<table>
<thead>
<tr>
<th>Citation</th>
<th>Thummalapalli, Rohit. 2018. EZH2 Inhibition Induces Endogenous Retroviral Elements and Primes Immunogenicity in Mesenchymal Small Cell Lung Cancer. Doctoral dissertation, Harvard Medical School.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessed</td>
<td>June 29, 2018 4:15:54 AM EDT</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:37006471">http://nrs.harvard.edu/urn-3:HUL.InstRepos:37006471</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>

(Article begins on next page)
# Table of Contents

**Abstract** ................................................................................................................................................. 3

**Glossary** ................................................................................................................................................... 4

**Introduction** ........................................................................................................................................... 6
  Tumor heterogeneity in cancer ......................................................................................................................... 6
  Mesenchymal tumor subclones as mediators of treatment resistance ....................................................... 7
  Mesenchymal subclones in small cell lung cancer ......................................................................................... 7
  Innate immune signaling in mesenchymal tumor subclones ...................................................................... 8
  Endogenous retroviruses as regulators of tumor immune signaling ............................................................ 9
  The SPARCS endogenous retroviral signature marks mesenchymal subclones in SCLC ............................. 11
  Epigenetic targeted therapy in the context of cancer immunotherapy .................................................... 14

**Methods** .................................................................................................................................................. 16

**Results** .................................................................................................................................................... 23
  SPARCS ERVs associate with DNA and histone methylation pathways across tumors ............................ 23
  SPARCS ERV loci display more open chromatin profiles marked by loss of histone methylation in mesenchymal tumor subclones ................................................................................ 24
  EZH2 inhibition induces SPARCS expression associated with the mesenchymal phenotype ................. 26
  Use of a synthetic oligonucleotide to characterize dsRNA sensing pathways in H69 cells ...................... 28
  EZH2 inhibition induces SPARCS-associated cytokine responses and markers of immunogenicity ........ 30
  SPARCS-high tumors reveal associations with key immune sensing pathways, cytokines, and epigenetic regulators .................................................................................................................. 33

**Discussion** ............................................................................................................................................... 36
  Regulation of a novel ERV subclass .............................................................................................................. 36
  Refinement of our understanding of how EZH2 blockade enhances immunogenicity .............................. 37
  Exploring how EZH2 inhibition may synergize with immune checkpoint blockade ............................ 38
  Exploring SPARCS ERV physiology more broadly across tumors .......................................................... 40

**References** ............................................................................................................................................... 42

**Figures and Tables** ................................................................................................................................. 51
  Figures ....................................................................................................................................................... 51
  Tables ........................................................................................................................................................ 60
  Figure Legends .......................................................................................................................................... 61
Abstract

Small cell lung cancer (SCLC) is characterized by heterogeneous mesenchymal and neuroendocrine cell states. In particular, aggressive mesenchymal subclones in SCLC harbor characteristic cytokine profiles and activation of innate immune signaling pathways. Recent work has shown that exposure of mesenchymal SCLC tumor cells to IFN-γ leads to production of double-stranded RNA (dsRNA) through transcription of a novel subclass of endogenous retroviruses (ERVs) located in the 3’ UTR of IFN-γ-inducible genes, with this subclass being termed SPARCS (Stimulated 3 Prime Antisense Retroviral Coding Sequences). However, the mechanism of SPARCS de-repression or the consequences of SPARCS dsRNA production for immune signaling and response to immunotherapy in mesenchymal subclones remained unclear. Chromatin profiling using ATAC-seq and ChIP-seq demonstrated increased accessibility of SPARCS loci marked by loss of H3K27 trimethylation in SCLC mesenchymal subclones, suggesting core regulation by the histone methyltransferase EZH2. Neuroendocrine SCLC cells treated with EZH2 inhibition displayed activated SPARCS expression, phenotypic changes suggestive of a mesenchymal state transition, and secretion of mesenchymal-associated cytokines with important roles in immune cell recruitment. EZH2 inhibition also induced endogenous production of IFN-γ and upregulation of antigen presentation machinery, suggesting the development of increased immunogenicity in this state. In addition, SPARCS-high samples across tumor types displayed loss of EZH2, highlighting the potential role of EZH2 in ERV regulation across tumor types. Taken together, these suggest that mesenchymal-associated alterations in histone demethylation lead to induction of ERVs and priming of immunogenicity in SCLC tumor subclones, identifying unique points of intervention which may have important consequences for immunotherapy in small cell lung cancer and other tumors.
Glossary

ARID1A  AT-rich interactive domain-containing protein 1A
ARID1B  AT-rich interactive domain-containing protein 1B
ARID2  AT-rich interactive domain-containing protein 2
ATAC-seq  assay for transposase accessible chromatin followed by sequencing
B2M  beta-2-microglobulin
CCL2  C-C motif chemokine ligand 2
CCL5  C-C motif chemokine ligand 5
CCLE  Cancer Cell Line Encyclopedia
CDK4  Cyclin-dependent kinase 4
CDK6  Cyclin-dependent kinase 6
cGAS  cyclic GMP-AMP synthase
ChIP-seq  chromatin immunoprecipitation followed by sequencing
CTLA-4  Cytotoxic T-Lymphocyte Associated Protein 4
CX3CL1  C-X3-C motif chemokine ligand 1
CX3CR1  C-X3-C chemokine receptor 1
CXCL9  C-X-C motif chemokine 9
CXCL10  C-X-C motif chemokine 10
CXCL11  C-X-C motif chemokine 11
CXCR3  C-X-C chemokine receptor 3
dAC  decitabine
DMSO  dimethyl sulfoxide
DNMT  DNA methyltransferase
dsDNA  double stranded DNA
dsRNA  double stranded RNA
EMT  epithelial-mesenchymal transition
ENCODE  Encyclopedia of DNA Elements
ERV  endogenous retrovirus
EZH2  Enhancer of zeste homolog 2
GSEA  gene set enrichment analysis
H3K27Ac  Histone 3 Lysine 27 acetylation
H3K27Me3  Histone 3 Lysine 27 trimethylation
H3K36  Histone 3 Lysine 36
H69M-  H69M-PD-L1<sub>low</sub>
H69M+  H69M-PD-L1<sub>high</sub>
HDAC  histone deacetylase
HGF  hepatocyte growth factor
HLA  human leukocyte antigen
HMW  high molecular weight
IFN-α  interferon alpha
IFN-β  interferon beta
IFN-γ  interferon gamma
IKKe  IκB kinase epsilon
IL-2  interleukin-2
IL-6  interleukin-6
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-8</td>
<td>interleukin-8</td>
</tr>
<tr>
<td>IRF3</td>
<td>Interferon Regulatory Factor 3</td>
</tr>
<tr>
<td>JAK</td>
<td>Janus tyrosine kinase</td>
</tr>
<tr>
<td>JMJD3</td>
<td>Jumonji domain containing 3</td>
</tr>
<tr>
<td>LINE</td>
<td>long interspersed nuclear element</td>
</tr>
<tr>
<td>LMW</td>
<td>low molecular weight</td>
</tr>
<tr>
<td>LTR</td>
<td>long terminal repeat</td>
</tr>
<tr>
<td>MAVS</td>
<td>mitochondrial antiviral signaling protein</td>
</tr>
<tr>
<td>MDA5</td>
<td>Melanoma Differentiation-Associated Protein 5</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility antigen</td>
</tr>
<tr>
<td>MSigDB</td>
<td>Molecular Signatures Database</td>
</tr>
<tr>
<td>NF-κB</td>
<td>nuclear factor kappa-light-chain-enhancer of activated B cells</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>PBRM1</td>
<td>polybromo 1</td>
</tr>
<tr>
<td>PD-1</td>
<td>Programmed cell death protein 1</td>
</tr>
<tr>
<td>PD-L1</td>
<td>Programmed death-ligand 1</td>
</tr>
<tr>
<td>PD-L2</td>
<td>Programmed death-ligand 2</td>
</tr>
<tr>
<td>Poly(I:C)</td>
<td>polyinosinic-polycytidylic acid</td>
</tr>
<tr>
<td>PRC2</td>
<td>Polycomb repressive complex 2</td>
</tr>
<tr>
<td>Rb</td>
<td>Retinoblastoma protein</td>
</tr>
<tr>
<td>RCC</td>
<td>renal cell carcinoma</td>
</tr>
<tr>
<td>RefSeq</td>
<td>NCBI Reference Sequence Database</td>
</tr>
<tr>
<td>RIG-I</td>
<td>Retinoic Acid Inducible Gene 1</td>
</tr>
<tr>
<td>RNA-seq</td>
<td>RNA sequencing</td>
</tr>
<tr>
<td>RPKM</td>
<td>Reads per Kilobase per Million mapped reads</td>
</tr>
<tr>
<td>RT</td>
<td>reverse transcriptase</td>
</tr>
<tr>
<td>SCLC</td>
<td>small cell lung cancer</td>
</tr>
<tr>
<td>SETD2</td>
<td>SET domain containing 2</td>
</tr>
<tr>
<td>SINE</td>
<td>short interspersed nuclear element</td>
</tr>
<tr>
<td>SMARCA4</td>
<td>SWI/SNF related, matrix associated, actin-dependent regulator of chromatin subfamily a, member 4</td>
</tr>
<tr>
<td>SPARCS</td>
<td>Stimulated 3 Prime Antisense Retroviral Coding Sequences</td>
</tr>
<tr>
<td>ssRNA</td>
<td>single stranded RNA</td>
</tr>
<tr>
<td>STAT1</td>
<td>signal transducer and activator of transcription 1</td>
</tr>
<tr>
<td>STING</td>
<td>Stimulator of Interferon Genes</td>
</tr>
<tr>
<td>STR</td>
<td>short tandem repeat</td>
</tr>
<tr>
<td>SWI/SNF</td>
<td>SWItch/Sucrose Non-Fermentable</td>
</tr>
<tr>
<td>TAP1</td>
<td>Transporter associated with Antigen Processing 1</td>
</tr>
<tr>
<td>TAP2</td>
<td>Transporter associated with Antigen Processing 2</td>
</tr>
<tr>
<td>TBK1</td>
<td>TANK-binding kinase 1</td>
</tr>
<tr>
<td>TCGA</td>
<td>The Cancer Genome Atlas</td>
</tr>
<tr>
<td>T_h1</td>
<td>Type 1 T helper</td>
</tr>
<tr>
<td>TLR3</td>
<td>Toll-Like Receptor 3</td>
</tr>
<tr>
<td>TME</td>
<td>tumor microenvironment</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>UTX</td>
<td>Ubiquitously transcribed tetratricopeptide repeat, X chromosome</td>
</tr>
</tbody>
</table>
Introduction

Tumor heterogeneity in cancer

Tumor cell heterogeneity has emerged as a major determinant of response to cancer therapy. Genomic and transcriptomic profiling of heterogeneous clonal landscapes of tumors have demonstrated the impact of intratumoral heterogeneity on response or resistance to cancer therapy, progression of cancer, and risk of relapse \(^1\), likely through multiple mechanisms. For example, heterogeneity likely confers clonal populations of cells with unique growth and metastatic properties, differential drug sensitivities, and can dictate the composition of cancer cells which are resistant to therapy or relapse earlier after therapy \(^2\). In addition, it is likely that unique subpopulations of tumor cells can induce changes in the immune cell composition of the tumor microenvironment through differential paracrine signaling behaviors \(^3\), likely further playing a role in shaping response to immune surveillance and immunotherapy. While much attention has been paid to the impact of progressive mutations on shaping the evolution of cancer cell subpopulations in response to therapy or in the tumor microenvironment, it is also becoming clear that key epigenetic changes in chromatin state can play a strong role in determining the behaviors of tumor cell subpopulations by inducing broad transcriptional programs with actions beyond those of individually mutated genes \(^4\). Hence, apart from cell subpopulations marked by mutations in particular oncogenic drivers, the role of distinct chromatin states as potential key drivers of treatment response and resistance in cancer is starting to be appreciated \(^4\). This idea has been supported by recent efforts in single-cell transcriptomic profiling of individual tumor and immune cells across multiple cancer types which have particularly highlighted transcriptomic and epigenomic heterogeneity among cancer cell subpopulations \(^5\).
Mesenchymal tumor subclones as mediators of treatment resistance

One of the major axes on which to grade tumor cell heterogeneity which has been established across multiple cancer types is the epithelial-mesenchymal spectrum. The epithelial-mesenchymal transition (EMT) is the process by which epithelial cells lose polarity and cell-cell adhesiveness and transition toward a mesenchymal phenotype classically marked by increased invasiveness and capacity for migration. It is increasingly becoming recognized that the mesenchymal cell state enriches for drug resistance across multiple tumor types, including pancreatic, breast, and bladder cancer. A mesenchymal cell state characterized by high expression of the tyrosine kinase AXL and engagement of NF-κB-related innate immune signaling pathways has been shown to enrich for distinct transcriptional programs in melanoma and may underlie resistance to BRAF/MEK-directed targeted therapy, hinting at the possibility of altered innate immune signaling pathways as potential mediators of tumor cell aggressiveness in mesenchymal cancer. However, the mechanisms by which the mesenchymal cell state promotes aggressiveness and treatment resistance more broadly across cancers have remained poorly defined.

Mesenchymal subclones in small cell lung cancer

A major malignancy which has been characterized by distinct mesenchymal cell populations is small cell lung cancer (SCLC). SCLC is an aggressive cancer with a median survival time of 15-20 months (limited stage) and 9-13 months (extensive stage). Although often highly chemosensitive and radiosensitive on initial presentation, current targeted therapy options in SCLC are mostly lacking. Hence, most SCLC patients experience relapse with systemic metastasis within 12-24 months of treatment initiation, resulting in a poor two-year
survival rate of 5-20% \(^{11}\). In addition, despite the recent successes of anti-PD-1- and anti-CTLA-4-based immunotherapies entering the first line of treatment for renal cell, melanoma, and non-small cell lung cancers \(^{12}\), immunotherapy has yet to make significant inroads in the treatment of SCLC \(^{12}\). Given the devastating outcome of most SCLC patients, identifying drivers of immunogenicity among heterogeneous subpopulations of SCLC cells resistant to therapy could help sensitize this aggressive tumor to immune checkpoint blockade-based approaches.

The presence of intratumoral heterogeneity in small cell lung cancer has been well established and may underlie its resistance to treatment. It has been shown that SCLC tumor cells display strong heterogeneity across the EMT spectrum, with well-described populations of neuroendocrine and mesenchymal (non-neuroendocrine) cells present \(^{13}\). This mesenchymal state switch has been shown to be induced by activation of RAS signaling \(^{13}\), HGF/MET signaling \(^{14}\), or chemotherapy with chronic doxorubicin \(^{14}\), with these mesenchymal subclones helping confer the neuroendocrine component with metastatic capacity through paracrine cytokine signaling such as the tumor-promoting CCL5 \(^{13}\). In addition, blockade of MET-induced EMT has been shown to sensitize cells to chemotherapy \(^{14}\). Hence, understanding the mechanisms underlying how this mesenchymal cell state promotes aggressiveness in the tumor microenvironment, and subsequently targeting these mesenchymal subclones in SCLC, have emerged as major goals.

**Innate immune signaling in mesenchymal tumor subclones**

Given that small cell lung cancer is almost ubiquitously defined by the loss of both \(p53\) and \(Rb\) \(^{15}\), previous efforts in the Barbie lab have used the \(Rb/p53\)-null H69 cell line to investigate heterogeneity between tumor subclones. Whereas parental cells from the H69 cell line are enriched for a neuroendocrine phenotype, the H69M subclone (induced by MET
activation) and chemoresistant H69AR subclone display strong mesenchymal patterns characterized by an adherent phenotype and activation of innate immune signaling pathways. H69M and H69AR display activated cytokine signaling (IL-6, IL-8, CCL2, CXCL10) and innate immune pathways including TBK1, NF-κB, and STAT1. In addition, H69 mesenchymal subclones uniquely attract T cells and monocytes in microfluidic co-culture compared to parental H69 cells. Furthermore, in contrast to parental H69, the H69M line contains a subclone enriched for high expression of PD-L1, the major ligand for the inhibitory T cell receptor PD-1, and CD44, a marker which has recently begun to characterize stem-like tumor cells associated with evasion of anti-tumor immune responses – suggesting important roles in resistance to tumor immune surveillance and immunotherapy.

In particular, we sought to understand the mechanisms underlying this “H69M PD-L1 high” state given that high tumoral PD-L1 expression has been shown to be a favorable prognostic marker for anti-PD-1 based therapy in non-small cell lung cancer, melanoma, and other solid tumors. Genomic sequencing of parental H69 cells and its mesenchymal subclones revealed no clear mutation or copy number alteration associating with the mesenchymal state, and H69M PD-L1 high reverted morphologically toward the parental H69 state through serial passages over time. Hence, these findings suggested a potential epigenetic mechanism of immunogenicity, and PD-L1 induction, marking the mesenchymal state in SCLC subclones, and prompted us to explore approaches in chromatin profiling to potentially address this state change.

**Endogenous retroviruses as regulators of tumor immune signaling**

A major mechanism of immunogenicity which is starting to be explored as a potential determinant of tumor microenvironment composition and response to immunotherapy is the
activation of noncoding genomic elements. Endogenous retroelements are noncoding regions which comprise up to 40% of the human genome, with the vast majority consisting of elements lacking long terminal repeats (LTRs). These non-LTR retroelements consist of long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs), while LTR-bound retroelements include endogenous retroviruses (ERVs). Transcription of LINEs, SINEs, and ERV elements have been shown to play critical roles in the pathogenesis of immune-related diseases and immune system function, with sense and antisense transcription of endogenous retroelements creating complementary strands of single stranded RNA (ssRNA) which may hybridize to form double stranded RNA (dsRNA). In addition, it has been suggested that increased levels of reverse transcriptase (RT) activity in the cytoplasm in cells with high endogenous retroviral expression may lead to reverse transcription of retroelement dsRNA molecules and subsequent production of cytosolic dsDNA. Together, these have been shown to be detected by cytoplasmic RNA sensors, such as the RIG-I/MAVS dsRNA sensing pathway, or cytoplasmic DNA sensors, such as the cGAS-STING pathway. Since both MAVS and STING are involved in promoting IRF3 phosphorylation and activation of inflammatory gene transcription and cytokine production through activation of TBK1, we considered endogenous retroelement activation as an epigenetic mechanism underlying the pTBK1, PD-L1 high state in mesenchymal SCLC subclones.

In particular, we chose to focus on endogenous retroviruses (ERVs) given the recent attention which has been paid to the transcriptional patterns of ERVs and their roles in shaping tumor immunity. ERVs have been shown to comprise up to 8-9% of the human genome and have been shown to be subject to genome-wide regulation, primarily by epigenetic silencing. It has been suggested that ERV silencing is a major function of DNA methylation in humans, with
over 90% of methylated CpG islands located in genomic regions harboring repetitive elements including ERVs. In concert with these findings, it has recently been demonstrated that broad induction of ERV elements genome-wide through disruption of DNA methyltransferase (DNMT) activity with decitabine (DAC) can augment anti-tumor immunity across multiple solid tumors. In particular, DNMT inhibition (DNMTi) was shown to induce ERV expression and anti-viral immunity in colorectal cancer by activating a type I interferon response dependent on MAVS and mimicked by transfection of RIG-I or MDA5 agonists. In addition, DNMTi was shown to induce a broad panel of ERVs and sensitize an ovarian cancer model to anti-CTLA-4 therapy, with response dependent on the dsRNA sensors MAVS and TLR3. However, these studies focused on broad de-repression of ERVs through DNMT inhibition, whereas it is thought that across individual tumors, distinct silencing mechanisms likely operate on different groups of endogenous retroelements. Furthermore, recent computational efforts have uncovered that specific tumor subtypes are enriched for particular subsets of ERVs, suggesting distinct tumor-specific mechanisms of ERV activation and further illuminating heterogeneity that warrants further investigation. Hence, potentially identifying a tumor-specific subset of ERVs in SCLC mesenchymal subclones and exploring strategies to induce their expression represented to us an opportunity to apply epigenetic approaches targeting unique ERV signatures to enhance immunotherapy.

The SPARCS endogenous retroviral signature marks mesenchymal subclones in SCLC

To potentially identify mesenchymal-associated ERVs in SCLC, previous work in the Barbie lab evaluated recently described ERV panels against gene expression profiling comparing the H69 cell line and H69M mesenchymal subclones, identifying that the ERV
MLT1C49 is upregulated in H69M PD-L1$^{\text{high}}$ compared to H69. A unique feature our group observed is the antisense orientation of this ERV in 3’ UTR of TRIM22, an IFN-γ stimulated gene. Given that previous groups have operated under the idea that RNA polymerase III-induced bidirectional transcription of repetitive elements can contribute to dsRNA production from ERV sequences, our group chose to explore 3’ UTR antisense ERVs in the genome, hypothesizing that antisense-oriented ERVs might be more prone to produce dsRNA from bidirectional transcription. Intersection of all 3’ UTR repeat elements from RefSeq with genes upregulated in H69M revealed TRIM22 and 14 additional genes containing antisense ERVs, including TRIM38 which contains two (MLT1A, MLT1J). Given the potential of this antisense ERV signature to generate dsRNA and “spark” immune signaling, our group has termed these mesenchymal-associated ERVs Stimulated 3 Prime Antisense Retroviral Coding Sequences (“SPARCS”) (Table 1). Notably, published ENCODE ChIP-seq data from HeLa cells suggested all that 15 genes in this ERV signature contain prominent STAT1 transcription factor binding peaks at their promoters, suggesting broad inducibility of SPARCS ERVs by IFN-γ, and pointing at a strong link between the mesenchymal state and IFN-γ signaling.

Further work in the Barbie lab confirmed that H69 mesenchymal subclones strongly express SPARCS ERVs in response to IFN-γ exposure and that these ERV loci produce high levels of dsRNA. In addition, our group observed strong links between SPARCS expression and secretion of CXCL10, a key T$_h$1-promoting chemokine involved in T cell recruitment to sites of inflammation which is secreted in response to MAVS sensing of dsRNA, as well as CCL2, a chemokine involved in recruitment of monocytes and myeloid-derived cells. In addition, deletion of the nucleic acid sensors MAVS and STING abrogated this SPARC-associated cytokine response and pTBK1 induction in the mesenchymal state. Computational analysis of the
SPARCS signature across the Cancer Genome Atlas (TCGA) and the Cancer Cell Line Encyclopedia (CCLE) RNA-seq databases revealed genome-wide associations of SPARCS with markers of antigen presentation/processing machinery, cytosolic RNA sensors, and immunosuppressive factors including \textit{CD274} (PD-L1), \textit{PDCD1LG2} (PD-L2), as well as additional immune checkpoints – all markers of a chronic virally infected state. In addition, SPARCS were enriched in multiple cancer histologies beyond SCLC, including melanoma, clear cell renal cancer, lung adenocarcinoma, and head and neck squamous cell carcinoma. Finally, immune cell gene set enrichment analysis (GSEA) revealed clear associations of SPARCS expression with signatures of the innate immune response, adaptive immune response, and macrophages, consistent with the association of SPARCS high state with T cell chemokine CXCL10 and the myeloid cell chemokine CCL2.

Our group has therefore identified SPARCS as a novel subclass of mesenchymal-associated endogenous retroviruses which promotes both immunogenicity and markers of immune suppression in cancer cells by inducing a virally-infected state. Given the association of T cell cytotoxic markers with response to immune therapy\textsuperscript{35, 39} and the idea that immune checkpoint expression, although functionally immunosuppressive, may represent a vulnerability to immune checkpoint-based therapy and may stratify potential responders\textsuperscript{20, 21}, SPARCS physiology may have potential implications for immunotherapies across a variety of tumors. However, little is known about the transcriptional mechanisms responsible for the SPARCS-high state in SCLC mesenchymal subclones. Given that epigenetic regulation of ERVs is starting to be explored across a variety of tumor types and transcriptional states\textsuperscript{24, 31, 32}, gaining an understanding of potential epigenetic mechanisms responsible for SPARCS induction is critical.
in order to explore how to leverage the SPARCS-high state when considering drug combinations with PD-1/CTLA-4 axis blockade.

**Epigenetic targeted therapy in the context of cancer immunotherapy**

A number of recent efforts have begun to focus on applying principles of epigenetics to cancer immunotherapy, given the potential for epigenetic agents to enhance multiple underlying mechanisms of the anti-tumor immune response both in tumor cells and the tumor microenvironment (TME)\(^{40,41}\). Broadly, epigenetic changes include DNA methylation and post-translational histone modifications including acetylation, methylation, and phosphorylation, which help to determine accessibility of chromatin to transcriptional activators including transcription factors and RNA polymerases\(^{40}\) and ultimately influence transcriptional levels of local genes. Tumor cell-autonomous effects of epigenetic agents on immunotherapy include induction of tumor-associated antigens such as germline cancer-testis antigens, which are often epigenetically silenced in somatic cells, by histone deacetylase (HDAC) inhibition and DNMT inhibition\(^{42}\). In addition to broad induction of endogenous retroviruses by DNMT inhibition as described previously\(^{33,34}\), the chromatin-binding functions of CDK4 and CDK6 have also been explored as regulators of endogenous retrovirus expression\(^ {43}\), antigen presentation, interferon signaling, and tumor immunogenicity\(^ {44}\). With regard to the roles of chromatin regulators on the tumor microenvironment, it has been suggested that the histone methyltransferase EZH2 and related components of the repressive PRC2 complex may directly silence the loci of the T\(_h\)1 chemokines CXCL9 and CXCL10 by histone H3 lysine 27 trimethylation (H3K27Me3) resulting in decreased local T cell infiltration, ultimately contributing to resistance to immunotherapy\(^ {45,46}\).
Similarly, HDAC inhibition has also been shown to enhance T cell-related chemokine expression and TME infiltration in lung cancer\(^4^7\).

Given (1) the growing literature supporting the idea that epigenetic modulation may influence tumor cell-intrinsic and tumor cell-extrinsic variables in the context of cancer immunotherapy, (2) the well-established roles of DNA and histone regulation on genome-wide patterns of endogenous retroviral expression, and (3) the association of SPARCS with multiple markers of immunogenicity which have also been associated with perturbation of key epigenetic regulators, we hypothesized that SPARCS ERV physiology may be a mechanism underlying these observations. To this end, we aimed to use approaches in chromatin profiling following by pharmacologic and genetic approaches targeting candidate chromatin regulators to understand the key epigenetic pathways upstream of the SPARCS-high state, and how they may be leveraged to augment responses to immunotherapy. In particular, we were interested in studying how depletion of the histone methyltransferase EZH2 may promote a T\(_h\)1-type chemokine response dependent on SPARCS dsRNA generation\(^4^5,4^6\), especially given the links between the SPARCS-high state and CXCL10\(^1^6\), a known downstream effector of dsRNA sensing\(^3^7\), and the established role of EZH2 in silencing of repetitive sequences in cells with intact Rb function\(^4^8\).

In doing so, we hoped to characterize a chromatin “state switch” in mesenchymal SCLC which might help explain not only its aggressiveness but also a potential vulnerability to PD-1/PD-L1-based immunotherapy.
Methods

Gene set enrichment analysis

The SPARCS signature was created by overlapping 3’UTR repeat elements from RefSeq with genes upregulated in H69M when compared to parental H69\textsuperscript{14,16}. The SPARCS gene set scores across CCLE (www.broadinstitute.org/CCLE) or TCGA (Pancan12) were derived by using an ssGSEA algorithm\textsuperscript{49}. These scores were used to identify top associated features from the matching datasets containing profiles of mRNA and pathways (MsigDB). To quantify the degree of association, an information-theoretic measure Information Coefficient (IC) was used and an empirical permutation test for statistical significance calculations.

OncoPanel assay

Somatic mutations, copy number variations and structural variants in parental H69 cells and H69M-PD-L1\textsuperscript{low}/H69M-PD-L1\textsuperscript{high} subclones were evaluated by performing the OncoPanel assay from the Center for Advanced Molecular Diagnostics from Brigham and Women’s Hospital. This OncoPanel assay surveys exonic DNA sequences of 300 cancer genes and 113 introns across 35 genes for rearrangement detection. DNA was isolated from cell lines and analyzed by massively parallel sequencing using a solution-phase Agilent SureSelect hybrid capture kit and an Illumina HiSeq 2500 sequencer.

Cell lines

The human SCLC cell lines H69, H69M, and H69AR were obtained from the laboratory of Dr. Joan Albanell. H82, H1436 and H2081 were obtained from the American Type Culture Collection (ATCC) (Rockville, MD). 786OsgCTRL and 786OsgBAF180 (PBRM1) were
obtained from the laboratory of William G. Kaelin, Jr. (Dana-Farber Cancer Institute). All were authenticated following Short Tandem Repeat (STR) genotyping. All cell lines were cultured in RPMI-1640 (Thermo Fisher Scientific, #11875-119) containing 10% FBS (Gemini Bio-products, #100-106) and 1X pen-strep (Gemini Bio-products, #400-109).

**ATAC-seq and analysis**

We performed assay for transposase-accessible chromatin followed by sequencing (ATAC-seq) on H69, H69M, and H69AR cells according to ⁵⁰. Briefly, we sorted 40-50,000 cells per biological replicate, which were then washed once in cold PBS and lysed in 50µL cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% IGEPAL CA-630). Lysed nuclei were incubated in Tn5 transposition reaction mix and purified using MinElute Reaction Cleanup kit (Qiagen). ATAC-seq fragments from one set of replicates for H69 and H69AR cells were size selected for fragments between 115 and 600 bp using Pippin Prep 2% Agarose Gel Cassettes and the Pippin Prep DNA Size Selection System (Sage Science). Post size-selection, ATAC libraries were amplified and Nextera sequencing primers ligated using Polymerase Chain Reaction (PCR). Finally, PCR primers were removed using Agencourt AMPure XP bead cleanup (Beckman Coulter/Agencourt) and library quality was verified using a Tapestation machine. High quality ‘multiplexed’ DNA libraries were sequenced on the Illumina HiSeq2000. The ends of the paired-end fragments are used as cut sites and enriched peaks were called with MACS2 with following parameters (--nomodel --extsize 200 --shift -100 -g hs -B --nolambda). For IGV visualization, shifted bedGraph were converted to wig files at 10bp resolution and normalized to read counts by wigmath tool of JavaGenomic toolkit.

**ChIP-seq and analysis**
We performed chromatin immunoprecipitation followed by sequencing (ChIP-seq) on H69, H69M, and H69AR cells for the H3K27Ac and H3K27Me3 marks. Briefly, approximately 10 million cells per condition were fixed in paraformaldehyde, lysed, and sheared on wet ice. The sheared chromatin was cleared and incubated overnight at 4°C with H3K27ac (Abcam, 4729) and H3K27Me3 (Abcam, 6002) ChIP-grade antibodies. Antibody-chromatin complexes were immunoprecipitated with protein G magnetic Dynal beads (Life Technologies), washed, eluted, reverse crosslinked, and treated with RNAse A followed by proteinase K. ChIP DNA was purified using Ampure XP beads (Beckmann Coulter) and then used to prepare sequencing libraries for sequencing with the Next-Seq Illumina genome analyzer. Reads were aligned to hg19 using Burrows-Wheeler Aligner (BWA) and identical ChIP-seq sequence reads were collapsed to a single read to avoid PCR duplicates. Enriched peaks were called with MACS2 with following parameters (--nomodel --extsize 200 --shift -100 -g hs -B --nolambda). For IGV visualization, shifted bedGraph were converted to wig files at 10bp resolution and normalized to read counts by wigmath tool of JavaGenomic toolkit.

**Compounds and treatments**

Recombinant human IFN-γ (#285-IF) protein was purchased from R&D Systems (Minneapolis, MN) and reconstituted in sterile, deionized water. MRT67307 and Ruxolitinib were synthesized and purchased from Shanghai Haoyuan Chemexpress Co. Both drugs were reconstituted at 10 mM in DMSO and stored at -20°C. GSK126 (#S7061) was purchased from Selleck chemicals (Houston, TX) and reconstituted at 5 mM in DMSO and stored at -20°C. DAC/decitabine (#S1200) was purchased from Selleck chemicals (Houston, TX) and
reconstituted at 10 mM in DMSO and stored at -20°C. EPZ6438 and GSK343 were obtained from the laboratory of Kwok K. Wong (Dana-Farber Cancer Institute).

For IFN pulse experiments, cells were pulsed 10 minutes with IFN-γ (200 ng/mL), extensively washed, and chased in fresh media for an additional 24 or 72 hours. To test drug effects on gene expression or protein secretion, IFN-γ pulsed H69AR cells were treated with DMSO, 1μM MRT67307 or 100 nM Ruxolitinib for 24, 48 and 72 hours. For epigenetic experiments, H69 cells were treated for 6 days with with either 5 μM GSK126, 250 nM to 5 μM EPZ6438, 250 nM to 10 μM GSK343, or 100 nM DAC. Drug was replenished every 3 days with both suspension and adherent cells carried each time. After the drug treatment period, equal numbers of DMSO-treated and drug-treated cells were exposed to either H2O (ctrl) or 200 ng/mL IFN-γ for 24 hours before harvesting of RNA or conditioned media (CM).

**Western blotting, antibodies, and ELISA**

Protein was isolated from cell lines using RIPA lysis buffer and measured by BCA (Pierce Biotechnology). Protein extracts were subjected to polyacrylamide gel electrophoresis using the 4%–12% NuPAGE gel system (Invitrogen, Carlsbad, CA), transferred to PVDF (Millipore) membranes, and immunoblotted using antibodies that specifically recognize EZH2 (#5246), STING (#13647), H3K27Me3 (#9733), and β-Actin (#4970) (Cell Signaling Technologies, Danvers, MA). Secondary antibodies were from LICOR Biosciences (Lincoln, NE): IRDye 800CW Goat anti-Mouse IgG (H + L) (#926-32210), IRDye 800CW Goat anti-Rabbit IgG (H + L) (#926-32211). LICOR blocking buffer (#927-40000) was used to dilute primary and secondary antibodies, with the exception of phosho-specific antibodies, which were diluted in HIKARI Signal Enhancer Solutions 1 and Solution 2 (Nacalai USA, Inc. # NU00101). Imaging
of blots and quantitation of bands was performed using the LICOR Odyssey system. ELISA for CCL2 and CXCL10 were done using the Human CCL2/MCP-1 and Human CXCL10/IP-10 Quantikine ELISA kits (R&D systems) according to manufacturer’s instructions.

**Cytokine profiling**

Multiplex assays were performed utilizing the bead-based immunoassay approach Bio-Plex Pro™ Human Cytokine 40-plex Assay (Cat# 171AK99MR2) on a Bio-plex 200 system (Cat# 171000201) (Bio-Rad Laboratories, Hercules, CA) and the Human Cytokine/Chemokine Maganetic Bead Panel (Cat# HCYTMAG-60K-PX30) on a Luminex MAGPIX system (Merck Millipore, Billerica, MA). Conditioned media concentration levels [pg/ml] of each protein were derived from 5-parameter curve fitting models. Fold changes relative to the corresponding control were calculated and plotted as log2FC. Lower and upper limits of quantitation (LLOQ/ULOQ) were imputed from standard curves for cytokines above or below detection.

**PCR primer design**

PCR primers for the human CCL2, CXCL10, IFN-γ, IFN-β, PD-L1, HLA-A, HLA-B, DDX58, DHX58, TAP2, IRF4, IRF9, TMEM173, and B2M genes were obtained from MGH PrimerBank (https://pga.mgh.harvard.edu/primerbank/). PCR primers for SPARCS ERVs (MLT1C49_TRIM22, MLT1A_TRIM38, MLT1J_TRIM38, THE1D_IL32, MLT1K_SPATS2L, MER92B_ADAM19, LTR26_IFI44L, MLTI1_F3, PABL_B_BEND6, MLTIAO_AIG1, MSTA_MSRB2, and LTR79_ANTXR1) were designed using NCBI-PrimerBLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) and verified using UCSC Genome Browser in silico PCR (https://genome.ucsc.edu/cgi-bin/hgPer). PCR primers for non-SPARCS ERVs (ERV
MER34, THE1D, ERV Fb1, ERV 9.1, MER57B1, and ERVF) were obtained from previous studies [27, 28].

**Quantitative RT-PCR**

Total cellular RNA was extracted using the miRNeasy Mini Kit (Qiagen, Hilden, Germany) according to manufacturer’s instructions. After extraction, 1 µg total RNA was used to generate cDNA with the SuperScript III First-Strand Synthesis SuperMix for qRT-PCR kit (Thermo Fisher Scientific, Waltham, MA). Quantitative reverse transcription PCR (qRT-PCR) of the indicated genes was performed using SYBR green PCR Master Mix (Applied Biosystems, Foster City, CA) and the Applied Biosystems 7300 Fast real-time PCR system and software. The relative expression was normalized with the expression of the housekeeping gene 36B4.

**CRISPR-Cas9 gene editing and lentiviral infection**

Oligonucleotides coding for guide RNAs that target the EZH2 and MAVS genes were chosen from the Avana library and the Brunello library. A non-targeting sgRNA from the Gecko library v2 was used as a dummy sgRNA for control (sgCTRL). Lenti CRISPRv2 vectors were cloned as previously described. HEK-293T cells were transduced with lentiCRISPRv2 using X-treme Gene 9 (Roche, Basel, Switzerland) according to the manufacturer's instructions. On day 2, target cells were seeded, and allowed to adhere overnight. On day 3 the supernatant of transduced HEK293T cells was collected and added to the target cells through a 0.45 µm filter. Supernatant from transduced HEK293T cells was again collected and added to target cells on day 4. On day 5, puromycin or blasticidin was added to select infected cells (for four days).
**Poly(I:C) treatment**

For Poly(I:C) dsRNA treatment experiments, H69ARsgCTRL and H69ARsgMAVS cells were plated in RPMI media, transfected with 0.5 µg/mL Poly(I:C) HMW (InvivoGen, Sand Diego, CA) using XtremeGene HP transfection reagent (Sigma-Aldrich, St. Louis, MO) and cultured for 72 hours. On day 3 after transfection, conditioned media was recovered and CXCL10 protein expression was quantified. RNA was extracted and expression levels of relevant genes were analyzed by qRT-PCR.

**TCGA correlative analyses**

The SPARCS gene set score was used to rank TCGA tumor samples (PanCan12, n = 3602) in order of decreasing SPARCS expression, with the top 50 and bottom 50 SPARCS samples used for further analyses. TCGA RPKM values for *STING, CXCL10, CCL2, EZH2, PBRM1, ARID2, ARID1A, ARID1B,* and *SMARCA4* were plotted for SPARCS-high and SPARCS-low samples and visualized in GraphPad Prism7.

**Statistical analyses**

All graphs depict mean ± s.d. unless otherwise indicated. Tests for differences between two groups were performed using two-tailed unpaired Student’s *t*-test or Mann-Whitney’s two-tailed test, as specified in the figure legends. P values were considered significant if less than 0.05. Asterisks used to indicate significance correspond with: * = p < 0.05, ** = p < 0.01, *** = p < 0.001. GraphPad Prism7 was used for statistical analysis of experiments, data processing and presentation.
Results

SPARCS ERVs associate with DNA and histone methylation pathways across tumors

To begin addressing how SPARCS ERVs may be epigenetically regulated, we ranked expression of the 15 gene SPARCS signature across the Cancer Genome Atlas (TCGA) PanCan12 (n = 3602 tumors) using single sample gene set enrichment analysis (ssGSEA) (Figure 1) and identified top associated MSigDB gene sets. Top epigenetic pathways included associations with the actions of EZH2 and SUZ12, both components of the Polycomb repressive complex 2 (PRC2) involved in transcriptional silencing by deposition of H3K27Me3 on chromatin. Additional pathways suggested association of SPARCS with the action of DNA methyltransferases (“LIANG SILENCED BY METHYLATION”), as well as specific associations with DNMT1. Given that EZH2 is the core enzymatic component catalyzing H3K27 trimethylation and the primary pharmacologic target of the PRC2 complex, and since DNMT inhibition have been well-studied in the context of ERV de-repression, we chose to focus on inhibition of histone and DNA methylation as potential drivers of the SPARCS-high state. In addition, we performed genomic sequencing on the H69 cell line as well as two mesenchymal subclones isolated from H69M which were previously sorted by flow cytometry for PD-L1 expression (Figure 2). OncoPanel sequencing of 300 cancer-associated genes revealed no clear enrichment of either H69M mesenchymal subclone with any mutation or copy number variant in any gene tested. Taken together, these results suggested that epigenetic, rather than genomic, alterations characterize the mesenchymal state in H69 SCLC subclones, and prompted us to explore chromatin profiling as a next step.
SPARCS ERV loci display more open chromatin profiles marked by loss of histone methylation in mesenchymal tumor subclones

In order to assess whether H69 mesenchymal subclones display more accessible chromatin profiles around SPARCS gene loci, we performed Assay for Transposase Accessible Chromatin using sequencing (ATAC-seq) on parental H69 cells as well as the mesenchymal subclones H69M and chemoresistant H69AR. Compared to parental H69, the mesenchymal subclone H69AR displayed gain of accessible chromatin peaks around almost every SPARCS gene, while H69M displayed gain of accessibility around multiple SPARCS loci, although to a lower extent (Figure 3a). Notably, ATAC-seq around the CD274 (PD-L1), CXCL10, and CCL2 loci did not show gain of accessibility in either H69 mesenchymal subclone (Figure 3b), despite the fact that these genes are more highly expressed in response to IFN-γ in H69 mesenchymal subclones compared to H69. These suggested that the activation of these genes in mesenchymal subclones is not necessarily due to altered chromatin state at these loci, and suggests that alternative factors may be driving their expression, such as SPARCS dsRNA sensing.

After establishing that H69 mesenchymal subclones display gain of accessibility around SPARCS ERV loci, we next sought to determine which specific epigenetic alterations might be most strongly linked to this permissive chromatin state. Given that EZH2 is responsible for H3K27 trimethylation, we performed H3K27Me3 ChIP-seq on H69, H69AR, and H69M cells (Figure 4). In concert with the ATAC-seq data, the H69M subclone displayed lower H3K27Me3 ChIP signal at a majority of SPARCS loci including TRIM22, TRIM38, SPATS2L, ADAM19, SERPINB9, F3, BEND6, AIG1, and TNFRSF9. Similar results were observed for H69AR, with significant loss of signal at TRIM22, SPATS2L, EPHA3, HERC3, ADAM19, F3, BEND6, and
ANTXR1. It should be noted that two SPARCS displayed higher signal in H69M (EPHA3, MSRB2) and H69AR (IFI44L, MSRB2) compared to H69 respectively, suggesting that there may be variability in regulation of individual SPARCS. However, in general, SPARCS appeared to be enriched for loss of H3K27Me3 signal corresponding to a more open chromatin state in mesenchymal subclones.

To further confirm that the open chromatin state of SPARCS loci in mesenchymal subclones corresponded to the development of active enhancers, we performed H3K27Ac ChIP-seq on H69, H69M, and H69AR (Figure 5). H3K27 acetylation is associated with active transcription and antagonizes H3K27Me3-mediated transcriptional silencing of Polycomb target genes; in addition, H3K27 acetylation marks active regulatory elements and can be used to understand how transcription factor activity can help promote the development of active enhancers. Interestingly, ChIP-seq showed gain of H3K27Ac signal around more SPARCS loci in H69M than in H69AR, despite the ATAC-seq results which suggested that SPARCS loci were more open in H69AR. Given that H69M expresses SPARCS in basal conditions, presumably due to endogenous IFN-γ secretion, whereas H69AR requires IFN-γ treatment to hyperactivate SPARCS expression, these results suggest that in the mesenchymal state, SPARCS may be marked by latent enhancers which require IFN-γ-mediated STAT1 transcription factor activity in order to set up active enhancers marked by H3K27Ac.

Taken together, these ATAC-seq and ChIP-seq results suggest that SPARCS transcriptional activation may be promoted by both the loss of H3K27Me3 and the presence of IFN-γ activity resulting in active SPARCS loci marked by H3K27Ac. Together with the ssGSEA results suggesting that SPARCS associates with histone methylation pathways across cancers
(Figure 1), these provided our rationale for inducing SPARCS expression in non-mesenchymal cells by depleting EZH2 activity and exposing cells to IFN-γ.

**EZH2 inhibition induces SPARCS expression associated with the mesenchymal phenotype**

Given our evidence that H69 mesenchymal subclones display less H3K27 trimethylation at SPARCS loci which potentially underlie their induction, we next sought to understand which factors may be involved in this process. H3K27 methylation is thought to be governed by a continuous interplay between the action of H3K27 methyltransferases, such as EZH2 and other PRC2 complex members, and the action of H3K27 demethylases, such as the JMJD3/UTX family. Clinically, modulation of H3K27 methylation has been dominated by progress in targeting of EZH2, particularly in hematologic malignancies marked by EZH2 gain of function mutations or overexpression. Given that SPARCS-high mesenchymal subclones display loss of K27Me3 at SPARCS loci, we expected that alterations of EZH2 targeting to ERV loci, as has been suggested in Rb-mutant tumors which display deranged EZH2 silencing of repetitive elements, may underlie this phenomenon. However, to our surprise, we found that H69M and H69AR also clearly displayed lower EZH2 expression compared to H69 (Figure 6a), suggesting both lower total levels of EZH2 in addition to potential differences in EZH2 targeting of genomic loci in these subclones.

Next, we aimed to use genetic and/or pharmacological manipulation to investigate how EZH2 may control SPARCS expression. CRISPR/Cas9-mediated deletion of EZH2 was unsuccessful in H69 (and multiple other SCLC lines with high EZH2 expression and neuroendocrine features including H82, H1436, and H2081), suggesting that complete depletion of EZH2 may not be compatible with cell viability in SCLC. Hence, we instead chose to focus
on the action of GSK126, a well-validated small molecule inhibitor of EZH2 which has been used in multiple cell- and murine-based immunotherapy studies\textsuperscript{40,45}. Indeed, we confirmed that GSK126 doses which have been used in prior studies\textsuperscript{45} efficiently depleted total H3K27Me3 levels in H69 cells (Figure 6b).

Given that the SPARCS signature was derived directly from gene expression data from the mesenchymal state (H69M), we next aimed to demonstrate how regulation of ERVs in the SPARCS subclass may be unique and distinct from previous studies who described broad ERV upregulation in response to DNMT inhibition\textsuperscript{33,34}. To do so, we pre-treated H69 cells with either EZH2 inhibition for 6 days using 5 µM GSK126 or DNMT inhibition for 6 days using 100 nM DAC (decitabine), followed by either control treatment or a 10 minute 200 ng/ml IFN-γ pulse and recovery after 24 hours (Figure 6c). To address whether our interventions were SPARCS-specific or could be applied to ERVs more generally, we assessed for relative induction of a panel of 12 SPARCS genes as well as a control panel of six non-SPARCS ERVs we obtained randomly from previous studies\textsuperscript{33,34}. In control conditions without IFN-γ pulse, neither EZH2 nor DNMT inhibition significantly induced ERVs from either the SPARCS (Figure 6d) or non-SPARCS (Figure 6e) panels. However, in IFN-γ pulse conditions, pre-treatment with EZH2 inhibition significantly induced a majority of SPARCS (9 of 12) compared to DNMT inhibition which induced none (Figure 6f), whereas only one of six non-SPARCS ERVs was induced by EZH2 inhibition (Figure 6g). Taken together, these results suggest that in the presence of IFN-γ, EZH2 inhibition specifically induces SPARCS ERVs and largely does not induce ERVs more generally, whereas DNMT inhibition is unable to strongly induce SPARCS ERVs – suggesting strong links between EZH2 activity and specific regulation of SPARCS ERV loci.
In addition, we noticed through these experiments that treatment of H69 cells with EZH2 inhibition significantly altered their morphologies in culture. SCLC cell models display phenotypic heterogeneity mainly governed by changes along the EMT spectrum, with neuroendocrine-type cells often growing as suspensions in culture, while SCLC lines displaying mesenchymal features such as expression of vimentin and fibronectin tend to grow as adherent cells. In basal conditions, the neuroendocrine H69 cell line grows as aggregates in suspension; however, when grown in the presence of EZH2 inhibition, we noted the development of progressively larger fractions of adherent cells compared to control (Figure 6h) – further suggesting potential activation of EMT-related pathways.

Finally, we aimed to confirm that the specific upregulation of SPARCS ERVs in response to EZH2 inhibition in H69 cells matched the profiles of SPARCS expression in H69 mesenchymal subclones. Notably, in response to IFN-γ, pre-treatment with EZH2 inhibition in H69 cells (Figure 7a) exactly matched the profile of SPARCS expression in H69AR cells (Figure 7b), with 9 of 12 SPARCS ERVs showing significant upregulation in both experiments, and the same three ERVs (MSTA from the MSRB2 gene, MER92B from ADAM19, and LTR26 from IFI44L) showing no induction in either H69/EZH2i conditions or H69AR conditions. Taken together, these implied strong concordance between the transcriptional state of H69 cells treated with EZH2 inhibition and the mesenchymal H69AR state (Figure 7c), and were further suggestive of histone methylation as a key regulator of mesenchymal-associated ERVs.

**Use of a synthetic oligonucleotide to characterize dsRNA sensing pathways in H69 cells**

After establishing that EZH2 inhibition specifically promotes SPARCS ERV expression, we next aimed to learn more about the immune signaling pathways related to dsRNA sensing in
H69 cells in order to predict the potential downstream effects of EZH2 inhibition and the development of a SPARCS-high state. To do so, we used poly(I:C), a synthetic dsRNA oligonucleotide which is known to activate multiple canonical innate immune pathways downstream of dsRNA sensing including activation of TLR3 and MAVS signaling \(^5\). We first confirmed that high molecular weight (HMW), rather than low molecular weight (LMW) poly(I:C) induced strong expression of \(CXCL10\) in H69 cells (data not shown), and therefore used 0.5 µg/ml HMW poly(I:C) for all subsequent experiments. This \(CXCL10\) induction was particularly noteworthy, given its production downstream of dsRNA sensing \(^3\) and its documented roles in potentiating the effects of immune checkpoint blockade as a \(T_h\)1 chemokine \(^4, 46\). With regard to cytokine production, poly(I:C) treatment significantly induced the expression of \(CXCL10\) and \(CCL2\) (Figures 8a and 8b). This was accompanied by an increase in \(IFN-\beta\) expression, which was expected given the role of the type I interferons \(IFN-\alpha\) and \(IFN-\beta\) in anti-viral responses \(^6\). More curiously, poly(I:C) also led to a modest induction of \(IFN-\gamma\) expression, which was noteworthy given that SPARCS are \(IFN-\gamma\)-inducible and suggested a potential mechanism of positive feedback which was later confirmed in our group \(^16\). Finally, poly(I:C) induced high levels of \(PD-L1\) expression (Figure 8a), further suggesting potential roles the SPARCS-high state may play in response to immunotherapy.

Given the dual roles of both MAVS and TLR3 in intracytoplasmic dsRNA sensing, we aimed to understand which of these pathways was critical to transducing the response to poly(I:C). Surprisingly, CRISPR/Cas9-mediated deletion of the dsRNA sensor MAVS alone almost completely abrogated the induction of these factors in response to poly(I:C), with the strongest effects seen on \(CXCL10\) and \(IFN-\beta\) expression (Figures 8a and 8b). In addition, Luminex cytokine profiling of H69AR cells with MAVS deletion (H69AR sgMAVS) confirmed
downregulation of all 28 cytokines tested, confirming the critical role of MAVS sensing in dsRNA-induced cytokine production (Figure 8c). We then used the poly(I:C) model to understand which downstream innate immune signaling pathways were critical to the dsRNA-induced cytokine response. Most strikingly, exposure of poly(I:C)-treated H69AR cells to 1 µM ruxolitinib, a JAK1/2 inhibitor, also completely abrogated CXC10, CCL2, IFN-γ, IFN-β and PD-L1 expression (Figure 8d), CXCL10 secretion (Figure 8e), and every cytokine tested by Luminex profiling (Figure 8f) – highlighting the essential role of activation of JAK/STAT signaling downstream of dsRNA sensing in these responses. We also confirmed that treatment of these cells with 1 µM MRT67307, a dual TBK1/IKKε inhibitor, led to significant decreases in poly(I:C)-induced CCL2, IFN-β, and PD-L1 expression (Figure 8d) and CXCL10 secretion (Figure 8e), which was noteworthy given the role of MAVS in phosphorylation of TBK1 and activation of downstream innate immune responses including activation of IRF3 signaling and IFN-β production. Taken together, these set of experiments confirmed the strong links between dsRNA sensing and CXCL10 and CCL2 production, IFN-β signaling, and PD-L1 expression, which is dependent on MAVS activation, TBK1 activity, and JAK/STAT signaling – suggesting that many of these factors may play important roles in the SPARCS-high state.

EZH2 inhibition induces SPARCS-associated cytokine responses and markers of immunogenicity

After establishing that EZH2 inhibition in H69 cells can induce the same SPARCS upregulated in response to IFN-γ in mesenchymal subclones, and after characterizing likely downstream effectors linked to dsRNA sensing, we next asked whether EZH2 inhibition can promote a cytokine response linked to the SPARCS-high state. In concert with the role of EZH2
inhibition upstream of SPARCS compared to DNMT inhibition, pre-treatment of H69 cells with GSK126 led to significant increases in CXCL10 expression compared to pre-treatment with DAC in response to IFN-γ stimulation (Figure 6i). Given that H69AR cells also secrete CXCL10, we next used Luminex cytokine profiling to broadly compare the cytokine response to IFN-γ treatment between H69 cells pre-treated with EZH2 inhibition (Figure 9a), H69AR cells (Figure 9b), and H69M-PD-L1^{high} cells (H69M+), which secrete IFN-γ at baseline. H69 cells pre-treated with EZH2 inhibition hyperactivated expression of a number of IFN-γ-induced cytokines, including CXCL10, CXCL11, CX3CL1, CXCL9, and CCL2 (Figure 9a). Also, notably, treatment with EZH2 inhibition alone without IFN-γ stimulation led to endogenous IFN-γ production (Figure 9a), a noteworthy finding given the modest induction of IFN-γ in response to poly(I:C) dsRNA stimulation (Figures 8a and 8d) despite the fact that IFN-γ is classically thought to be exclusively secreted by immune cells rather than tumor cells.

To explore how these cytokine changes in response to EZH2 inhibition in H69 cells compared to cytokine expression in H69 mesenchymal subclones, we intersected the cytokines that were significantly upregulated in H69/EZH2i conditions, H69AR, and H69M-PD-L1^{high} (H69M+) cells (Figure 9c). Strikingly, almost every cytokine which was induced by EZH2 inhibition in H69 cells was also induced either in response to IFN-γ in H69AR cells or in basal conditions in H69M+ cells. When we compared the intersection of all three conditions, two main functional families of induced cytokines became clear: CXCL10 and CXCL11, both Th1 chemokines and ligands of the receptor CXCR3 responsible for T cell chemoattraction, CCL2 and CCL8, both involved in monocyte chemoattraction, CX3CL1, a ligand for the CX3CR1 receptor involved in attraction of both T cells and monocytes. Furthermore, we confirmed that two additional EZH2 inhibitors, EPZ6438 and GSK343, induced CXCL10 and CCL2 in
response to IFN-γ in H69 cells (Figure 10c), as well as SPARCS ERVs and PD-L1 (Figures 10a and 10b). Taken together, these results suggested that EZH2 blockade-induced SPARCS activation leads to mesenchymal-associated changes in cytokine secretion which may underlie the attraction of both T cells and monocytes into the tumor microenvironment.

In addition to characterizing the cytokine responses downstream of EZH2 inhibition and dsRNA sensing, we also aimed to characterize whether EZH2 inhibition in H69 cells can induce tumor cell-autonomous markers of immunogenicity associated with the SPARCS-high state. EZH2 inhibition induced components of the antigen presentation machinery, including the MHC class I genes *HLA-A* and *HLA-B* and the antigen processing transporter *TAP2*, all of which are significantly associated with the SPARCS-high state across tumors in TCGA. The underlying mechanism behind these associations was not immediately clear; although we considered potential translation of ERV sequences into peptides with homology to viral proteins which might be detected as foreign and presented to the cell surface, the discovery that EZH2 inhibition promotes endogenous IFN-γ production (Figure 9a) suggests that activated JAK/STAT signaling in these conditions may directly promote antigen presentation pathways. Regardless, these results suggest that dsRNA generation may contribute to the observation by prior groups that EZH2 blockade-mediated induction of antigen presentation contributes to response to anti-CTLA-4 and IL-2 immunotherapy in melanoma. Taken together, the observation that tumors with activated MHC class I expression may display improved responsiveness to immunotherapies highlight the potential translational relevance of increased immunogenicity following EZH2 inhibition.

In addition to markers of antigen presentation, EZH2 inhibition was associated with upregulation of multiple cytoplasmic nucleic acid sensing pathways. As was expected given
likely dsRNA generation in these conditions, EZH2 inhibition led to upregulation of *DDX58* and *DHX58*, which encode RIG-I RNA helicase family members functioning upstream of MAVS (Figure 11a). However, more surprisingly, we also noted that EZH2 inhibition led to increased expression of *TMEM173*, which codes for the cytoplasmic dsRNA sensor STING, both at the RNA and protein levels (Figures 11a and 11b). Multiple groups have observed cross-talk between RIG-I/MAVS-mediated RNA sensing and STING-mediated cytoplasmic DNA sensing; whether these are due to direct activation of RIG-I agonists on STING induction or through potential reverse transcription of ERV dsRNA into cytoplasmic DNA remains to be seen. Regardless, multiple groups have begun to explore STING agonists as potential adjuvants to immune checkpoint blockade and cancer vaccines in murine models, further supporting the role of EZH2 blockade-induced immunogenicity in response to immunotherapy.

**SPARCS-high tumors reveal associations with key immune sensing pathways, cytokines, and epigenetic regulators**

Finally, after using the H69 system to understand the relationships between EZH2 loss and induction of immunogenicity in mesenchymal subclones, we wanted to explore how these observations may extend more broadly to tumors with SPARCS-high or SPARCS-low features. To this end, we derived an aggregate SPARCS gene set score by incorporating relative expression of all 15 SPARCS genes using an ssGSEA algorithm, and applied this score to TCGA PanCan12 RNA-seq data (n = 3602 tumors) to rank all TCGA samples from high to low SPARCS expression. We labeled the top 50 as “SPARCS-high” samples and the bottom 50 as “SPARCS-low,” and used the expression patterns of these groups of tumor samples for further analysis. With regard to immune signaling and cytokine production, SPARCS-high and
SPARCS-low tumors across TCGA showed excellent concordance with the results we
discovered from EZH2 inhibition in H69 mesenchymal subclones. SPARCS-high samples
revealed significantly higher expression of TMEM173 (STING) (Figure 12a), as well as CXCL10
and CCL2 (Figures 12b and 12c) – suggesting that these important physiological links with
potential consequences for immunotherapy may be extended to other tumors with SPARCS-high
features, not just mesenchymal SCLC.

We also uncovered interesting correlations of the SPARCS-high state with patterns of
epigenetic regulation across tumors. As expected given the results seen in H69, the SPARCS-
high state was strongly enriched for loss of EZH2 (Figure 12d), suggesting that EZH2 inhibition
may play important roles in SPARCS de-repression and dsRNA generation in a number of other
tumor types. More surprisingly, however, apart from EZH2, we also noted strong correlation of
the SPARCS-high state with loss of PBRM1, a member of the SWI/SNF chromatin remodeling
complex 84 often inactivated in renal cell carcinoma 85. PBRM1 loss been recently reported to
promote responsiveness to anti-CTLA-4 therapy in RCC 86, and inactivation of SWI/SNF
complex members including PBRM1 and ARID2 have been linked to secretion of IFN-γ-induced
T cell effector cytokines and improved response to immunotherapy in murine melanoma models
87. Hence, we wondered whether SWI/SNF complex members may be associated with SPARCS
regulation more broadly, and chose to explore other members of the SWI/SNF complex
commonly mutated across cancers 84. To our surprise, SPARCS associated not only with loss of
PBRM1, but also with loss of multiple other SWI/SNF components, including ARID2, ARID1A,
ARID1B, and SMARCA4 (Figures 12f-12i). To further explore these findings, we tested SPARCS
expression and CXCL10 production from renal cell lines displaying PBRM1 (BAF180) loss.
BAF180-deficient 786-O cells displayed clear induction of CXCL10 secretion in response to
IFN-γ (Figure 12j) and de-repressed multiple SPARCS (Figure 12k). However, these changes were less striking compared to EZH2 inhibition (Figure 7a), suggesting that there may be some role for other epigenetic regulators upstream of the SPARCS axis, but also further supporting the stronger role of EZH2 in driving these changes. Taken together, these results further support the role of EZH2 in silencing of SPARCS loci across tumors as a mechanism of evading immunogenicity, but also introduce the possibility of multiple other mutant epigenetic regulators known to induce response to immunotherapy in other tumors which may operate partly through the SPARCS axis to exert their effects.
Discussion

Through these sets of experiments, we have characterized the critical role of EZH2 in silencing a novel subclass of endogenous retroviruses with potential consequences for cancer immunotherapy. By identifying an epigenetic state switch in mesenchymal subclones of small cell lung cancer marked by loss of histone trimethylation at SPARCS loci, we have uncovered how EZH2 inhibition may induce immunogenicity in this cell state by promoting SPARCS dsRNA production in response to IFN-γ. Further exploration of how the SPARCS-high state can be leveraged both in SCLC and related tumors may reveal new insights into immune signaling and immunotherapy in cancers with mesenchymal features.

Regulation of a novel ERV subclass

Primarily, our study adds to the growing literature surrounding epigenetic control as a major mechanism of regulation of endogenous retroviral elements. With regard to the role of ERV induction on immune signaling and response to immune checkpoint blockade, our study builds on the work of previous groups who primarily characterized DNMT inhibition as a non-specific inducer of ERVs genome-wide. Given that multiple groups are starting to explore tumor-specific ERVs and their associated mechanisms of induction, we hope to shed more light on the role of EZH2 inhibition on endogenous retroviral control going forward. Whereas we have uncovered its role in neuroendocrine SCLC cell lines which display high EZH2 expression, the observation that EZH2 inhibition has been shown to promote increased immunogenicity and T cell-mediated responses in diverse tumors including melanoma and ovarian cancer, as well as the association of SPARCS expression with EZH2 loss across TCGA PanCan12 (Figure
12d), raise the intriguing possibility that SPARCS control by EZH2 may be observed in other tumors as well.

**Refinement of our understanding of how EZH2 blockade enhances immunogenicity**

Our study also certainly adds to the understanding of how EZH2 may silence immunogenicity in cancer as reported by multiple groups\(^{45,46,77}\). Whereas previous groups have argued that the main immune-modulating functions of EZH2 and DNMT1 center around direct silencing of the promoters of the CXCL family of T\(_h\)1 chemokines\(^{45,46}\), our observation that EZH2 blockade likely generates dsRNA through SPARCS de-repression suggests that EZH2 may regulate the induction of CXCLs through additional pathways. In fact, our observation that H69 mesenchymal subclones do not display increased accessibility at the CXCL10 locus by ATAC-seq (Figure 3b) suggest that direct epigenetic effects at the loci of these cytokines are unlikely to fully explain the action of EZH2 in promoting local T cell responses. Going forward, we will aim to repeat these EZH2 inhibition experiments in a dual STING/MAVS double knockout cell line, to further confirm that induction of these cytokines is at least partially dependent on cytoplasmic nucleic acid sensing rather than direct epigenetic silencing of cytokine loci.

Apart from effects on the tumor microenvironment by influencing local cytokine secretion, the tumor cell-autonomous effects of EZH2 inhibition in the context of cancer immunotherapy remain less clear. Similar to the findings of other groups\(^{77}\), we have recapitulated the idea that EZH2 blockade may help promote tumor cell antigen presentation marked by upregulation of MHC class I members, β-2-microglobulin, and the TAP family of antigen transporters. Two possible mechanisms could underlie the development of this observed
immunogenicity following EZH2 inhibition. Although ERVs are thought to be highly mutated and are unlikely to produce functional proteins \(^{24, 26, 88}\), it is possible that some SPARCS RNAs are able to be translated into short peptides with viral homology and presented to the cell surface through class I MHC activity. To this end, understanding which SPARCS ERVs may encode functional open reading frames (ORFs), or exploring by mass spectrometric approaches whether H69 mesenchymal subclones (or other SPARCS-high/EZH2-low tumors) express viral protein signatures, could help explore whether SPARCS activation is associated with the production of potential virally-derived neoantigens which may promote response to T cell therapy.

On the other hand, perhaps a more likely explanation centers around our observation that EZH2 inhibition promotes endogenous tumor-derived IFN-\(\gamma\). Given that MAVS and STING activation are more classically associated with production of type I interferons (IFN-\(\alpha\) and IFN-\(\beta\)) \(^{30, 89}\), the mechanism underlying the production of IFN-\(\gamma\) in these conditions was not immediately clear to us. However, given that IFN-\(\gamma\) is known to upregulate MHC class I expression \(^{76}\), it is possible that direct production of IFN-\(\gamma\) from EZH2 inhibition in these condition is responsible for this increased immunogenicity. Repeating these experiments in the presence of a neutralizing antibody to IFN-\(\gamma\) may shed more light on this mechanism.

**Exploring how EZH2 inhibition may synergize with immune checkpoint blockade**

Nevertheless, whether EZH2 inhibition promotes immunogenicity within the tumor cell compartment or the tumor microenvironment, multiple groups have demonstrated in murine models that anti-EZH2 therapy may synergize with immune checkpoint blockade. We believe that SPARCS dsRNA induction may at least partly explain some of these observations, and so will aim to build on these findings. First, we will use 3D microfluidic co-culture to demonstrate
that H69 cells treated with EZH2 inhibition successfully attract Jurkat T cells and THP-1 monocytes, as has been demonstrated in H69 mesenchymal subclones. After establishing these behaviors, our goal will be to use a novel \textit{ex vivo} platform developed by the Barbie lab to study the interactions of patient-derived tumor cells and autologous immune cells in response to EZH2 inhibition. Understanding whether EZH2 blockade-induced endogenous retroviral expression in patient-derived tumor samples may synergize with either anti-PD-1 or anti-CTLA-4 therapies is an important goal and would further highlight the translational relevance of our study.

More broadly, our discovery that EZH2 regulates a dsRNA-STING/MAVS axis may have other important consequences for immunotherapy. Prior to our discovery that EZH2 regulates SPARCS expression, our group focused on understanding immune signaling pathways in H69 mesenchymal subclones given the known activation of TBK1 in this state. Multiple studies have highlighted how activation of MAVS and STING sensing can promote TBK1 activation and how STING agonism can promote responses to cancer immunotherapy. In concert with these findings, our group has also begun to explore how dual TBK1/IKKe inhibition can synergize with anti-PD-1 therapy in murine models potentially due to beneficial effects on the tumor microenvironment. Our discovery that EZH2 controls SPARCS expression and that EZH2 blockade can promote a PD-L1\textsuperscript{high}, STING\textsuperscript{high} state potentially through dsRNA induction and sensing helps shed light on potential transcriptional states of tumors which may be more responsive to TBK1-based immunotherapies. Understanding whether STING\textsuperscript{high} and pTBK1\textsuperscript{high} tumors display loss of EZH2 expression, and whether these mark potential responders to immunotherapy, is an important goal going forward.
Exploring SPARCS ERV physiology more broadly across tumors

Finally, a current limitation of our study is that we chose to explore EZH2 and H3K27Me3 primarily due to the association of the SPARCS-high state with histone methylation pathways by ssGSEA (Figure 1). However, there may be other chromatin regulators upstream of SPARCS expression, especially given our surprising observation that SPARCS expression was associated with loss of multiple other chromatin regulators, including SWI/SNF family members (Figures 12e-12i), as well as the H3K36 methyltransferase SETD2 (data not shown). The observation that alterations in SWI/SNF components may play strong roles in response to immunotherapy in renal cell carcinoma and that ERV expression may predict responsiveness in related groups of patients highlight the importance of rigorously identifying potential regulators of SPARCS expression in an unbiased manner. To this end, future work includes performing a genome-wide CRISPR screen followed by either single-cell RNA-seq to analyze SPARCS expression or flow cytometry-based readouts to assess cell surface SPARCS induction in order to either confirm EZH2 as the strongest regulator of SPARCS expression, or identify other core regulators which may be clinically actionable and complementary to immunotherapy. Indeed, preliminary evidence from our group suggesting that increased SPARCS expression promotes improved survival in melanoma (unpublished) highlights the importance of finding and validating novel SPARCS regulators across an array of tumor types.

Summary

Here we identify EZH2 as a core regulator of a novel subclass of endogenous retroviruses (SPARCS) which are primed to produce dsRNA in response to IFN-γ in mesenchymal subclones of small cell lung cancer. EZH2 inhibition induces SPARCS expression, secretion of key
cytokines linked to dsRNA sensing with demonstrated roles in response to immunotherapy, and improves tumor cell immunogenicity through induction of endogenous IFN-γ secretion and antigen presentation pathways. In addition, loss of EZH2 as well as additional chromatin regulators correlate with SPARCS expression across tumors. Taken together, these results highlight an important connection between epigenetic regulation of non-coding elements and tumor immune signaling which may have important translational implications for cancer immunotherapy.
References


Figures and Tables

Figure 1. Single sample GSEA reveals association of the SPARCS-high state with histone and DNA methylation pathways across TCGA RNA-seq.

<table>
<thead>
<tr>
<th>SPARCS Signature</th>
<th>IC</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUYTSEN EZH2 TARGETS UP</td>
<td>0.8</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>PASINI SUZ12 TARGETS DN</td>
<td>0.565</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>LIANG SILENCED BY METHYLATION</td>
<td>0.605</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>MIKKELSEN MEF LCP WITH H3K4ME3</td>
<td>0.701</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>JACKSON DNMT1 TARGETS UP</td>
<td>0.453</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

Figure 2. Genomic sequencing of H69 and mesenchymal subclones reveals no association with mutations or copy number alterations across a panel of oncogenes.
Figure 3. ATAC-seq reveals gain of accessible chromatin sites around SPARCS genes in H69 mesenchymal subclones, but no difference in accessibility at key SPARCS-associated genes.
Figure 4. H3K27Me3 ChIP-seq reveals loss of H3K27Me3 signal around SPARCS loci in H69 mesenchymal subclones.
Figure 5. H3K27Ac ChIP-seq reveals gain of H3K27Ac signal around SPARCS loci in H69 mesenchymal subclones, particularly in H69M.
Figure 6. EZH2 inhibition is more strongly linked to SPARCS ERV expression than DNMT inhibition.
Figure 7. EZH2 inhibition followed by IFN-γ treatment de-represses the same SPARCS loci in H69 cells that are activated in H69AR following IFN-γ treatment.

Figure 8. The synthetic dsRNA oligonucleotide poly(I:C) induces production of key cytokines, interferons, and immunosuppressive factors which are dependent on MAVS activity and JAK/STAT signaling.
Figure 9. Luminex cytokine profiling reveals strong concordance in cytokine production between H69 cells treated with EZH2 inhibition and H69 mesenchymal subclones in response to IFN-γ.
Figure 10. Additional EZH2 inhibitors induce SPARCS expression, PD-L1 expression, and SPARCS-linked cytokine production in H69 cells.

Figure 11. EZH2 inhibition induces antigen presentation and cytoplasmic nucleic acid sensing pathways.
Figure 12. Comparison of SPARCS-high and SPARCS-low cell lines in TCGA reveals key associations with nucleic acid sensors, cytokines, and epigenetic regulators.
<table>
<thead>
<tr>
<th>Gene</th>
<th>ERV</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRIM22</td>
<td>MLT1C49</td>
</tr>
<tr>
<td>TRIM38</td>
<td>MLT1A</td>
</tr>
<tr>
<td>TRIM38</td>
<td>MLT1J</td>
</tr>
<tr>
<td>IL32</td>
<td>THE1D</td>
</tr>
<tr>
<td>SPATS2L</td>
<td>MLT1K</td>
</tr>
<tr>
<td>EPHA3</td>
<td>LTR37A</td>
</tr>
<tr>
<td>HERC3</td>
<td>MSTB</td>
</tr>
<tr>
<td>ADAM19</td>
<td>MER92B</td>
</tr>
<tr>
<td>SERPINB9</td>
<td>MER50</td>
</tr>
<tr>
<td>IFI44L</td>
<td>LTR26</td>
</tr>
<tr>
<td>F3</td>
<td>MLT1I</td>
</tr>
<tr>
<td>BEND6</td>
<td>PABL_B</td>
</tr>
<tr>
<td>AIG1</td>
<td>MLT1A0</td>
</tr>
<tr>
<td>MSRB2</td>
<td>MSTA</td>
</tr>
<tr>
<td>TNFRSF9</td>
<td>MLT1I</td>
</tr>
<tr>
<td>ANTXR1</td>
<td>LTR79</td>
</tr>
</tbody>
</table>

Table 1. List of SPARCS genes and corresponding 3’ UTR ERVs.
Figure Legends

Figure 1. Single sample GSEA reveals association of the SPARCS-high state with histone and DNA methylation pathways across TCGA RNA-seq.
Among gene sets involving epigenetic pathways, ranking of TCGA tumor samples from high to low SPARCS expression reveals association of SPARCS with components of the transcriptional repressive complex PRC2 including the histone methyltransferase EZH2 and the polycomb protein SUZ12. SPARCS is also associated with targets of DNA methyltransferases, specifically DNMT1.

Figure 2. Genomic sequencing of H69 and mesenchymal subclones reveals no association with mutations or copy number alterations across a panel of oncogenes.
OncoPanel Assay of 300 cancer genes and 113 introns across 35 genes for mutations and copy number detection was performed for the H69 parental cell line, H69M-PD-L1 high, and H69M-PD-L1 low subclones. Mutation and copy number profiles showed no significant differences between the three cell lines.

Figure 3. ATAC-seq reveals gain of accessible chromatin sites around SPARCS genes in H69 mesenchymal subclones, but no difference in accessibility at key SPARCS-associated genes.
ATAC-seq was performed in basal conditions on H69, H69AR, and H69M cells. Arrows indicate ATAC-seq peaks present in mesenchymal subclones not present in parental H69. (A) ATAC-seq around SPARCS loci reveals gain of accessible sites around a majority of SPARCS loci in
H69AR and gain of accessibility at multiple SPARCS in H69M. (B) ATAC-seq around the loci of key genes associated with the SPARCS-high state, including CD274 (PD-L1), CXCL10, and CCL2, reveal no gain of accessible sites in H69 mesenchymal subclones. ATAC-seq tracks visualized on IGV genome browser software.

**Figure 4.** H3K27Me3 ChIP-seq reveals loss of H3K27Me3 signal around SPARCS loci in H69 mesenchymal subclones.

H3K27Me3 ChIP-seq was performed in basal conditions on H69, H69AR, and H69M cells, showing loss of signal in mesenchymal subclones particularly in H69M. ChIP-seq tracks were visualized on IGV genome browser software.

**Figure 5.** H3K27Ac ChIP-seq reveals gain of H3K27Ac signal around SPARCS loci in H69 mesenchymal subclones, particularly in H69M.

H3K27Ac ChIP-seq was performed in basal conditions on H69, H69AR, and H69M cells, showing gain of signal in mesenchymal subclones, particularly in H69M. ChIP-seq tracks visualized on IGV genome browser software.

**Figure 6.** EZH2 inhibition is more strongly linked to SPARCS ERV expression than DNMT inhibition.

(A) Western blot of EZH2 expression in H69, H69M, and H69AR cells in basal conditions. (B) Western blot of total H3K27Me3 levels in H69 cells treated with 2, 5, and 10 µM GSK126 for 6 days. (C) qPCR of ERV expression from SPARCS and non-SPARCS genes in H69 cells treated with 100 nM DAC or 5 µM GSK126 for 6 days followed by 10 min 200 ng/ml IFN-γ pulse and
recovery after 24 hours. (D) Plot of relative expression of SPARCS genes over DMSO from H69 cells treated with 100 nM DAC or 5 µM GSK126 for 6 days followed by ctrl treatment for 24 hours. Statistics indicate mean relative expression of genes. (E) Plot of relative expression of non-SPARCS genes over DMSO from H69 cells treated with 100 nM DAC or 5 µM GSK126 for 6 days followed by ctrl treatment for 24 hours. (F) Plot of relative expression of SPARCS genes over DMSO from H69 cells treated with 100 nM DAC or 5 µM GSK126 for 6 days followed by 10 min 200 ng/ml IFN-γ pulse and recovery after 24 hours. (G) Plot of relative expression of non-SPARCS genes over DMSO from H69 cells treated with 100 nM DAC or 5 µM GSK126 for 6 days followed by 10 min 200 ng/ml IFN-γ pulse and recovery after 24 hours. (H) Light microscopy images of adherent fraction of H69 cells treated with either DMSO or 2 µM GSK126 for 12 days followed by ctrl 10 min 200 ng/ml IFN-γ pulse, with pictures taken after 24 hours. (I) CXCL10 ELISA from H69 cells treated with 100 nM DAC or 5 µM GSK126 for 6 days followed by ctrl treatment or 10 min 200 ng/ml IFN-γ pulse and recovery after 24 hours. Bar graph in D corresponds to the plots in G and H. n.s. = p > 0.05, * = p < 0.05, ** = p < 0.01, *** = p < 0.001.

**Figure 7. EZH2 inhibition followed by IFN-γ treatment de-represses the same SPARCS loci in H69 cells that are activated in H69AR following IFN-γ treatment.**

(A) qPCR of relative expression compared to DMSO 6d/ctrl 24h of SPARCS genes for H69 cells treated with either DMSO or 5 µM GSK126 for 6 days followed by 10 min 200 ng/ml IFN-γ pulse and recovery after 24 hours. (B) qPCR of relative expression compared to ctrl 24h within each cell line of SPARCS genes for H69 or H69AR cells treated with 10 min 200 ng/ml IFN-γ
pulse and recovery after 24 hours. (C) Log₂ fold change comparison of IFN-γ-induced expression of SPARCS genes in EZH2i treated H69 cells (Fig. 7a) versus H69AR cells (Fig 7b).

Figure 8. The synthetic dsRNA oligonucleotide poly(I:C) induces production of key cytokines, interferons, and immunosuppressive factors which are dependent on MAVS activity and JAK/STAT signaling.

(A) qPCR of CCL2, CXCL10, IFN-γ, IFN-β, and PD-L1 genes from H69AR sgCTRL and H69AR sgMAVS cells treated with 0.5 µg/ml HMW poly(I:C) for 72 hours. (B) CXCL10 ELISA from same conditions as Fig. 8a. (C) Luminex cytokine profiling of H69AR sgCTRL and sgMAVS cells treated with 10 min 200 ng/ml IFN-γ pulse and recovery after 72 hours. (D) qPCR of CXCL10, CCL2, IFN-β, IFN-γ, and PD-L1 genes from H69AR cells treated with 0.5 µg/ml HMW poly(I:C) and either DMSO, 1 µM MRT67307, or 1 µM ruxolitinib for 72 hours. (E) CXCL10 ELISA from same conditions as Fig. 8d. (F) Luminex cytokine profiling of H69AR cells treated with 10 min 200 ng/ml IFN-γ pulse and recovery after 72 hours in DMSO or 1 µM ruxolitinib conditions.

Figure 9. Luminex cytokine profiling reveals strong concordance in cytokine production between H69 cells treated with EZH2 inhibition and H69 mesenchymal subclones in response to IFN-γ.

(A) Luminex cytokine profiling of H69 cells treated with EZH2i (5 µM GSK126 for 6 days followed by ctrl treatment for 24h), EZH2i+IFN-γ (5 µM GSK126 for 6 days followed by 10 min 200 ng/ml IFN-γ pulse and recovery after 24h), or IFN-γ alone (DMSO for 6 days followed by 10 min 200 ng/ml IFN-γ pulse and recovery after 24h). Heat map shows log₂ fold change (L2FC)
of cytokine levels compared to DMSO control (DMSO for 6 days followed by ctrl treatment for 24h). (B) Luminex cytokine profiling of H69 and H69AR cells treated with 10 min 200 ng/ml IFN-γ pulse and recovery after 24h. Heat map shows L2FC cytokine levels compared to 24h ctrl treatment within each respective cell line. (C) Venn diagram displaying differential cytokine production between H69M-PD-L1<sup>high</sup> (H69M+) subclone and H69 in basal conditions (data not shown), between H69AR and H69 in IFN-γ pulse conditions (Fig. 9b), and between H69 EZH2i+IFN-γ and IFN-γ only conditions (Fig. 9a).

**Figure 10. Additional EZH2 inhibitors induce SPARCS expression, PD-L1 expression, and SPARCS-linked cytokine production in H69 cells.**

(A) qPCR indicating relative expression of the SPARCS ERVs MLT1C49 and MLT1K and the PD-L1 gene in H69 cells treated with DMSO, 250 nM, 1 μM, 2 μM, and 5 μM EPZ6438 for 6 days followed by 10 min 200 ng/ml IFN-γ pulse and recovery after 24h. (B) qPCR indicating relative expression of the SPARCS ERVs MLT1C49 and MLT1K and the PD-L1 gene in H69 cells treated with DMSO, 250 nM, 1 μM, 2 μM, 5 μM, and 10 μM GSK343 for 6 days followed by 10 min 200 ng/ml IFN-γ pulse and recovery after 24h. (C) CCL2 and CXCL10 ELISA from same conditions in Fig. 10a and Fig. 10b.

**Figure 11. EZH2 inhibition induces antigen presentation and cytoplasmic nucleic acid sensing pathways.**

(A) qPCR of relative expression of HLA-A, HLA-B, DDX58, DHX58, TAP2, IRF4, IRF9, TMEM173 (STING), and B2M genes from H69 cells pre-treated with either DMSO or 5 μM GSK126 followed by 10 min 200 ng/ml IFN-γ pulse and recovery after 24 hours compared to
DMSO 6d + ctrl 24h samples. (B) Western blot of STING expression in H69 cells pre-treated with DMSO or 2, 5, or 10 µM GSK126 for 6 days followed by 1 or 3 additional days in culture.

**Figure 12. Comparison of SPARCS-high and SPARCS-low cell lines in TCGA reveals key associations with nucleic acid sensors, cytokines, and epigenetic regulators.**

Plots of TCGA RPKM values of (A) TMEM173 (STING), (B) CXCL10, (C) CCL2, (D) EZH2, (E) PBRM1, (F) ARID2, (G) ARID1A, (H) ARID1B, and (I) SMARCA4 from RNA-seq data in primary tumors grouped in SPARCS\textsuperscript{high} (n=50) and SPARCS\textsuperscript{low} (n=50) categories. (J) CXCL10 ELISA from 786-O wild-type (sgCTRL) and 786-O sgBAF180 cells treated with 10 min 200 ng/ml IFN-γ pulse and recovery after 24h. (K) qPCR of SPARCS expression from same conditions as Fig. 12j.