KDM5 demethylases suppress R-loop-mediated "viral mimicry" and DNA damage in breast cancer cells

Authors

Lena Lau[1], Kurt Henderson[1], Ahu Turkoz[1], Sara Linker[2], Dörte Schlesinger[4], Brad Townsley[5], Brian Egan[5], Shoba Ragunathan, Robert Rollins[2], Xianju Bi[6], Zhijian Chen[6], Oleg Brodsky[2], Clifford Restaino[2], Murali Gururajan[2], Kristen Jensen-Pergakes[2], Anders Malarstig[3], Chames Kermi*[1], Paul Moore*[1] & Marie Classon*[#1]

*Contributed equally

[#] Corresponding Author

[1]Pfizer Center for Therapeutic Innovation, South San Francisco, CA

[2]Pfizer Oncology, La Jolla, CA

[3]Pfizer, Stockholm, Sweden

[4] Karolinska Institute, Stockholm, Sweden

[5] Active Motif, Carlsbad, CA

[6] UT SouthWestern, Dallas, TX

Keywords: RNA:DNA hybrids, R-loop, dsRNA, ADAR, pattern recognition receptor agonists, viral mimicry, innate immunity, surface MHC, DNA damage, compensatory epigenetic regulation, cancer therapy

Abstract

Tumors with low expression of Interferon-Stimulated Genes (ISG) and Antigen Presentation (AP) genes respond relatively poorly to current immunotherapies. One of the early hallmarks of cancer is DNA hypomethylation in genomic repeat regions, resulting in the expression of normally silenced endogenous “viral” elements. Such epigenetic changes have the potential to augment anti-tumor immune responses as well as reduce tumor cell fitness through the generation of aberrant nucleic acid species (NAS) and consequent activation of NAS-sensing pathways. Therefore, tumor evolution should favor additional selective events that suppress NAS generation, possibly yielding specific therapeutic vulnerabilities. Here, we show that the Lysine Demethylase 5 (KDM5) family of epigenetic regulatory enzymes suppress R-loop formation in genomic repeat regions in cancer cells. We find that KDM5 inhibition in luminal breast cancer cells results in R-loop-mediated DNA damage, reduced cell fitness and an increase in ISG and AP signatures as well as cell surface Major Histocompatibility Complex (MHC) class I, mediated by RNA:DNA hybrid activation of the CGAS/STING pathway. KDM5 inhibition does not result in DNA damage or activation of the CGAS/STING pathway in normal breast epithelial cells, suggesting that KDM5 inhibitors may enable a wide therapeutic window in this setting, as compared to STING agonists or Type I Interferons. These findings provide new insights into the interplay between epigenetic regulation of genomic repeats, R-loop formation, innate immunity, and cell fitness in the context of cancer evolution and therapeutic vulnerability.

Introduction

Cancer treatments have evolved along three main paradigms; cytotoxic chemotherapy, rationally-targeted agents and immune-based therapies. Still, innate and acquired drug resistance remain key limitations to improving outcomes in advanced cancer patients. With immune-based therapies, it is well documented that responses depend on an appropriate tumor microenvironment (TME). In recent years it has been shown that tumor-intrinsic expression of endogenous, virus-like elements can affect tumor cell viability and the TME by mimicking viral infection and activating pattern recognition receptors (PRR) such as TMEM173/STING and melanoma differentiation associated protein 5 (MDA5) – a process coined "viral mimicry" (1-4). Suppression of the expression of these repeat elements and viral mimicry may also mediate drug resistance to other therapies (5).

Throughout evolution, the human genome has been modified by waves of viral insertions, and a large fraction of the human genome is therefore composed of repetitive elements and retrotransposons, including interspersed long-terminal repeat (LTR)-based endogenous retroviruses (HERVs), non-LTR-based Short- and Long-Interspersed Nuclear Elements (SINEs and LINEs), as well as centromeric, telomeric and satellite repeats (6, 7). HERV, LINE and SINE elements are mostly silenced in healthy somatic cells but can be deregulated in cancer cells due to DNA hypomethylation in regions containing such elements (8, 9). Notably, aberrant expression of these elements is increasingly recognized as playing important roles, not only in cancer evolution and drug resistance but also in the development of other diseases (10).

Mechanistically, it has been shown that expression of HERVs, LINEs and SINEs can cause cells to appear virally-infected to the immune system (termed "viral mimicry"), in part through the formation of double-stranded RNAs (dsRNAs) that are detected by MDA5, thereby triggering a Type I Interferon (IFN-I) response that can decrease tumor cell fitness and increase immunogenicity. Consequently, during tumor evolution such a response may need to be blunted by the acquisition of mutations in genes that promote

IFN-I signaling, or by the possible deployment of viral restriction factors or compensatory epigenetic repressive mechanisms (3, 11, 12). Interestingly, the latter two non-mutational mechanisms may serendipitously present unique tumor-specific therapeutic opportunities. For example, it has been shown that the A-to-I-editing viral restriction factor ADAR (13) destabilizes inverted repeat SINE elements (IR-Alus) and prevents their detection by MDA5 (14). As a result, tumors that harbor increased expression of SINE elements (and consequently ISGs) are highly sensitive to ADAR loss (15). Notably, ADAR disruption has also been shown to lead to both intrinsic anti-tumor activity and enhanced immunotherapy response in preclinical mouse models (16, 17) and Figure S2a). Previous studies have also demonstrated that regulators of DNA methylation (1, 2, 18), as well as inhibitors of other epigenetic regulators such as histone methyltransferases (HMTs) (19-22) and histone lysine demethylases (KDMs) (23-27), can induce viral mimicry by reactivating repeat elements and increasing ISG signatures beyond a threshold level of tolerance--thereby affecting cell survival.

Notably, suppression of repeat transcription may also be responsible for the low levels of MHC class I on the surface of tumor cells (28). Reduced surface MHC I on tumor cells will decrease their intrinsic immunity but also limit responsiveness to therapies such as adoptive T cell therapies. In addition, peptides derived from the repeat genome have also been shown to be presented by tumor cells (29-32).

In addition to dsRNA activation of MDA5, viral mimicry can also be mediated through the CGAS/STING pathway, which is initiated by the sensing of DNA or RNA:DNA hybrid species (33, 34). Agonists of the STING pathway are currently being evaluated in clinical trials. The lysine demethylase 5 family (KDM5A-D), which removes histone 3 lysine 4 (H3K4) di- and tri-methyl modifications (35), has been suggested to blunt viral mimicry in tumor cells through the repression of STING expression or modification of signaling components downstream of STING (24-26). Here, we demonstrate that disruption of KDM5 activates viral mimicry in luminal breast cancer cell lines that are responsive to STING agonism. We also show that KDM5 inhibition specifically promotes viral mimicry in tumor cells and not in normal primary breast epithelial cells.

Mechanistically, we demonstrate that inhibition of KDM5 and the resulting increase in H3K4 tri-methylation in repeat regions lead to an increase in R-loop formation. R-loops are three-stranded nucleic acid structures that naturally form during cellular processes such as transcription, replication, and DNA repair when a nascent RNA transcript hybridizes to its DNA template, leaving a loop of single-stranded DNA (36). Although R-loops are crucial intermediates in normal cellular processes, if not properly regulated, they can also compromise genomic stability, induce DNA damage and trigger inflammatory responses through upregulation of interferon signaling. As such, dysregulation of R-loops and the cGAS-STING pathway have been implicated in both cancer and autoimmune diseases (37, 38). As an example, it has recently been demonstrated that disruption of the RNA:DNA helicase Senataxin (SETX), which can resolve R-loops, in cancer cells results in accumulation of R-loop-derived cytoplasmic RNA:DNA hybrids that can activate an innate immune response (39). Increases in transcription-replication conflicts (TRCs) and subsequent R-loop formation have also been linked to DNA damage and heightened sensitivity to PARP inhibitors (40-42). Here, we show that KDM5 disruption not only stimulates STING-mediated ISG and AP signatures as well as MHC class I presentation, but also induces a DNA damage response that is independent of the cGAS-STING-Type I IFN axis. In summary, our studies suggest that KDM5 inhibitors hold promise as stand-alone cancer therapies or in combination with immune or DNA damaging cancer therapies.

Results

Luminal breast cancer cell lines display features of "cold" tumors, but viral mimicry can be induced through the STING, MAVS and IFN I pathways

Immune-directed therapies can produce significant clinical benefit for some, but not all, cancer patients. Breast cancers are among the tumor types that are least responsive to immune-based therapies and display low ISG and AP gene expression signatures

compared to tumor types that respond well. Among the breast cancer molecular subtypes, luminal breast cancers not only display the lowest ISG and AP gene signatures (Figure 1 a-c, Figure S1a-b and table 1) but also exhibit a lower number of MHC class I molecules on their surface in comparison to triple negative breast cancer cell lines (Figure 1d) and tumor types that are more immunogenic (data not shown). Collectively, these features are indicative of "cold" tumors and most likely contribute to the poor response to immune-based therapies seen for this breast cancer subtype.

As mentioned above, loss of DNA methylation in genomic repeat regions of cancer cells is an early event in tumor development that potentially contributes to mutations, translocations and expression of oncogenes. However, de-repression of LINE, SINEs, HERVs and other repeat elements also promotes viral mimicry: generation of cytoplasmic NAS and activation of PRRs and IFN-I signaling pathways. Therefore, during tumor evolution various mechanisms can be engaged to counter the negative effects of viral mimicry and create an equilibrium that favors the tumor (43). To better understand the viral mimicry state in luminal breast cancer cells, we first examined a panel of cell lines for their ability to respond to PRR agonists or to IFN-I stimulation (see figure 1e for a simplified schematic). Using MHC class I surface levels as well as activation of an integrated Interferon-Stimulated Response Element (ISRE) reporter as read-outs, these studies showed that all tested luminal breast tumor cell lines respond to IFNβ (a type I IFN) to varying degrees (Figure 1f, upper panels and 1g, left panel, as well as S1c, left panel and S1e, upper panel). Most of the luminal lines also respond to MDA5/RIGI agonists, such as Poly-IC, or an *in vitro* transcribed inverted repeat SINE element (IR-Alu) (Figures 1f, lower panel and 1g, right panel as well as S1d, right panel and S1e, bottom 2 panels), but demonstrate a varied response to the STING agonist di-ABZI (Figures 1f, middle panel and 1g, middle panel, as well as S1c, right panel and S1e, 2nd panel). Collectively, these data show that most of the luminal breast tumor cell lines in this panel have intact signaling pathways downstream of PRRs, suggesting that PRR agonists could be used to increase the ISG and AP signatures and decrease cell fitness in luminal breast cancer cells. However, like Type I IFNs, PRR agonists are likely to indiscriminately induce such phenotypes in both tumor and normal epithelial cells, thereby

limiting the therapeutic window for such agents. This is exemplified by an increase in MHC class I surface expression on both tumor cells and normal human mammary epithelial cells (HMECs) in response to STING or RIGI/MDA5 agonists as well as to Type I IFN (Figure 1H).

Next, we evaluated factors implicated in upstream suppression of STING or MAVS signaling in human tumor cell lines or in suppression of tumor immune responses in mouse models (17), Figure S2a for summary). These included the A-to-I editing enzyme ADAR and the histone demethylase KDM5, shown to suppress signaling through MDA5/MAVS and STING, respectively (14, 24). The HCC1428 luminal breast cancer cell line was chosen as a model for most of these studies because it responds well to STING and MDA5 agonists as well as to IFN-I (Figure 1f-h and S1c-e). Consistent with the effects of MDA5 agonists, CRISPR-mediated disruption of ADAR in HCC1428 cells (figure S2b) results in an increase in both ISRE-reporter activity (Figure 2a, upper panel) and surface presentation of MHC class I molecules (Figure 2a, lower panel). Similar results were also observed in T47D and MCF7 luminal breast cancer lines (data not shown). Likewise, the KDM5-specific inhibitor C48 (44) increases ISRE reporter expression (Figure 2b, upper panel) as well as ISG and AP signatures (Figure 2c) in HCC1428 cells. MHC class I surface expression is also increased following C48 exposure in HCC1428 cells (figure 2b lower panel) and other luminal breast lines (data not shown). As previously shown in other cell lines, inhibition of KDM5 results in viral mimicry mediated through the CGAS/STING pathway (Figures 2d-e and S2c-e). Consistent with other studies (24), KDM5-dependent effects on viral mimicry in HCC1428 cells is primarily mediated by KDM5C (Figures S2f- and h). We also observe that the effect on surface levels of MHC class I is mostly mediated by the KDM5C paralog in the HCC1428 cell line model (Figure S2g).

To evaluate whether viral mimicry activation and the subsequent increase in surface MHC class I can affect TCR-driven T cell activation, we developed a MCF7:Jurkat T cell co-culture system (Figure 2f). Briefly, luminal MCF7 breast cancer cells engineered to over-express the NY-ESO1 antigen were cultured with Jurkat cells expressing a NY-ESO1-specific T-cell receptor (TCR) and a luciferase reporter under control of the Nuclear Factor of Activated T cells (NFAT) promoter that is activated upon TCR activation. Using this

system, we demonstrated that disruption of ADAR in MCF7:NY-ESO1 cells results in robust activation of the NFAT reporter in co-cultured Jurkat:NY-ESO1 TCR cells (Figure 2g). Treatment with the KDM5 inhibitor C48 prior to mixing these cells with NY-ESO1 TCR Jurkat cells results in a dose-dependent increase in NFAT reporter activity (Figure 2h), which is reversed upon CGAS or STING depletion in MCF7:NYESO1 cells (Figure 2i). Similarly to the effects on ISG and AP signatures, induction of the Jurkat NFAT reporter is largely dependent on the disruption of the KDM5B and C paralogs in the MCF7:NY-ESO1 cell line (Figure 2j and S2f). Together, these experiments demonstrate that PRR and IFN-I signaling can be augmented in luminal breast cancer cells by direct agonists or by the disruption or inhibition of factors that may repress these pathways upstream.

The effect of KDM5i on cell fitness is not mediated through activation of CGAS/STING or Type I IFN signaling

In addition to increasing the immunogenicity of tumor cells, viral mimicry can also reduce cell fitness. This relationship is well-documented in the MDA5/MAVS dsRNA-sensing pathway. In brief, MDA5/MAVS activation by dsRNA induces Type I IFN signaling through the JAK/STAT pathway, resulting in ISG induction. This includes an increase in Protein Kinase R (PKR), which binds dsRNA and signals through eIF2α to attenuate global translation and drive apoptosis (see Figure S2i for model). Consistent with this mechanism of action, treatment with Ruxolitinib, a Janus Kinase (JAK) inhibitor that blocks the IFN-I response, rescues the effects of ADAR disruption not only on ISG induction, MHC Class I presentation, but also overall cell fitness (Figures 2k-l and S2j). This direct relationship between ADAR, dsRNA levels, IFN signaling and cell fitness is further illustrated by the observation that cancer cell dependencies on ADAR, as reported in DepMap, correlate strongly with baseline ISG and AP signatures (figure S2k for summary) (15).

In contrast, Ruxolitinib treatment does not reverse KDM5 inhibitor-mediated loss of cell fitness (Figure 2o) despite blunting induction of ISG/AP signatures and activation of the NFAT reporter in the MCF7:NYESO1/Jurkat co-culture system (Figures 2m-n; S2l-m).

Similarly, genetic disruption of CGAS or STING upstream of Type I IFN signaling does not reverse C48-mediated cell fitness loss in HCC1428 cells (Figure S2n) but does abrogate ISG induction (Figure 2d and S2c-e). In summary, these data demonstrate that KDM5, unlike ADAR, regulates cancer cell fitness and viral mimicry through distinct mechanisms.

KDM5 inhibition induces viral mimicry and a type I IFN response in luminal breast cancer cells, but not in normal breast epithelial cells

Upstream regulators of PRRs may more selectively affect tumor cells compared to direct PRR agonists. For example, sensitivity to ADAR disruption is observed in tumors that exhibit relatively high ISG signatures (15) and Figure S2k), yielding a therapeutic window. To investigate whether KDM5 inhibition would provide a larger therapeutic window than STING agonists, we compared the effect of C48 on ISG and AP signatures as well as cell fitness phenotypes in breast tumor cells and normal breast epithelial cells (HMECs). Although KDM5 inhibition increases H3K4me3 and CGAS/STING levels in both HCC1428 tumor cells and HMECs (Figures 3a and b, S3a), treatment of HMECs with increasing concentrations of C48 did not induce a dose-dependent increase of MHC class I surface expression in HMECs as compared to luminal breast tumor lines (Figure 3c, Figure S3b and c). Furthermore, KDM5 inhibition did not elevate ISG and AP signatures in HMECs as compared to HCC1428 cells, or other KDM5 inhibitor-sensitive luminal breast cancer (Figures 3d , S3d-e and data not shown). Together, these data show that the effect of KDM5 inhibition on ISG and AP signatures is tumor cell-specific in this tissue and not solely dependent on increased STING levels as previously suggested (24).

The STING-independent, adverse effects of KDM5 disruption on cell fitness is also specific to tumor cells and is largely dependent on KDM5C (Figures 3e, S3f-i and data not shown). KDM5s have been suggested to play a role in DNA damage and repair (DDR) (45, 46). To investigate whether a DDR phenotype contributed to the reduced cell fitness phenotype seen following KDM5 inhibition, we used gamma (γ)-H2AX as a

measure of DNA damage. These experiments showed that γ-H2AX staining increases in the HCC1428 tumor cells, but not in HMECs, following exposure to C48 (Figure 3f). Taken together with the data presented in figure 2 suggesting that the effect of KDM5 inhibition on cell fitness is independent of STING and IFN-I, these data suggest that the cell fitness phenotype is driven by DNA damage rather than an IFN-I response. Collectively, these data also show that disruption of upstream regulators of PRR signaling, such as KDM5, selectively affects tumor cells, and as such provides a therapeutic window for cancer treatment.

KDM5 inhibition induces R-loop formation in areas of the repetitive genome that harbor increases in H3K4me3

As mentioned, previous studies have suggested that disruption of KDM5 family members activates the STING pathway partially by upregulating the expression of STING itself (24). However, our experiments show that STING levels are induced by KDM5 inhibition in both normal and tumor cells (figure 3b), suggesting that increased levels of STING cannot be the only mechanism by which KDM5 disruption results in induction of innate immune phenotypes. Furthermore, our experiments also suggest that the effect of KDM5 on cell fitness is not mediated through STING or type I IFN (Figure 2o, Figure S2n), but rather through a more complex mechanism involving DDR.

De-regulated control of genomic R-loops (36, 37) has been shown to result in DNA damage and stimulation of the cGAS-STING pathway via excised RNA:DNA hybrids (39) Figure 4a for schematic). While R-loops are crucial regulators of specific biological processes, if not properly regulated, they can also pose a threat to genomic integrity, contributing to cancer progression and chemotherapy resistance; thus, R-loops represent a potential target for novel cancer therapeutics (47). In order to evaluate whether accumulation of R-loops can explain the phenotypes observed following KDM5 disruption in luminal breast cancer cells, we first stained control or C48 treated cells with GFP-tagged RNaseH1 protein, which selectively binds RNA:DNA hybrids (48). These experiments show that KDM5 inhibition results in an increase of GFP+ cells, indicative of

increased RNA:DNA hybrid formation in HCC1428 cells (Figure 4b for representative images and quantification). This increase is not observed in HMEC cells (data not shown). Furthermore, a partial depletion of XPF, an endonuclease involved in nucleotide excision repair that facilitate the excision of RNA:DNA hybrids from R-loops (39), attenuates C48-mediated induction of viral mimicry, as measured by ISRE-GFP reporter activity and ISG mRNA expression in HCC1428 cells (Figures 4c-e and S4a left panel). Partial loss of XPF also attenuates the dose-dependent increase of MHCI component mRNA expression and surface levels induced by C48 (Figures 4d, right panel, 4f, and S4a right panel). Collectively, these data suggest that the viral mimicry phenotypes induced by KDM5 inhibition in sensitive breast cancer cell lines are initiated by an increase in the generation of R-loop-derived RNA:DNA hybrid species.

Transcription replication conflicts (TRCs), often magnified in cancer, occur when the processes of DNA replication and transcription co-occur on the same DNA template and are intricately linked to the initiation/persistence of R loops (49). Transcriptional output is linked to epigenetic modifications, and it has previously been shown that the *de novo* DNA methyltransferase DNMT3b and other epigenetic regulators may mechanistically restrict R-loop mediated DNA damage (50, 51). Increases in H3K4 methylation is linked to increases in transcriptional output (52), possibly creating an epigenetic environment conducive to an increase in R-loops in tumor cells that already display an increase in TRC. To explore a possible relationship between C48-induced increases in H3K4 trimethylation and genomic R-loops, we first performed Cleavage Under Targets and Tagmentation (CUT-and-Tag) experiments (53) using an H3K4me3-specific antibody. This method combines antibody-targeted cleavage by Tn5 transposase with parallel DNA sequencing to evaluate regions in the genome that display changes as detected by the antibody used. KDM5 inhibition induces global changes in H3K4me3; therefore, we added equal numbers of *Drosophila melanogaster* spike-in nuclei to both DMSO and C48-treated samples to enable downstream normalization (see Figure S4b for experimental set-up). Indeed, mapping reads to the *Drosophila* genome reveals that C48-treated samples consistently harbor less *Drosophila* reads as compared to DMSO-treated samples, instead of the expected equal counts (Figure S4c); therefore, H3K4me3 peaks

would be undercounted in C48-treated samples without spike-in normalization (Figure S4d and e as examples). Therefore, peak counts were normalized by the spike-in reads prior to differential peak calling. As expected, we observed an increase in H3K4me3 in promoter regions upon C48 treatment of HCC1428 cells (Figures 4g and S4e, yellow bar). Additionally, we observed an increase in H3K4me3 in intergenic and intronic regions, which often harbor insertions of repetitive elements (54), in C48-treated HCC1428 samples (Figures 4g and s4e, blue and orange bars). Given that R-loop generation has been described in telomeric, centromeric and other repeats (55, 56), we expanded the analysis of H3K4 trimethylation to repeat regions. This analysis revealed increased H3K4me3 reads mapping to both sub-telomeric and centromeric regions following C48 treatment of HCC1428 and other luminal breast cancer cell lines (Figures 4h-i and data not shown). Interestingly, KDM5 has previously been identified as a telomere-associated protein (57).

In addition to observing C48-mediated increases of H3K4me3 specifically in telomeric repeats, when subdividing the genome into annotated repetitive regions and non-repeat regions, we observed that there was a significant increase in H3K4me3 reads mapping to repeat regions in general in HCC1428 cells (Fig 4j). The results of these CUT-and-Tag experiments are consistent with those derived from the analysis of H3K4me3 reads from chromatin immunoprecipitation (ChIPseq) experiments (data not shown). We subsequently separated the analysis by repeat classes present in the human genome, and observed that SINE elements exhibited the highest coverage of H3K4me3 in C48-treated HCC1428 samples and were significantly increased in comparison to untreated samples (Figures 4k and S4f, pink, Figure 4l). We also observed increases in H3K4me3 in other genomic repeats such as LINEs and LTRs in HCC1428 cells treated with C48 (Figures 4k and S4f, yellow and dark blue respectively, Figure 4l) and in other luminal breast cancer cell lines following C48 treatment (data not shown). Further analysis distillation of repeat classes into SINE, LINE and LTR subfamilies revealed that KDM5 inhibition results in increased H3K4me3 mainly in the AluY subfamily, as well as LINE1/LINE2 subfamilies and a few LTR subfamilies (Fig S4g).

To investigate a possible relationship between genomic regions that harbor increased H3K4me3 upon KDM5 inhibition and R-loop prevalence, we performed CUT-and-Tag experiments using an antibody (S9.6) that detects RNA:DNA hybrids in the genome. As in the H3K4me3 CUT-and-Tag experiments, *Drosophila* "spike-ins" were added and used for normalization of the data (Figure S4b for outline). Similar to what was observed for H3K4me3, KDM5 inhibition increases S9.6 reads mapping to telomeric and centromeric repeats in HCC1428 cells (Figure S4h). Since we also observed H3K4me3 increases mapping to other repeat classes following KDM5 inhibitor exposure, we expanded the S9.6 CUT-and-Tag analysis to the rest of the repeat genome. As a group, repeats that harbor increased H3K4me3 also display a significant increase in S9.6 reads upon C48 treatment (Figure S4i, middle panel) and this increase is greater than what is observed in repeats that do not display increases in H3K4me3 (Figure S4i, left panel). Collectively, these data suggest that KDM5 inhibition increases R-loop formation in repetitive regions of breast cancer genomes. Repeats that harbor KDM5 inhibitor-mediated increases in both H3K4me3 and S9.6 reads traverse multiple repeat classes, including SINEs, LINEs and LTRs (Figure 4m). Additional analysis of RNA sequencing data also suggest that the most significant increases of S9.6 reads are observed in repeat classes that not only display elevated H3K4me3 levels, but also exhibit higher transcriptional activity following KDM5 inhibition (Figure S4i, right panel). Together, these analyses demonstrate that KDM5 inhibition and subsequent increases in H3K4me3 correlate with an increased prevalence of R-loops in genomic repeats in tumor cells.

Collectively, we show that the KDM5 demethylases regulate viral mimicry and DNA damage in luminal breast cancer cells. Mechanistically, KDM5 inhibition increases R-loop prevalence in repetitive genomic regions, triggering activation of the CGAS/STING pathway and consequent upregulation of ISGs and MHC class I surface expression. KDM5 inhibitors also reduce cell fitness, not mediated through induction of type I IFN signaling, but rather because of R-loop-generated DNA damage (Figure 5). As mentioned above, R-loops can form during transcription and have the potential to impede replication fork progression, leading to replication stress. This interference can result in DNA damage, necessitating repair mechanisms to maintain genomic integrity. Conversely,

replication stress itself might induce conditions conducive to R-loop formation, suggesting a bidirectional relationship. The study of chromatin dynamics and its impact on DNA damage and repair mechanisms has garnered significant attention in recent years. Several studies have implicated KDM5 demethylases in DDR (46, 58) and a recent study revealed a potential role for KDM5 in maintaining the delicate balance between transcriptional regulation and replication to preserve DNA integrity (52). Collectively, such observations suggest that KDM5 inhibitors, or other factors whose disruption or inhibition results in tumor-specific increases in R-loops, could pair well with PARP inhibitors and other DNA damage-promoting drugs.

Importantly, we also show that KDM5 inhibition selectively affects cell fitness and immunogenicity phenotypes of cancer cells without having similar effects in primary breast epithelial cells, potentially resulting in a therapeutic window for KDM5 inhibitors or degraders as single agents. Our data also show that KDM5 disruption or inhibition can activate T cells in a co-culture system. Important for possible combinations with TME enhancers, preliminary data suggest that there are no major negative effects of KDM5 inhibitors on primary T cells or macrophages (data not shown). The STING-independent KDM5 inhibitor effect on cell fitness in luminal breast cancer cells and previous observations that KDM5 disruption affects the survival of drug-tolerant persister cells (59, 60) suggest that there could be combinatorial potential, not only with agents that affect the TME, but also with other standard-of-care agents. Future studies will have to explore the relevance of these findings in other cancer subtypes. Preliminary studies suggest that hematologic cancer cells that are sensitive to KDM5 inhibitors also display markers of type I IFN signaling activation (data not shown). Collectively, these findings expand on KDM5 inhibitors or degraders as promising candidates for single-agent cancer therapy, with further potential to enhance the efficacy of immunotherapies as well as DNA damaging agents.

Materials and Methods

Cell Lines

HCC1428, SKBR3, and MCF7 were purchased from ATCC. Jurkat expressing a luciferase reporter construct driven by an NFAT-response element were purchased from Promega (cat J1621). HMEC were purchased from Lonza.

Cell Culture Conditions

HCC1428, SKBR3, MCF7, and Jurkat cells were maintained at 37°C with 5% CO_2 RPMI1640 containing 10% heat-inactivated fetal bovine serum (FBS, Sigma) and 4.5g/L glucose. HMEC cells were cultured in MEBM Basal Medium supplemented with MEGM SingleQuots (Lonza CC-3150). Experiments using HMEC cells were completed before 15 passages.

Inhibitors/Agonists Used

KDM5 inhibitor C48 was synthesized at Pfizer. KDM5 inhibitor C70 was purchased from XcessBio (cat M60192). Recombinant IFN-β was purchased from PBL Assay Sciences (cat 11410-2). STING agonist di-ABZI was purchased from Selleck Chemicals (cat S8796). JAK inhibitor ruxolitinib (cat tlrl-rux) and low molecular weight polyIC (cat tlrl-picw) were purchased from InvivoGen. IR-Alu was synthesized by reverse transcription.

Cell Line Generation

ISRE-GFP containing cell lines: Lentivirus containing ISRE-GFP were purchased from GenTarget (cat LVP937-N). Cells were seeded to 90-100% confluence in 6-well plates. Lentiviral particles were combined with 2ml media at ~15-20% MOI. Viral media was removed the day after and replaced with fresh media. Selection in media supplemented with 200ug/ml geneticin (Gibco) was performed the day after, for 7 days. Cell line pools were used for experiments. NY-ESO1: MCF7 cells were transduced with lentiviral particles expressing NY-ESO1, followed by stable selection in the presence of 1 ug/ml puromycin. NY-ESO1 TCR/CD3: Jurkat cells expressing NFAT-response element-driven luciferase reporter were cultured as described above supplemented with 200 ug/ml

hygromycin B. Cells were transduced with a lentiviral construct expressing NY-ESO1 TCR and selected with 2 ug/ml puromycin to generate a stable pool of Jurkat NY-ESO1 TCR/CD3 cells. Pools were used for experiments that were run with no selection media.

Immunoblotting

Cells were lysed in RIPA buffer (50mM Tris HCl, pH 7.4, 150mM NaCl, 0.5% deoxycholate, 0.1% sodium dodecyl sulfate, 1% NP-40; Thermo 89900) with proteinase/phosphatase inhibitor cocktail (Thermo) and benzonase (Sigma). Whole cell extracts were sonicated via Active Motif Pixul multi sample sonicator, then assayed for protein concentration via BCA (Thermo). Samples were leveled and denatured for 10 minutes at 70ºC, then mixed with 4X LDS sample buffer (Thermo NP0008) and and 10X sample reducing agent (Thermo NP0009). Samples were loaded at 1 µg for H3 and H3K4me3, and at 20 µg for all other proteins, then separated by electrophoresis on a 4-20% precast polyacrylamide gel (Bio-Rad 4561096), transferred using Bio-Rad TransBlot Turbo to nitrocellulose membranes, and blocked in blocking buffer (Rockland NC0168431) for 20 minutes. Membranes were incubated with primary antibodies overnight at 4ºC on a rocking platform, followed by washing in 0.1% Tween/PBS. Membranes were incubated with appropriate HRP-linked secondary antibodies at room temperature (RT) for 1h and washed 3X before signal detection. Membranes were developed by chemiluminescence using Bio-Rad ECL reagent (cat 1705062). Images were acquired with the Bio-Rad ChemiDoc imaging system. Primary antibodies used were: ADAR (Invitrogen MAB-31609), Actin (Millipore A5541), H3K4me3 (Cell Signaling 9751), H3 (Cell Signaling 4499), STING (Cell Signaling 13647), Tubulin (Cell Signaling 2128), cGAS (Cell Signaling 15102, 79978), KDM5A (Bethyl Labs A300-897A), KDM5B (Bethyl Labs A301-813A), KDM5C (Bethyl Labs A301-034A), XPF (Cell Signaling 13465), and Calnexin (Cell Signaling 2679). Secondary antibodies used were anti-mouse IgG, HRP linked and anti-rabbit IgG (Cell Signaling 7076), HRP linked (Cell Signaling 7074).

Immunofluorescence

Cells were plated in 96-well black Cell Carrier plates (Perkin Elmer 6055302) 2-3 days before harvest. On day of harvest, plates were washed with PBS and fixed in ice-cold

methanol at -20°C for 10 minutes. After 3x PBS wash, cells were permeabilized with 0.24% Triton X-100 in PBS, 3 minutes on ice. Cells were then incubated in 50 µl of 1x RNaseH buffer supplemented with 20mM $MgCl_2$, with or without RNaseH diluted 1:50 (NEB M0297), for 3 hours at 37°C. After 3x PBS wash, cells were incubated in blocking solution (3% BSA in PBS) for 1 hour at room temperature. Cells were then incubated with GFP-RNaseH1::D210N protein at 0.05 µg/ml or anti-γH2AX antibody (Cell Signaling 2577) at 1:500, diluted in blocking solution, overnight at 4°C. The day after, plate was incubated at 37°C for an additional 1.5 hrs to intensify signal. Cells were washed 3x in PBS, 5 min rocking, then incubated in secondary antibody at 1:1000 to probe γH2AX (AlexaFluor 488 anti-rabbit IgG, Thermo A-21206) and Hoescht (Invitrogen H3570) at 1:2000, diluted in blocking solution, for 1 hr at room temperature. After final 3 washes in PBS (5 min each rocking), 100 µl PBS was left in wells for image acquisition.

Image acquisition and analysis

Images were acquired on a PerkinElmer Operetta CLS equipped with 63x water immersion objective, using Two Peak autofocus and confocal optic mode. Image analysis was performed using Harmony (version 4.9) software and exported with equal contrast adjustment. The HOECHST 33258 channel was used to identify nuclei using the Find Nuclei module in the analysis sequence, using method C with a common threshold of 0.40, area of 30µm, splitting coefficient of 9.8, individual threshold of 0.30, and contrast of 0. The EGFP channel was used to identify the cytoplasm region using the Find Cytoplasm module in the analysis sequence, using the nuclei identified in the previous step, via method A, and an individual threshold of 0.15. Cell regions of Cell and Cytoplasm were output in this step. Micronuclei were identified using the HOECHST 33258 channel and the module Find Micronuclei, using nuclei identified in the previous step, cell region Cell, and method B with a detection sensitivity of 0.10, a splitting sensitivity of 0.50, default background correction coefficients, typical diameter of 2.5 µm, and max diameter of 5 µm. Mean fluorescence intensities for GFP-dRNH1 and yH2AX were calculated for each subcellular region and channel of interest using the Calculate Intensity Properties module, and were exported along with micronucleus and nucleus

counts using the Define Results module. For each experiment, 9 fields were analyzed in triplicate for 2 independent experiments.

Cellular fitness/growth curve analysis

Cells were plated into 24-well, 12-well or 6-well plates and treated as indicated. For culture over 7 days, cells were split on day 7 of experiment. All cells were split at the same ratios regardless of confluence. Whole-well phase-only images were acquired on IncuCyte (Sartorius) at the indicated times. Confluence was calculated by IncuCyte software after setting the following parameters: segmentation adjustment = 1.5, area min = 600 μm^2, eccentricity max = 0.990.

To calculate doubling times for statistical analysis, confluence for each day was normalized to that of day 0 in each sample. Log_2 of each normalized confluence timepoint was calculated to generate a linear curve (excluding day 0 timepoint), and slope was calculated from each line. Doubling time was calculated as 1/slope of each line.

Growth Inhibition Assays

Assays were performed by ChemPartner Co., Ltd (Shanghai, China). Briefly, breast cancer cell lines were treated with varying concentrations of C48, and viability was determined via CyQUANT assays, IC_{50} values were calculated, and inhibition was calculated as area under the IC_{50} curve (AUC) after 21 days of treatment.

Cell Assays/Drug Treatments

For KDM5 inhibitor experiments, cells were treated with KDM5-C48 or C-70 at the indicated concentrations for 7 days (unless otherwise noted). In experiments using siXPF, cells were treated with C48 at the indicated concentrations 1 day after transfection for a total of 5 days. In JAK inhibition experiments, cells were treated with 1 μM ruxolitinib for the duration of the experiment. Media and drugs were changed every 2-3 days. For agonist experiments, cells were treated with recombinant IFN-β at 1000 U/ml, di-ABZI at 1 μM, or transfected with 2.5 μg/ml low molecular weight polyIC or 300 ng/ml IR-Alu using RNAiMAX for the last 24 hours.

siRNA Knock-Down

For individual knockdown, HCC1428 cells in 6-well plates were transfected at 25 nM final concentration with 5 µl RNAiMAX (Thermo) per well in OptiMEM. siRNAs used (39)) were:
siLuciferase (siLuc): 5'- CUUACGCUGAGUACUUCGA-3'
siXPF: 5'- ACAAGACAAUCCGCCAUUA-3'

CRISPR-mediated Gene Disruption

crRNAs (ordered from IDT) were combined with tracrRNA (IDT 1072533) at equal concentration (200 µM each), heated to 95°C for 5 min and allowed to cool to room temperature to anneal. 0.3 µl Cas9 protein (IDT 1081059) was combined with 0.2µl buffer R from Neon Transfection System 10 µl Kit (Thermo MPK1096); 0.5 µl of the Cas9 mix was then combined with 0.5µl of annealed RNA duplex or sgRNA (IDT) for 10-20 min at room temperature. 2 µl of Alt-R Electroporation Enhancer, 10.8 µM stock concentration (IDT 1075916) was added to the mixture, followed by 9ul of 1.2x10^5 cells resuspended in buffer R. 10 µl of cell/sgRNA/Cas9 mixture (approx. 1x10^5 cells) was then electroporated using Neon Transfection System (Thermo), setting #3, and plated into 6-well plates. For electroporation of 1x10^6 million cells, all components were multiplied by 10 and Neon Transfection System 100 µl Kit (Thermo MPK10096) was used, and cells were plated in 10cm plates. For CRISPR disruption of ADAR , cells were harvested for immunoblot 4 days post-electroporation, and for flow cytometry and RNA isolation 7 days post-electroporation. For KDM5 paralog CRISPR disruptions, cells were harvested 7 days post-electroporation for downstream analysis. For cell grow growth analyses via IncuCyte, the day of electroporation was designated day 0. For STING/CGAS CRISPR mediated disruptions, cells were cultured for 3 days before KDM5 inhibition. Individual Cas9-mediated Gene disruptions were performed for each experimental replicate. crRNA/sgRNA sequences are the following:

crRNA:
sgHPRT1i (intronic cutting control): 5'-AATTATGGGGATTACTAGGA-3'
sgADAR #3: 5'-ATACCTGAACACCAACCCTG-3'
sgADAR #4: 5'-AGCCGAATGCCATCCCACGT-3'

sgTMEM173 #1: 5'-GGCAAACAAAGTCTGCAAGG-3'
sgTMEM173 #3: 5'-CCTCACCCTGGTAGGCAATG-3'
sgMB21D1 #2: 5'-TAATAAGAAGTGTTACAGCA-3'
sgMB21D1 #3: 5'-GAGCTACTATGAGCACGTGA-3'

sgRNA:
sgKDM5Ai (intronic cutting control): 5'-GAGGTAATGGAGTACCTAAG-3'
sgKDM5A #1: 5'-ATCTACTCTGAAGATCCCTG-3'
sgKDM5A #4: 5'-TCTGTGAACTCCTCCCAACT-3'
sgKDM5Bi (intronic cutting control): 5'-AGGACTGGGCGATTAGACTG-3'
sgKDM5B #1: 5'-AACCCTGGTTAAAACCACTG-3'
sgKDM5B #5: 5'-AACCCCATGATATTCCCCAG-3'
sgKDM5Ci (intronic cutting control): 5'-AGGACTGGGCGATTAGACTG-3'
sgKDM5C #3: 5'-CGCCTTGAGCGCATACACAG-3'
sgKDM5C #4: 5'-TGGCACACCATGGGACATGA-3'

RNA isolation and qRT-PCR

Cells were harvested on the indicated days and total RNA was isolated using Qiagen RNeasy kit following manufacturer's protocol and subjected to DNase treatment. For qRT-PCR analysis, 10ng of RNA was subjected to one-step cDNA synthesis and amplification using 1x Taqman Fast Virus 1-Step Master Mix (Applied Biosystems 4444436), along with 1x Taqman probes. Samples were run in duplicate in 384-well plates (Applied Biosystems 4483321) on ViiA7 (Applied Biosystems) using the following conditions: 48°C for 15 min (cDNA synthesis), 95C for 20 sec (inactivation/denaturation), followed by 95°C for 3 sec (denaturation) and 60°C for 1 min (elongation) for a total of 40 cycles. Delta Ct was calculated from subtracting Ct values of gene of interest from B-actin Ct values. Raw values are expressed as 2^delta Ct. Probes used were purchased from Thermo and are as follows: B-actin (Hs01060665_g1), OAS2 (Hs04185073_m1), IFI44L (Hs0019915_m1), IFI44 (Hs00951349_m1), ISG15 (Hs00192713_m1), IFIT1 (Hs01911452_s1), TMEM173 (Hs00736955_g1), MB21D1 (Hs00403553_m1), IFNL1

(Hs00601677_g1), CXCL10 (Hs00171042_m1), B2M (Hs00187842_m1), HLA-B (Hs00818803_g1), ERCC4 (Hs01063530_m1).

RNASeq Sample Generation and Analysis

Cells were treated with DMSO or 2.5 μM C48 for 2 or 7 days. RNA was prepared as described above. Samples were generated in triplicate. RNA was randomly primed (not poly(A)-selected), rRNA depleted, and sequenced via Illumina NovaSeq using 150bp paired-end read sequencing with a sequencing depth of 100 million reads (by Azenta Life Sciences). RNAseq fastq files were aligned to the human GENCODE v28 transcriptome using STAR aligner (v2.6.0c) and gene expression was quantified using RSEM (v.1.3.0). Data4Cure was utilized for analysis (https://www.data4cure.com). ISG and AP gene lists were generated as described (see GSVA analysis).

CUT-and-Tag Sample Generation and Analysis

Cells were treated with DMSO or 2.5 μM C48 for 5 days. Samples were generated in duplicate, cryopreserved, and sent to Active Motif for downstream preparation. Cryopreserved cells were thawed from -80 ^{o}C in a 37 ^{o}C water bath with gentle movement until no ice was visible. Cell samples were then gently pipetted up and down to resuspend cells and all liquid was transferred to 2 ml round bottom snap-cap microcentrifuge tubes (Fisher 14-666-315). Samples were centrifuged at 600 G in a centrifuge at 4°C for 3 minutes. Liquid was carefully removed with a pipettor, taking care to avoid cell pellets. Samples were then resuspended in DIG-Wash buffer with Proteinase Inhibitor cocktail and processed according to the CUT&Tag-IT Assay Kit, Anti-Rabbit (Active Motif 53160) protocol for H3K4me3 CUT&Tag samples and CUT&Tag-IT® R-loop Assay Kit (Active Motif 53167) for R-loop samples, with the addition of Drosophila spike-in nuclei and spike-in antibody as follows: Drosophila spike-in nuclei CUT&Tag-IT® Spike-In Control, Anti-Rabbit (Active Motif 53168) for H3K4me3 CUT&Tag and CUT&Tag-IT® Spike-In Control, R-loop (Active Motif 53174) were thawed from -80°C in a 37°C water bath with gentle movement until no ice was visible. Spike-in nuclei were gently pipetted up and down to resuspend and homogenize the suspension. Spike-in nuclei were added to each sample at a ratio of 1:10 spike-in nuclei to cells, and each were added to ConA beads. Cells,

nuclei and ConA beads were mixed by gentle pipetting and strip tubes containing the samples were placed on a room temperature nutator for 10 minutes. Spike-in antibody was added to each sample according to the kit instructions at the same time as the primary antibody, H3K4me3 (Active Motif 39160) or S9.6 (included in R-loop CUT&Tag kit) was added to the cell sample/spike-in nuclei mix. The CUT&Tag-IT R-loop Assay Kit protocol was followed as per the manufacturer's instructions for all remaining steps for R-loop samples and CUT&Tag-IT Assay Kit, Anti-Rabbit for the H3K4me3 samples. NGS libraries were evaluated on a TapeStation 4150 (Agilent, Santa Clara, CA) using High Sensitivity D1000 ScreenTapes (Agilent 5067-5584) and pooled for equal sequencing depths targeting approximately 30 million reads per sample. Sequencing was conducted on a NovaSeq6000 (Illumina, San Diego, CA).

Samples were aligned to hg38 and dm6 genome using Bowtie2 (v2.3.4.1). Samples were normalized according to Active Motif's normalization protocol. Briefly, uniquely aligning *Drosophila* tags pulled down by H2Av were counted. Normalization factor was generated by dividing *Drosophila* counts of sample with lowest tag count by the *Drosophila* count of each sample. To find location of peaks in the genome, peak positions were first called per sample using Homer (http://homer.ucsd.edu/homer/), then merged into a set of consensus genomic positions. They were then annotated, and peak scores were calculated per sample. To count reads aligning to repetitive elements, RepEnrich2 was utilized (https://github.com/nerettilab/RepEnrich2). Repeat reference was downloaded from UCSC Genome Browser (hg38; https://genome.ucsc.edu/). To calculate percent reads mapping to repeat regions, sum of total counts output by RepEnrich2 was divided by number of reads that align concurrently at least once. Differential analysis was performed using edgeR, using normalization factors as the library size. Counts per million (CPM) was also calculated using normalization factors as the library size. To calculate percentage of reads in subtelomeric regions, samples were aligned to subtelomeric sequences (61) using Bowtie2. Centromeric sequences were downloaded from UCSC Genome Browser and Bowtie2 was used to determine percentage of reads that align to centromeric regions.

GEO numbers

GSVA Analysis

To obtain enrichment scores, gene set variation analysis (GSVA) was performed using the GSVA R package (62)version 3.16 with Gaussian kernels. IFN/AP gene sets are provided in supplementary table 1. The DepMap 22Q2 CCLE dataset (https://depmap.org/portal/) were utilized as the gene expression data. GSVA scores were plotted by cancer type, as well as breast cancer subtype as defined by(63).

Gene dependency

Gene dependency (CERES) scores were obtained from DepMap, 22Q2 (https://depmap.org/portal/).

Co-Culture Experiments

The day before co-culture, 40,000 MCF7::NY-ESO1 cells were plated in white 96-well plates (Corning 3917) in triplicate. The next day, 100,000 Jurkat::NY-ESO1 TCR cells were plated in each well in 75 µl fresh media supplemented with 25mM HEPES (Gibco 15630-080). After 5 hrs of co-culture, plate was taken out of the incubator and rested at room temperature for 5 min. 75 µl of BioGlo Luciferase Reagent (Promega G7940) was added to each well incubated for 5 min shaking before reading on EnVision 2104 (Perkin Elmer) on default luminescence settings.

Flow Cytometry

Cells (at least 100,000) were dissociated using accutase (Gibco A1110501) and transferred to 96-well 1ml plates (Greiner Bio-One 780261). After spinning down 5 min at 400g, cells were incubated in 50 µl FACS buffer (1x PBS + 2% FBS) containing antibody (5 µl) and live/dead stain (1 µl of 1:5 dilution, Thermo L34975) for 30 min at 4°C. Cells were then washed with 1ml FACS buffer, centrifuged as above, and resuspended in 100 µl FACS buffer. Cells were then moved to 96-well plates (Falcon 353263) and run through LSR Fortessa (BD Biosciences). Antibodies used are the following: AlexaFluor 488 anti-

human HLA-A, B, C clone W6/32 (Biolegend 311413), and APC anti-human HLA-A, B, C clone W6/32 (Biolegend 311410).

KDM5 paralog/ADAR dependency in mouse

Data was extracted from the Genome data provided in https://tumorimmunity.org/#/(17), comparing sgKDM5 and sgADAR dependency between immune-competent (ICB) mice and immune-incompetent (NSG) mice.

Statistical Analysis

For all experiments, p-values are calculated as indicated in figure legends. Symbols are as follows: n.s.>0.05, *<0.05, **<0.1, ***<0.001, ****<0.0001.

Models

Models were made on BioRender (https://www.biorender.com),

Acknowledgements: We would like to that Thomas Paul, Fred Derheimer, Jon Oyer, David Shields, Scott Tria, Michelle Wagner, William Snyder and many other colleagues from the Pfizer CTI and Oncology organizations for discussions throughout the course of this work. We also thank Jeff Settleman for reading of the manuscript.

References

1. Roulois D, Loo Yau H, Singhania R, Wang Y, Danesh A, Shen SY, Han H, Liang G, Jones PA, Pugh TJ, O'Brien C, De Carvalho DD. 2015. DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. Cell 162:961-73.
2. Chiappinelli KB, Strissel PL, Desrichard A, Li H, Henke C, Akman B, Hein A, Rote NS, Cope LM, Snyder A, Makarov V, Budhu S, Slamon DJ, Wolchok JD, Pardoll DM, Beckmann MW, Zahnow CA, Merghoub T, Chan TA, Baylin SB, Strick R. 2015. Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. Cell 162:974-86.
3. Chen R, Ishak CA, De Carvalho DD. 2021. Endogenous Retroelements and the Viral Mimicry Response in Cancer Therapy and Cellular Homeostasis. Cancer Discov 11:2707-2725.
4. Jansz N, Faulkner GJ. 2021. Endogenous retroviruses in the origins and treatment of cancer. Genome Biol 22:147.
5. Guler GD, Tindell CA, Pitti R, Wilson C, Nichols K, KaiWai Cheung T, Kim HJ, Wongchenko M, Yan Y, Haley B, Cuellar T, Webster J, Alag N, Hegde G, Jackson E, Nance TL, Giresi PG, Chen KB, Liu J, Jhunjhunwala S, Settleman J, Stephan JP, Arnott D, Classon M. 2017. Repression of Stress-Induced LINE-1 Expression Protects Cancer Cell Subpopulations from Lethal Drug Exposure. Cancer Cell 32:221-237 e13.
6. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, et al. 2001. Initial sequencing and analysis of the human genome. Nature 409:860-921.
7. Wells JN, Feschotte C. 2020. A Field Guide to Eukaryotic Transposable Elements. Annu Rev Genet 54:539-561.
8. Feinberg AP, Vogelstein B. 1983. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature 301:89-92.
9. Ehrlich M. 2009. DNA hypomethylation in cancer cells. Epigenomics 1:239-59.
10. Burns KH. 2020. Our Conflict with Transposable Elements and Its Implications for Human Disease. Annu Rev Pathol 15:51-70.
11. Vashi N, Bakhoum SF. 2021. The Evolution of STING Signaling and Its Involvement in Cancer. Trends Biochem Sci 46:446-460.
12. Cheon H, Wang Y, Wightman SM, Jackson MW, Stark GR. 2023. How cancer cells make and respond to interferon-I. Trends Cancer 9:83-92.
13. Nishikura K. 2016. A-to-I editing of coding and non-coding RNAs by ADARs. Nat Rev Mol Cell Biol 17:83-96.
14. Mehdipour P, Marhon SA, Ettayebi I, Chakravarthy A, Hosseini A, Wang Y, de Castro FA, Loo Yau H, Ishak C, Abelson S, O'Brien CA, De Carvalho DD. 2020. Epigenetic therapy induces transcription of inverted SINEs and ADAR1 dependency. Nature 588:169-173.

15. Gannon HS, Zou T, Kiessling MK, Gao GF, Cai D, Choi PS, Ivan AP, Buchumenski I, Berger AC, Goldstein JT, Cherniack AD, Vazquez F, Tsherniak A, Levanon EY, Hahn WC, Meyerson M. 2018. Identification of ADAR1 adenosine deaminase dependency in a subset of cancer cells. Nat Commun 9:5450.
16. Ishizuka JJ, Manguso RT, Cheruiyot CK, Bi K, Panda A, Iracheta-Vellve A, Miller BC, Du PP, Yates KB, Dubrot J, Buchumenski I, Comstock DE, Brown FD, Ayer A, Kohnle IC, Pope HW, Zimmer MD, Sen DR, Lane-Reticker SK, Robitschek EJ, Griffin GK, Collins NB, Long AH, Doench JG, Kozono D, Levanon EY, Haining WN. 2019. Loss of ADAR1 in tumours overcomes resistance to immune checkpoint blockade. Nature 565:43-48.
17. Dubrot J, Du PP, Lane-Reticker SK, Kessler EA, Muscato AJ, Mehta A, Freeman SS, Allen PM, Olander KE, Ockerman KM, Wolfe CH, Wiesmann F, Knudsen NH, Tsao HW, Iracheta-Vellve A, Schneider EM, Rivera-Rosario AN, Kohnle IC, Pope HW, Ayer A, Mishra G, Zimmer MD, Kim SY, Mahapatra A, Ebrahimi-Nik H, Frederick DT, Boland GM, Haining WN, Root DE, Doench JG, Hacohen N, Yates KB, Manguso RT. 2022. In vivo CRISPR screens reveal the landscape of immune evasion pathways across cancer. Nat Immunol 23:1495-1506.
18. Wu MJ, Kondo H, Kammula AV, Shi L, Xiao Y, Dhiab S, Xu Q, Slater CJ, Avila OI, Merritt J, Kato H, Kattel P, Sussman J, Gritti I, Eccleston J, Sun Y, Cho HM, Olander K, Katsuda T, Shi DD, Savani MR, Smith BC, Cleary JM, Mostoslavsky R, Vijay V, Kitagawa Y, Wakimoto H, Jenkins RW, Yates KB, Paik J, Tassinari A, Saatcioglu DH, Tron AE, Haas W, Cahill D, McBrayer SK, Manguso RT, Bardeesy N. 2024. Mutant IDH1 inhibition induces dsDNA sensing to activate tumor immunity. Science 385:eadl6173.
19. Shen JZ, Qiu Z, Wu Q, Finlay D, Garcia G, Sun D, Rantala J, Barshop W, Hope JL, Gimple RC, Sangfelt O, Bradley LM, Wohlschlegel J, Rich JN, Spruck C. 2021. FBXO44 promotes DNA replication-coupled repetitive element silencing in cancer cells. Cell 184:352-369 e23.
20. Cuellar TL, Herzner AM, Zhang X, Goyal Y, Watanabe C, Friedman BA, Janakiraman V, Durinck S, Stinson J, Arnott D, Cheung TK, Chaudhuri S, Modrusan Z, Doerr JM, Classon M, Haley B. 2017. Silencing of retrotransposons by SETDB1 inhibits the interferon response in acute myeloid leukemia. J Cell Biol 216:3535-3549.
21. Morel KL, Sheahan AV, Burkhart DL, Baca SC, Boufaied N, Liu Y, Qiu X, Canadas I, Roehle K, Heckler M, Calagua C, Ye H, Pantelidou C, Galbo P, Panja S, Mitrofanova A, Wilkinson S, Whitlock NC, Trostel SY, Hamid AA, Kibel AS, Barbie DA, Choudhury AD, Pomerantz MM, Sweeney CJ, Long HW, Einstein DJ, Shapiro GI, Dougan SK, Sowalsky AG, He HH, Freedman ML, Balk SP, Loda M, Labbe DP, Olson BM, Ellis L. 2021. EZH2 inhibition activates a dsRNA-STING-interferon stress axis that potentiates response to PD-1 checkpoint blockade in prostate cancer. Nat Cancer 2:444-456.
22. Deblois G, Tonekaboni SAM, Grillo G, Martinez C, Kao YI, Tai F, Ettayebi I, Fortier AM, Savage P, Fedor AN, Liu X, Guilhamon P, Lima-Fernandes E, Murison A, Kuasne H, Ba-Alawi W, Cescon DW, Arrowsmith CH, De Carvalho DD, Haibe-Kains B, Locasale JW, Park M, Lupien M. 2020. Epigenetic Switch-Induced Viral Mimicry Evasion in Chemotherapy-Resistant Breast Cancer. Cancer Discov 10:1312-1329.
23. Sheng W, LaFleur MW, Nguyen TH, Chen S, Chakravarthy A, Conway JR, Li Y, Chen H, Yang H, Hsu PH, Van Allen EM, Freeman GJ, De Carvalho DD, He HH, Sharpe AH, Shi Y. 2018. LSD1 Ablation Stimulates Anti-tumor Immunity and Enables Checkpoint Blockade. Cell 174:549-563 e19.

24. Wu L, Cao J, Cai WL, Lang SM, Horton JR, Jansen DJ, Liu ZZ, Chen JF, Zhang M, Mott BT, Pohida K, Rai G, Kales SC, Henderson MJ, Hu X, Jadhav A, Maloney DJ, Simeonov A, Zhu S, Iwasaki A, Hall MD, Cheng X, Shadel GS, Yan Q. 2018. KDM5 histone demethylases repress immune response via suppression of STING. PLoS Biol 16:e2006134.
25. Zhang SM, Cai WL, Liu X, Thakral D, Luo J, Chan LH, McGeary MK, Song E, Blenman KRM, Micevic G, Jessel S, Zhang Y, Yin M, Booth CJ, Jilaveanu LB, Damsky W, Sznol M, Kluger HM, Iwasaki A, Bosenberg MW, Yan Q. 2021. KDM5B promotes immune evasion by recruiting SETDB1 to silence retroelements. Nature 598:682-687.
26. Shen HF, Zhang WJ, Huang Y, He YH, Hu GS, Wang L, Peng BL, Yi J, Li TT, Rong R, Chen XY, Liu JY, Li WJ, Ohgi K, Li SW, Rosenfeld MG, Liu W. 2021. The Dual Function of KDM5C in Both Gene Transcriptional Activation and Repression Promotes Breast Cancer Cell Growth and Tumorigenesis. Adv Sci (Weinh) 8:2004635.
27. Leadem BR, Kagiampakis I, Wilson C, Cheung TK, Arnott D, Trojer P, Classon M, Easwaran H, Baylin SB. 2018. A KDM5 Inhibitor Increases Global H3K4 Trimethylation Occupancy and Enhances the Biological Efficacy of 5-Aza-2'-Deoxycytidine. Cancer Res 78:1127-1139.
28. Dhatchinamoorthy K, Colbert JD, Rock KL. 2021. Cancer Immune Evasion Through Loss of MHC Class I Antigen Presentation. Front Immunol 12:636568.
29. Saffern M, Samstein R. 2023. MHCing the tumour's dark genome. Nat Rev Immunol 23:140.
30. Griffin GK, Wu J, Iracheta-Vellve A, Patti JC, Hsu J, Davis T, Dele-Oni D, Du PP, Halawi AG, Ishizuka JJ, Kim SY, Klaeger S, Knudsen NH, Miller BC, Nguyen TH, Olander KE, Papanastasiou M, Rachimi S, Robitschek EJ, Schneider EM, Yeary MD, Zimmer MD, Jaffe JD, Carr SA, Doench JG, Haining WN, Yates KB, Manguso RT, Bernstein BE. 2021. Epigenetic silencing by SETDB1 suppresses tumour intrinsic immunogenicity. Nature 595:309-314.
31. Bonte PE, Arribas YA, Merlotti A, Carrascal M, Zhang JV, Zueva E, Binder ZA, Alanio C, Goudot C, Amigorena S. 2022. Single-cell RNA-seq-based proteogenomics identifies glioblastoma-specific transposable elements encoding HLA-I-presented peptides. Cell Rep 39:110916.
32. Kong Y, Rose CM, Cass AA, Williams AG, Darwish M, Lianoglou S, Haverty PM, Tong AJ, Blanchette C, Albert ML, Mellman I, Bourgon R, Greally J, Jhunjhunwala S, Chen-Harris H. 2019. Transposable element expression in tumors is associated with immune infiltration and increased antigenicity. Nat Commun 10:5228.
33. Chen Q, Sun L, Chen ZJ. 2016. Regulation and function of the cGAS-STING pathway of cytosolic DNA sensing. Nat Immunol 17:1142-9.
34. Mankan AK, Schmidt T, Chauhan D, Goldeck M, Honing K, Gaidt M, Kubarenko AV, Andreeva L, Hopfner KP, Hornung V. 2014. Cytosolic RNA:DNA hybrids activate the cGAS-STING axis. EMBO J 33:2937-46.
35. Pavlenko E, Ruengeler T, Engel P, Poepsel S. 2022. Functions and Interactions of Mammalian KDM5 Demethylases. Front Genet 13:906662.
36. Crossley MP, Bocek M, Cimprich KA. 2019. R-Loops as Cellular Regulators and Genomic Threats. Mol Cell 73:398-411.
37. Mackay RP, Xu Q, Weinberger PM. 2020. R-Loop Physiology and Pathology: A Brief Review. DNA Cell Biol 39:1914-1925.

38. Cristini A, Tellier M, Constantinescu F, Accalai C, Albulescu LO, Heiringhoff R, Bery N, Sordet O, Murphy S, Gromak N. 2022. RNase H2, mutated in Aicardi-Goutieres syndrome, resolves co-transcriptional R-loops to prevent DNA breaks and inflammation. Nat Commun 13:2961.
39. Crossley MP, Song C, Bocek MJ, Choi JH, Kousouros JN, Sathirachinda A, Lin C, Brickner JR, Bai G, Lans H, Vermeulen W, Abu-Remaileh M, Cimprich KA. 2023. R-loop-derived cytoplasmic RNA-DNA hybrids activate an immune response. Nature 613:187-194.
40. Petropoulos M, Karamichali A, Rossetti GG, Freudenmann A, Iacovino LG, Dionellis VS, Sotiriou SK, Halazonetis TD. 2024. Transcription-replication conflicts underlie sensitivity to PARP inhibitors. Nature 628:433-441.
41. Laspata N, Kaur P, Mersaoui SY, Muoio D, Liu ZS, Bannister MH, Nguyen HD, Curry C, Pascal JM, Poirier GG, Wang H, Masson JY, Fouquerel E. 2023. PARP1 associates with R-loops to promote their resolution and genome stability. Nucleic Acids Res 51:2215-2237.
42. Kemiha S, Poli J, Lin YL, Lengronne A, Pasero P. 2021. Toxic R-loops: Cause or consequence of replication stress? DNA Repair (Amst) 107:103199.
43. Kermi C, Lau L, Asadi Shahmirzadi A, Classon M. 2022. Disrupting Mechanisms that Regulate Genomic Repeat Elements to Combat Cancer and Drug Resistance. Front Cell Dev Biol 10:826461.
44. Liang J, Labadie S, Zhang B, Ortwine DF, Patel S, Vinogradova M, Kiefer JR, Mauer T, Gehling VS, Harmange JC, Cummings R, Lai T, Liao J, Zheng X, Liu Y, Gustafson A, Van der Porten E, Mao W, Liederer BM, Deshmukh G, An L, Ran Y, Classon M, Trojer P, Dragovich PS, Murray L. 2017. From a novel HTS hit to potent, selective, and orally bioavailable KDM5 inhibitors. Bioorg Med Chem Lett 27:2974-2981.
45. Gong F, Clouaire T, Aguirrebengoa M, Legube G, Miller KM. 2017. Histone demethylase KDM5A regulates the ZMYND8-NuRD chromatin remodeler to promote DNA repair. J Cell Biol 216:1959-1974.
46. Gaillard S, Charasson V, Ribeyre C, Salifou K, Pillaire MJ, Hoffmann JS, Constantinou A, Trouche D, Vandromme M. 2021. KDM5A and KDM5B histone-demethylases contribute to HU-induced replication stress response and tolerance. Biol Open 10.
47. Elsakrmy N, Cui H. 2023. R-Loops and R-Loop-Binding Proteins in Cancer Progression and Drug Resistance. Int J Mol Sci 24.
48. Crossley MP, Brickner JR, Song C, Zar SMT, Maw SS, Chedin F, Tsai MS, Cimprich KA. 2021. Catalytically inactive, purified RNase H1: A specific and sensitive probe for RNA-DNA hybrid imaging. J Cell Biol 220.
49. Goehring L, Huang TT, Smith DJ. 2023. Transcription-Replication Conflicts as a Source of Genome Instability. Annu Rev Genet 57:157-179.
50. Shih HT, Chen WY, Wang HY, Chao T, Huang HD, Chou CH, Chang ZF. 2022. DNMT3b protects centromere integrity by restricting R-loop-mediated DNA damage. Cell Death Dis 13:546.
51. Bayona-Feliu A, Herrera-Moyano E, Badra-Fajardo N, Galvan-Femenia I, Soler-Oliva ME, Aguilera A. 2023. The chromatin network helps prevent cancer-associated mutagenesis at transcription-replication conflicts. Nat Commun 14:6890.
52. Wang H, Fan Z, Shliaha PV, Miele M, Hendrickson RC, Jiang X, Helin K. 2023. H3K4me3 regulates RNA polymerase II promoter-proximal pause-release. Nature 615:339-348.

53. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S. 2019. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 10:1930.
54. Elbarbary RA, Lucas BA, Maquat LE. 2016. Retrotransposons as regulators of gene expression. Science 351:aac7247.
55. Gong Y, Liu Y. 2023. R-Loops at Chromosome Ends: From Formation, Regulation, and Cellular Consequence. Cancers (Basel) 15.
56. Liu Y, Liu Q, Su H, Liu K, Xiao X, Li W, Sun Q, Birchler JA, Han F. 2021. Genome-wide mapping reveals R-loops associated with centromeric repeats in maize. Genome Res 31:1409-1418.
57. Fujita T, Asano Y, Ohtsuka J, Takada Y, Saito K, Ohki R, Fujii H. 2013. Identification of telomere-associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP). Sci Rep 3:3171.
58. Kumbhar R, Sanchez A, Perren J, Gong F, Corujo D, Medina F, Devanathan SK, Xhemalce B, Matouschek A, Buschbeck M, Buck-Koehntop BA, Miller KM. 2021. Poly(ADP-ribose) binding and macroH2A mediate recruitment and functions of KDM5A at DNA lesions. J Cell Biol 220.
59. Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, McDermott U, Azizian N, Zou L, Fischbach MA, Wong KK, Brandstetter K, Wittner B, Ramaswamy S, Classon M, Settleman J. 2010. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell 141:69-80.
60. Vinogradova M, Gehling VS, Gustafson A, Arora S, Tindell CA, Wilson C, Williamson KE, Guler GD, Gangurde P, Manieri W, Busby J, Flynn EM, Lan F, Kim HJ, Odate S, Cochran AG, Liu Y, Wongchenko M, Yang Y, Cheung TK, Maile TM, Lau T, Costa M, Hegde GV, Jackson E, Pitti R, Arnott D, Bailey C, Bellon S, Cummings RT, Albrecht BK, Harmange JC, Kiefer JR, Trojer P, Classon M. 2016. An inhibitor of KDM5 demethylases reduces survival of drug-tolerant cancer cells. Nat Chem Biol 12:531-8.
61. Stong N, Deng Z, Gupta R, Hu S, Paul S, Weiner AK, Eichler EE, Graves T, Fronick CC, Courtney L, Wilson RK, Lieberman PM, Davuluri RV, Riethman H. 2014. Subtelomeric CTCF and cohesin binding site organization using improved subtelomere assemblies and a novel annotation pipeline. Genome Res 24:1039-50.
62. Hanzelmann S, Castelo R, Guinney J. 2013. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 14:7.
63. Dai X, Cheng H, Bai Z, Li J. 2017. Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. J Cancer 8:3131-3141.

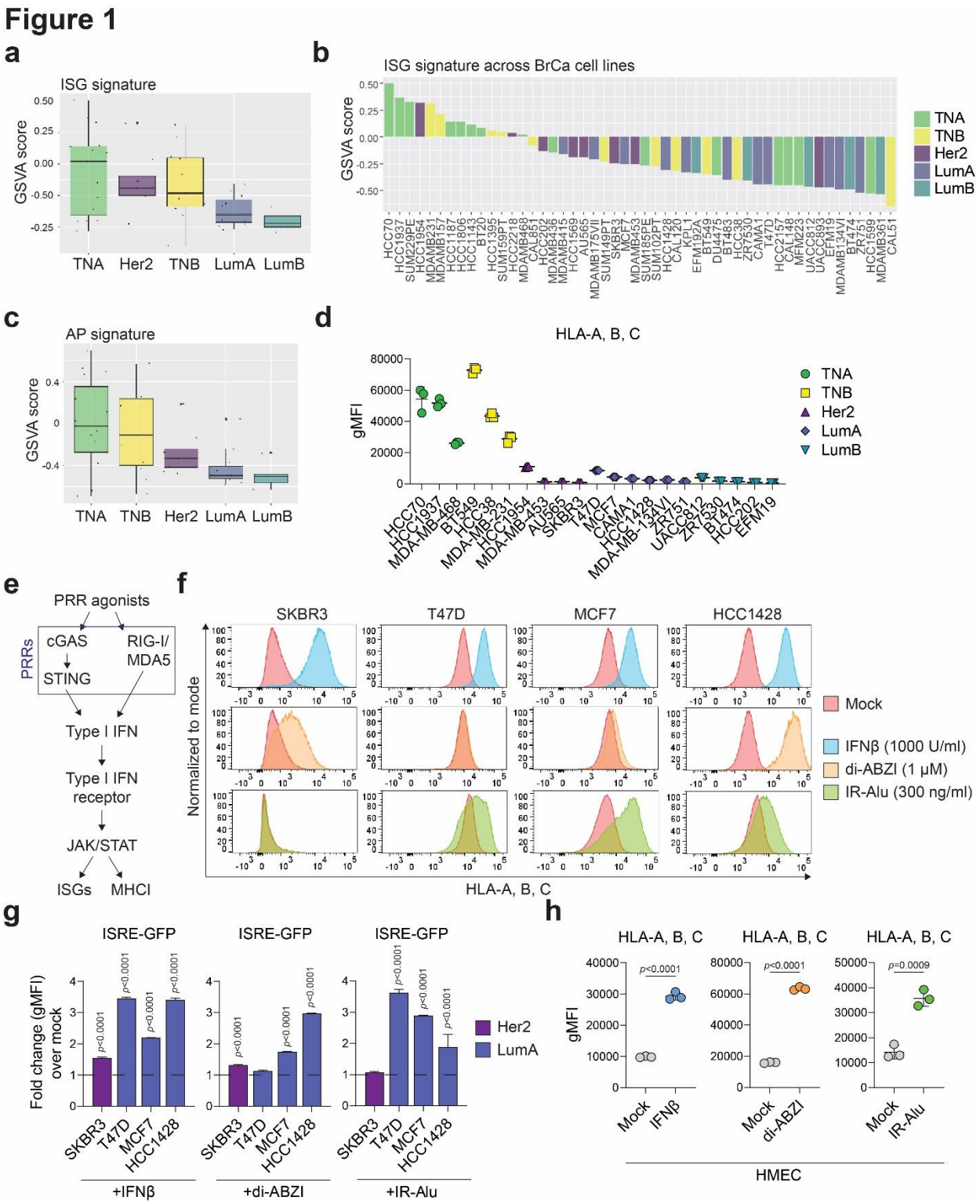

Figure 1: Low immune signatures in luminal breast cancer cells can be increased by agonists that activate Type I IFN. a, Gene Set Variation Analysis (GSVA) enrichment scores of IFN stimulated gene (ISG) set expression in breast cancer cell lines grouped by subtype and **b,** across individual breast cancer cell lines. TNA/B: triple negative A/B, Her2: Her2-positive,

LumA/B: luminal A/B. **c,** GSVA enrichment scores of Antigen Presentation (AP) gene set expression in breast cancer cell lines grouped by subtype. **d,** Quantification of HLA-A, B, C surface levels in the indicated breast cancer cell lines, measured as geometric mean fluorescence intensity (gMFI) by flow cytometry, n=3 for each cell line. **e,** Simplified schematic of Type I IFN signaling activation by activation of pattern recognition receptors (PRRs). **f,** Flow cytometry histograms of HLA-A, B, C surface levels in indicated cell lines following treatment as indicated for 24 hours. Plots are representative of 3 separate experiments. **g,** FACS quantification of ISRE-GFP fluorescence in indicated cell lines, colored by subtype, following treatment as indicated for 24 hours. Data is represented as fold change over mock-treated control, n=3, data are mean ± sd, p-values are from 2-way ANOVA test using Sidak's multiple comparisons, comparing each treatment vs mock treatment in all cell lines. **h,** gMFI of HLA-A, B, C surface levels in HMEC cells, as determined by flow cytometry. Cells were treated with 1000 U/ml IFNβ, 1 μM di-ABZI, or transfected with 300 ng/ml IR-Alu for 24 hours. n=3, data are mean ± sd, p-values are from unpaired t-tests.

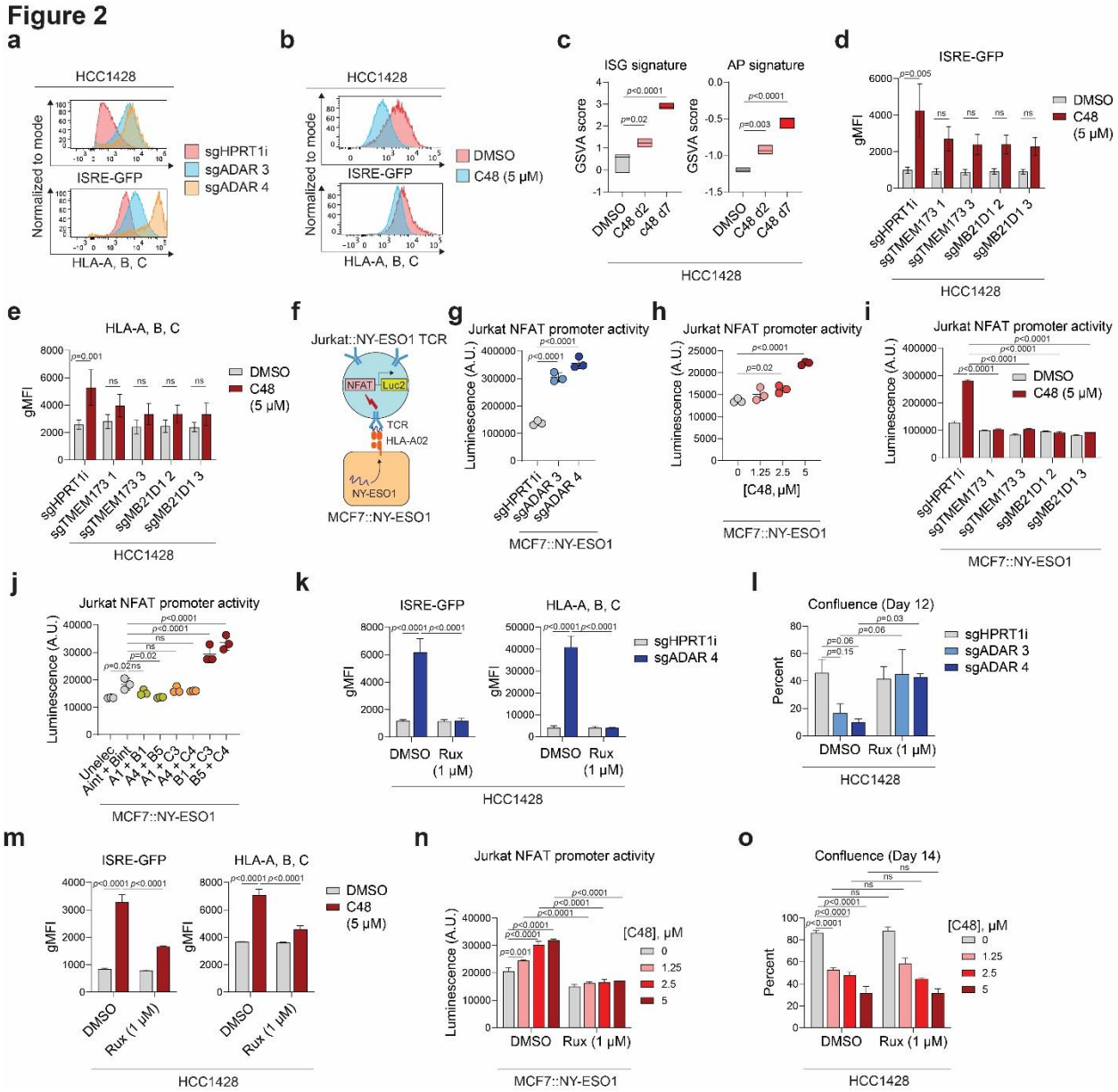

Figure 2: Disruption of upstream regulators of PRR signaling activates viral mimicry through distinct mechanisms a, Representative flow cytometry histograms of ISRE-GFP fluorescence and HLA-A, B, C surface expression (geometric mean fluorescence intensity, gMFI, normalized to mode) in CRISPR disrupted HCC1428 control (sgHPRT1i) or ADAR (sgADAR) cells. **b,** Representative flow cytometry histograms of ISRE-GFP fluorescence and HLA-A, B, C surface expression (geometric mean fluorescence intensity, gMFI, normalized to mode) in HCC1428 cells treated mock or C48 treated as indicated for 7 days. **c**, GSVA enrichment scores of ISG and AP gene set expression in HCC1428 cells treated with DMSO for 7 days as a control, or 2.5 μM C48 for 2 or 7 days as indicated. n=3, data are derived from RNAseq analysis, boxplots show min to max, line at median. p-values are from ordinary 1-way ANOVA using Dunnett's multiple comparisons test. **d,** Geometric mean fluorescence intensity (gMFI) of ISRE-GFP activity and **e,** HLA-A, B, C surface levels, as determined by flow cytometry, in HCC1428 cells CRISPR disrupted using an intronic cutting control (sgHPRT1i) or 2

independent sgRNAs against STING (TMEM173) or cGAS (MB21D1) and treated as indicated. Samples were harvested 10 days post-electroporation and 7 days post-treatment. n=3, data are mean ± sem, p-values are from 2-way ANOVA using Sidak's multiple comparisons test. **f,** Schematic of MCF7/Jurkat generated co-cultured system. **g,** Luminescence from NFAT-luciferase reporter (in arbitrary units A.U.) of engineered Jurkat cells co-cultured with engineered MCF7 cells: 4 days post-CRISPR disruption with the indicated sgRNAs, **h,** 7 days post treatment with mock and the indicated concentrations of C48, **i,** 10 days post CRISPR disruption with the indicated sgRNAs and treated as indicated 7 days, and **j,** 7 days post-electroporation with sgRNAs against 2 of 3 KDM5 paralogs**. g-j** Luciferase activity was measured 5 hours after co-culture, n=3, data are mean ± sd, p-values are ordinary 1-way ANOVA using Dunnett's multiple comparisons test in **g-h**, ordinary 1-way ANOVA using Sidak's multiple comparisons test in **i**, and 2-way ANOVA using Tukey's multiple comparisons test in **j**. **k.** Quantification of ISRE-GFP activity and HLA-A, B, C surface expression, measured as gMFI, in HCC1428 cells 7 days post-CRISPR disruption using sgRNAs against intronic HPRT1 or ADAR as indicated. Cells are treated with Ruxolitinib (Rux) as indicated throughout the 7 days of the experiment. n=3, data are mean ± sd, p-values are from 2-way ANOVA using Tukey's multiple comparisons test. **l,** Percent confluence of HCC1428 cells 12 days post CRISPR disruption using sgRNAs against intronic HPRT1 or ADAR. Cells are treated with DMSO or Rux as indicated throughout the 12 days. n=3, data are mean ± sem, p-values are from 2-way ANOVA using Sidak's multiple comparisons test. **m,** Quantification of ISRE-GFP and HLA-A, B, C surface levels, determined by flow cytometry and measured as gMFI, in HCC1428 cells treated as indicated for 7 days. n=3, data are mean ± sd, p-values calculated as in **k**. **n,** Luminescence of engineered Jurkat cells co-cultured with MCF7:NY-ESO1 treated as indicated for 7 days, measured 5 hours after co-culture. n=3, data are mean ± sd, p-values are from 2-way ANOVA using Tukey's multiple comparisons test. **o.** Percent confluence of HCC1428 cells 14 days post-treatment as indicated. Samples were split at the same ratios on day 7 of treatment. n=3, data are mean ± sd, p-values are from 2-way ANOVA using Tukey's multiple comparisons test.

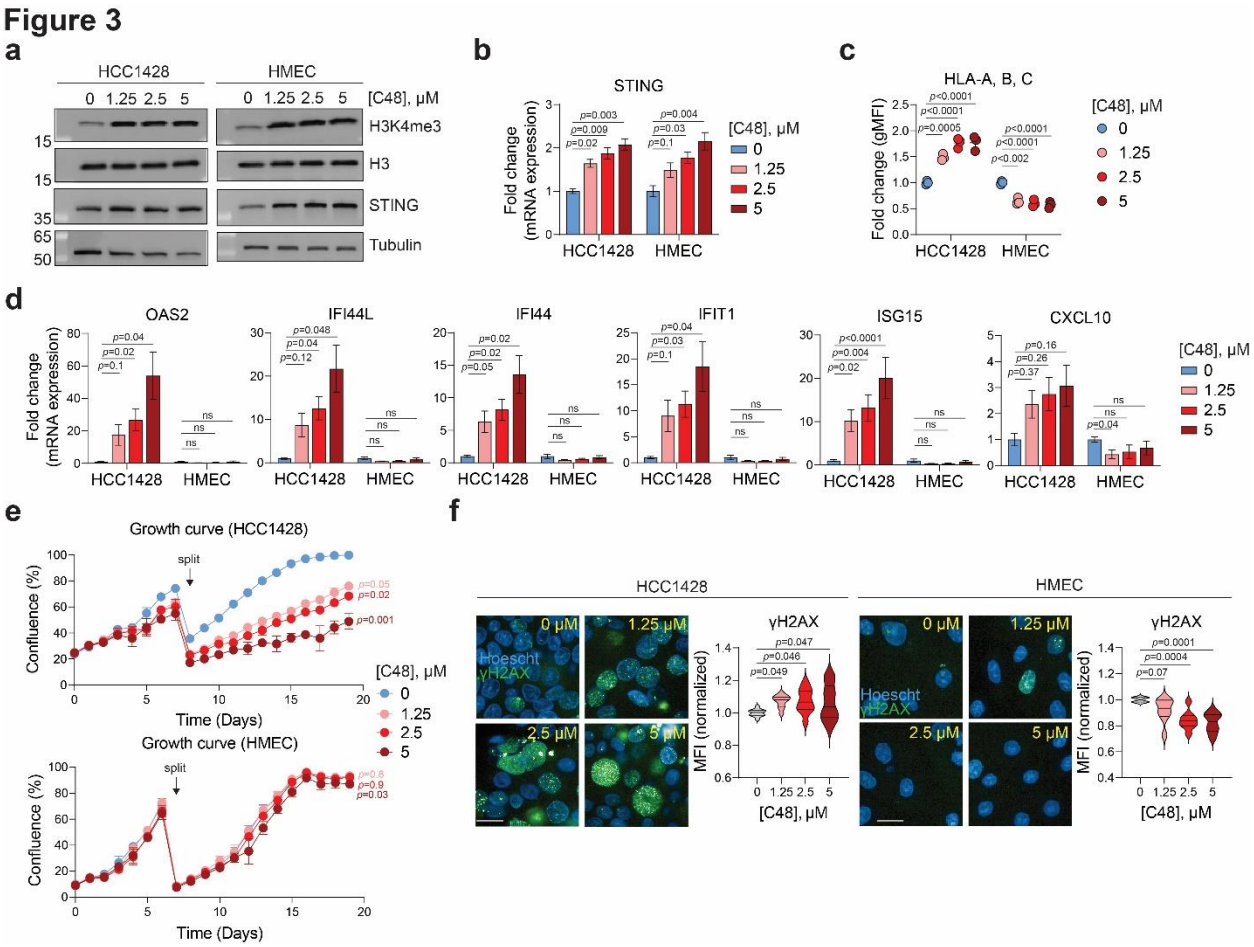

Figure 3: KDM5 inhibition induces "viral mimicry" and DNA damage specifically in tumor cells, but not in normal epithelial breast cells. a, Representative immunoblot of the indicated proteins in HCC1428 tumor cells and normal HMEC cells treated with mock or the indicated concentrations of C48 for 7 days. Histone H3 and tubulin were used as loading controls. **b,** qRT-PCR analysis of STING mRNA expression levels, measured as 2^(Ct values of indicated gene subtracted from Ct value of β-actin), in normal and tumor cells treated as indicated for 7 days. Data are represented as fold change expression over DMSO-treated (0 μM) control. n=3-5, data are mean ± sem, p-values are from repeated measures 1-way ANOVA using Dunnett's multiple comparisons test. **c,** Quantification HLA-A, B, C surface levels, determined by flow cytometry and measured as geometry mean fluorescence intensity (gMFI), in normal and tumor cells treated as indicated for 7 days. Data are represented as fold change gMFI over DMSO-treated (0 μM) control. n=3, data are mean ± sd, p-values are from ordinary 1-way ANOVA using Dunnett's multiple comparisons test. **d,** qRT-PCR analysis of ISG signature genes from cells treated as indicated for 7 days. Data is analyzed as in **b,**. **e,** Percent confluence of HCC1428 and HMEC cells treated with DMSO or the indicated concentrations of C48 over 19 days. Samples were split at the same ratios on day 7 of treatment. n=3, data are mean ± sd, p-values are from ordinary 1-way ANOVA using Tukey's multiple comparisons test, comparing the calculated doubling time of the first 7 days derived from each curve to that of DMSO-treated (0 μM) control. **f,** Representative immunofluorescence (IF) images and quantification of γ-H2AX staining (green) in cells treated with mock or the indicated concentrations of C48 for 7 days. Blue = Hoescht. Scale bar = 50 μM. Violin plots (with lines depicting quartiles) are derived from

mean fluorescence intensity (MFI) of γ-H2AX staining, normalized to DMSO-treated (0 μM) control. n=8-13 wells, p-values are from mixed effects analysis with using Dunnett's multiple comparisons test for HCC1428, repeated measures 1-way ANOVA using Dunnett's multiple comparisons test for HMEC.

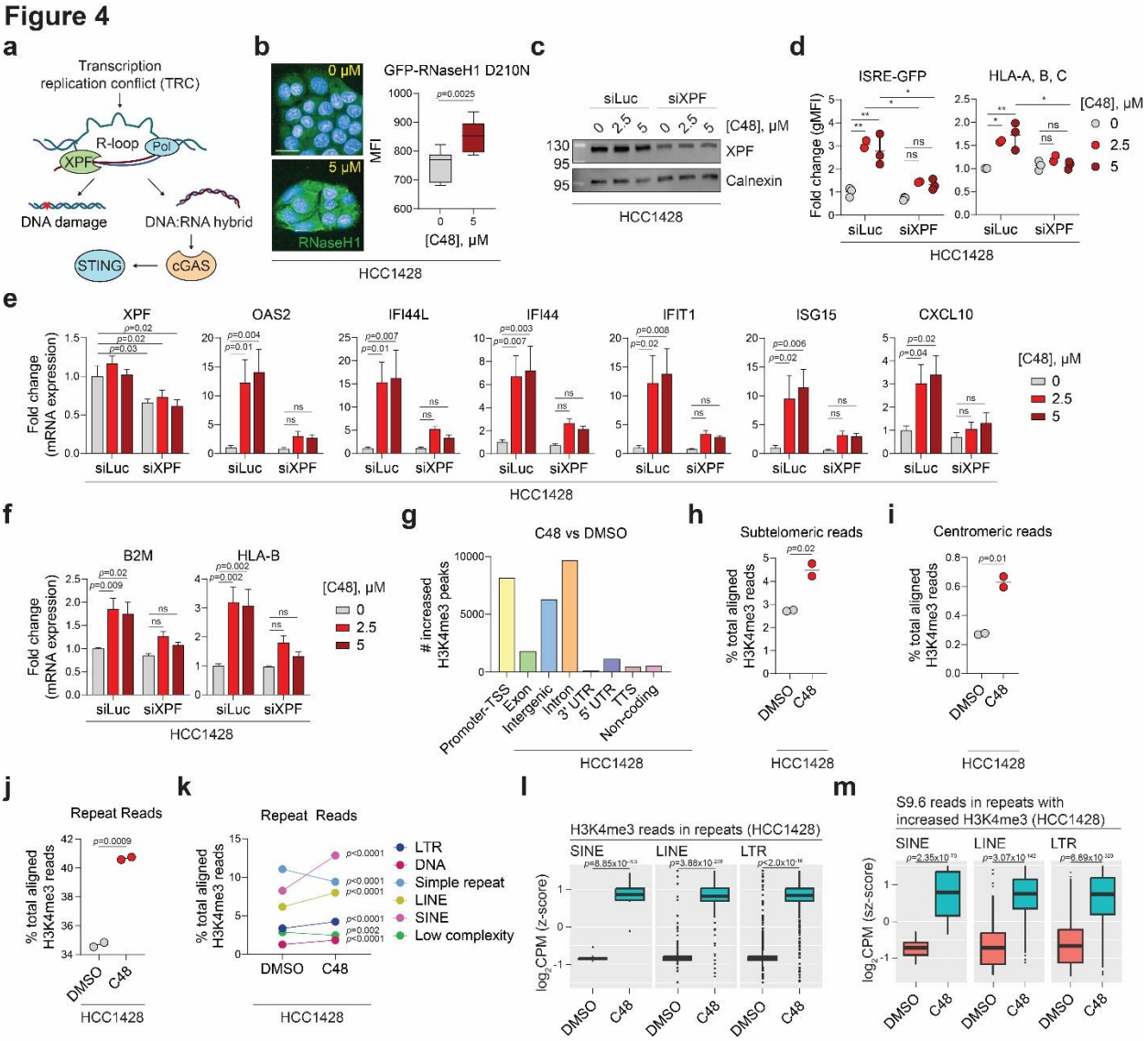

Figure 4. KDM5 inhibition induces R-loop formation in repeat regions that harbor increased H3K4me3. a, Schematic of R-loops and their potential deleterious consequences. **b,** Left: representative immunofluorescence images of cells incubated with GFP-RNaseH1 D210N protein (green), which detects RNA:DNA hybrids, in HCC1428 cells treated with mock and C48 as indicated. Blue = Hoescht. Scale bar = 50 μM. Right: Mean fluorescence intensity quantification of GFP positivity. n=8 wells (2 independent experiments where 9 fields /well were counted in triplicate), whiskers are from min to max, p-values are from unpaired t-test. **c,** Representative immunoblot of XPF protein levels in HCC1428 cells transfected with siRNAs against luciferase (control) or XPF and treated as indicated. Samples were harvested 6 days post-transfection and after 5 days of treatment. Calnexin was used as a loading control. **d,** FACS quantification of ISRE-GFP fluorescence and HLA-A, B, C surface levels in HCC1428

cells transfected and treated as in **c,**. Data are represented as fold change of gMFI over control (siLuciferase-transfected cells treated with DMSO, labeled as 0 μM C48). n=3, data are mean ± SEM, p-values are from mixed-effects analysis with multiple comparisons. **e-f,** qRT-PCR analysis of mRNA expression of XPF and genes representing ISG and AP signatures, measured as 2^(Ct values of indicated gene subtracted from Ct value of β-actin), of the indicated genes in HCC1428 cells transfected and treated as in **c,**. Data are represented as fold change of mRNA expression levels over control. n=3-4, data are mean ± SEM, p-values are from mixed-effects analysis with multiple comparisons. ns = not significant. **g,** Genomic annotations of spike-in-normalized H3K4me3 peaks that are significantly increased ($\log_2$ fold change > 0, $p_{adj} < 0.05$) in C48-treated compared to DMSO-treated CUT-and-Tag generated data from HCC1428 cells. **h,** Percent of total aligned H3K4me3 reads that map to subtelomeric regions, as defined in (61)or **i,** to centromeric regions, as defined in UCSC Genome Browser in H3K4me3 CUT-and-Tag data from cells treated as indicated. n=2, p-values are from unpaired t-tests. **j,** Percent of total aligned H3K4me3 reads that map to repeat regions in CUT-and-Tag data from DMSO or C48-treated HCC1428 cells. n=2, p-value is from unpaired t-test. **k,** as in **j,** but separated into individual repeat classes. Only classes that represent over 3% of total aligned H3K4me3 reads are graphed. p-values are from Sidak's multiple comparisons test. **l,** Log_2 counts per million (CPM) of H3K4me3 reads that map to the indicated class of repeats in CUT-and-Tag data from HCC1428 cells treated as indicated. Values are z-score scaled. Statistical test is a multiple regression model, and p-values are from the interaction terms of "DMSO" and "C48". **m,** Log_2CPM of S9.6 CUT-and-Tag generated reads in subsets of repeats that harbor significantly increased H3K4me3 after C48 treatment, separated into the indicated repeat classes. Values are z-score scaled. p-values are calculated as in **l**.

Figure 5

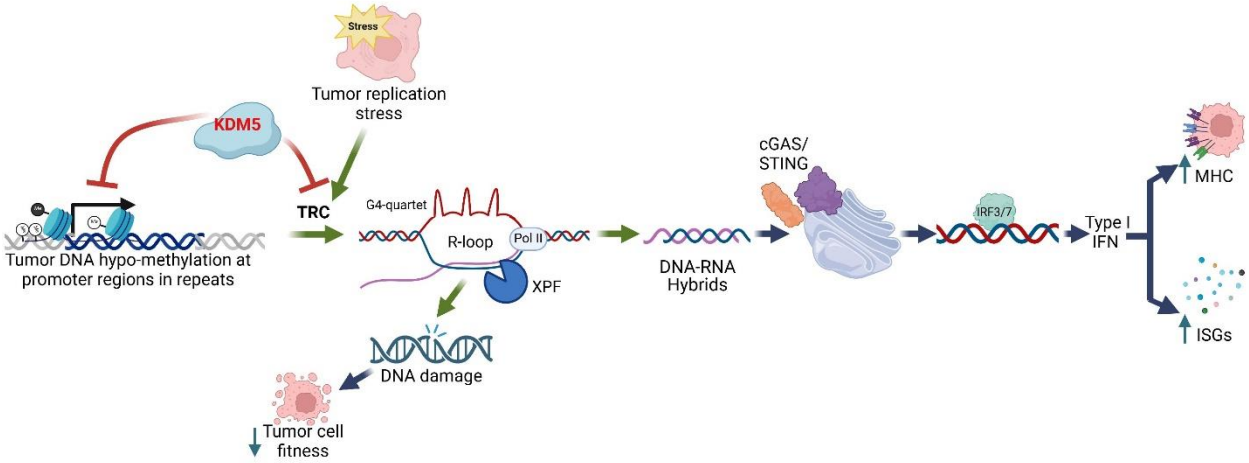

Figure 5. Model. The KDM5 family of epigenetic regulatory enzymes reduces R-loop formation in genomic repeat regions in luminal breast cancer cells. Genetic disruption or chemical inhibition of KDM5 causes R-loop-mediated DNA damage in luminal breast cancer cells. R-loop generated RNA:DNA hybrids also result in the activation of the CGAS/STING pathway, increasing ISG and AP signatures as well as cell surface MHCI, which may enhance interactions with the TME. Importantly, KDM5 inhibition does not result in DNA damage or activation of the CGAS/STING pathway in normal breast epithelial cells, suggesting that KDM5

inhibitors present a wider therapeutic window in this setting than STING agonists or type I Interferons.

Supplementary Figure 1

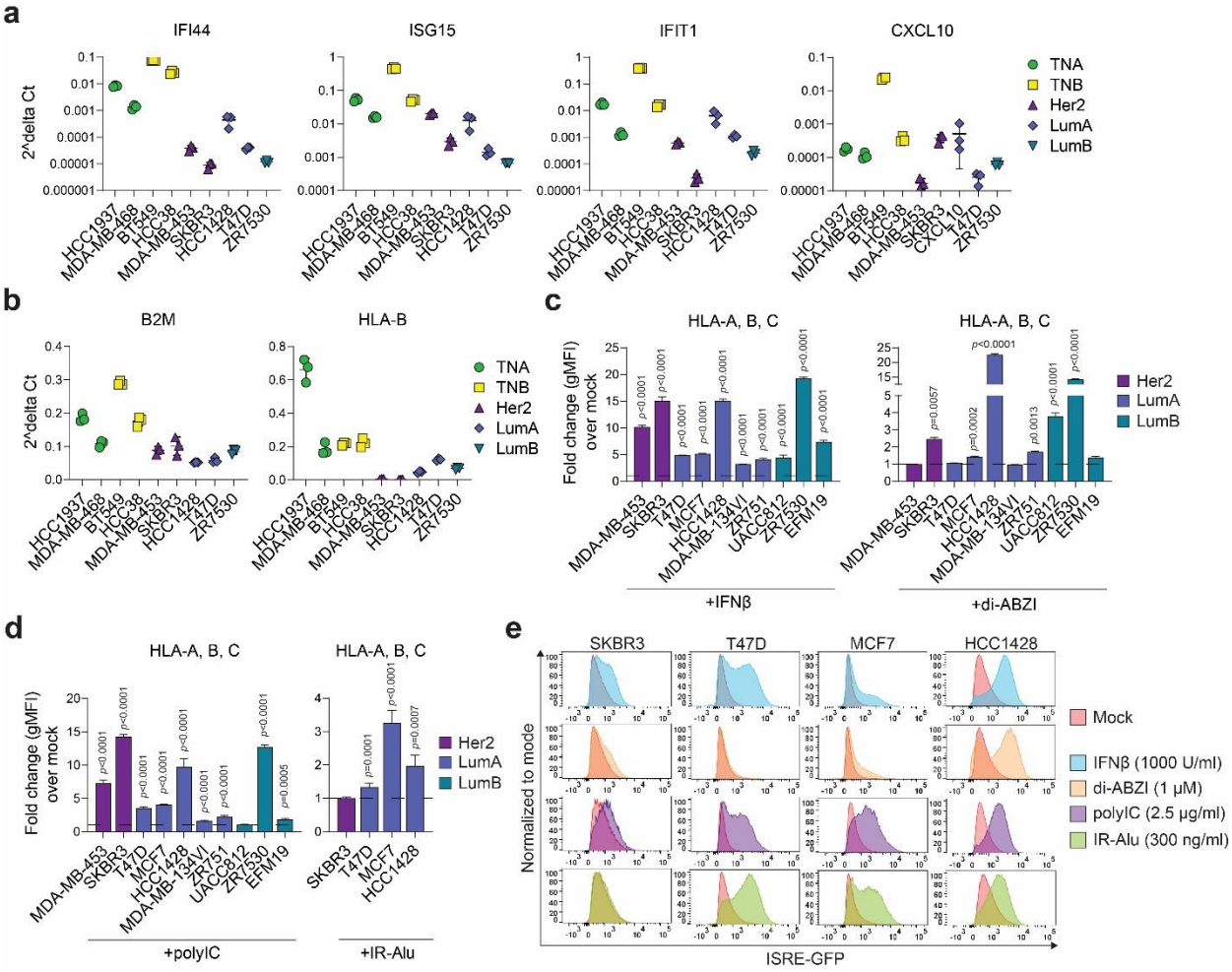

Figure S1. Low immune signatures in luminal breast cancer cells can be increased by agonists that activate type I IFN. a, qRT-PCR analysis of mRNA expression levels of indicated ISG signature and **b,** AP signature genes in the indicated cell lines at baseline. Delta Ct values are Ct values of indicated gene subtracted from Ct values of β-actin. Samples were run in triplicate and data are mean ± sd. **c-d,** Quantification of HLA-A, B, C surface levels, measured as geometric mean fluorescence intensity (gMFI), determined by flow cytometry, in the indicated cell lines following treatment with 1000 U/ml IFNβ, 1 μM di-ABZI, or transfected with 2.5 μg/ml polyIC or 300 ng/ml IR-Alu for 24 hours. Breast lines are colored by subtype as indicated in the figure. Data are represented as fold change gMFI of treatment over that of mock treatment. n=3, data are mean ± sd, p-values are from 1-way ANOVA test using Dunnett's multiple comparisons, comparing the different treatments vs mock treatment in each cell line. For IR-Alu experiment, p-values are from 2-way ANOVA test using Sidak's multiple comparisons test, comparing each treatment vs mock treatment in all cell lines. **e,** Examples of flow cytometry histograms of ISRE-GFP fluorescence in the indicated cell lines after 24h of the indicated treatment. Plots are representative examples from 3 independent experiments.

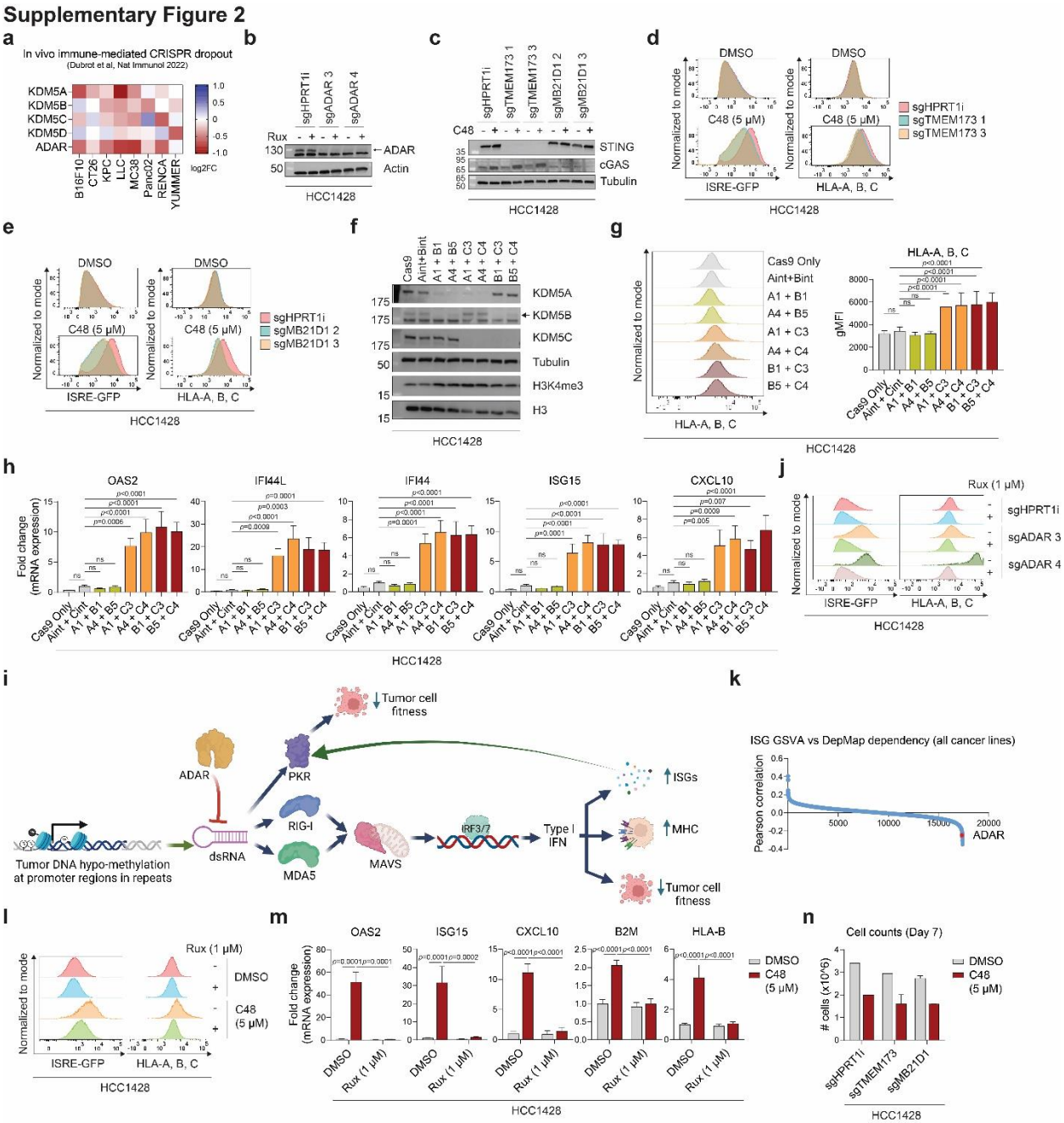

Figure S2. Perturbation of upstream regulators of PRR signaling activates viral mimicry through distinct mechanisms. a, Heatmap of $\log_2$ fold change dropout of sgRNAs targeting the indicated genes in the indicated cell lines injected into immune-competent (ICB) mice versus immune-incompetent (NSG) mice, summarized as reported (17). **b,** Representative immunoblot of ADAR protein levels in HCC1428 cells electroporated with sgRNAs against intronic HPRT1 (HPRT1i, control) or 2 independent sgRNAs against ADAR (sgADAR), with actin as loading control. Cells were either treated with DMSO (-) or 1 μM ruxolitinib (+) for duration of experiment and were harvested 4 days post-KO **c,** Representative immunoblots of the indicated proteins in HCC1428 cells electroporated with 2 independent sgRNAs against STING (TMEM173) or cGAS (MB21D1) and treated with DMSO (-) or 5 μM C48 (+). Samples were harvested 10 days post-

CRISPR disruption and 7 days post-C48 treatment. **d,** Representative flow cytometry histograms of ISRE-GFP levels and HLA-A, B, C surface levels in HCC1428 cells CRISPR disrupted with sgRNAs against STING and **e,** cGAS, and treated as in **c, f.** Representative immunoblot of the indicated proteins in HCC1428 electroporated with Cas9 only, sgRNAs targeting intronic regions of KDM5A and KDM5B (Aint+Bint) or 2 independent sets of sgRNAs targeting 2 of the 3 KDM5 paralogs found in female breast cancer cells. Samples were harvested 7 days post CRISPR disruption. **g,** Representative flow cytometry histograms and quantification of HLA-A, B, C surface levels, measured as gMFI, of samples treated as in **f,**. n=4-5, data are mean ± sd, p-values are from mixed-effects analysis using Sidak's multiple comparisons test. **h,** qRT-PCR of mRNA expression levels of representative ISG signature genes, calculated as 2^(Ct values of the indicated gene subtracted from β-actin Ct value)**.** Data are represented as fold change over Aint + Cint control. n=4-5. Data are mean ± sem. p-values are from mixed effects analysis using Dunnett's multiple comparisons test. **i.** Model of ADAR's role in ISG induction, MHCI surface expression, and tumor cell fitness. **j,** Representative flow cytometry histograms of ISRE-GFP fluorescence and HLA-A, B, C surface expression in HCC1428 cells electroporated with gRNAs against intronic HPRT1 (control) or ADAR, and treated with DMSO (-) or 1 μM ruxolitinib. **k.** Pearson correlations of ISG gene set expression GSVA scores and DepMap 22Q2 dependency (CERES) scores for all genes across all cell lines common to both data sets. Correlation value for ADAR is indicated. **l,** Representative flow cytometry histograms of ISRE-GFP fluorescence (left) and HLA-A, B, C surface levels (right) in HCC1428 cells treated as indicated. **m,** qRT-PCR analysis of mRNA expression levels, calculated as 2^(Ct values of the indicated gene subtracted from β-actin Ct value), in HCC1428 cells treated as indicated. n=3, data are mean ± sd, p-values are from ordinary 1-way ANOVA using Tukey's multiple comparisons test. **n,** Number of HCC1428 cells 10 days post-CRISPR disruption using sgRNA against intronic HPRT1, or 2 independent gRNAs against STING or cGAS, and treated as indicated for 7 days. n=1-2, data are mean ± sd.

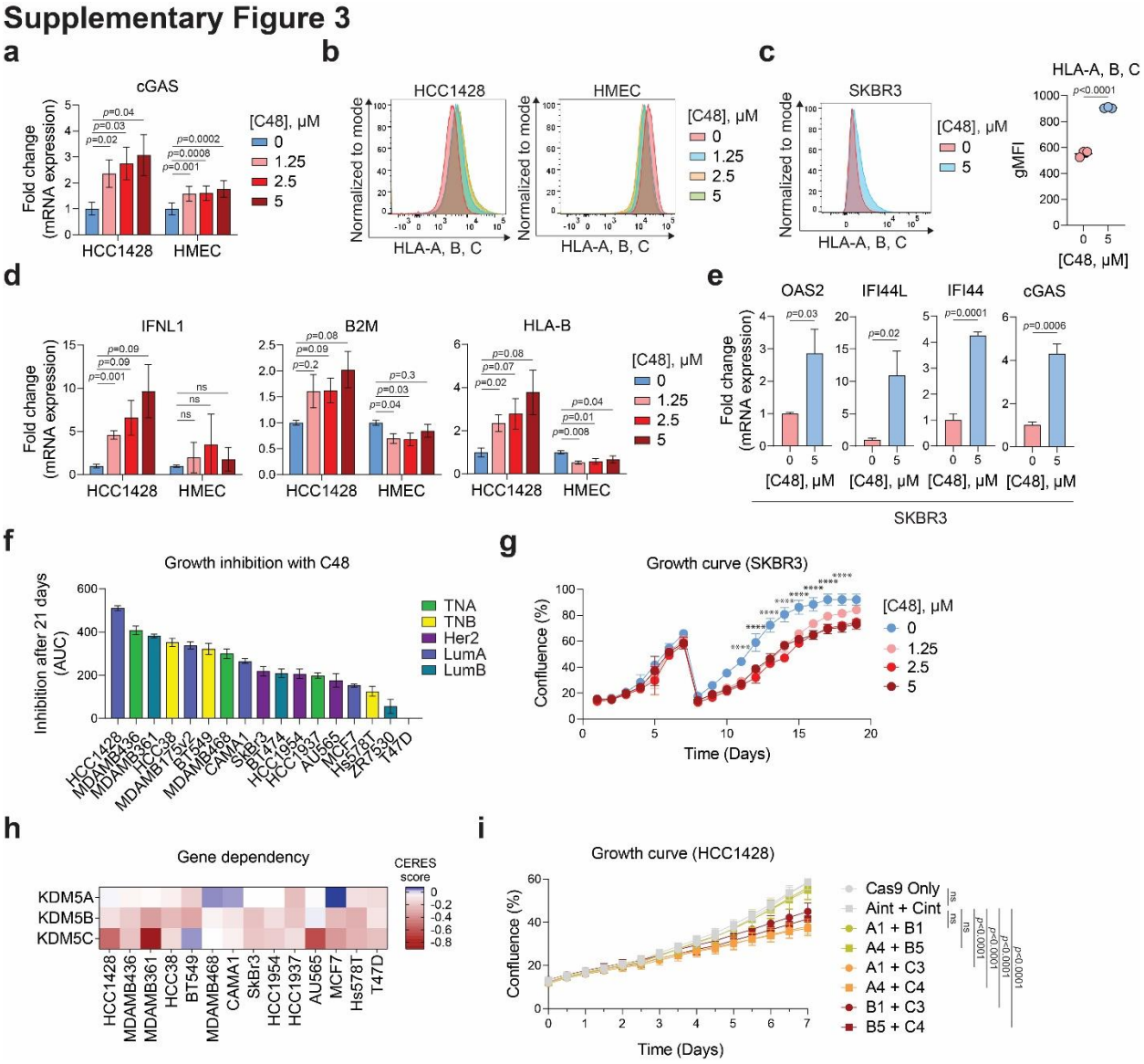

Figure S3: KDM5 inhibition or disruption induces viral mimicry and loss of cell fitness in luminal breast cancer cell lines. a, qRT-PCR analysis, measured as 2^(Ct values of indicated gene subtracted from Ct value of β-actin), of cGAS expression in cells treated as indicated for 7 days. Data are represented as fold change expression over DMSO-treated (0 μM) control. n=3-5, data are mean ± sem, p-values are from repeated measures 1-way ANOVA using Dunnett's multiple comparisons test. **b,** Representative flow cytometry histograms of HLA-A, B, C surface levels in the indicated cells treated with mock or the indicated concentrations of C48 for 7 days. **c,** Representative flow cytometry histograms of HLA-A, B, C surface levels (left) and quantification (right) as geometric mean fluorescence intensity (gMFI) in SKBR3 cells treated with DMSO (0 μM) or 5 μM C48 for 7 days. n=3, data are mean ± sd, p-values are from unpaired t-test. **d,** qRT-PCR analysis of the indicated IFN and AP genes in normal epithelial or tumor cell lines treated with the indicated concentrations of C48 for 7 days. Data is analyzed as in **a,**. **e,** qRT-PCR analysis of the indicated ISG signature genes and cGAS in SKBR3 cells treated with DMSO (0 μM) or 5 μM C48 for 7 days. Data are represented as fold change expression over DMSO-treated (0 μM) control. n=3-4, data are mean ± sem, p-values are from

unpaired t-tests. **f,** Growth inhibition after C48 treatment in the indicated cell lines for 21 days. Data are represented as area under the curve (AUC) of CyQuant quantification after treatment with a titration of C48 concentrations. **g,** Percent confluence of SKBR3 cells treated with the indicated concentrations of C48 over 19 days. Samples were split at the same ratios on day 7 of treatment. n=3, data are mean ± sd, p-values are from ordinary 2-way ANOVA using Tukey's multiple comparisons test, comparing the confluence of each sample to DMSO-treated (0 μM) control on the same day. ****p<0.0001 for all concentrations compared to 0 μM/DMSO control. **h,** Heatmap of CERES scores of the various KDM5 paralogs in the indicated breast cell lines, as reported in DepMap 22Q2. The more negative the score, the higher likelihood the gene of interest is essential. **i,** Percent confluence of HCC1428 cells electroporated with Cas9 only, sgRNAs targeting intronic regions of KDM5A and KDM5C (Aint+Cint) or 2 independent sets of gRNAs targeting 2 of the 3 KDM5 paralogs. n=4-5, p-values are from mixed-effects analysis using Sidak's multiple comparisons test comparing the calculated doubling time derived from each curve.

Supplementary Figure 4

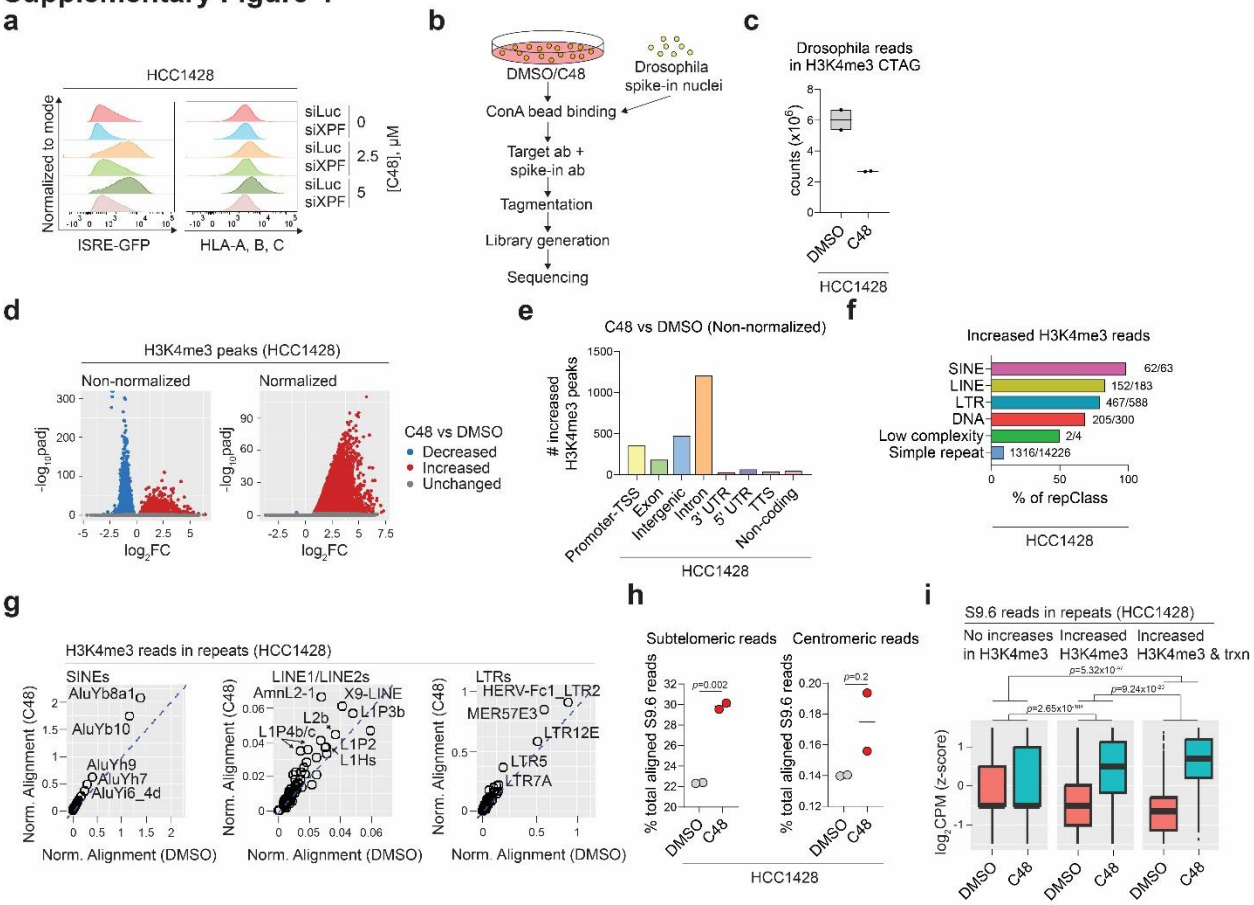

Figure S4. KDM5 inhibition increases R-loop abundance in repeat regions that harbor increased H3K4me3. a, Representative flow cytometry plots of ISRE-GFP fluorescence and HLA-A, B, C surface levels in HCC1428 cells transfected with siRNAs against Luciferase (siLuc) as a control or XPF (siXPF), and treated as indicated. **b,** Schematic of CUT-and-Tag experiment using Drosphila spike-ins. DMSO- or C48-treated (2.5 μM) cells and Drosophila spike-in nuclei were bound to Concanavalin A beads. Drosophila nuclei were added at a 1:10 ratio of spike-in

nuclei to experimental cells for downstream normalization. H3K4me3, S9.6, or spike-in antibodies were added to recruit pA-Tn5 to H3K4me3, RNA/DNA hybrids, or Drosophila H2Av respectively. Associated DNA were tagmented, and libraries were generated and sequenced. Each condition was run in duplicate (n=2). **c,** Number of reads from H3K4me3 CUT-and-Tag (CTAG) experiment that were successfully mapped to the *Drosophila melanogaster* genome. **d,** Volcano plots of H3K4me3 peaks mapped to the human genome without normalization (left) or after normalization to Drosophila spike-in reads (right). Red dots indicate peaks that are significantly increased (log_2 fold change > 0, $p_{adj} < 0.05$) in C48 vs DMSO-treated cells, while blue dots indicate peaks that are significantly decreased (log_2 fold change < 0, $p_{adj} < 0.05$). **e,** Genomic annotations of non-normalized H3K4me3 reads that are significantly increased (log_2 fold change > 0, $p_{adj} < 0.05$) in CUT-and-Tag data generated from HCC1428 cells, C48-treated compared to DMSO-treated cells if normalization to drosophila spike-in is not performed. **f,** Percent of repeat class (repClass) that have increased H3K4me3 reads in CUT-and-Tag data generated from HCC1428 cells treated with C48 vs DMSO-treated cells. The numbers at the right of the bars indicate the actual number of repeat subfamilies in each repeat class that have increased H3K4me3. **g.** Percent of total H3K4me3 reads mapping to individual SINE, LINE-1/LINE-2 and LTR subfamilies in HCC1428 CUT-and Tag data where the cells were treated with DMSO or C48 for 7 days. Values are an average of two replicates. Values are represented as normalized alignment (% total reads of TE subfamily divided by TE's prevalence in genome, in base pairs, x 10^{-6}). Dashed line represents where values are equal in both DMSO and C48-treated cells. **h,** Percent of total aligned S9.6 reads that map to subtelomeric regions, as definedin (61) (left), or centromeric regions, as defined in UCSC Genome Browser (right) in cells treated as indicated. n=2, p-values are from unpaired t-tests. **i,** Log_2CPM of S9.6 reads in subset of repeats that harbor no increases in H3K4me3 (left), significant increases in H3K4me3 (middle), or significant increases in both H3K4me3 and transcription (right) in HCC1428 data from cells treated with DMSO or C48 treatment. Values are z-score scaled. Statistical test is a multiple regression model, and p-value is from the interaction terms of "C48" and "increased H3K4me3", or "C48" and "increased H3K4me3 and transcription (trxn)".