JCI The Journal of Clinical Investigation

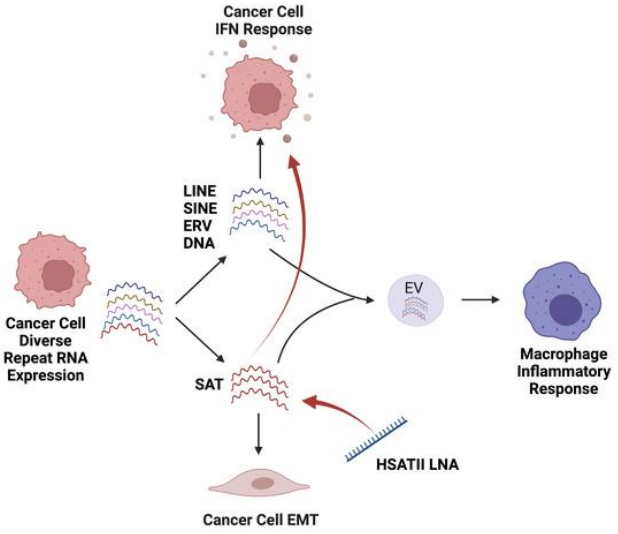
Satellite repeat RNA expression in epithelial ovarian cancer associates with a tumor immunosuppressive phenotype

Rebecca L. Porter, … , Benjamin D. Greenbaum, David T. Ting

J Clin Invest. 2022. https://doi.org/10.1172/JCI155931.

Research In-Press Preview Oncology

Graphical abstract



Satellite repeat RNA expression in epithelial ovarian cancer associates with a tumor immunosuppressive phenotype

Rebecca L. Porter[1,2,3,*], Siyu Sun[4,*], Micayla N. Flores[1], Emily Berzolla[1], Eunae You[1], Ildiko E. Phillips[1], Neelima KC[1], Niyati Desai[1], Eric C. Tai[1], Annamaria Szabolcs[1], Evan Lang[1], Amaya Pankaj[1,5], Michael J. Raabe[1], Vishal Thapar[1], Katherine H. Xu[1], Linda T. Nieman[1], Daniel Rabe[1], David Kolin[6], Elizabeth Stover[3], David Pepin[5], Shannon L. Stott[1], Vikram Deshpande[7], Joyce F. Liu[3], Alexander Solovyov[4], Ursula A. Matulonis[3], Benjamin D. Greenbaum[4,#], David T. Ting[1,2,#]

[1] Mass General Cancer Center, Harvard Medical School; Charlestown, MA, USA.
[2] Department of Medicine, Massachusetts General Hospital, Harvard Medical School; Boston, MA, USA.
[3] Division of Gynecologic Oncology, Dana-Farber Cancer Institute, Boston, MA
[4] Computational Oncology, Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center; New York, NY, USA.
[5] Department of Surgery, Massachusetts General Hospital Harvard Medical School; Boston, MA, USA.
[6] Department of Pathology, Brigham and Women's Hospital, Harvard Medical School; Boston, MA, USA.
[7] Department of Pathology, Massachusetts General Hospital, Harvard Medical School; Boston, MA, USA.

* These authors contributed equally to this work

Address correspondence to
Benjamin Greenbaum, PhD (greenbaB@mskcc.org)
Associate Attending, Computational Oncology
Program Director of Computational Immuno-Oncology
Department of Epidemiology & Biostatistics
Memorial Sloan Kettering Cancer Center
321 E 61st Street, Room 256
New York, NY 10065
(646) 608-7667

David T. Ting, MD (dting1@mgh.harvard.edu)
Associate Clinical Director for Innovation
Massachusetts General Hospital Cancer Center
Associate Professor of Medicine
Harvard Medical School
Building 149, Thirteenth Street, Rm 6-6003
Charlestown, Massachusetts 02129
Tel: 617-240-9402 Fax: 617 724-3676

ABSTRACT

Aberrant expression of viral-like repeat elements is a common feature in epithelial cancers, but the substantial diversity of repeat species provides a distinct view of the cancer transcriptome. Repeatome profiling across ovarian, pancreatic, and colorectal cell lines identifies distinct clustering that is independent of tissue of origin that is seen with coding gene analysis. Deeper analysis of ovarian cancer cell lines demonstrated that HSATII satellite repeat expression was highly associated with epithelial mesenchymal transition (EMT) and anti-correlated with interferon (IFN) response genes indicative of a more aggressive phenotype. This relationship of HSATII with high EMT and low IFN response genes was also found in RNA-seq of primary ovarian cancers and associated with significantly shorter survival in a second independent cohort of ovarian cancer patients. Repeat RNAs were also found enriched in tumor derived extracellular vesicles that were capable of stimulating monocyte derived macrophages demonstrating a mechanism of altering the tumor microenvironment with these viral-like sequences. Targeting of HSATII with anti-sense locked nucleic acids (LNAs) stimulated IFN response and induced MHC I expression in ovarian cancer cells lines, highlighting a potential strategy of modulating the repeatome to re-establish anti-tumor cell immune surveillance.

GRAPHICAL ABSTRACT

Repeatome transcriptional profiling of ovarian cancers identify HSATII satellite repeats linked with epithelial mesenchymal transition and innate immune suppression that can be reversed with repeat specific anti-sense nucleic acids.

INTRODUCTION

Repetitive elements make up ~50% of the human genome (1, 2), and their aberrant expression has been described across a wide range of cancers (3-5). Repeat suppression is achieved through a combination of epigenetic modifications (6) and the activity of tumor suppressors including TP53 (7, 8). The loss or mutation of these guardians of the epigenomic and genome leads to de-repression of repeat RNAs that are sensed by pattern recognition receptors and trigger an innate immune system-mediated interferon response analogous to anti-viral host response (8-16). However, the repeatome is diverse and exhibits differences in cellular and immunological response based on repeat RNA sequence motifs (17) and variation in co-expression of distinct clusters of repeats (5, 18, 19). In addition, some repeats (HERV) that have been associated with response to immune checkpoint inhibitors (19-21), while other repeats are associated with immune-depleted tumor microenvironments (19, 22). Collectively, these studies link different cancer cell intrinsic immune responses with expression of specific classes of repeat RNAs. To understand the relationship between repeat RNAs and immune response, we have focused on epithelial ovarian cancers (EOCs) given the role of TP53 in regulating repeat elements (7, 8) and the high rates of *TP53* mutations in EOC (23).

RESULTS

Total RNA-seq profiling reveals distinct clusters of repeat RNA expression across different epithelial cancers

In order to comprehensively define the expression of repeat RNAs in EOC and compare this with expression in other cancers, we applied our previously established computational alignment methods for total RNA-seq (19) to 31 patient-derived low passage EOC, 17 commercially available ovarian, 26 pancreatic ductal adenocarcinoma (PDAC), and 11 colorectal cancer (CRC) cell lines (**Figure 1A**). In EOC models, this confirmed high contribution of non-coding transcripts to the total transcriptome (**Figure 1B**) with all major subclasses of repeats represented (**Figure 1C**). Expression levels of individual repeat RNAs varied, with some repeat RNAs (e.g. L1HS and HERVH) expressed at high levels comparable to those of traditional housekeeping genes like *ACTB* and *GAPDH* (**Supplemental Figure 1A**). Across EOC, PDAC and CRC, clustering of cell lines by coding genes segregated the samples by cancer type with high accuracy (**Figure 1D**). As expected given unique transcriptional programs associated with different cancer types, major clusters comprised of nearly entirely EOC (C1 and C5), CRC (C3) and PDAC (C4) emerged, with one additional cluster comprised of samples of all three cancer types (C2). Notably, clustering by repeat RNA expression alone was able to similarly distinguish between cancer types with few exceptions (**Figure 1E**), suggesting that despite common overall repeat dysregulation across epithelial cancers, some repeat RNA species are cancer type-specific and may have important biological roles or consequences in these tumors.

Satellite repeat RNAs cluster distinctly from other repeat elements and display variable expression across cancer models

To identify subclasses of repeat RNAs with biological relevance across tissue types, consensus clustering analysis of repetitive elements across all cell line samples was performed (see Methods). We detected five distinct clusters of co-expressed repetitive elements with Cluster 2, denoted by the red asterisk, demonstrating the strongest consensus correlation across samples (**Figure 2A and Supplemental Figure S1B**). Subclass analysis revealed an enrichment for satellite (SAT) repeats in Cluster 2 (**Figure 2B**). Notably, SAT expression was found to be highly variable in EOC cell lines, with SAT RNAs representing the highest proportion of the 50 most variant transcripts (**Figure 2C**). In line with this, clustering of consensus expression profiles for each major subclass of repeats highlighted the distinct expression patterns of the SAT subclass across cell lines (**Figure 2D**). Further, hierarchical clustering of EOC, PDAC, and CRC cell lines based on SAT RNA expression demonstrated unique clustering of samples that was not driven by tissue of origin (**Figure 2E**). This indicated that SAT RNA expression patterns could have shared transcriptional programs across diverse epithelial cancers.

Satellite repeat RNA expression is linked with an immunosuppressive epithelial-mesenchymal transition gene expression pattern in EOC

To better characterize the relationship of repeatome profiles with coding gene behavior in EOC, we first applied gene set enrichment analysis (GSEA) with the HALLMARK gene set from the Broad Molecular Signature Database (MSigDB) (24, 25) to a gene list ranked based on the correlation with the consensus expression calculated for each repeat subclass across EOC cell lines. This demonstrated high positive

correlation of SAT repeats with the epithelial mesenchymal transition (EMT) gene set and anti-correlation with several immune and interferon response sets including the Interferon alpha (IFNα) and Interferon gamma (IFNγ) gene sets (**Figure 3A**). A parallel analysis of separating EOC cell lines into SAT-high and SAT-low based on median consensus expression also identified enrichment for the HALLMARK EMT gene set in SAT-high and IFNα, IFNγ and Inflammatory response gene sets in SAT-low cell lines (**Supplemental Figure 2A**), further validating these associations.

To futher investigate this observation, hierarchical clustering of EOC models by consensus expression of repeat RNA subclasses was performed, which separated EOC cell lines into three major clusters as depicted in **Figure 3B.** Repeat High (Rep-H) cell lines displayed high expression of all subclasses while Repeat Low (Rep-L) cell lines had relatively low repeat RNA expression in general. A third distinct cluster also emerged which exhibited high expression of all subclasses of repeat RNAs except for SAT RNAs, termed “SAT-deplete” (SAT-D). To further characterize the specific contribution of SAT repeats specifically, GSEA was performed between Rep-H and SAT-D EOC cell lines (**Supplemental Figure 2B**) and again demonstated enrichment of EMT-related genes and downregulation of genes related to innate immune and IFN response pathways in the Rep-H (including SATs) compared with SAT-D (lacking SATs) cell lines, confirming the association observed in the total cohort when analyzed by correlation with SAT expression (**Figure 3A, Supplemental Figure 2A**). Rep-L EOC had higher enrichment of cell cycle and replication related pathways (HALLMARK E2F, G2M, MYC Targets) indicating an anti-correlation of repeat expression with mitotic activity. Collectively, this refined repeat subtyping identifies unique characteristics

including high EMT expression in cell lines with high SAT expression (Rep-H), activation of IFN response genes in repeat-high without SAT (SAT-D), and high proliferative activity in repeat-low cell lines (Rep-L).

HSATII is a representative SAT repeat RNA that correlates with worsened clinical outcomes in human EOC

In order to further investigate the biological implications of high SAT expression in EOC, we selected the human satellite II (HSATII) as a representative repeat species within the SAT subclass, which we had previously found enriched across epithelial cancers (3). As expected, HSATII expression was significantly higher in Rep-H cell lines (**Supplemental Figure 3A**), which was validated in a subset of cell lines by RNA in situ hybridization (RNA-ISH) (**Figure 3C**). Consensus clustering analysis of repetitive elements across EOC samples repeated with HSATII removed (**Figure 3D**, **Supplemental Figure 3B**) yielded a similar SAT-driven cluster displaying the strongest consensus correlation (red asterisk), implying that HSATII is not the sole driver of the SAT subclass, but instead is a representative member. In line with this, differential gene expression analysis between HSATII-low and HSATII-high EOC cell lines to determine coding gene expression patterns linked with HSATII expression revelaed upregulation of genes related to EMT and down-regulation of genes related to IFN response and inflammatory pathways in HSATII-high samples (**Figure 3E and 3F**), similar to results from these comparisons based on total SAT expression (**Supplemental Figure 2A**).

To interrogate the association of SAT repeats with transcriptional programs in patients, we investigated patterns of HSATII expression in total RNA-seq from a cohort of 96 human primary ovarian carcinomas (26). Similar to EOC cell lines, tumors with high HSATII expression (**Supplemental Figure 3C**) demonstrated upregulation of genes related to EMT and downregulation of genes in the IFNα, IFNγ, and inflammatory pathways (**Figure 4A and 4B**) compared with those with low HSATII expression. We then performed HSATII RNA-ISH with quantitative image analysis in a separate cohort of patients with advanced HGSOC from the Dana-Farber Cancer Institute to segregate primary tumors into low or high HSATII expression (**Figure 4C**). Notably, separating tumors by HSATII expression revealed significantly shorter overall survival in HSATII-high tumors (**Figure 4D**). Taken together, this work highlights that repeats are a diverse set of RNA species; some are associated with tumor cell IFN response (9, 10, 12, 13, 27, 28) while others, such as SAT repeats, are associated with EMT and low IFN signaling that is typically seen in more aggressive tumors.

Given the correlation of HSATII with low IFN response, high EMT, and worsened survival, we next evaluated the relationship between HSATII expression and the immune microenvironment in ovarian cancer. Cellular deconvolution analysis of the 96 total RNA-seq ovarian tumors using xCell (29) was performed to estimate percentages of specific immune populations and then calculate their correlation with HSATII expression (**Figure 5A and Supplemental Figure 5A**). Immune cells positively correlated with HSATII included immature dendritic cells (iDC), regulatory T-cells (Treg), and myeloid cells (monocytes, macrophages, neutrophils) indicating an immune microenvironment dominated by innate immune cells. Given our prior work

demonstrating that some non-coding RNAs expressed in cancer cells can directly activate cells of the mononuclear phagocytic system (17), we hypothesized that extracellular vesicles (EVs) could serve as a vehicle to deliver HSATII and other repeats with the ability to modulate innate immune cells within the tumor microenvironment. To test this, we first collected EVs released by PDAC and EOC cell lines and confirmed that isolated EVs expressed typical EV-associated cell surface markers (**Supplemental Figure 5B**). RNA was then purified from tumor cell-derived EVs and subjected to total RNA-Seq. Compared with the RNA profile of each parental cell line, a robust enrichment of a diverse set of repeat RNAs was detected in EVs isolated from each cell line (**Figure 5B**) with HSATII being one of the most prevalent RNAs (**Figure 5C**). To then test the effect of HSATII-enriched EVs on human myeloid cells, we purified EOC-derived EVs and applied them to flow cytometry-sorted $CD14^+$ peripheral blood mononuclear cells (PBMCs) collected from healthy donor human subjects (**Figure 6A**). $CD14^+$ PBMCs exposed to EOC EVs demonstrated upregulation of genes related to the activation of the innate immune and interferon responses (HALLMARK IFNα, IFNγ, and Inflammatory Response) compared with unexposed CD14+ cells (**Figure 6, B and C**). A similar activation of genes within these pathways was observed in response to both PDAC and EOC tumor cell-derived EVs and in CD14+ cells from multiple individual healthy donors in separate experiments (**Supplemental Figure 5C**), suggesting a common response of monocyte derived cells to repeat RNA-enriched EVs in the tumor microenvironment.

Given the presence of multiple classes of repeat RNAs in EOC-derived EVs (**Figure 5B**) and the distinct effects on tumor and immune cells conferred by different

repeat RNA species, we sought to determine if HSATII specifically can stimulate myeloid cells. To test this, we used THP-1 monocytic cells to compare the relative response of these innate immune cells to EOC EVs and in vitro transcribed HSATII RNA. THP-1 cells treated with EOC EVs from two different cell lines had significant induction of IFN response genes including *DHX58*, *IFNB1*, *ISG15*, *OAS2*, *MX1*, *MX2*, and *IFI44* as measured by quantitative RT-PCR (**Figure 6D**). HSATII compared to GFP RNA transfection significantly induced expression of *IFNB1, OAS2, ISG15, MX1, MX2, and IFI44*, which indicates that HSATII RNA in EVs partially contributes to the IFN response seen by EOC EVs in monocyte derived cells (**Figure 6E**). This suggests that HSATII is sensed by and can generate an IFN response in immune cells that are enriched in the tumor microenvironment of HSATII-high tumors.

Modulating the repeatome with epigenetic drugs or repeat specific anti-sense oligos have diverse effects in EOC

Although we have shown that repeat RNAs can be transmitted to responding innate immune cells and drive an IFN response, our collective analysis in EOC cell line models and tumors indicated that tumor cells with high baseline levels of SAT RNAs lack IFN pathway activation, implying that they have developed an adaptation to suppress the IFN response to repeats. We therefore hypothesized that modulating different repeats in tumor cells may overcome this repeatome tolerance. Repetitive elements are known to be suppressed in the normal genome in part by DNA and histone methylation (7, 8) and epigenetic therapies have been shown to induce transcription of some repeat species (9, 12, 13) in ovarian cancer models. Thus, we

first tested the effect of DNA methyltansferase inhibitor (DNMTi - 5-azacytidine 500 nM) and histone deacetylase inhibitor (HDACi - Trichostatin A 250 nM) treatment of EOC cell lines. As expected, these drugs induced broad changes in repeat element expression, but there were notable differences with DNMTi promoting a greater induction of ERV, SINE and LINE elements and HDACi with consistent increased SAT elements across cell lines (**Supplementary Figure 5A,B**). Analysis of coding genes induced by these agents revealed enrichment of IFN response gene expression in cell lines treated with DNMTi whereas EMT pathway gene enrichment was appreciated in cell lines treated with HDACi (**Supplemental Figure 5C,D**), which was consistent with co-expression patterns of these distinct repeat subsets in our EOC cell lines (**Figure 3A**) and tumors (**Figure 4A and 4B**). These findings suggest that DNA methylation and histone acetylation have different contributions to the regulation of the repeatome profile in EOC, and importantly, the response to these drugs can have discordant pro- and anti- tumoral effects on cancer cells.

Given the consistent relationship between SAT repeat expression with EMT-high and IFN-low phenotype, we pursued direct targeting of the HSATII-specific locked nucleic acids (LNA) as an anti-sense oligo therapeutic. HSATII LNAs and control scramble LNAs were transfected into EOC cell lines followed by total RNA-Seq analysis at various times post-transfection, which revealed a specific and marked increase in HSATII RNA in the cells, peaking on days 2-3 (**Figure 7A**), with minimal off target effects on other repeat RNA species. Analysis of the coding gene transcripts in HSATII LNA-transfected cells over time revealed an upregulation of innate immune response genes and interferon-stimulated genes, indicating that HSATII LNAs could target

cancer-specific HSATII RNA and trigger an IFN response (**Figure 7B**). In addition, EOC cells grown in non-adherent culture following HSATII LNA transfection consistently demonstrated significant reduction in tumor sphere growth, and in the case of OVSAHO increased cell death, compared with cells transfected with control LNA (**Figure 7C**). Further investigation into the immune-related transcriptional changes in HSATII LNA-transfected tumor cells also revealed alterations in expression of genes related to MHC Class I antigen presentation. Similar to the anti-correlation observed between steady-state HSATII levels and innate immune and interferon response genes, we find that EOC cell lines (**Supplemental Figure 6A**) and primary human EOC tumors (**Supplemental Figure 6B**) with higher HSATII RNA levels at baseline have decreased expression of MHC Class I-related genes. However, EOC cells transfected with HSATII LNA revealed a striking upregulation of these MHC Class I-related genes (**Figure 6D**). Further, HSATII LNA-transfected EOC cell lines also demonstrated an increase in MHC Class I proteins on the cell surface compared with control LNA-transfected cells (**Figure 6E**). Taken together, increasing HSATII RNA levels with targeted LNA induces an interferon response associated with EOC tumor cell cytotoxicity and upregulation of MHC genes, which suggest the possibility that HSATII RNA modulation could sensitize EOC tumor cells to immunotherapy strategies.

DISCUSSION:

Repeat RNAs are commonly expressed in EOC yet despite their immunogenic potential, immune checkpoint inhibitors have had only modest activity in EOC (30-33). This apparent paradox reflects the marked diversity of repeat RNA species that

comprise the repeatome. Here, we demonstrate that various repeat species are co-regulated in distinct clusters which exhibit diverse expression patterns across epithelial cancers and likely reflect the inherent differences in tissue of origin and the genetic mutation background of each tumor. Furthermore, consensus clustering of repeat RNAs revealed a distinct pattern of co-expression of SAT RNAs as well as a high level of variation in SAT expression levels seen across EOC samples. Given this, we honed in on SAT repeats in EOC and strove to determine their patterns of expression and the tumor cell phenotypes and tumor microenvironmental characteristics associated with these patterns.

Although some repeat species, such as ERVs and LINE elements, can activate interferon signaling (9, 12, 13) and SAT repeats are able to stimulate immune cells (17), our current work reveals lower tumor cell intrinsic interferon pathway activation in EOC models with high SAT expression. This suggests that tumor cells with high steady state SAT repeat expression have the ability to tolerate this by suppressed interferon activation mechanisms that merit further investigation. Conversely, this work also unveiled the striking link between expression of SAT repeats, including HSATII, and an EMT transcriptional program in tumor cells. This is consistent with mounting evidence that the EMT state in tumor cells is characterized by an immunosuppressed phenotype (34). Further, HSATII-high cell lines also exhibit suppressed levels of MHC Class I (MHC-I) gene expression, a phenomenon also linked to the EMT state of a cell. In murine mammary carcinoma models, epithelial cells have high MHC-I expression whereas mesenchymal cells express low levels of MHC-I (35). This relationship was

also observed in lung cancer cell lines, where metastatic lines that had undergone EMT had lower expression of MHC-I genes than non-metastatic cell lines (36).

Our current work suggests that EMT activation in SAT-expressing cancers likely promotes a tumor cell intrinsic immunosuppression, possibly creating a permissive state in which tumor cells are tolerant to high levels of immunogenic repeats and resistant to immunotherapy. Indeed, prior transcriptional profiling of EOC has identified 4 molecular subtypes including mesenchymal, immunoreactive, proliferative, and differentiated, with worsened overall survival observed in the mesenchymal subtype (37). Consistent with this association, higher HSATII in primary ovarian tumors was found to be correlated with worsened outcomes in a cohort of ovarian cancer patients in our study, analogous to our prior observations linking increased HSATII copy numbers with lower survival rates in colorectal cancer (38).

Interestingly, a positive association with the presence of innate immune cells in primary tumor RNA-seq profiles is revealed in HSATII-high models. In particular, monocytes, macrophages, and iDCs were enriched in HSATII-high ovarian cancers, suggesting a potential functional relationship between HSATII and these innate immune cells in the tumor microenvironment. Similarly, a relationship between EMT and immunosuppressive tumor associated macrophages has been reported in many cancers including EOC (34, 39). Our prior work showed that HSATII uniquely displays pathogen-associated CpG motifs and that monocyte derived DCs and bone marrow derived macrophages can be stimulated with transfection of HSATII RNA in a CpG dependent manner (17). Taken together, this suggests a potential role for SAT RNAs in establishing or maintaining an immunosuppressive tumor microenvironment in EOC.

While further mechanistic studies are required to conclude this, our current findings that repeat RNA-enriched tumor-derived EVs can induce interferon response genes in human primary monocytic cells and a macrophage cell line suggests a direct mechanism for repeat RNA modulation of the tumor immune microenvironment. Although many different repeat RNA species are present in these EVs, including ERVs which are known to stimulate IFN responses, we also demonstrated the ability of HSATII RNA to directly stimulate IFN response genes in macrophages, implying their distinct role in modulating the tumor immune microenvironment. This is complementary to prior work showing HSATII enrichment in pancreatic cancer EVs (40) and more recent work demonstrating an association of repeat RNA expression with IFN response in fibroblasts in pancreatic cancer (22). A similar link between repeat-enriched EVs and alterations in the immune microenvironment of Ewing's Sarcoma has also been reported (41). Altogether, these collective studies imply that specific repeat RNA species are able to induce distinct transcriptional responses in tumor and microenvironmental cells, while SAT repeat expression leads to EMT and an immunosuppressed phenotype of tumor cells, when released outside the tumor cell they can induce a secondary IFN response in myeloid cells that generates a tumor permissive microenvironment (**Figure 8**). The strong correlation between HSATII RNA levels and clinical outcomes readily apparent in even a small cohort of ovarian cancer patients from our institution also highlights the potential of repeat RNA species as prognostic biomarkers in ovarian and likely other cancers.

Beyond this, our work also highlights the potential therapeutic opportunities associated with modulation of repeat RNAs from a tumor cell intrinsic perspective.

Although HSATII-high cell lines were found to exhibit a transcriptional profile consistent with suppression of the innate immune and IFN response pathways, we find that acutely elevating levels of HSATII RNA with HSATII-specific LNAs in the cell can overcome this suppression and induce significant cytotoxicity. While SAT RNAs were also upregulated in EOC cell lines treated with HDACi, this method of manipulation did not result in concomitant IFN pathway activation. This is likely a reflection of the HSATII specificity of an LNA compared to the broad effects of HDACi on SAT RNAs and other coding and non-coding RNAs. We note that TSA is a generalized HDAC inhibitor and more specific HDAC class I and class II inhibitors merit further investigation. Interestingly, most LNAs are thought to decrease target RNA through RNase H mediated degradation; however, our particular LNA design led to highly specific and robust elevation of HSATII RNA presumably through stabilization of HSATII RNA species and inhibition of reverse transcriptional machinery (38) or other undetermined repeat RNA processing proteins. In addition, HSATII perturbation was found to result in upregulation of MHC Class I genes and PD-L1 on tumor cells. Given that lack of MHC Class I expression is a common mechanism of resistance to immune checkpoint blockade (42-44), these findings raise the possibility that in an *in vivo* setting, SAT repeat RNA modulation could synergize with ICB by re-sensitizing cells to anti-PD-1 agents (**Figure 8**), similar to other strategies to re-express MHC Class I in tumor models (45). This is particularly intriguing in EOC, which is a disease that is known to have cytotoxic T lymphocytes present in the tumor (46) yet displays very low response rates to single agent checkpoint inhibitors (30, 31, 33, 47). Thus, strategies to render EOC tumors more susceptible to ICB while simultaneously limiting immune-related toxicities are critically needed, and further

investigation of repeat RNA modulation is warranted. Moreover, the demonstrated presence of HSATII and other SAT RNAs in tumor cell-derived EVs indicates the potential of SAT RNAs to be developed as blood-based prognostic and predictive biomarkers in EOC. Overall, in the context of a growing interest in enabling technologies for nucleic acid therapeutics (48), our current findings highlight a potential to translate the exponentially increasing understanding of repeat element dysregulation in cancer to clinical use, especially in cancers where immunotherapy has proven only modestly active.

METHODS

Cell Lines

EOC cell lines:

CAOV-4, IGROV1, JHOS-4, OAW28, OV90, OVCAR4, OVCAR8 and OVKATE were generous gifts from Dr. Cyril Benes (Massachusetts General Hospital Cancer Center, Boston, MA). PA-1, CAOV-3, SW626, SKOV3 and OVCAR3 were purchased from ATCC (TCP-1021, HTB-161). KURAMOCHI and OVSAHO were gifts from Dr. Kevin Elias (Dana-Farber Cancer Institute, Boston, MA) and COV362, ES2, JHOS-2, OC314, SNU8 and SNU119 were kindly gifted by Dr. Elizabeth Stover (Dana-Farber Cancer Institute, Boston, MA). Patient-derived EOC cell lines were generated in the laboratory of Dr. David Pepin (Massachusetts General Hospital Cancer Center, Boston, MA) as previously described (49) and generously gifted. Briefly, all deidentified ascites samples were collected after informed consent under a Massachusetts General Hospital Institutional Review Board (IRB) approved protocol (2007P001918) and freshly collected ascites was used to obtain a nucleated cell pellet containing immune cells, fibroblasts, mesothelial cells, and cancer cells. These were introduced into tissue culture under two different conditions (adherent and nonadherent) and preserved in liquid nitrogen. Low passages of both the adherent and suspension cultured cell lines were used in vitro experiments.

PDAC cell lines:

PDAC2, PDAC3, PDAC5, PDAC6, PDAC8 and PDAC9 were generated from metastatic ascites fluid of pancreatic adenocarcinoma patients at the Massachusetts General

Hospital under a discarded tissue protocol in accordance with the Massachusetts General Hospital IRB protocol 2011P001236 and Dana-Farber Harvard Cancer Center IRB protocol 02-240 as previously described (50). Cell lines MGH927-1611 were generous gifts from the laboratory of Dr. Andrew Liss under Massachusetts General Hospital IRB protocol 2003P001289.

CRC and THP-1 cell lines:

All CRC cell lines and THP-1 cell line were obtained from ATCC.

For all in vitro experiments using tumor cell lines except for LNA transfection, cell lines were grown as tumorspheres under non-adherent conditions in 3D media. For PDAC and CRC cell lines 3D media contained serum free RPMI supplemented with 20 ul/ml B27 (Invitrogen/Life Technologies), 20 ng/ml EGF (Invitrogen/Life Technologies), 20 ng/ml bFGF (Invitrogen/Life Technologies) and 1% Pen/Strep (Gibco/Life Technologies). For EOC cell lines, standard base growth media (see Supplementary Table 1) supplemented with 10% FBS (Thermo Scientific) and 1% Pen/Strep (Gibco/Life Technologies) was used for 2D adherent cultures and standard base growth media without FBS was used for 3D cultures in non-adherent tissue culture dishes. THP-1 cells were maintained with 1% penicillin-streptomycin, 10% FBS and 0.05 mM β-mercaptoethanol (Sigma-Aldrich, Cat#M3148) in RPMI 1640 medium.

Drug Treatments

IGROV1, OV90, OAW28 and CaOV3 cells were plated at 300,000 cells per well in 6 well ultra-low attachment cell culture dishes (Sigma-Aldrich, Cat #CLS3471) in standard growth media. Cells were incubated for 48-72 hours to allow tumorsphere formation, then treated with 500 nM 5-azacytidine (Sigma-Aldrich, Cat #A2385), 250 nM Trichostatin A (Sigma-Aldrich, Cat #T1952) or DMSO for vehicle control and incubated for 72 hours. Tumorspheres were then harvested and RNA was isolated for further analyses.

RNA Isolation and RNA-seq Library preparation

Cells ($2x10^5$ per well) were transferred to 6-well ultra-low attachment culture dishes in preferred growth media to allow tumorsphere formation. Tumorspheres were collected on days 3-5 of 3D culture depending on rate of growth. RNA was extracted using the miRNEasy Mini Kit (Qiagen) including the optional on column DNAse treatment (Qiagen). In some cases, RNA quality was analyzed using the Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA). To generate libraries for total RNA sequencing the Clontech-Takara Smarter Stranded Total RNA-Seq kit v2 (Cat No#634413) was used according to the manufacturer's instructions. Pooled libraries were sequenced on an Illumina NextSeq 500 sequencer.

RNA-Seq Data Analysis:

Reads alignments

Reads were trimmed and quality checked using $skewer$. Briefly, ends of the reads were trimmed to remove N's and bases with quality less than 20. After that, the quality scores

of the remaining bases were sorted, and the quality at the 20th percentile was computed. If the quality at the 20th percentile was less than 15, the whole read was discarded. Also, reads shorter than 40 bases after trimming were discarded. If at least one of the reads in the pair failed the quality check and had to be discarded, we discarded the mate as well. Quality filtered reads were mapped using STAR aligner and assigned to genes (Gencode annotation) and repeat elements (RepeatMasker annotation) using $featureCounts$ function of $Subread$ package with the external Ensembl annotation. Unassigned reads were then remapped to Repbase consensus sequence. Repeat counts from RepeatMasker annotation and Repbase were added together.

Counts filtering, normalization and differential expression

Gene expression in terms of $\log_2$-CPM (counts per million reads) was computed and normalized across samples using the TMM (trimmed-mean of M-values) method as implemented in the calcNormFactors function of edgeR (19). These low-count values (CPM< 2) that likely due to sequencing errors were removed before calculating the size factor for each sample. Then, filtered CPM was log2 transformed and used in heat-map visualization and pearson correlation analysis. On the heatmap, genes (rows) were scaled by *z*-score scaling. Heat maps were generated by the $pheatmap$ R statistical programming package. The adjusted p-value was calculated using Benjamini & Hochberg correction. Differential expression analysis were carried out using $limma$ in R Differential expression analysis was performed using *limma (51)*.

Consensus expression analysis

Consensus expression of each repeat classes was generated using gene set variation analysis (GSVA). The input matrix is normalized $\log_2$-CPM expression of repeat elements, and a gene list containing predefined gene sets assignment, e.g. SAT, LINE, SINE, ERV, DNA. The $gsva()$ function employs the method described by Hänzelmann, Castelo, and Guinney (52) was applied.

Gene set enrichment analysis (GSEA)

GSEA ranks all of the genes in the dataset based on either differential expression calculated by *limma* or based on the Pearson correlation coefficients between consensus expression of repeats and coding genes. To test the gene set significance, an enrichment score is defined as the maximum distance from the middle of the ranked list. Thus, the enrichment score indicates whether the genes contained in a gene set are clustered towards the beginning (upregulated/positively correlated) or the end (downregulated/negatively correlated) of the ranked list. The GSEA was applied for searching HALLMARK pathways and was accomplished using the function $GSEA$ in package $clusterprofiler$ (53).

Immune Infiltration Analysis

The populations of major types of infiltrating immune cells were evaluated through "xCell" (R package "xCell")(29). The xCell algorithm was used to specifically infer 64 immune and stromal cell types in each sample, based on mRNA expression profiles.

The gene length normalized expression profiles of 96 human early ovarian samples were prepared and uploaded to the xCell web. Analysis was performed by xCell signature (*N* = 64) with 1,000 permutations, based on the parameter settings. Analysis was performed by xCell signature (*N* = 64) with 1,000 permutations, based on the parameter settings. Q values for Pearson correlation between normalized HSATII expression and different immune infiltration signatures were calculated using function $qvalue$ in R with bootstrap method.

Ovarian Carcinoma Tumors

Dana-Farber Harvard Cancer Center IRB protocol 15-015 was used to obtain human ovarian carcinoma tissues for RNA-ISH analysis.

RNA-ISH

For the detection of HSATII RNA levels, automated RNA-ISH assay was performed using Advanced Cell Diagnostics (ACD) probes against HSATII (ACD 512018) and the RNAscope 2.5 LS Reagent Kit-Brown from (Catalogue No.322100) on the BondRx 6.0 platform (Leica Biosystems Inc., Buffalo Grove, IL). 5 µm sections of FFPE tissue (human colorectal cancer tissue, or cell blocks) were mounted on Surgipath X-tra glass slides, baked for 1 hour at 60°C, and placed on the BOND RX for processing. On the BOND RX, the staining protocol used was the ACD ISH DAB Protocol. The RNA unmasking conditions for the tissue consisted of a 15-minute incubation at 95°C in Bond Epitope Retrieval Solution 2 (Leica Biosystems) followed by 15-minute incubation with

Proteinase K which was provided in the kit. Probe hybridization was done for 2 hours with RNAscope probes which were provided by ACD.

The signal was visualized by sequential addition of Red substrate which binds to Amp 6 and Green substrate which binds to Amp 10 producing Red and Green precipitates (dots). The target mRNAs were then visualized using a standard brightfield microscope which showed CYP24A1 signal as Red and FN1 signal as Green.

RNA-ISH quantification:

Slides were imaged using a Leica Aperio CS-O slide scanning microscope at 40x magnification. To determine the relative levels of RNA and protein markers, the images were quantified using Halo software by Indica Labs. The color components for cell nuclei (blue, hematoxylin) and RNA-ISH probe (brown, HSATII) were extracted using color deconvolution. The hematoxylin and probe areas were quantified within representative regions that were annotated by a trained pathologist. Stained RNA-ISH slides were scored according to the fractional area of probe staining in the annotated regions. The fractional probe area was defined as the total probe area divided by total cellular area (sum of hematoxylin and probe areas).

EV Isolations

PDAC and EOC cell lines were grown non-adherently in ULA flasks (Corning) for 5 days prior to EV isolation. The PDAC3 and PDAC6 cells were grown in serum free media containing high glucose DMEM with 20 µl/ml B27 (Invitrogen/Life Technologies), 20ng/ml EGF (Invitrogen/Life Technologies), and 1% Pen/Strep. IGROV1 cells were grown in RPMI medium and OAW28 cells were grown in DMEM/F12, both containing

1% Pen/Strep and 10% FBS. After 3 days in culture, the IGROV1 and OAW28 media was replaced with serum free media and the cells were cultured for an additional 48 hours. The conditioned medium for each cell line was then centrifuged at 2000 x g for 10 min at room temperature to remove larger particles, cells, and debris. Then, 45 µl of media was concentrated, 15 µl at a time, using Amicon Ultra-15 filters (Milipore/Sigma) at 4000 x g at 4°C for approximately 15 minutes per 15 µl. The concentrated media was diluted to 500 µl with filtered PBS and loaded onto 70nm qEV columns (iZon). An additional 2.35 ml of PBS was added to reach the void volume of 2.85 ml following the manufacturer's instructions. After allowing the void volume to pass through the column, the EV-rich fractions were eluted with 1.5 ml of PBS and collected. The EVs were quantified using a NanoSight LM10 (Malvern Panalytical) and stored at -80°C.

EV Protein Analysis:

Concentrated EVs were lysed with 4X Laemmli sample buffer (Bio-Rad, Cat#161-0747) without reducing reagents and boiled for 5 min at 95^0C. Samples were subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and then transferred to a polyvinylidene difluoride (PVDF) membrane. Membranes were blocked with 3% of bovine serum albumin (BSA, Sigma-Aldrich, Cat# A2058) for 1 hour and incubated with primary antibodies overnight at 4^0C. Membranes were washed with 1X phosphate-buffered saline (PBS) containing 0.1% tween-20 (Sigma-Aldrich, Cat#P1379) (PBST) for 10 minutes 3 times and incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies for 1-2 hr at room temperature. After washing, signal was detected with enhanced chemiluminescence (SuperSignal™ West Pico

PLUS Chemiluminescent Substrate, Thermo Scientific Cat# 34577), and images were developed using G:BOX (Syngene).

Antibodies and concentrations for western blot analysis:

CD63 (BioLegend, Cat# 353039) 1:1000

Flotillin-1 (Cell Signaling, Cat# 3253S) 1:1000

CD81 (BioLegend, Cat# 349502) 1:1000

CD9 (Cell Signaling, Cat# 13174S) 1:1000

HRP-conjugated goat anti-rabbit (Cell Signaling, Cat# 7074S) 1:5000

HRP-conjugated goat anti-mouse (Cell Signaling, Cat# 7076S) 1:5000

Co-Culture of Human Monocytes and THP-1 Cells with Tumor-Derived EVs

PBMCs from healthy volunteers were isolated from buffy coats by Ficoll density gradient centrifugation and washed in PBS containing 2% FBS. Following red blood cell lysis, PBMCs were stained with APC-conjugated anti-human CD14 antibody (BD Biosciences cat#555399) and sorted on a BD FACSAria II Cell Sorter. CD14+ PBMCs were seeded at 15-20,000 cells/well in a 48 well plate in high serum RPMI media containing 20% FBS and 1% Pen/Strep. After 24 hours, the media was replaced with serum-free media and cells were dosed with 1000x EVs that were isolated from cancer cells and quantified as described previously. After 48 hours, the cells were harvested and RNA was extracted using the Qiagen miRNeasy Mini Kit (cat#217004).

THP-1 cells were plated at $1x10^5$ cells/well in ultra-low attachment (ULA) 24-well plate (Coring, Cat#3473) without β-mercaptoethanol and FBS. Then, cells were treated with $1x10^8$ (for low concentration) and $1x10^9$ (for high concertation) particles of EVs derived

from OAW28 and IGROV1. After 48 hrs, cells were collected, and RNA extraction was performed

HSATII in vitro transcription (IVT) and transfection

HSATII containing fragment on chromosome 10 was amplified by PCR with M13 forward and reverse primers as described before (38) and subjected to in vitro transcription with SP6 RNA polymerase following the MAXIscript™ SP6 Transcription Kit recommendation (Thermo Fisher Scientific, Cat# AM1308). Then, in vitro transcribed HSATII was purified with Megaclear™ Transcription Clean-Up Kit (Thermo Fisher Scientific, Cat# AM1908). Each RNA (500 fmol) was then transfected using jetMESSENGER (Polyplus-transfection, Cat# 15001) in accordance with the manufacturer's instructions. After 48 hours, the cells were harvested, and gene expression was analyzed by RT-qPCR.

qRT-PCR

RNA extraction was performed using the miRNEasy Mini Kit (Qiagen, Cat# 217004) according to the manufacturer's instructions. Total RNA (1 µg) was reverse transcribed using TaqMan™ Reverse Transcription Reagents (Invitrogen™, Cat# N8080234). qRT-PCR was conducted using the PowerUp™ SYBR™ Green Master mix (Applied Biosystems, Cat# A25742). The following primers were used: *GAPDH* Forward 5′-ACATCATCCCTGCCTCTACT-3′, Reverse 5′-TCCACCACTGACACGTTG-3′; *DHX58* Forward 5'-ATGACCACCTGGAGATGCCTGA-3', Reverse 5'-CATTGTAGCGCCTCAGGTGAAG-3'; *IFNB1* Forward 5'-

ATGGGAGGCTTGAATACTGC-3', Reverse 5'-TCATAGATGGTCAATGCGGC-3'; *OAS2* Forward 5'-CCGTTGGTGTTGGCATCTTC-3', Reverse 5'-GCATTGTCGGCACTTTCCAA-3'; *ISG15* Forward 5'-CTCTGAGCATCCTGGTGAGGAA-3', Reverse 5'-AAGGTCAGCCAGAACAGGTCGT-3'; *MX1* Forward 5'-TCATAGATGGTCAATGCGGC-3', Reverse 5'-TCATAGATGGTCAATGCGGC-3'; *MX2* Forward 5'-GCCCTTAGCATGCTCCAGAA-3', Reverse 5'-ATCGTGCTCTGAACAGTTTGG-3'; *IFI44* Forward 5'-GTGAGGTCTGTTTTCCAAGGGC-3', Reverse 5'-CGGCAGGTATTTGCCATCTTTCC-3'. Reactions were performed on a QuantStudio 3 (Applied Biosystems) thermocycler. The level of gene expression was calculated based on the $2^{-\Delta\Delta CT}$ method and normalized to the Ct value of *GAPDH* as endogenous control. The reference samples were the PBS or no transfection conditions, respectively.

HSATII LNA Transfection

OVCAR4, IGROV1, OAW28, and OVSAHO cells were plated at a density of 120,000 cells/well in 6 well tissue culture treated plates (Corning). OVCAR4 and IGROV1 cells were grown in RPMI medium and OAW28 and OVSAHO cells were grown in DMEM/F12, both containing 1% Pen/Strep and 10% FBS. After 2 days, 500 nM Negative control (Scramble) or HSATII custom designed LNA were transfected into the cells with Lipofectamine 2000 (Invitrogen/Life Technologies) following the manufacturer's recommendations. One day post transfection, 2 out of 3 wells for each experimental condition were trypsinized, counted, and plated for proliferation assays and the remaining wells were media changed for flow cytometry. For proliferation assays, cells were seeded at a density of 1000 cells/well into a 96-well ULA plate

(Corning) and quantified using CellTiter-Glo 3D luminescent cell viability assay (Promega) with a SpectraMax microplate reader (Molecular Devices). For flow cytometry, cells were harvested 2 days after transfection and incubated with anti-MHCI and MHCII antibodies before quantification on a cytometer.

Flow Cytometric Analysis

Following transfection with HSATII-specific or control LNA, IGROV1 (top panels) and OAW28 (bottom panels) were grown in standard adherent culture for 48 hours. Cells were then trypsinized, washed with PBS, counted and resuspended at 2.5e6 cells/ml in flow cytometry buffer (PBS + 2% FBS). For each condition, 2.5e5 cells were stained with PE-conjugated anti Human HLA Class I (R&D Systems, FAB7098P) and APC-conjugated anti-human HLA-DR (Biolegend, 307609) at a concentration of 1:100. After staining, cells were washed and resuspended in flow cytometry buffer with addition of 4',6-diamino-2-phenylindole, dihydrochloride (DAPI; Sigma Aldrich, D9542) to a final concentration of 0.1μg/ml. Single stained control samples and experimental samples were run on a Fortessa X-20 flow cytometer (BD Biosciences) and signal was analyzed using FlowJo software.

Quantification and Statistical Analyses

For all experiments p values were calculated using PRISM9 GraphPad paired t-test or two-way ANOVA, unless otherwise noted (*<0.05, **<0.01, ***<0.001, ****<0.0001). Kaplan Meier curve plotted and log rank analysis performed with GraphPad PRISM9. Additional computational analysis statistics are detailed in computational methods

sections above. For RNA sequencing data, statistical analyses were performed with R/Rstudio software (version 4.0.3). False discovery rate (FDR) of 0.05 was used for GSEA and Q values of 0.1 for correlation significance.

Study approval

Patient derived cell lines and tumor materials were obtained under Massachusetts General Hospital IRB protocols 2003P001289, 2007P001918, and 2011P001236 and Dana-Farber Harvard Cancer Center IRB protocols 02-240 and 15-015.

Data Availability Statement

All RNA-seq data has been uploaded to NCBI GEO GSE205430.

Code Availability Statement

There was no custom code developed for this project, but all code and statistical packages are detailed in the methods above.

All software for RNA-seq and digital image data analysis is described in the methods above and all software will be provided upon request.

AUTHOR CONTRIBUTIONS

RLP and SS contributed equally to this work. RLP wrote the original draft, had the initial conception of the project, and obtained funding for this project, which was used to assign RLP as the first author on this manuscript.

Conceptualization: RLP, SS, BDG, DTT

Methodology: SS, DP, A. Solovyov

Software: SS, MJR, VT, A. Solovyov, BDG

Formal analysis: RLP, SS, EY, KX, LTN, A. Solovyov, BDG

Investigation: RLP, EY, IEP, MNF, EB, NKM, ND, ECT, A. Szabolcs, ERL, AP, EY, MJR, VT, KHX

Resources: DK, ES, DP, SLS, VD, JFL, UAM

Data Curation: ES, RLP

Writing - Original Draft: RLP, DTT

Writing - Review & Editing: RLP, SS, BDG, DTT

Visualization: RLP, SS, EY, MNF, EB

Supervision: RLP, A. Solovyov, BDG, DTT

Project administration: RLP, BDG, DTT

Funding acquisition: RLP, BDG, DTT

ACKNOWLEDGEMENTS

We thank Danielle Bestoso for laboratory and administrative support. We are thankful for the patients that provided tumor material for cell lines and tissue analysis for this project.

This work was supported by research funding from the National Institutes of Health R01CA240924 (DTT, BDG), U01CA228963 (BDG, DTT); Department of Defense OC180106 (RLP); ACD-Biotechne (DTT, VD); Robert L. Fine Cancer Research Foundation (DTT).

COMPETING INTERESTS STATEMENT

DTT and BDG are founders and consultants for ROME therapeutics. None of this work has been supported by the company, but the company has licensing rights to HSATII LNAs related to colon cancer.

BDG has received honoraria for speaking engagements from Merck, Bristol Meyers Squibb, and Chugai Pharmaceuticals; has received research funding from Bristol Meyers Squibb and Merck; and has been a compensated consultant for Darwin Health, Merck, and PMV Pharma.

DTT has received consulting fees from Tekla Capital, Ikena Oncology, Foundation Medicine, Inc., NanoString Technologies, EMD Millipore Sigma, and Pfizer that are not

related to this work. DTT is a founder and has equity PanTher Therapeutics and TellBio, Inc., which is not related to this work. DTT receives research support from ACD-Biotechne, PureTech Health LLC, and Ribon Therapeutics, which was not used in this work. DTT's interests are managed by Mass General Brigham in accordance with their conflict of interest policies.

REFERENCES

1. Britten RJ, and Kohne DE. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science.* 1968;161(3841):529-40.
2. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature.* 2001;409(6822):860-921.
3. Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S, Coffman EJ, et al. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science.* 2011;331(6017):593-6.
4. Rodic N, Sharma R, Sharma R, Zampella J, Dai L, Taylor MS, et al. Long interspersed element-1 protein expression is a hallmark of many human cancers. *The American journal of pathology.* 2014;184(5):1280-6.
5. Rooney MS, Shukla SA, Wu CJ, Getz G, and Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell.* 2015;160(1-2):48-61.
6. Slotkin RK, and Martienssen R. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet.* 2007;8(4):272-85.
7. Wylie A, Jones AE, D'Brot A, Lu WJ, Kurtz P, Moran JV, et al. p53 genes function to restrain mobile elements. *Genes & development.* 2016;30(1):64-77.
8. Leonova KI, Brodsky L, Lipchick B, Pal M, Novototskaya L, Chenchik AA, et al. p53 cooperates with DNA methylation and a suicidal interferon response to maintain epigenetic silencing of repeats and noncoding RNAs. *Proceedings of the National Academy of Sciences of the United States of America.* 2013;110(1):E89-98.
9. Chiappinelli KB, Strissel PL, Desrichard A, Li H, Henke C, Akman B, et al. Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. *Cell.* 2015;162(5):974-86.
10. Roulois D, Loo Yau H, Singhania R, Wang Y, Danesh A, Shen SY, et al. DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. *Cell.* 2015;162(5):961-73.
11. Mehdipour P, Marhon SA, Ettayebi I, Chakravarthy A, Hosseini A, Wang Y, et al. Epigenetic therapy induces transcription of inverted SINEs and ADAR1 dependency. *Nature.* 2020;588(7836):169-73.
12. Moufarrij S, Srivastava A, Gomez S, Hadley M, Palmer E, Austin PT, et al. Combining DNMT and HDAC6 inhibitors increases anti-tumor immune signaling and decreases tumor burden in ovarian cancer. *Sci Rep.* 2020;10(1):3470.
13. Stone ML, Chiappinelli KB, Li H, Murphy LM, Travers ME, Topper MJ, et al. Epigenetic therapy activates type I interferon signaling in murine ovarian cancer to reduce immunosuppression and tumor burden. *Proceedings of the National Academy of Sciences of the United States of America.* 2017;114(51):E10981-E90.
14. Sheng W, LaFleur MW, Nguyen TH, Chen S, Chakravarthy A, Conway JR, et al. LSD1 Ablation Stimulates Anti-tumor Immunity and Enables Checkpoint Blockade. *Cell.* 2018;174(3):549-63 e19.

15. Canadas I, Thummalapalli R, Kim JW, Kitajima S, Jenkins RW, Christensen CL, et al. Tumor innate immunity primed by specific interferon-stimulated endogenous retroviruses. *Nat Med.* 2018;24(8):1143-50.
16. Griffin GK, Wu J, Iracheta-Vellve A, Patti JC, Hsu J, Davis T, et al. Epigenetic silencing by SETDB1 suppresses tumour intrinsic immunogenicity. *Nature.* 2021.
17. Tanne A, Muniz LR, Puzio-Kuter A, Leonova KI, Gudkov AV, Ting DT, et al. Distinguishing the immunostimulatory properties of noncoding RNAs expressed in cancer cells. *Proceedings of the National Academy of Sciences of the United States of America.* 2015;112(49):15154-9.
18. Desai N, Sajed D, Arora KS, Solovyov A, Rajurkar M, Bledsoe JR, et al. Diverse repetitive element RNA expression defines epigenetic and immunologic features of colon cancer. *JCI insight.* 2017;2(3):e91078.
19. Solovyov A, Vabret N, Arora KS, Snyder A, Funt SA, Bajorin DF, et al. Global Cancer Transcriptome Quantifies Repeat Element Polarization between Immunotherapy Responsive and T Cell Suppressive Classes. *Cell Rep.* 2018;23(2):512-21.
20. Panda A, de Cubas AA, Stein M, Riedlinger G, Kra J, Mayer T, et al. Endogenous retrovirus expression is associated with response to immune checkpoint blockade in clear cell renal cell carcinoma. *JCI insight.* 2018;3(16).
21. Zapatka M, Borozan I, Brewer DS, Iskar M, Grundhoff A, Alawi M, et al. The landscape of viral associations in human cancers. *Nat Genet.* 2020;52(3):320-30.
22. Espinet E, Gu Z, Imbusch CD, Giese NA, Buscher M, Safavi M, et al. Aggressive PDACs Show Hypomethylation of Repetitive Elements and the Execution of an Intrinsic IFN Program Linked to a Ductal Cell of Origin. *Cancer Discov.* 2021;11(3):638-59.
23. Network CGAR. Integrated genomic analyses of ovarian carcinoma. *Nature.* 2011;474(7353):609-15.
24. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America.* 2005;102(43):15545-50.
25. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, and Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 2015;1(6):417-25.
26. Engqvist H, Parris TZ, Rönnerman EW, Söderberg EMV, Biermann J, Mateoiu C, et al. Transcriptomic and genomic profiling of early-stage ovarian carcinomas associated with histotype and overall survival. *Oncotarget.* 2018;9(80):35162-80.
27. Jayabal P, Ma X, and Shiio Y. EZH2 suppresses endogenous retroviruses and an interferon response in cancers. *Genes & cancer.* 2021;12:96-105.
28. Sun S, Frontini F, Qi W, Hariharan A, Ronner M, Wipplinger M, et al. Endogenous retrovirus expression activates type-I interferon signaling in an experimental mouse model of mesothelioma development. *Cancer Letters.* 2021;507:26-38.
29. Aran D, Hu Z, and Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 2017;18(1):220.
30. Disis ML, Taylor MH, Kelly K, Beck JT, Gordon M, Moore KM, et al. Efficacy and Safety of Avelumab for Patients With Recurrent or Refractory Ovarian Cancer:

Phase 1b Results From the JAVELIN Solid Tumor Trial. *JAMA Oncol.* 2019;5(3):393-401.
31. Matulonis UA, Shapira-Frommer R, Santin AD, Lisyanskaya AS, Pignata S, Vergote I, et al. Antitumor activity and safety of pembrolizumab in patients with advanced recurrent ovarian cancer: results from the phase II KEYNOTE-100 study. *Ann Oncol.* 2019;30(7):1080-7.
32. Konstantinopoulos PA, and Cannistra SA. Immune Checkpoint Inhibitors in Ovarian Cancer: Can We Bridge the Gap Between IMagynation and Reality? *J Clin Oncol.* 2021;39(17):1833-8.
33. Porter RL, and Matulonis UA. Checkpoint Blockade: Not Yet NINJA Status in Ovarian Cancer. *J Clin Oncol.* 2021:JCO2101886.
34. Taki M, Abiko K, Ukita M, Murakami R, Yamanoi K, Yamaguchi K, et al. Tumor Immune Microenvironment during Epithelial–Mesenchymal Transition. *Clinical Cancer Research.* 2021;27(17):4669-79.
35. Dongre A, Rashidian M, Reinhardt F, Bagnato A, Keckesova Z, Ploegh HL, et al. Epithelial-to-Mesenchymal Transition Contributes to Immunosuppression in Breast Carcinomas. *Cancer Res.* 2017;77(15):3982-9.
36. Pires PRL, Xavier PLP, and Fukumasu H. Abstract B180: Effects of EMT process under MHC class I and TAP1 gene expression related to antigen presentation. *Cancer Immunology Research.* 2019;7(2_Supplement):B180-B.
37. Murakami R, Matsumura N, Mandai M, Yoshihara K, Tanabe H, Nakai H, et al. Establishment of a Novel Histopathological Classification of High-Grade Serous Ovarian Carcinoma Correlated with Prognostically Distinct Gene Expression Subtypes. *The American journal of pathology.* 2016;186(5):1103-13.
38. Bersani F, Lee E, Kharchenko PV, Xu AW, Liu M, Xega K, et al. Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. *Proceedings of the National Academy of Sciences of the United States of America.* 2015;112(49):15148-53.
39. Wang G, Xu D, Zhang Z, Li X, Shi J, Sun J, et al. The pan-cancer landscape of crosstalk between epithelial-mesenchymal transition and immune evasion relevant to prognosis and immunotherapy response. *npj Precision Oncology.* 2021;5(1):56.
40. Kishikawa T, Otsuka M, Yoshikawa T, Ohno M, Yamamoto K, Yamamoto N, et al. Quantitation of circulating satellite RNAs in pancreatic cancer patients. *JCI Insight.* 2016;1(8).
41. Evdokimova V, Ruzanov P, Gassmann H, Zaidi SH, Peltekova V, Heisler LE, et al. Exosomes transmit retroelement RNAs to drive inflammation and immunosuppression in Ewing Sarcoma. *bioRxiv.* 2019:806851.
42. Goodman AM, Castro A, Pyke RM, Okamura R, Kato S, Riviere P, et al. MHC-I genotype and tumor mutational burden predict response to immunotherapy. *Genome Medicine.* 2020;12(1):45.
43. Lee JH, Shklovskaya E, Lim SY, Carlino MS, Menzies AM, Stewart A, et al. Transcriptional downregulation of MHC class I and melanoma de- differentiation in resistance to PD-1 inhibition. *Nature Communications.* 2020;11(1):1897.
44. Rodig SJ, Gusenleitner D, Jackson DG, Gjini E, Giobbie-Hurder A, Jin C, et al. MHC proteins confer differential sensitivity to CTLA-4 and PD-1 blockade in untreated metastatic melanoma. *Sci Transl Med.* 2018;10(450).

45. Gu SS, Zhang W, Wang X, Jiang P, Traugh N, Li Z, et al. Therapeutically Increasing MHC-I Expression Potentiates Immune Checkpoint Blockade. *Cancer Discov.* 2021;11(6):1524-41.
46. Zhang L, Conejo-Garcia JR, Katsaros D, Gimotty PA, Massobrio M, Regnani G, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N Engl J Med.* 2003;348(3):203-13.
47. Hamanishi J, Takeshima N, Katsumata N, Ushijima K, Kimura T, Takeuchi S, et al. Nivolumab Versus Gemcitabine or Pegylated Liposomal Doxorubicin for Patients With Platinum-Resistant Ovarian Cancer: Open-Label, Randomized Trial in Japan (NINJA). *J Clin Oncol.* 2021:JCO2100334.
48. Kulkarni JA, Witzigmann D, Thomson SB, Chen S, Leavitt BR, Cullis PR, et al. The current landscape of nucleic acid therapeutics. *Nature Nanotechnology.* 2021;16(6):630-43.
49. Pépin D, Sosulski A, Zhang L, Wang D, Vathipadiekal V, Hendren K, et al. AAV9 delivering a modified human Mullerian inhibiting substance as a gene therapy in patient-derived xenografts of ovarian cancer. *Proceedings of the National Academy of Sciences of the United States of America.* 2015;112(32):E4418-27.
50. Indolfi L, Ligorio M, Ting DT, Xega K, Tzafriri AR, Bersani F, et al. A tunable delivery platform to provide local chemotherapy for pancreatic ductal adenocarcinoma. *Biomaterials.* 2016;93:71-82.
51. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.
52. Hänzelmann S, Castelo R, and Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics.* 2013;14:7.
53. Yu G, Wang LG, Han Y, and He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284-7.

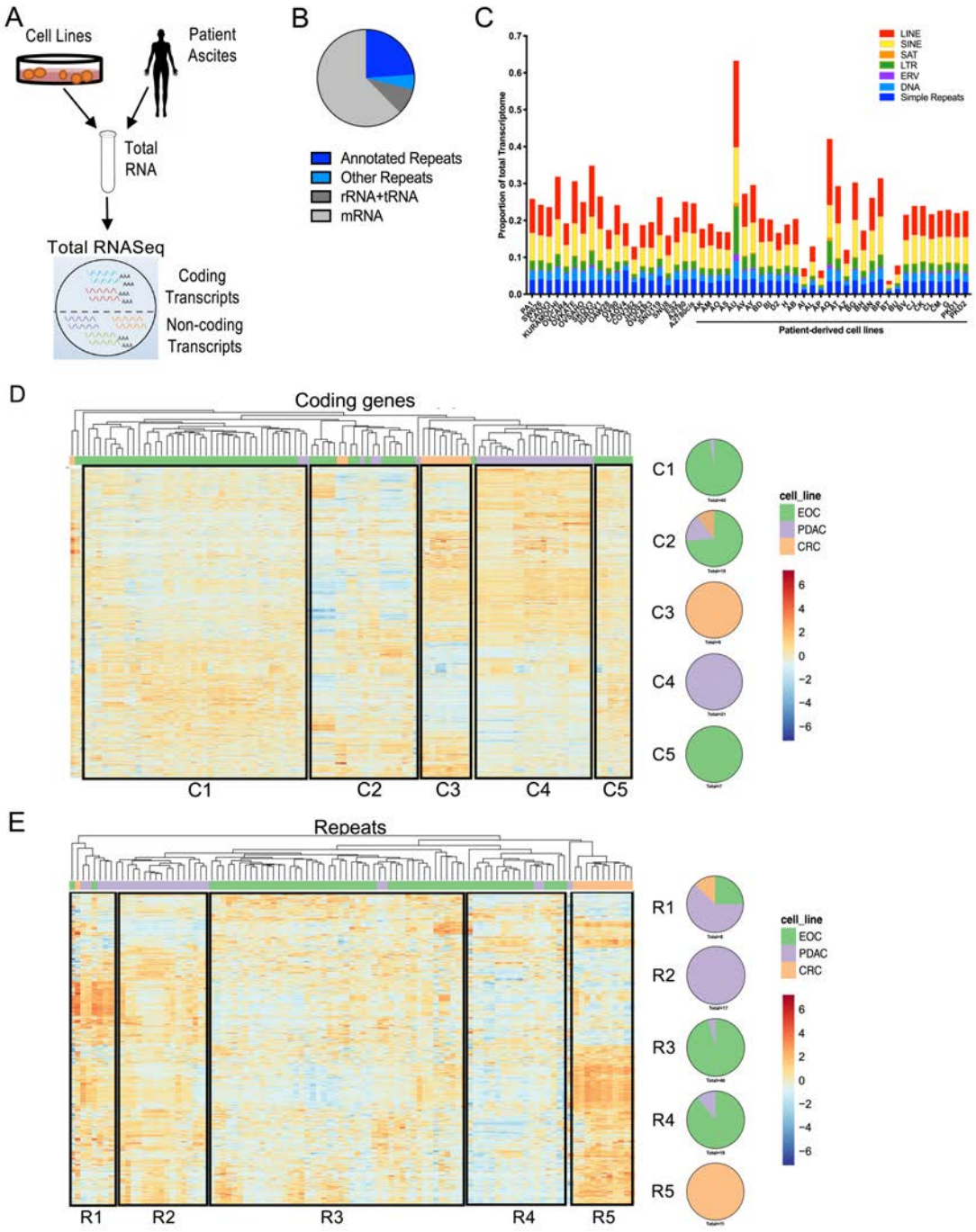


Figure 1. Diverse repeat RNA expression profiles are present in epithelial cancers and cluster tumors by tissue of origin distinctly from coding gene-based clustering. (**A**) Graphical abstract of experimental strategy. (**B**) Proportion of the total transcriptome represented by mRNA, rRNA/tRNA, annotated repeats and non-annotated repeats, averaged across all EOC cell lines. (**C**) Quantification of sub-classes of repeat RNAs across EOC models using total RNA-Seq. Expressed as proportion of total transcription, including coding and non-coding reads in each cell line or patient-derived cells. (**D**) Heatmap and hierarchical clustering of epithelial ovarian cancer (EOC; green), pancreatic ductal adenocarcinoma (PDAC; purple) and colorectal cancer (CRC; gold) cell lines by coding gene expression including all coding genes that were differentially expressed between any two cell lines (p.adj < 0.05 & |log2FC| >1). Expression is plotted as scaled $\log_2$(normalized counts per million). Pie graphs C1-C5 depict the cancer-type composition of each cluster as labeled. (**E**) Heatmap and hierarchical clustering of epithelial ovarian cancer (EOC; green), pancreatic ductal adenocarcinoma (PDAC; purple) and colorectal cancer (CRC; gold) cell lines by repeat RNA expression including all repeat species that were differentially expressed between any two cell lines (p.adj < 0.05 & |$\log_2$FC| >1). Expression is plotted as scaled $\log_2$(normalized counts per million). Major clusters defined by similar repeat expression profiles are outlined in black boxes. Pie graphs R1-R5 depict the cancer-type composition of each cluster as labeled.

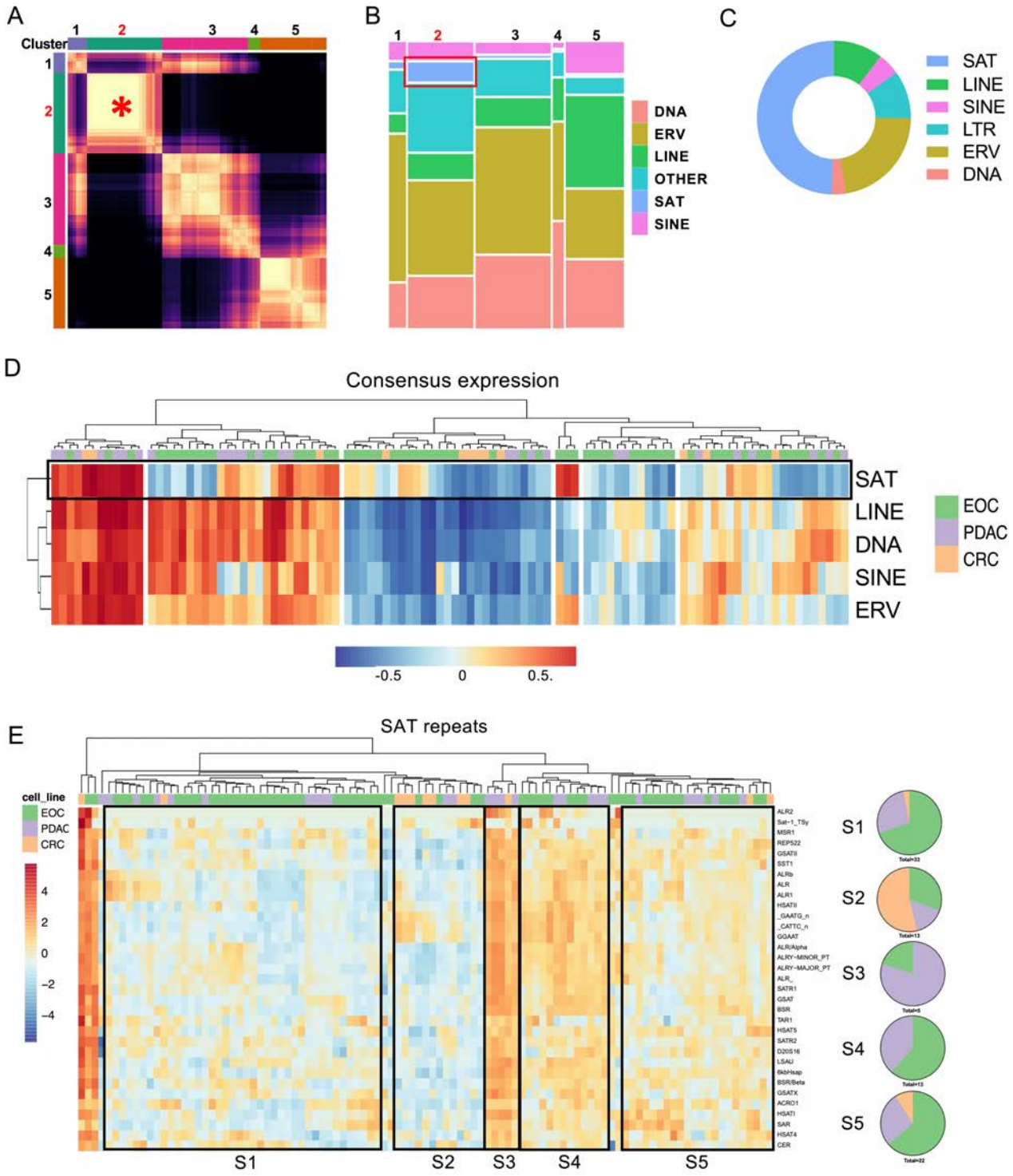


Figure 2. Repeat RNAs are coregulated in discrete clusters with SAT RNAs exhibiting unique expression patterns in epithelial cancers. (**A**) Heatmap for consensus clustering of repeat elements based on normalized expression. Red asterisk highlights SAT-driven Cluster 2 showing strongest consensus correlation. (**B**) Mosaic plot demonstrating relative repeat element subclass composition of each consensus cluster from panel A. (**C**) Proportion of total repeat expression for each subclass within the top 50 variant repeat RNAs across cell lines. (**D**) Hierarchical clustering of consensus expression of each repeat subclass across EOC (green), PDAC (purple) and CRC (gold) cell lines depicting SAT consensus expression distinct from consensus expression of other repeats subclasses. (**E**) Heatmap and hierarchical clustering of EOC (green), PDAC (purple) and CRC (gold) cell lines by satellite repeat (SATs) RNA expression. Expression is plotted as scaled $\log_2$(normalized counts per million). Major clusters defined by similar SAT expression profiles are outlined in black boxes. Pie graphs S1-S5 below depict the cancer-type composition of each cluster, highlighting clusters distinct from tissue of origin.

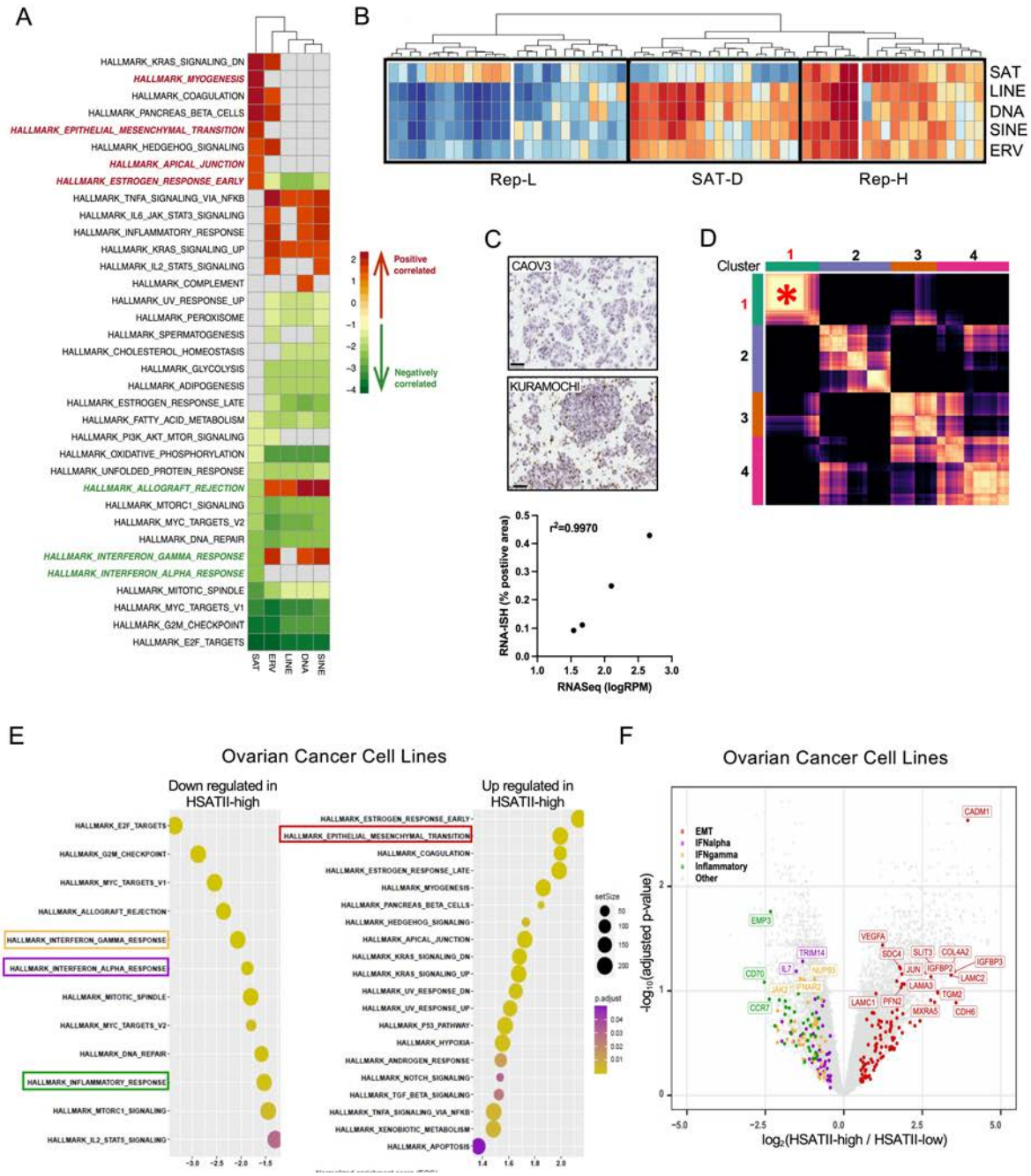


Figure 3. Satellite repeat expression is associated with upregulation of epithelial-mesenchymal transition and downregulation of innate immune response genes in EOC models. (**A**) Heatmap of enriched GO terms identified using Gene Set Enrichment Analysis (GSEA) plotted based on Normalized Enrichment Score (NES). GSEA was applied to a ranked gene list based on their correlation with the consensus expression calculated for each repeat subclass with the false discovery rate (FDR) set at 0.05. Positive enrichment scores (red) indicate functions that positively correlate with repeat subclass expression. Negative enrichment scores (green) indicate functions that negatively correlate with repeat expression. (**B**) Hierarchical clustering of consensus expression calculated for each repeat subclass in EOC cell lines. Major clusters are outlined in black boxes. (**C**) Representative RNA-ISH images with HSATII-specific probes in two EOC cell lines (top panels) and correlation (Pearson r^2 shown) between HSATII RNA expression as determined with RNA-seq by log(reads per million) and percent positive tumor cells by RNA-ISH. (**D**) Heatmap for consensus clustering of all repeat elements except HSATII, which was removed from analysis, based on normalized expression. Asterisk highlights SAT-driven Cluster 1 showing highest consensus correlation, similarly to clustering with HSATII included. (**E**) GSEA of HALLMARK terms based on the log_2FC of coding genes for samples in HSATII-high versus low, based on highest (Q4) and lowest (Q1) quantile (see **Supplemental Figure 3C**). (**F**) Volcano plot depicting differentially expressed coding genes between HSATII-high and HSATII-low EOC cell lines. EMT, IFNa, IFNg and Inflammatory HALLMARK pathways are highlighted.

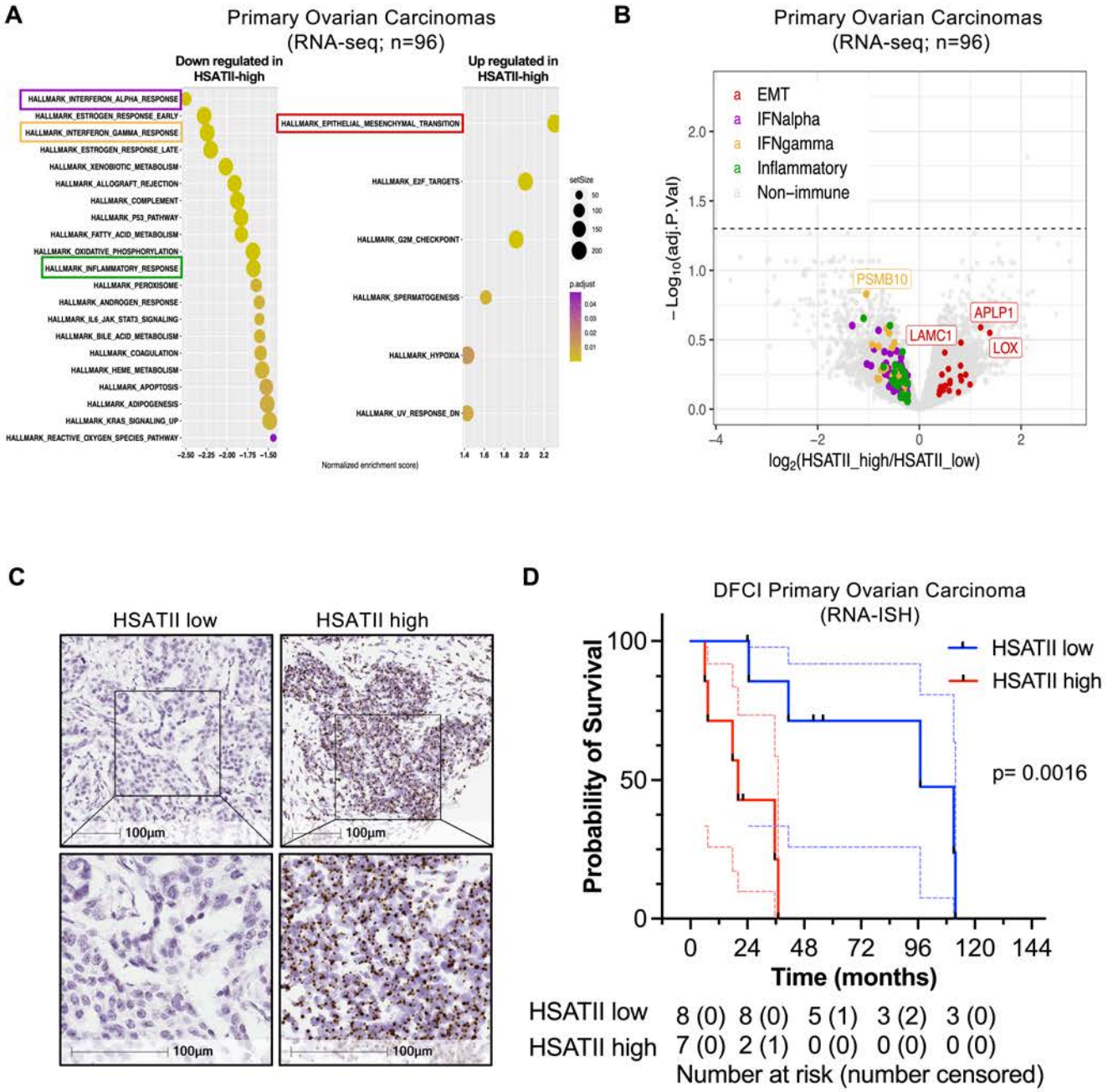


Figure 4. High satellite repeat expression is linked with upregulation of epithelial-mesenchymal transition, suppressed immune response and worsened clinical outcomes in primary human EOC. (**A**) GSEA results ranked by normalized enrichment score for pathways containing genes that are upregulated (right) and downregulated (left) in HSATII-high compared with HSATII-low early-stage human ovarian carcinoma samples (n=96) (**B**) Volcano plot depicting differentially expressed coding genes between HSATII-high and HSATII-low early-stage human ovarian carcinoma samples (n=96). Genes driving the enrichment in EMT, IFNa, IFNg and Inflammatory Hallmark pathways are highlighted. (**C**) Representative images of RNA-ISH with HSATII-specific probe depicting an example of an HSATII-low (left panels) and HSATII-high (right panels) primary human EOC tumor. (**D**) Kaplan-Meier survival curves for HSATII-high (red) and HSATII-low (blue) in a cohort of 16 primary human EOC tumors using quantified RNA-ISH. All data points shown and 95% CI shown (dotted lines). Log-Rank p-value 0.0016.

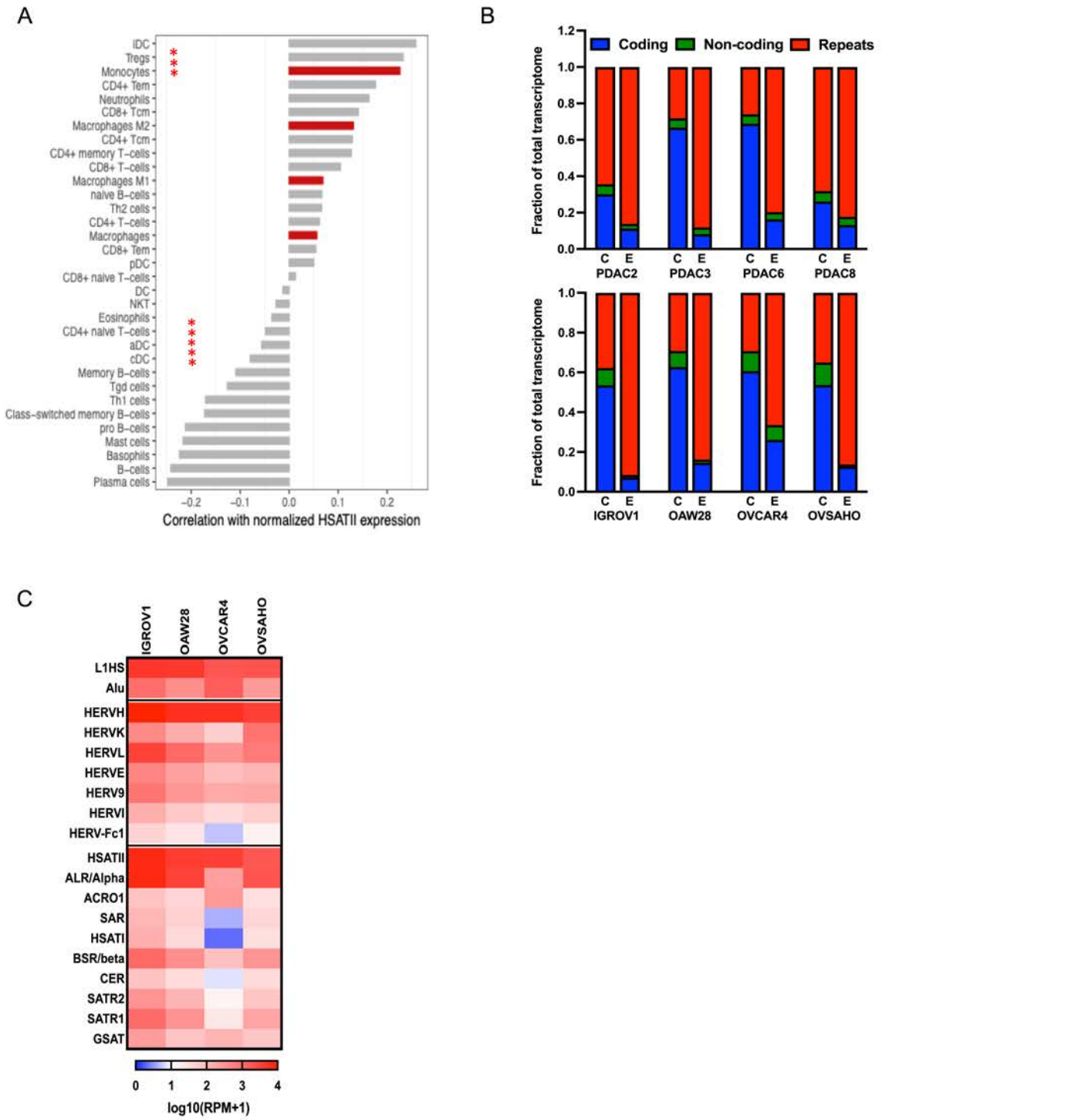


Figure 5. Repeat RNAs are enriched in tumor cell derived extracellular vesicles can induce changes in the tumor immune microenvironment. (**A**) Pearson correlation coefficients between normalized HSATII expression and the relative frequency of immune cell types in 96 human early stage ovarian carcinoma tumor samples as identified by the "xCell" algorithm. Red * indicates correlations with Q value < 0.1. (**B**) RNA content of tumor cells (C) and tumor cell-derived EVs (E) in PDAC (top) and EOC (bottom) cell lines as determined by total RNA-seq and plotted as fraction of total transcriptome. (**C**) Expression heatmap of representative repetitive elements in EVs released by EOC cell lines. * $p < 0.05$

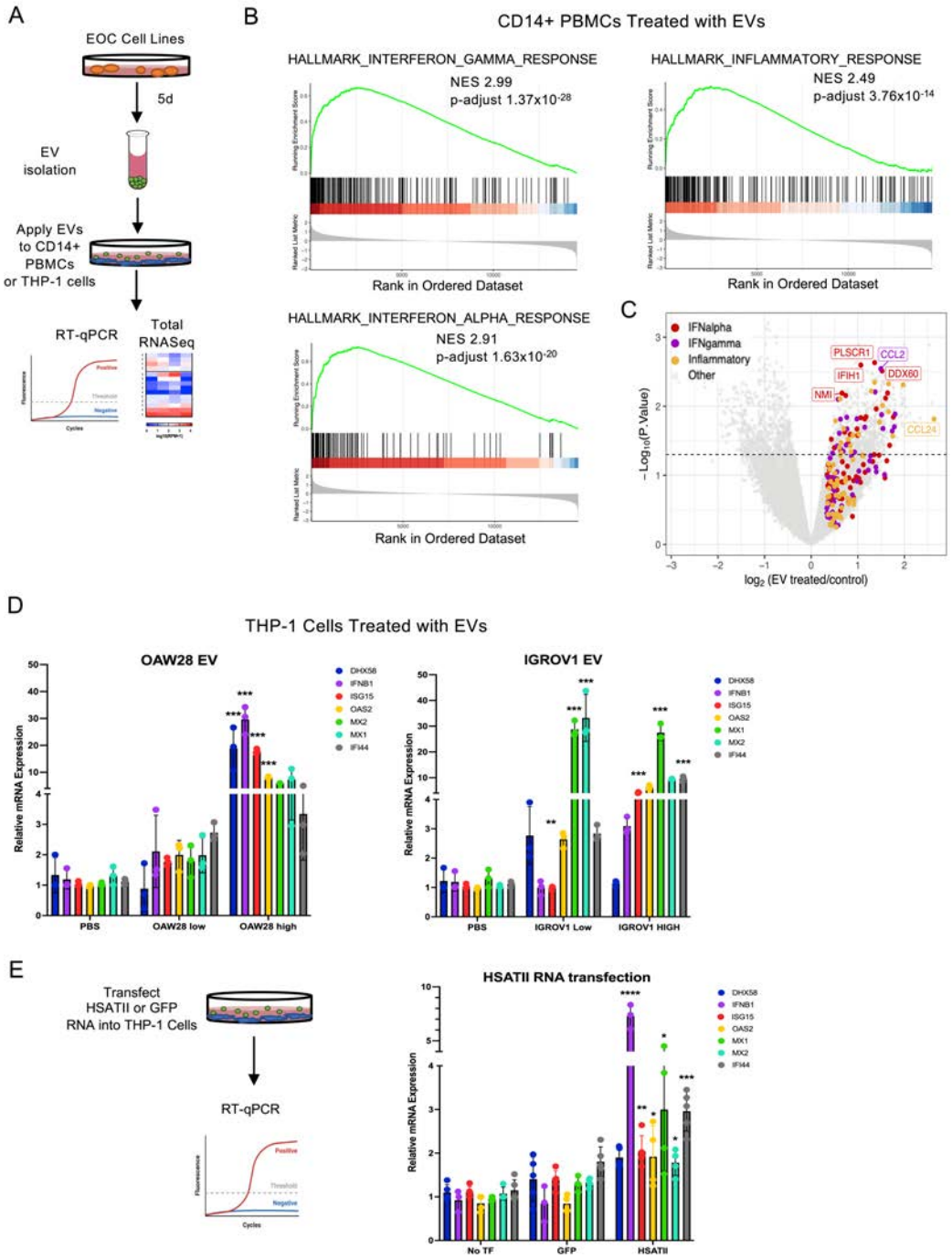


Figure 6. Repeat RNA-enriched extracellular vesicles can induce changes in the tumor immune microenvironment. (**A**) Schema of experimental design relating to data in panels E-H. (**B**) Gene set enrichment analysis of IFN-response signatures and inflammatory response in EV treated versus untreated samples. NES, normalized enrichment score. (**C**) Volcano plot depicting differentially expressing coding genes between EV-treated and untreated EOC cell lines. Genes driving the enrichment in IFNa, IFNg and Inflammatory HALLMARK pathways are highlighted. (**D**) Quantitative RT-PCR of IFN response genes from THP-1 monocyte cell line treated with high dose or low dose EVs from ovarian cell lines OAW28 (left) and IGROV1 (right). (**E**) Schema of THP-1 cells treated with HSATII or GFP RNA transfection (left) and quantitative RT-PCR of IFN response genes from transfection (right) for no transfection (TF), GFP RNA, and HSATII RNA. For RT-PCR, all data points are shown with mean and STDEV (error bars). One-way ANOVA analysis performed with Tukey's multiple comparisons test with p-value significance shown between EV treatment and PBS or HSATII and GFP RNA (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).

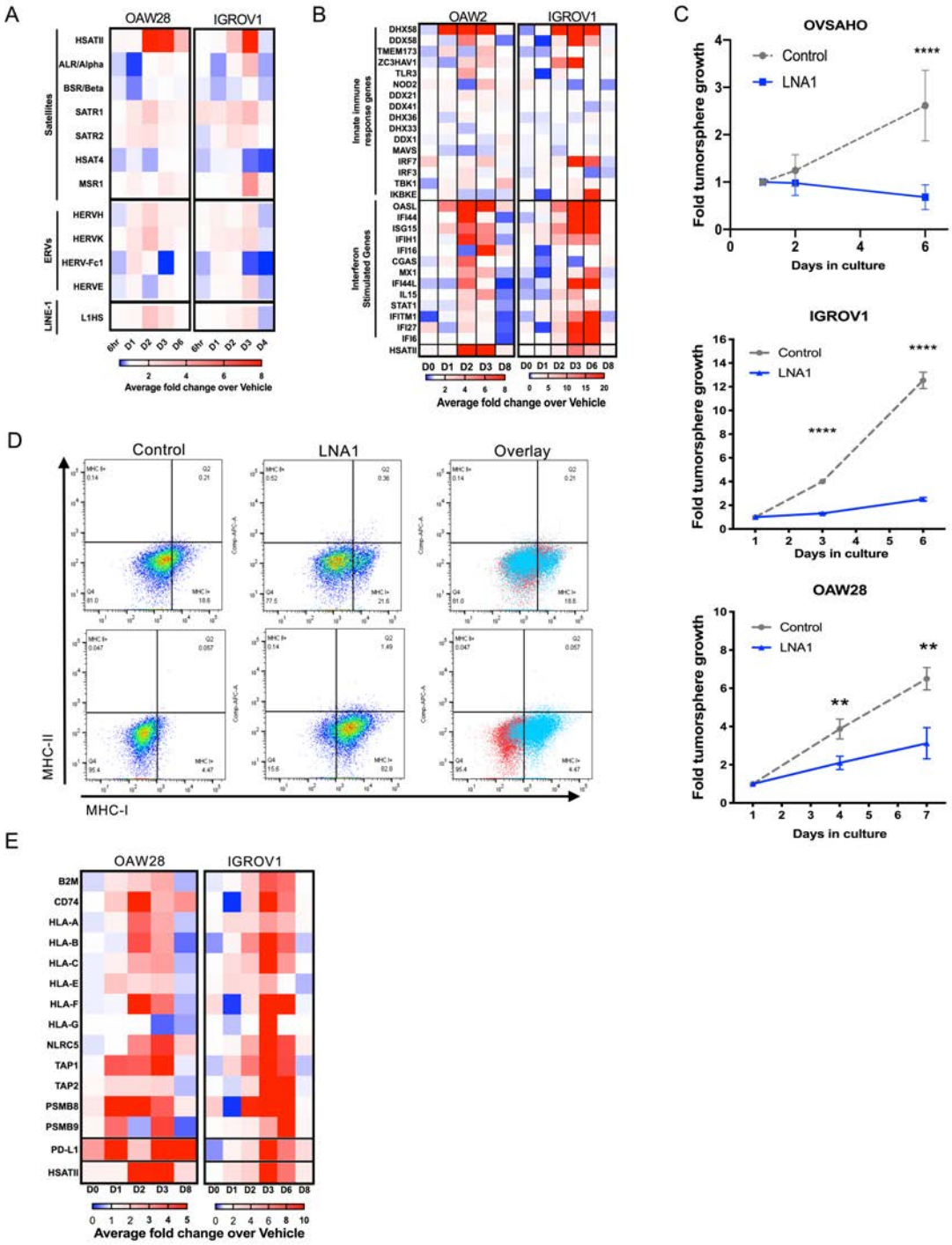


Figure 7. Modulation of HSATII RNA with LNA is cytotoxic and induces interferon response and increases MHC Class I expression. (**A**) Expression levels of HSATII and other repeat RNAs in EOC cell lines transfected with HSATII-specific LNA relative to scramble control LNA over time, plotted as fold change over control on days 0 through 6 post-transfection. (**B**) Expression heatmap depicting relative expression of innate immune response genes and interferon-stimulated genes (ISGs) in EOC cell lines transfected with HSATII-specific LNA relative to scramble control LNA over time. (**C**) Effect of HSATII-specific LNA (LNA1) on tumorsphere growth in EOC cell lines as determined by 3D CellTiterGlo viability assays. Plots represent 4 separate experiments for each cell line, with two-tailed unpaired t test performed at each time point (** $p < 0.01$, **** $p < 0.0001$). (**D**) Expression heatmap depicting relative expression of MHC-Class I genes and PD-L1 in EOC cell lines transfected with HSATII-specific LNA relative to scramble control LNA over time. (**E**) Flow cytometric analysis of MHC-I and MHC-II cell surface protein expression on the cell surface of EOC cell lines transfected with HSATII-specific LNA (LNA1) compared with control LNA.

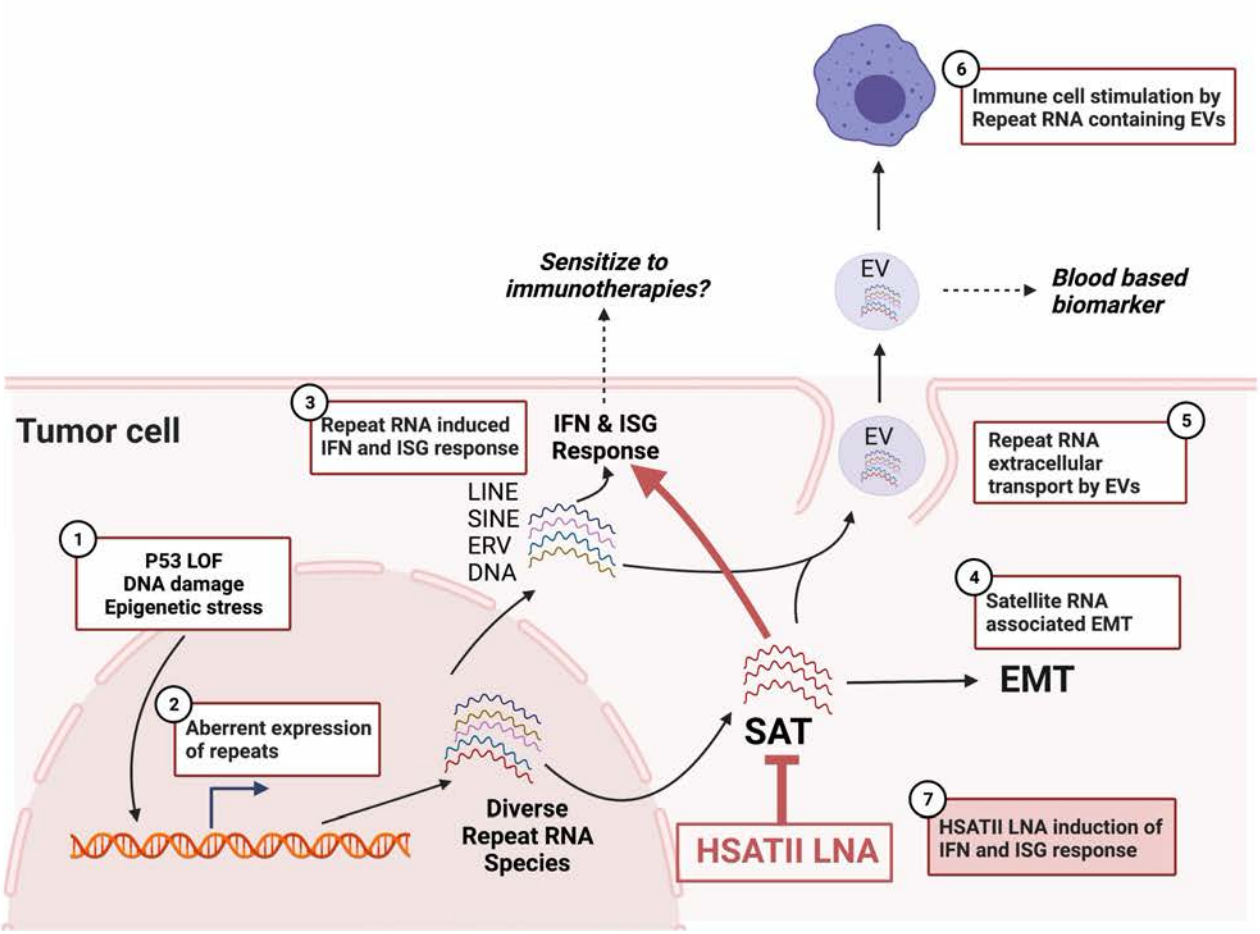


Figure 8. Working model of potential biologic and therapeutic implications of SAT RNA expression in epithelial tumor cells. The reported and hypothesized tumor cell autonomous and tumor microenvironmental effects of the aberrant expression of SAT RNAs are depicted in a working model. This model highlights the potential utility of SAT repeat RNAs as biomarkers and therapeutic targets in EOC.