERVmap_1248* and *ERVmap_k48* was investigated in human mesothelial cells LP9/TERT-1, SDM104 and in human PM cells Mero82, ACC-Meso4, NCI-H226, and SDM103T2, used as surrogate for normal tissue versus mesothelioma, respectively. The expression of *ERVmap_1248* is on average 16-fold more enriched in PM cells compared to *ERVmap_k48*, supporting the concept of *ERVmap_1248* activation in mesothelioma (**Fig. 3A**). We had previously shown that ERVs which are expressed in mouse mesothelioma cells form dsRNA [9] and *ERVmap_1248* is also predicted to form dsRNA (**Supplementary Fig. 3**). This was confirmed (**Fig. 3B**) by RNA pull-down experiments using J2 anti-dsRNA antibody. *RNA Binding Motif 8A* (*RBM8A*)-3'UTR which we have previously shown to be a substrate for dsRNA editing [31] was used as positive control (**Fig. 3B**).

Next, we aimed at investigating whether promoter demethylation was a possible cause for increased *ERVmap_1248* expression in PM as previously observed with murine mesothelioma ERV [9]. Therefore, DNA demethylation was induced in human mesothelial cells LP9/TERT-1 and SDM104 by treatment with 5-Aza-2'-deoxycytidine (5-Aza-CdR), which is a DNA methylthransferase inhibitor (DNMTi). This treatment resulted in a significantly 8-fold increase in *ERVmap_1248* expression while the control ERV *ERVmap_k48* increased by about 50% in LP9/TERT1 but not in SDM104 cells (**Fig. 3C**). The expression of two cancer-associated testis antigens *CTAG1B* and *MAGE-C1* genes was used as positive controls as we previously described [32].

Basal levels of interferon-stimulated genes (ISG) expression is higher in PM cells with intact IFNB1 coding gene and 5-Aza-CdR increases ISGs expression in mesothelial cells and it is associated with ERVmap_1248 promoter demethylation

To assess whether differential *ERVmap_1248* expression between mesothelial and PM cells is associated with a differential type-I IFN signaling activation as observed in mouse model [9], we investigated the basal expression levels of various interferon-stimulated genes (ISG)[33]. We observed that melanoma differentiation-associated protein 5 (MDA5), retinoic acid inducible gene I (RIG-I), signal transducer and activator of transcription 1 (STAT1), and interferon-stimulated gene 15 (ISG15) levels are higher in most PM cells when compared to mesothelial cells (**Fig. 4A**). Low levels of RIG-I in H226 cells are likely due to the mutation in RIG-I encoding *DDX58* gene resulting in G191_K193 duplication at the end of the second caspase activation and recruitment domain (http://www.cbioportal.org/). In addition, in PM cell line SDM103T2, ISGs levels were at the same level as mesothelial cells (**Fig. 4A)**. We have previously shown

that basal type-I IFN signaling is present in cells with intact *IFNB1* gene in PM [34]. Accordingly, deficiency of *IFNB1* gene was observed in SDM103T2 but not the other cell lines (**Fig. 4B**). A PCR fragment covering exon 5 of *BAP1* was used as control for genomic DNA input during PCR (**Fig. 4B**). BAP1 is known to be deleted in NCI-H226 cells [35].

We next assessed whether demethylation, which results in increased *ERVmap_1248* expression in mesothelial cells, is also associated with increased ISG levels. As expected, we observed increased levels of RIG-I, STAT1, interferon regulatory factor 3 (IRF3) and ISG15 in mesothelial cells LP9/TERT-1 and SDM104 upon 5-Aza-CdR treatment. *IFNB1*-deficient PM cell line SDM103T2 was used as control. In this cell line, the increase of gene expression of *CTAG1B* and *MAGE-C1* was observed as expected (**Supplementary Fig. 4**), however, we observed no significant increase of IFN-dependent STAT1, RIG-I or IRF3. Only the expression of IFN-independent [36] ISG15 was significantly increased (**Fig. 4C**).

To confirm that the increase in *ERVmap_1248* expression observed upon treatment with 5-Aza-CdR is due to promoter demethylation, we first identified the target region based on the analysis of CpG islands (**Supplementary Fig. 5**) to design methylation-specific primer "M" and unmethylation-specific primer "U" for the region of the promoter of *ERVmap_1248*. These primers were used on sodium bisulfite treated DNA, where all methyl-free cytosines are converted into uracils, whereas methylated cytosines remain unchanged allowing the use of quantitative methylation specific PCR (qMSP).

In SDM104 cells, *ERVmap_1248* promoter methylation decreased significantly upon 5-Aza-CdR treatment, and we also observed that *ERVmap_1248* promoter is significantly more demethylated in tumor tissues compared to non-tumor tissues (**Fig. 4D**).

Expression of ISG can be decreased in mesothelioma cells by treatment with JAK inhibitor Ruxolitinib and MAVS silencing

To verify the activation of type-I IFN signaling, PM cells were treated with Ruxolitinib, a JAK1/2 inhibitor blocking the type-I IFN signaling. This treatment resulted in decreased levels of ISG in PM cell lines NCI-H226, Mero82 and ACC-Meso4 with an intact *IFNB1* gene but not in *IFNB1* gene-deficient cell line SDM103T2 (**Fig. 5A and Supplementary Fig. 6**). In order to confirm the involvement of dsRNA sensing, we silenced mitochondrial antiviral signaling protein (MAVS), which is downstream of the activation of dsRNA sensors RIG-I and MDA5. Silencing MAVS in PM cells Mero82 and ACC-Meso4 resulted in a significant decrease of RIG-I, STAT1 and ISG15 (**Fig. 5B**).

Altogether this data indicates that increased ISGs expression is dependent on type-I IFN receptor signaling.

ISGs expression is higher in ERVmap_1248 high PM tissues and patients with higher ERVmap1248 expression have better survival

PM patients were distributed into four groups (Q1-Q4) based on their *ERVmap_1248* expression (**Fig. 6A**). Q1 has lowest and Q4 has highest *ERVmap_1248* expression. The expression of ISGs, which we previously have associated with better OS, [9] was higher in the tumor tissues with high expression levels of *ERVmap_1248* (**Fig. 6B**). Furthermore, *ERVmap_1248* expression was analyzed in relation to survival. Overall survival (OS) was longer among patients with higher *ERVmap_1248* expression (Q4 or Q2-4) compared to patients whose tumors had lower *ERVmap_1248* expression (Q1) (Q4 versus Q1: median OS 25.16 versus 13.55 months, HR = 1.57, p = 0.019, Q2-4 versus Q1: median OS 24.51 versus 13.55 months, HR = 1.761, p = 0.003) (**Fig. 6C**). This observation was confirmed using the TCGA and Bueno datasets (**Supplementary Fig. 7A)** and is consistent with the TCGA clusters classification where the cluster with the best survival is related with type-I IFN signaling and *BAP1* mutation is associated with a longer overall patient survival rate [15]. No significant association with histotype was observed (**Supplementary Fig. 7B**). To increase the clinical relevance of *ERV-map1248* expression, we next assessed how its expression was associated with the four-gene (*CD8A*, *STAT1*, *LAG3* and *CD274*) inflammatory signature score, which has recently been described to predict response to checkpoint inhibitors in the CheckMate743 trial [37]. The four-gene signature score was significantly higher in Q4 compared to the Q1 group (**Supplementary Fig. 7C**).

Finally, using plasma collected in the SAKK 17/04 (NCT00334594)[22] trial we determined that circulating *ERVmap_1248* (**Fig. 6D**) can be detected in a small fraction (13%) of PM patients while only one normal control was positive out of 42 normal plasma (2%). Relative levels could not be assessed due to difference in levels of the two normalizers used (**Supplementary Fig. 8A**). Circulating *ERVmap_1248* was independent of the age of the patients (**Supplementary Fig. 8B**). The quality of RNA was assessed by quantification of relative levels of miR-625-3p where the levels of miR-16 used as normalizer were not different between healthy volunteers and PM patients (**Supplementary Fig. 8C**). The bimodal detection of *miR-625-3p* previously described [38] was observed only in PM patients (**Fig. 6D**).

We reasoned that the low sensitivity of ERV detection might be due to ERVmap being full length proviral sequences where the size is longer compared to solo-LTR. Therefore, we investigated whether *LTR7Y* expression could be used as circulating ERV instead of *ERVmap_1248*. We first determined that *LTR7Y* expression is enriched in PM tissues with high *ERVmap_1248* expression and that *LTR7Y* levels are significantly correlated with *ERVmap1248* (**Fig. 6E**). *LTR7Y* was detected in 75% of patients, however it was less specific since it was observed in 24% of healthy controls (**Fig. 6F**). As observed for *ERVmap_1248*, circulating *LTR7Y* was independent of the age and of gender (**Supplementary Fig. 8D and 8E**). The limitation of this investigation is that plasma samples did not match tumor samples.

Nevertheless, altogether our data demonstrate that mesothelioma-specific ERV levels are associated with a better clinical outcome and can be detected in blood.

Discussion

In this study we provide evidence that ERV expression is associated with clinical outcome and viral mimicry response in human mesothelioma, consistent with our observation in a mouse model of mesothelioma development in mice exposed to asbestos [9].

This is the first time that a specific ERV is associated with clinical outcome in human PM. A previous study [17] had identified the location of mesothelioma-specific expressed ERVs based on the comparison of TCGA data with the average expression of ERV in GTEX dataset. Interestingly the location of some of these ERV is near genes relevant for mesothelioma such as *UPK3B* and *MSLN*. For example, *UPK3B* is a marker for mesothelial cells [39] and high levels of *UPK3B* expression are associated with better overall survival [5, 6]. In addition, *UPK3B* and *MSLN* are significantly positively correlated with mesothelioma E-score [8]. In our study, we used three different datasets and validated the enrichment of specific ERVs, including in blood from PM patients.

Both selected solo-LTR and near full-length proviral sequences show elevated expression levels in mesothelioma. Three solo-LTR were commonly enriched in the tumor tissues and mesothelioma primary cultures. Solo-LTR are known to act as enhancers and it is noteworthy that *LTR7Y* acts as stage-specific promoter in pluripotent epiblasts, one of three major cell types in the pre-implantation blastocyst [25]. *LTR7Y* expression is controlled by methylation [40] and is frequently observed in enhancer regions in naïve human embryonic stem cells [41], while it is not present in adult tissue [25]. The observation that mesothelioma cells express pluripotent cell enhancers is consistent with our previous data where we had used a lentiviral fluorescence-based reporter construct sensing high SOX2 and OCT4 levels to identify and isolate a subpopulation of mesothelioma cells with cancer stem cell properties, characterized by chemoresistance and a higher tumor-initiating capacity in orthotopic xenograft and allograft mouse models [42]. Future studies could take advantage of the new knowledge and use the recently described LTR7Y-driven reporter [43] to further explore pluripotent mesothelioma cells.

Although not further investigated in this study, *LTR7Y* is in genomic regions enriched in retroposed genes, or genes linked to mesothelioma biology (e.g. S-score[8]), which is consistent with promoter or proximal enhancer effect of TE [12]. Not much is known about the other two commonly expressed solo-LTR, with the exception that a subset of *LTR48B* elements acquired enhancer activity in pluripotent cells [44].

Consistent with the enrichment in *LTR7Y*, the three ERVmap genes enriched in mesothelioma belong to the ERVH family, which is characterized by LTR7 promoter family and is expressed early embryo [45].

In addition to embryogenesis, subsets of *LTR7* and *LTR7Y* elements are known to be upregulated in oncogenic states due to promoter demethylation [46].

A correlation of the *LTR7* transcriptional regulatory signals with human embryonic stem cells (hESC)-specific expression of lncRNAs has been reported [47], including *linc-ROR* which is enriched in PM translatome [26]. This is consistent with the observation that high level of transcription of several ERV loci promotes the expression of lncRNA [48], which appear important in controlling cell identity [49, 50].

Silencing of ERV in adult tissues occurs through binding of HERV-targeting KRAB-Zinc Finger Proteins (KZFP), which recruit KAP1/TRIM28 co-repressor to induce heterochromatin formation [41]. Therefore, the level of variation of HERVH-associated KZFP can potentially be the reason of differential expression of HERVH in PM. For example, the potential repressor ZNF534, which is particularly enriched in *LTR7* and which is associated with pluripotency [45], is upregulated in sarcomatoid compared to epithelioid mesothelioma [5].

ERVK family were less enriched in mesothelioma. ERVKs are the only ERVs that are human specific with intact open reading frames, reported to generate viral-like proteins in teratocarcinoma cell lines and human blastocysts [51]. *ERVmap-k48* used as control has a sequence of approx. 3900 bp encoding for *Gag* and is located near the housekeeping gene *SSBP1*, which has been hypothesized to drive its transcription and possibly explains the reason for stable levels in normal *vs.* tumor tissue [29].

KAP1/TRIM28 recruits chromatin modifiers including SETDB1, which is mutated in a subset of mesothelioma [5, 7, 16], thereby possibly also contributing to differential HERVH and HERVK expression.

According to the knowledge about epigenetic control of ERV expression [52], we observed that *ERVmap_1248* expression increases upon inhibition of methylation in normal mesothelial cells. Induction of the expression of ERV has been documented in studies supporting the use of viral mimicry in clinical trials, where effects of immune checkpoint inhibitor are tested in combination with demethylating agents [52, 53]. Basal ERV expression was correlated with low methylation pattern in a pan-cancer analysis [54]. Changes in DNA methylation have been documented in human mesothelioma (reviewed in [55]), and we recently discussed [9] other factors such as control of DNA methylation that are dysregulated in mesothelioma besides KZFP.

Consistent with the association between expression of transposable elements and the viral mimicry response observed in cancer in general [28] and our own observation in a mesothelioma development model [9], we

observed a basal activation of type-I IFN signaling in tumors expressing high levels of *ERVmap_1248*. Of note, based on the mRNA expression profile, mesothelioma tumors have been clustered into four groups [5, 7]. Pathway enriched analysis of genes expressed in the clusters revealed among others, enrichment of reactome antiviral mechanism by ISG in one of the TCGA clusters, and this is confirmed in the epithelioid group of Bueno *et al* [5]. Patients with this profile have a better clinical outcome [7, 9, 15] and BAP1 mutations [15], consistent with our observations that tumors with high levels of *ERVmap_1248* are associated with BAP1 mutations.

Mesothelioma cells were previously shown to maintain the activation of the type-I IFN signaling pathway [9, 56]. In addition, mesothelioma was described as being a cancer highly enriched for the 38-ISG signature, not always justified by the presence of immune cells in the microenvironment [57]. We put forward the hypothesis that ERV expression reactivation in selected samples is a possible cause for type-I IFN activation. These tumors are most likely associated with *BAP1* but not *CDKN2A/B* mutations if we consider that *IFNB1* and all the 15 other type-I IFN genes are co-deleted in a large fraction of tumors bearing *CDKN2A/B* deletions [34].

Activation of type-I IFN was associated with response to immune checkpoint blockade in clear cell renal cancer [58]. ERV expression is a predictor of patient response to immunotherapy in a urothelial cancer cohort [30] and, interestingly, in that study it was a better predictor compared to type-I IFN signature. Investigation of the expression of 66 [59] ERVs showed that some ERVs are associated with both immune activation and checkpoint pathway upregulation in clear cell renal cell carcinoma [60] and expression levels of one of those ERVs predicted response to checkpoint blockade. In addition, high ERV expression was associated with better overall clinical outcome in a cohort of melanoma patients while repression of ERV was observed in the cohort with worst outcome [61]. Furthermore, *ERVmap_2637* expression was higher in melanoma patients with complete response to anti-PD1 treatment [62] and negatively correlates with *KDM5B* expression, which recruits SETDB1. Therefore, our observations are also important for mesothelioma therapy. Indeed, therapeutic approaches exploiting type-I IFN pathway signaling have already been implemented in the clinic [63] or proposed on the basis of preclinical studies [64, 65]. Future studies may investigate whether ERV expression could be a predictor of sensitivity to those therapeutic approaches and immune checkpoint inhibition, although it should be taken into account for therapies inducing type-I IFN signaling that some mesothelioma have lost type-I IFN genes [34] and might therefore not be able to activate such signaling. ERV expression could be e.g. be helpful to stratify patients with epithelioid histotype, which overall respond less well to immune checkpoint inhibition compared to sarcomatoid histotype [66].

CRediT

Conceptualization: EFB, HR, JFF

Methodology: HR, WQ, DJ, JFF, EFB

Software: WQ, HR

Validation: SS, WQ, HR

Formal analysis: WQ, SS, EFB, CM

Investigation: SS, MR, AH, MW

Resources: RS, MF, FC

Data Curation: WQ, SS

Writing - Original Draft: SS, EFB

Writing - Review & Editing: all

Visualization: SS

Supervision: EFB

Funding acquisition: EFB

Funding

This work was supported by Swiss National Science Foundation, grant 320030_182690, Walter-Bruckerhoff Stiftung and Stiftung für Angewandte Krebsforchung; SS is supported by China Scholarship Council (CSC). DJ is supported by Inserm, the Ligue Contre le Cancer (Ile de France committee), the Chancellerie des Universités de Paris (Legs POIX), and Cancéropôle Île-de-France (Emergence). JFF is supported by La Ligue Régionale Grand Ouest contre le Cancer, ARSMESO44, Fondation ARC, Institut National du Cancer (INCA-PLBIO21-050)" and LabEX IGO program.

Acknowledgments

We are grateful to SAKK 17/04 patients for donating their blood for translational investigation.

References

1. Carbone M, Adusumilli PS, Alexander HR, Jr., et al. Mesothelioma: Scientific clues for prevention, diagnosis, and therapy. *CA Cancer J Clin* 2019;69:402-429.
2. Meiller C, Montagne F, Hirsch TZ, et al. Multi-site tumor sampling highlights molecular intra-tumor heterogeneity in malignant pleural mesothelioma. *Genome Med* 2021;13:113.

3. Zhang M, Luo JL, Sun Q, et al. Clonal architecture in mesothelioma is prognostic and shapes the tumour microenvironment. *Nat Commun* 2021;12:1751.
4. Sauter JL, Dacic S, Galateau-Salle F, et al. The 2021 WHO Classification of Tumors of the Pleura: Advances Since the 2015 Classification. *J Thorac Oncol* 2022;17:608-622.
5. Bueno R, Stawiski EW, Goldstein LD, et al. Comprehensive genomic analysis of malignant pleural mesothelioma identifies recurrent mutations, gene fusions and splicing alterations. *Nat Genet* 2016;48:407-416.
6. de Reynies A, Jaurand MC, Renier A, et al. Molecular classification of malignant pleural mesothelioma: identification of a poor prognosis subgroup linked to the epithelial-to-mesenchymal transition. *Clin Cancer Res* 2014;20:1323-1334.
7. Hmeljak J, Sanchez-Vega F, Hoadley KA, et al. Integrative Molecular Characterization of Malignant Pleural Mesothelioma. *Cancer discovery* 2018;8:1548-1565.
8. Blum Y, Meiller C, Quetel L, et al. Dissecting heterogeneity in malignant pleural mesothelioma through histo-molecular gradients for clinical applications. *Nat Commun* 2019;10:1333.
9. Sun S, Frontini F, Qi W, et al. Endogenous retrovirus expression activates type-I interferon signaling in an experimental mouse model of mesothelioma development. *Cancer Lett* 2021;507:26-38.
10. Hoyt SJ, Storer JM, Hartley GA, et al. From telomere to telomere: The transcriptional and epigenetic state of human repeat elements. *Science* 2022;376:eabk3112.
11. Wells JN, Feschotte C. A Field Guide to Eukaryotic Transposable Elements. *Annu Rev Genet* 2020;54:539-561.
12. Friedli M, Trono D. The developmental control of transposable elements and the evolution of higher species. *Annu Rev Cell Dev Biol* 2015;31:429-451.
13. Jansz N, Faulkner GJ. Endogenous retroviruses in the origins and treatment of cancer. *Genome biology* 2021;22:147.
14. Fueyo R, Judd J, Feschotte C, et al. Roles of transposable elements in the regulation of mammalian transcription. *Nat Rev Mol Cell Biol* 2022.
15. Osmanbeyoglu HU, Palmer GE, Sagan C, et al. Isolated BAP1 loss in malignant pleural mesothelioma predicts immunogenicity with

implications for immunotherapeutic response. *BioRxiv* 2022.
16. Creaney J, Patch AM, Addala V, et al. Comprehensive genomic and tumour immune profiling reveals potential therapeutic targets in malignant pleural mesothelioma. *Genome Med* 2022;14:58.
17. Attig J, Young GR, Hosie L, et al. LTR retroelement expansion of the human cancer transcriptome and immunopeptidome revealed by de novo transcript assembly. *Genome research* 2019;29:1578-1590.
18. Jin Y, Tam OH, Paniagua E, et al. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* 2015;31:3593-3599.
19. Tokuyama M, Kong Y, Song E, et al. ERVmap analysis reveals genome-wide transcription of human endogenous retroviruses. *Proc Natl Acad Sci U S A* 2018;115:12565-12572.
20. Oehl K, Kresoja-Rakic J, Opitz I, et al. Live-Cell Mesothelioma Biobank to Explore Mechanisms of Tumor Progression. *Frontiers in oncology* 2018;8:40.
21. Sidi R, Pasello G, Opitz I, et al. Induction of senescence markers after neo-adjuvant chemotherapy of malignant pleural mesothelioma and association with clinical outcome: an exploratory analysis. *Eur J Cancer* 2011;47:326-332.
22. Stahel RA, Riesterer O, Xyrafas A, et al. Neoadjuvant chemotherapy and extrapleural pneumonectomy of malignant pleural mesothelioma with or without hemithoracic radiotherapy (SAKK 17/04): a randomised, international, multicentre phase 2 trial. *Lancet Oncol* 2015;16:1651-1658.
23. Quetel L, Meiller C, Assié JB, et al. Genetic alterations of malignant pleural mesothelioma: association with tumor heterogeneity and overall survival. *Mol Oncol* 2020;14:1207-1223.
24. Colunga T, Hayworth M, Kress S, et al. Human Pluripotent Stem Cell-Derived Multipotent Vascular Progenitors of the Mesothelium Lineage Have Utility in Tissue Engineering and Repair. *Cell Rep* 2019;26:2566-2579 e2510.
25. Goke J, Lu X, Chan YS, et al. Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells. *Cell Stem Cell* 2015;16:135-141.
26. Grosso S, Marini A, Gyuraszova K, et al. The pathogenesis of mesothelioma is driven by a dysregulated translatome. *Nat Commun* 2021;12:4920.

27. Vargiu L, Rodriguez-Tome P, Sperber GO, et al. Classification and characterization of human endogenous retroviruses; mosaic forms are common. *Retrovirology* 2016;13:7.
28. Chen R, Ishak CA, De Carvalho DD. Endogenous Retroelements and the Viral Mimicry Response in Cancer Therapy and Cellular Homeostasis. *Cancer discovery* 2021;11:2707-2725.
29. Schmitt K, Reichrath J, Roesch A, et al. Transcriptional profiling of human endogenous retrovirus group HERV-K(HML-2) loci in melanoma. *Genome Biol Evol* 2013;5:307-328.
30. Solovyov A, Vabret N, Arora KS, et al. Global Cancer Transcriptome Quantifies Repeat Element Polarization between Immunotherapy Responsive and T Cell Suppressive Classes. *Cell Rep* 2018;23:512-521.
31. Abukar A, Wipplinger M, Hariharan A, et al. Double-Stranded RNA Structural Elements Holding the Key to Translational Regulation in Cancer: The Case of Editing in RNA-Binding Motif Protein 8A. *Cells* 2021;10.
32. Kresoja-Rakic J, Kapaklikaya E, Ziltener G, et al. Identification of cis- and trans-acting elements regulating calretinin expression in mesothelioma cells. *Oncotarget* 2016;7:21272-21286.
33. Schneider WM, Chevillotte MD, Rice CM. Interferon-stimulated genes: a complex web of host defenses. *Annu Rev Immunol* 2014;32:513-545.
34. Delaunay T, Achard C, Boisgerault N, et al. Frequent Homozygous Deletions of Type I Interferon Genes in Pleural Mesothelioma Confer Sensitivity to Oncolytic Measles Virus. *J Thorac Oncol* 2020;15:827-842.
35. Jensen DE, Proctor M, Marquis ST, et al. BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression. *Oncogene* 1998;16:1097-1112.
36. Grandvaux N, Servant MJ, tenOever B, et al. Transcriptional profiling of interferon regulatory factor 3 target genes: direct involvement in the regulation of interferon-stimulated genes. *J Virol* 2002;76:5532-5539.
37. Peters S, Scherpereel A, Cornelissen R, et al. First-line nivolumab plus ipilimumab versus chemotherapy in patients with unresectable malignant pleural mesothelioma: 3-year outcomes from CheckMate 743. *Ann Oncol* 2022;33:488-499.
38. Kresoja-Rakic J, Szpechcinski A, Kirschner MB, et al. miR-625-3p and lncRNA GAS5 in Liquid Biopsies for Predicting the Outcome of Malignant Pleural Mesothelioma Patients Treated with Neo-Adjuvant Chemotherapy and Surgery. *Noncoding RNA* 2019;5.
39. Kanamori-Katayama M, Kaiho A, Ishizu Y, et al. LRRN4 and UPK3B are markers of primary mesothelial cells. *PLoS One* 2011;6:e25391.
40. Theunissen TW, Friedli M, He Y, et al. Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem Cell* 2016;19:502-515.
41. Pontis J, Planet E, Offner S, et al. Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic Genome Activation and Control Transcription in Naive Human ESCs. *Cell Stem Cell* 2019;24:724-735 e725.
42. Blum W, Pecze L, Felley-Bosco E, et al. Stem Cell Factor-Based Identification and Functional Properties of In Vitro-Selected Subpopulations of Malignant Mesothelioma Cells. *Stem Cell Reports* 2017;8:1005-1017.
43. Szczerbinska I, Gonzales KAU, Cukuroglu E, et al. A Chemically Defined Feeder-free System for the Establishment and Maintenance of the Human Naive Pluripotent State. *Stem Cell Reports* 2019;13:612-626.
44. Casanova M, Moscatelli M, Chauviere LE, et al. A primate-specific retroviral enhancer wires the XACT lncRNA into the core pluripotency network in humans. *Nat Commun* 2019;10:5652.
45. Carter TA, Singh M, Dumbovic G, et al. Mosaic cis-regulatory evolution drives transcriptional partitioning of HERVH endogenous retrovirus in the human embryo. *Elife* 2022;11.
46. Kong Y, Rose CM, Cass AA, et al. Transposable element expression in tumors is associated with immune infiltration and increased antigenicity. *Nat Commun* 2019;10:5228.
47. Kelley D, Rinn J. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome biology* 2012;13:R107.
48. Kapusta A, Kronenberg Z, Lynch VJ, et al. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet* 2013;9:e1003470.
49. Lu X, Sachs F, Ramsay L, et al. The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nat Struct Mol Biol* 2014;21:423-425.

50. Wang J, Xie G, Singh M, et al. Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature* 2014;516:405-409.
51. Grow EJ, Flynn RA, Chavez SL, et al. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature* 2015;522:221-225.
52. Chiappinelli KB, Strissel PL, Desrichard A, et al. Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. *Cell* 2015;162:974-986.
53. Roulois D, Loo Yau H, Singhania R, et al. DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. *Cell* 2015;162:961-973.
54. Smith CC, Beckermann KE, Bortone DS, et al. Endogenous retroviral signatures predict immunotherapy response in clear cell renal cell carcinoma. *J Clin Invest* 2018;128:4804-4820.
55. Vandenhoeck J, van Meerbeeck JP, Fransen E, et al. DNA Methylation as a Diagnostic Biomarker for Malignant Mesothelioma: A Systematic Review and Meta-Analysis. *J Thorac Oncol* 2021;16:1461-1478.
56. Chernova T, Sun XM, Powley IR, et al. Molecular profiling reveals primary mesothelioma cell lines recapitulate human disease. *Cell Death Differ* 2016;23:1152-1164.
57. Liu H, Golji J, Brodeur LK, et al. Tumor-derived IFN triggers chronic pathway agonism and sensitivity to ADAR loss. *Nat Med* 2019;25:95-102.
58. de Cubas AA, Dunker W, Zaninovich A, et al. DNA hypomethylation promotes transposable element expression and activation of immune signaling in renal cell cancer. *JCI Insight* 2020;5.
59. Mayer J, Blomberg J, Seal RL. A revised nomenclature for transcribed human endogenous retroviral loci. *Mob DNA* 2011;2:7.
60. Panda A, de Cubas AA, Stein M, et al. Endogenous retrovirus expression is associated with response to immune checkpoint blockade in clear cell renal cell carcinoma. *JCI Insight* 2018;3.
61. Badal B, Solovyov A, Di Cecilia S, et al. Transcriptional dissection of melanoma identifies a high-risk subtype underlying TP53 family genes and epigenome deregulation. *JCI Insight* 2017;2.
62. Zhang SM, Cai WL, Liu X, et al. KDM5B promotes immune evasion by recruiting SETDB1 to silence retroelements. *Nature* 2021;598:682-687.
63. Sterman DH, Recio A, Haas AR, et al. A phase I trial of repeated intrapleural adenoviral-mediated interferon-beta gene transfer for mesothelioma and metastatic pleural effusions. *Mol Ther* 2010;18:852-860.
64. Vanbervliet-Defrance B, Delaunay T, Daunizeau T, et al. Cisplatin unleashes Toll-like receptor 3-mediated apoptosis through the downregulation of c-FLIP in malignant mesothelioma. *Cancer Lett* 2020;472:29-39.
65. Achard C, Boisgerault N, Delaunay T, et al. Sensitivity of human pleural mesothelioma to oncolytic measles virus depends on defects of the type I interferon response. *Oncotarget* 2015;6:44892-44904.
66. Baas P, Scherpereel A, Nowak AK, et al. First-line nivolumab plus ipilimumab in unresectable malignant pleural mesothelioma (CheckMate 743): a multicentre, randomised, open-label, phase 3 trial. *Lancet* 2021;397:375-386.

Legends to the Figures:

Figure 1. LTRs represent the most abundant TE upregulated or downregulated in PM patients' datasets. (A) Pie charts showing up vs downregulated TEs loci. (B) TE loci up or downregulated in TCGA, Bueno or FunGeST vs mesothelial precursor. Results were filtered by padj < 0.01 and abs (log2Fold changes) > 1. Results are shown for loci within the stated LTR families. ns = expression not significantly changed. (C) Overlap of the significantly upregulated genes in the three comparisons visualized as Venn diagram. Significance was defined by padj < 0.01.

Figure 2. PM-specific ERV. (A) Expression of two representative ERV sequences *ERVmap_1248* and *ERVmap_1064* increases in tumors but not the control *ERVmap_k-48*. (*) P<0.05, (**) P<0.01, (***) P<0.001, (****) P<0.0001 ns = not significant, Mann-Whitney test. (B) RT-qPCR validation of ERV expression in samples from human non tumor tissue and tumor tissue. (*) P<0.05, (**) P<0.01, (***) P<0.001, ns = not significant, Mann-Whitney test. (C) Location of ERVmaps and other ERVs with enriched expression in PM. The ideogram was created using http://visualization.ritchielab.org/phenograms/plot.

Figure 3. ERVmap_1248 expression is lower in normal mesothelial cells compared to PM cells, it can form dsRNA and its expression can be induced by 5-Aza-CdR in mesothelial cells. (A) Basal expression of *ERVmap_1248* and *ERVmap_k48* in human PM cells and human mesothelial cells. (B) RT-qPCR analysis of *ERVmap_1248* and *RNA Binding Motif 8A* (*RBM8A*)-3'UTR transcripts captured by J2 antibody in pulldown assay, (*) P<0.05, (**) P<0.01, two-tailed paired t-tests. RBM8A-3'UTR is used as positive control (C) ERVmap_1248 expression increased significantly with 5-Aza-CdR treatment in LP9/TERT-1 and SDM104 mesothelial cells, while ERVmap_k48 expression increased by about 50% in LP9/TERT1 but not in SDM104 cells. (*) P<0.05, (****) P<0.0001 ns = not significant, two-tailed paired t test. MAGE-C1 and CTAG1B are used as positive controls (36).

Figure 4. Basal ISGs expression is higher in PM cells with intact IFN coding genes and and 5-Aza-CdR increased ISGs in mesothelial cells and it is associated with promoter demethylation. (A) The expression of ISGs MDA5, RIG-I, STAT1 and ISG15 increased in PM cells except SDM103T2 cells on protein level. (B) PCR mediated detection of IFNB1 and BAP1 genes using genomic DNA from mesothelial and PM cells. The PCR fragment of BAP1 was used as DNA input control. (C) The expression of RIG-I, STAT1, IRF3 and ISG15 all increased upon 5-Aza-CdR treatment in LP9/TERT-1 and SDM104 mesothelial cells but not all in SDM103T2 PM cells. WB quantification: n=4-12. (*) P<0.05, (**) P<0.01, (***) P<0.001, two-tailed paired t test. (D) The fraction of ERVmap_1248 promoter methylation decreased after 5-Aza-CdR treatment of SDM104 cells (left panel).The human PM tissue have less methylation

percentage in ERVmap_1248 promoter compared to normal tissue. (*) P<0.05, (**) P<0.01, Mann Whitney test.

Figure 5. Basal levels of ISG can be decreased by treatment with JAK inhibitor Ruxolitinib and MAVS silencing. (A) The expression of ISGs RIG-I, STAT1 and ISG15 decreased after Ruxolitinib (Ruxo), a JAK1/2 inhibitor treatment in PM cells except SDM103T2 cells. WB quantification: n=3-4. (*) P<0.05, (**) P<0.01, (****) P<0.0001, two-tailed paired t-test. (B) Silencing MAVS decreased ISGs RIG-I, STAT1 and ISG15 levels in Mero82 and ACC-Meso4 PM cells. WB quantification: n=3. (*) P<0.05, (**) P<0.01, (***) P<0.001, (****) P<0.0001, two-tailed paired t-test.

Figure 6. *ERVmap_1248* high expression is associated with clinical outcome in PM (A) *ERVmap_1248* expression distribution in tumor tissues from PM patients. Q=quarter. (B) The expression of ISGs *DDX58*, *IFITM1* and *IFIT2* are higher in *ERVmap_1248* high tumor tissues. (*) P<0.05, (**) P<0.01, (****) P<0.0001, two-tailed unpaired t test. (C) Kaplan-Meier curves of overall survival according to *ERVmap_1248* expression in PM patients. Red and black curves represent lower (Q1) or higher *ERVmap_1248* (Q4 or Q2-4) expression respectively. Overall survival was calculated from date of diagnosis. PM patients from Q2 or Q2-4 have better survival rate than Q1. Gehan-Breslow-Wilcoxon test.(D) *ERVmap_1248* detection in the plasma of PM patients and healthy volunteers. Detection of *miR-625-3p* is used as comparison. (E) *LTR7Y* and *ERVmap_1248* expression are correlated. (*) P<0.05, (**) P<0.01, two-tailed unpaired t test. Correlation: R=0.6798, P>0.0001, Pearson correlation analysis. (F) *LTR7Y* detection in the plasma of PM patients and healthy volunteers. (****) P<0.0001, Chi-square test.

Supplementary Data:

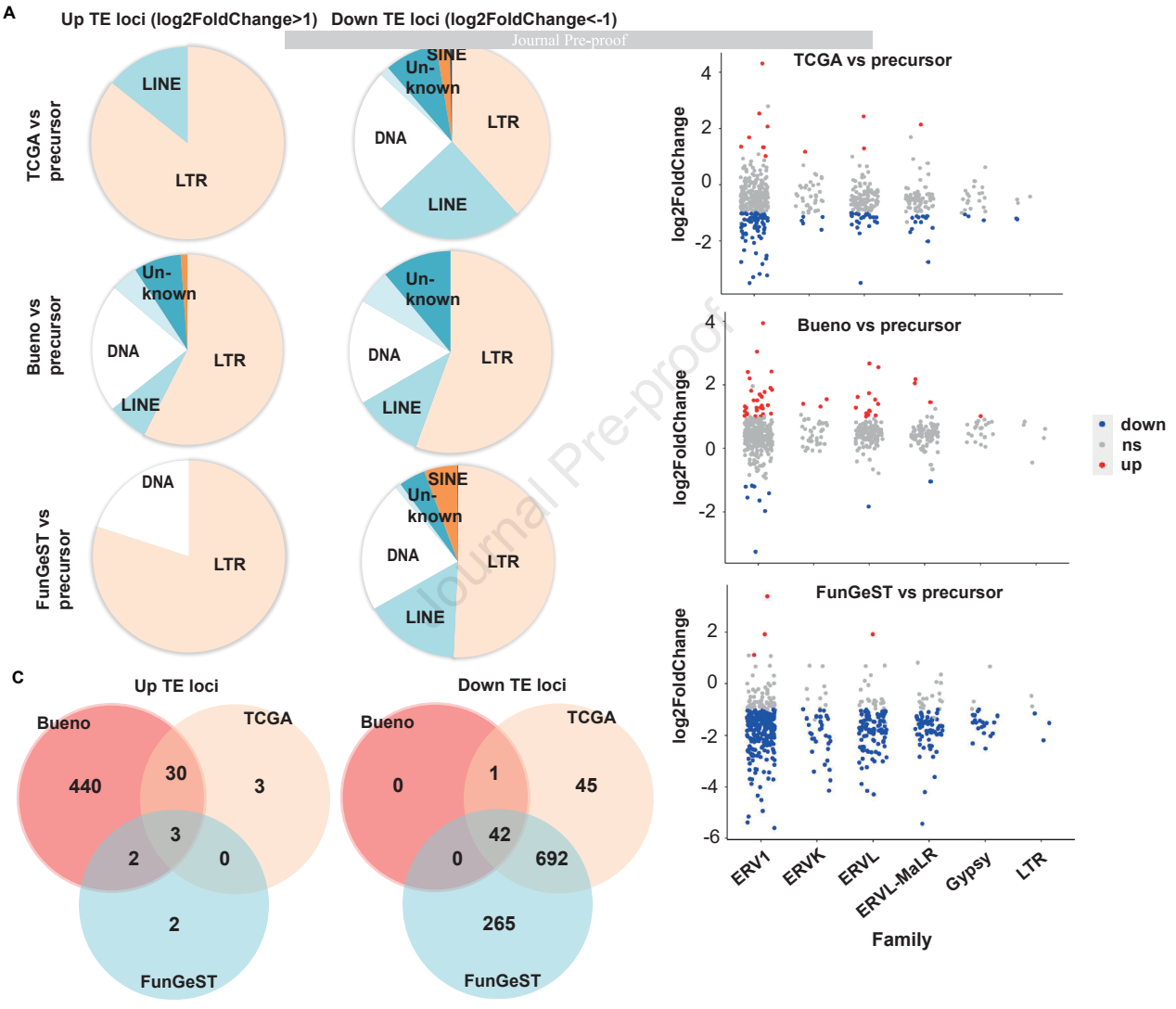
Supplementary Methods.docx

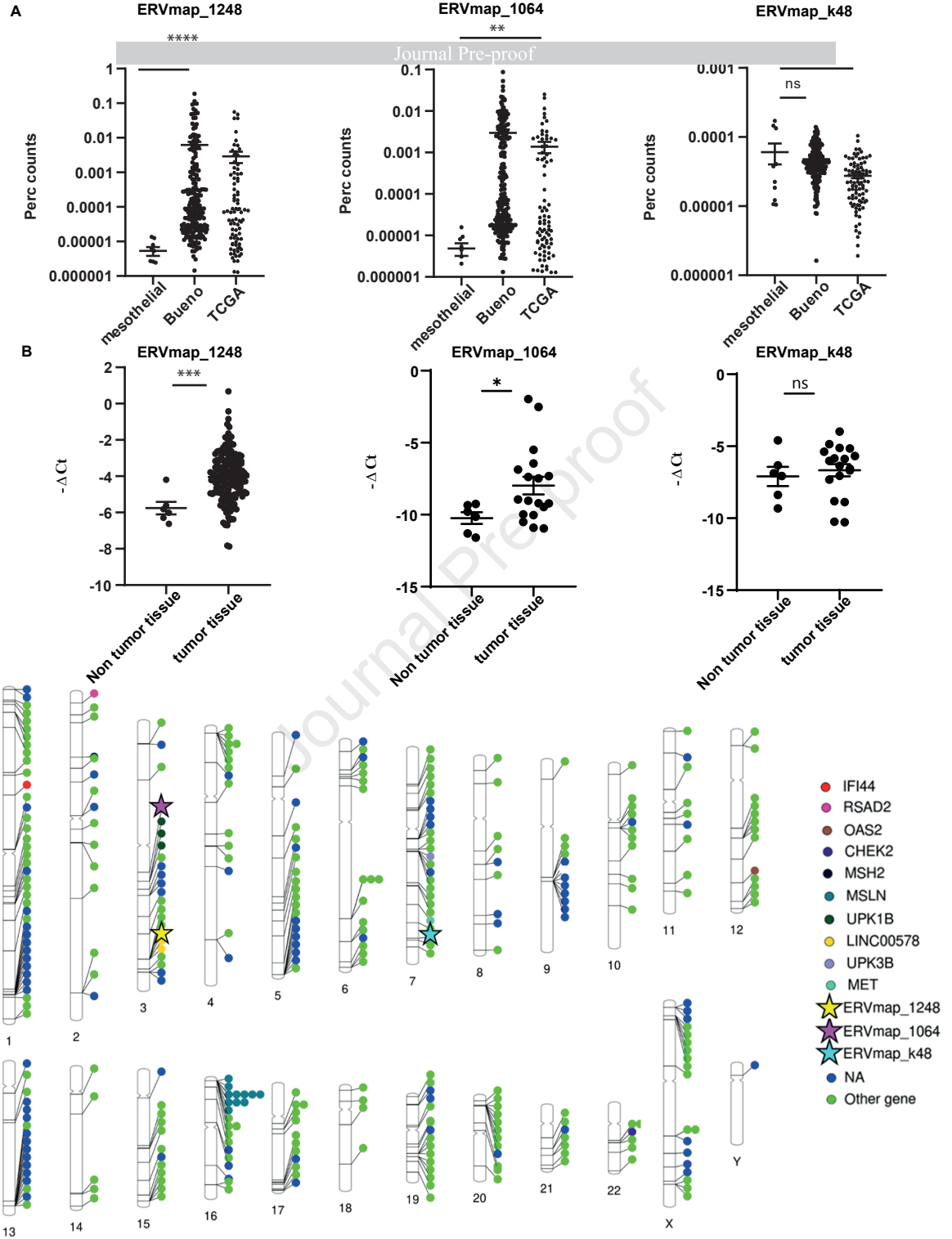
Supplementary Table 1.xlsx

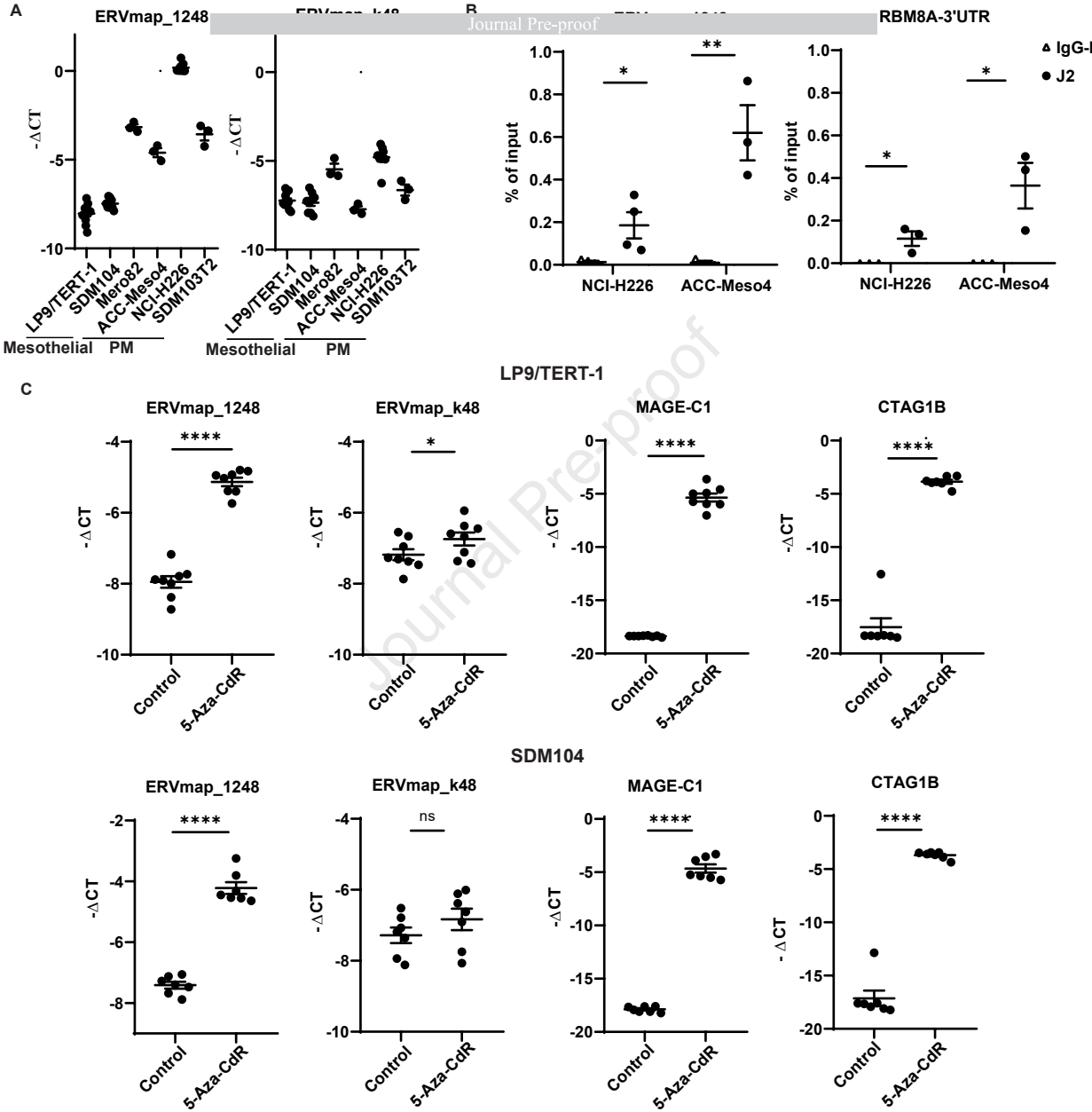
Supplementary Table 2.xlsx

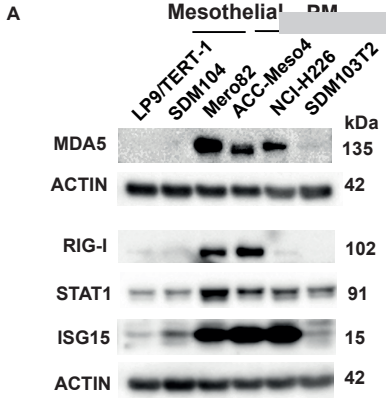
Supplementary Table 3.xlsx

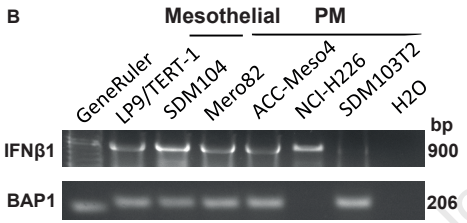
Supplementary Fig. 1-8.pdf

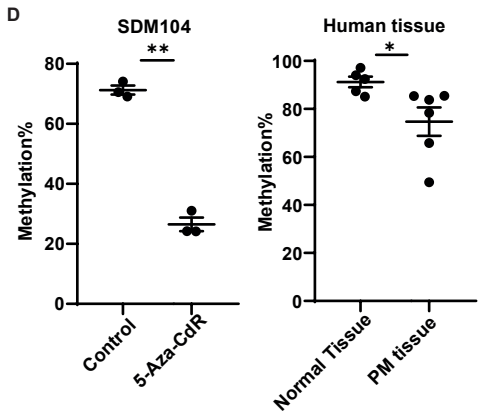


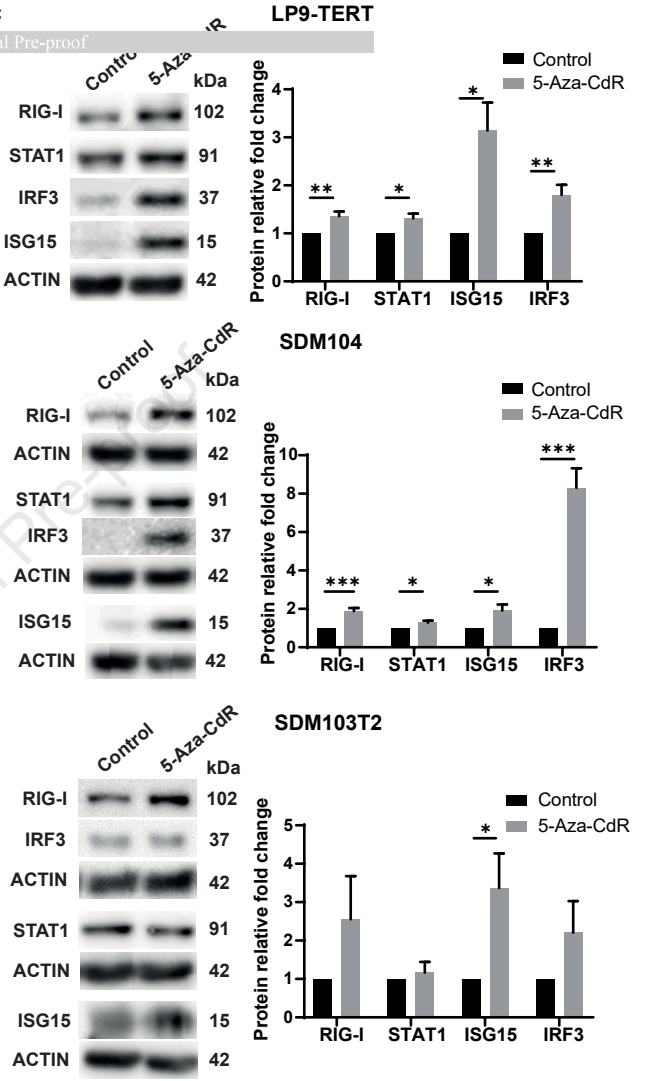


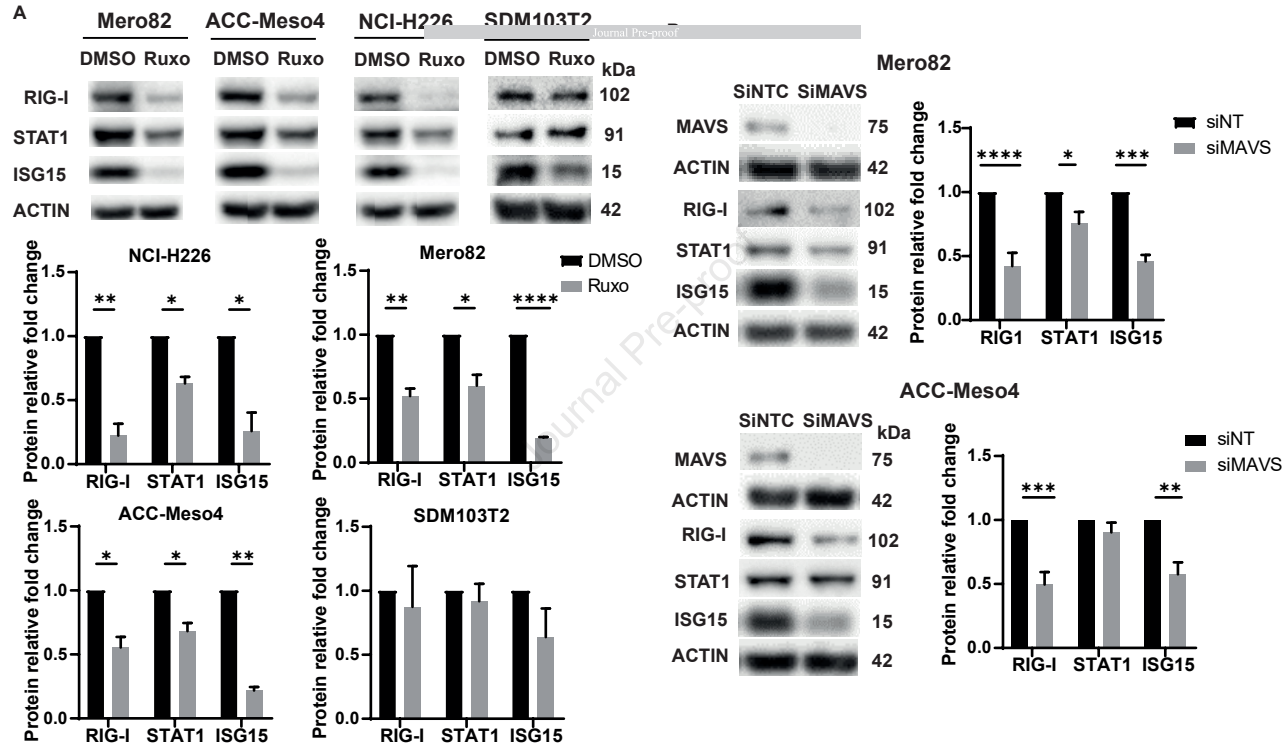











A
Mero82
ACC-Meso4
NCI-H226
SDM103T2
DMSO Ruxo
DMSO Ruxo
DMSO Ruxo
DMSO Ruxo
kDa
RIG-I
STAT1
ISG15
ACTIN
102
91
15
42
NCI-H226
Mero82
DMSO
Ruxo
Protein relative fold change
RIG-I STAT1 ISG15
ACC-Meso4
SDM103T2
Mero82
SiNTC SiMAVS
MAVS
ACTIN
RIG-I
STAT1
ISG15
75
42
102
91
15
42
RIG1 STAT1 ISG15
siNT
siMAVS
ACC-Meso4
SiNTC SiMAVS
kDa
RIG-I STAT1 ISG15
siNT
siMAVS