RESEARCH ARTICLE

Reverse Transcriptase Inhibition Disrupts Repeat Element Life Cycle in Colorectal Cancer

Mihir Rajurkar[1], Aparna R. Parikh[1,2], Alexander Solovyov[3], Eunae You[1], Anupriya S. Kulkarni[1], Chong Chu[4], Katherine H. Xu[1], Christopher Jaicks[1], Martin S. Taylor[5], Connie Wu[6,7], Katherine A. Alexander[8], Charly R. Good[8], Annamaria Szabolcs[1], Stefanie Gerstberger[2], Antuan V. Tran[4], Nova Xu[1], Richard Y. Ebright[1], Emily E. Van Seventer[1], Kevin D. Vo[1], Eric C. Tai[1], Chenyue Lu[1], Jasmin Joseph-Chazan[1], Michael J. Raabe[1], Linda T. Nieman[1], Niyati Desai[1], Kshitij S. Arora[1,5], Matteo Ligorio[1,9], Vishal Thapar[1], Limor Cohen[6,7], Padric M. Garden[6,7], Yasmeen Senussi[6,7], Hui Zheng[10], Jill N. Allen[1,2], Lawrence S. Blaszkowsky[1,2], Jeffrey W. Clark[1,2], Lipika Goyal[1,2], Jennifer Y. Wo[1,11], David P. Ryan[1,2], Ryan B. Corcoran[1,2], Vikram Deshpande[1,5], Miguel N. Rivera[1,5], Martin J. Aryee[1,5], Theodore S. Hong[1,11], Shelley L. Berger[8], David R. Walt[6,7], Kathleen H. Burns[6,12], Peter J. Park[4], Benjamin D. Greenbaum[3,13], and David T. Ting[1,2]

Cancer Res 2024;84:2...
16 July 2024

Illustration by Bianca Dunn

ABSTRACT Altered RNA expression of repetitive sequences and retrotransposition are frequently seen in colorectal cancer, implicating a functional importance of repeat activity in cancer progression. We show the nucleoside reverse transcriptase inhibitor 3TC targets activities of these repeat elements in colorectal cancer preclinical models with a preferential effect in p53-mutant cell lines linked with direct binding of p53 to repeat elements. We translate these findings to a human phase II trial of single-agent 3TC treatment in metastatic colorectal cancer with demonstration of clinical benefit in 9 of 32 patients. Analysis of 3TC effects on colorectal cancer tumorspheres demonstrates accumulation of immunogenic RNA:DNA hybrids linked with induction of interferon response genes and DNA damage response. Epigenetic and DNA-damaging agents induce repeat RNAs and have enhanced cytotoxicity with 3TC. These findings identify a vulnerability in colorectal cancer by targeting the viral mimicry of repeat elements.

SIGNIFICANCE: Colorectal cancers express abundant repeat elements that have a viral-like life cycle that can be therapeutically targeted with nucleoside reverse transcriptase inhibitors (NRTI) commonly used for viral diseases. NRTIs induce DNA damage and interferon response that provide a new anticancer therapeutic strategy.

INTRODUCTION

The repeat transcriptome or "repeatome" consists of the expression of a variety of repeat RNA species including long interspersed nuclear element-1 (LINE-1) retrotransposons, human endogenous retroviruses (HERV), and satellite repeats. In normal development, repeat RNAs are expressed in the setting of embryogenesis (1, 2) and are epigenetically suppressed in differentiated tissues. They are awakened in the setting of tumorigenesis and aberrantly expressed in a spectrum of cancers (3–5), which has been shown to engage the innate (6–9) and adaptive (10–12) immune response through pathogen-associated sequence features found in immunostimulatory viral RNA and DNA (13). Analogous to retroviruses, repeat RNAs are also frequently reverse transcribed into DNA (14–16) that can alter the distribution and composition of these immunostimulatory nucleic acids. The presence of repeat DNA products from reverse transcription (RT) can provide a source of cytosolic DNA in cancer cells that has been shown to lead to increased metastatic behavior in cancers through cGAS–STING activation (17, 18). This suggests that inhibition of the RT activity of these retroelements with nucleoside RT inhibitors (NRTI; refs.16, 19, 20), a class of agents commonly used in HIV and hepatitis B virus (HBV), would affect tumor progression and metastatic dissemination. Our prior work demonstrated the anticancer effects of the NRTIs ddC and d4T *in vitro* and *in vivo* in colorectal cancer cell lines (16), but these early NRTIs are known to have significant side effects in patients. Looking toward clinical feasibility of repeat RT inhibition, we chose lamivudine (3TC) given previous literature demonstrating potency against endogenous RT activity (19–23) and the well-established clinical tolerability in patients with HIV and HBV, unlike ddC and d4T. Here, we present the use of 3TC in preclinical functional colorectal cancer models, mechanistic studies to understand the effects of 3TC on cytosolic DNA, and the first clinical trial of 3TC in patients with metastatic colorectal cancer.

RESULTS

3TC Functional Effects on Colorectal Cancer Preclinical Models

We evaluated 10 colorectal cancer cell lines (DLD1, HCT15, HT29, C2BBe1, LS123, SW948, HCT8, HCT116, LOVO, and RKO) with 3TC in Transwell migration and in soft-agar colony formation as *in vitro* functional assays of motility

[1]Mass General Cancer Center, Harvard Medical School, Charlestown, Massachusetts. [2]Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts. [3]Computational Oncology, Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York. [4]Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts. [5]Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts. [6]Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. [7]Wyss Institute for Biologically Inspired Engineering, Harvard Medical School, Boston, Massachusetts. [8]Epigenetics Institute, Departments of Cell and Developmental Biology, Genetics, and Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania. [9]Department of Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts. [10]Biostatistics Center, Massachusetts General Hospital, Boston, Massachusetts. [11]Department of Radiation Oncology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts. [12]Department of Oncologic Pathology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts. [13]Physiology, Biophysics and Systems Biology, Weill Cornell Medicine, Weill Cornell Medical College, New York, New York.

Note: Supplementary data for this article are available at Cancer Discovery Online (http://cancerdiscovery.aacrjournals.org/).

M. Rajurkar, A.R. Parikh, and A. Solovyov contributed equally to this article.

Corresponding Authors: David T. Ting, Mass General Cancer Center, 149 13th Street, Room 6003, Charlestown, MA 02129. Phone: 617-240-9402; E-mail: dting1@mgh.harvard.edu; and Benjamin D. Greenbaum, Memorial Sloan Kettering Cancer Center, 321 E 61st Street, Room 256, New York, NY 10065. Phone: 646-608-7667; E-mail: greenbaB@mskcc.org

Cancer Discov 2022;12:1462–81

doi: 10.1158/2159-8290.CD-21-1117

and anoikis resistance that are essential components of the metastatic cascade. Boyden chamber Transwell migration assays (Fig. 1A) noted marked effects of 3TC on migratory capability in DLD1, HCT15, C2Bbe1, LS123, and RKO cell lines (Fig. 1B–E; Supplementary Fig. S1A and S1B). Notably, we found that 3TC functional effects on migration were more prevalent among p53-mutant (p53-Mut; DLD1, HCT15, HT29, C2BBe1, LS123, and SW948) compared with p53-wild-type (p53-WT; HCT8, HCT116, LOVO, and RKO) cell lines. Similarly, treatment with 3TC demonstrated significant reduction in anchorage-independent growth in soft agar in p53-Mut cell lines (Fig. 1F–H; Supplementary Fig. S2A and S2B) compared with p53-WT cell lines (Fig. 1I–K). We extended these *in vitro* findings to xenograft tumors using a sensitive (SW620) and resistant (HCT116) cell line. The p53-Mut SW620 xenograft demonstrated significant response to 3TC alone (ANOVA $P < 0.0001$) compared with PBS vehicle control-treated tumors (Fig. 1L), whereas the p53-WT HCT116 xenograft did not respond to single-agent 3TC (Fig. 1M).

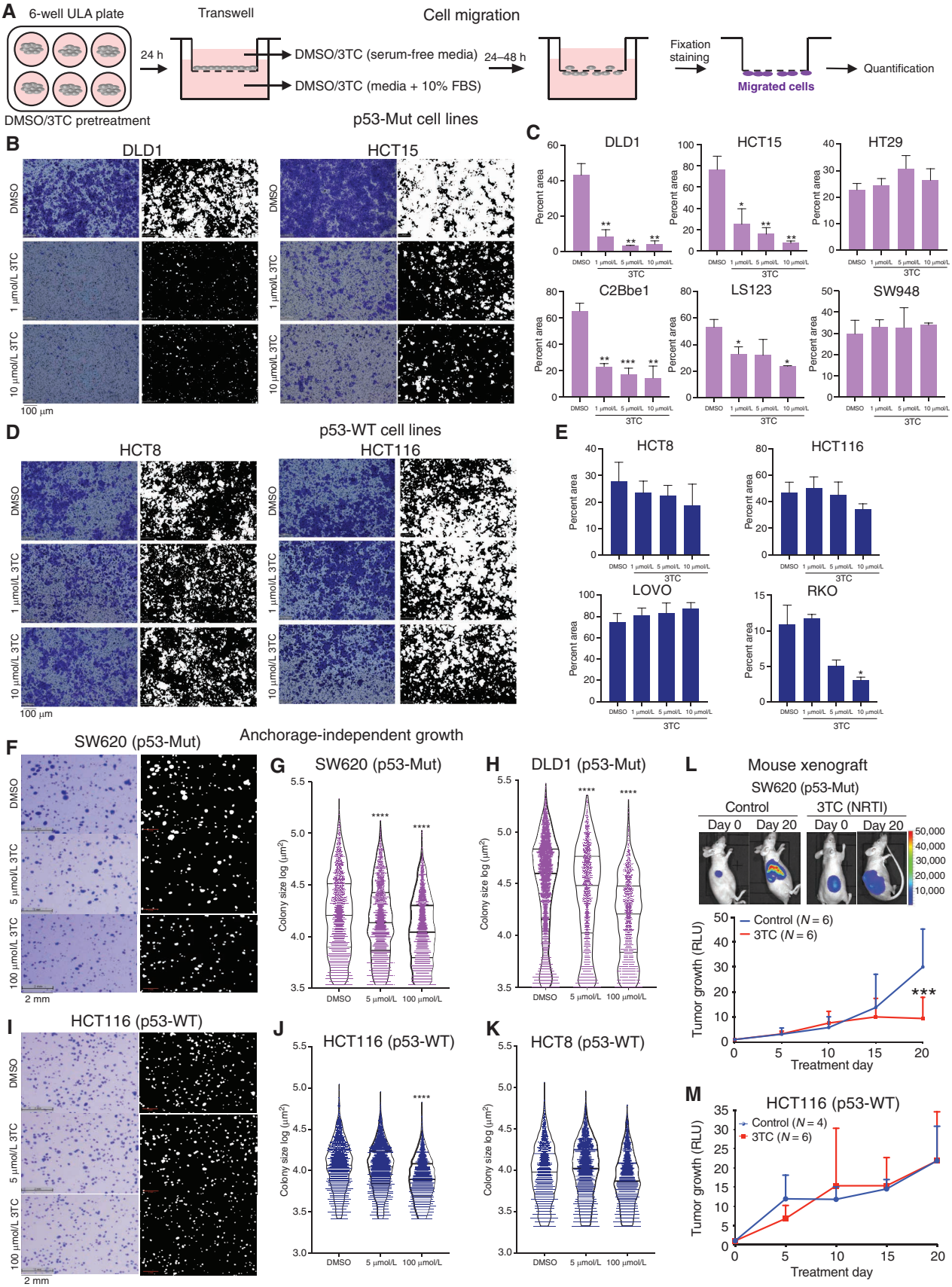
P53 Is a Direct Repressor of Repeat RNA Expression

Given the observed differential response to 3TC based on p53 mutation status, we wanted to better understand the relationship of repeat RNA expression levels with p53. We evaluated whether p53 has direct interactions with genomic repeats, given the association of p53 mutation or loss with derepression of repeat RNA expression (6, 7, 24) and LINE-1 retrotransposition (15, 25). We performed p53 immunoprecipitation followed by DNA sequencing (p53-IP-seq) in p53-Mut (SW620, DLD1) and p53-WT (HCT116, HCT8) cell lines done in biological duplicates (Fig. 2A–C; Supplementary Fig. S3A and S3B; Supplementary Data S1 and S2). Differential enrichment analysis of p53-bound repeat elements (FDR < 0.2) demonstrated markedly different proportions of specific repeat classes with notably higher satellite (SAT; 34% of repeats) and LINE-1 (L1; 31% of repeats) elements in p53-WT compared with p53-Mut cell lines (Fig. 2C). Notably, the HSATII satellite repeat was enriched, which we have previously shown to be reverse-transcribed in colorectal cancer (16). Both p53-Mut cell lines have DNA binding domain mutations (SW620–R273H; DLD1–S241F), which suggests that loss of function of DNA binding is associated with diminished p53 interaction with these repeat DNA sequences. Total RNA sequencing (RNA-seq) of colorectal cancer cell lines ($n = 3$ biological replicates) comparing p53-Mut (SW620, DLD1) and p53-WT (HCT116, HCT8) cells noted significant differential expression of the L1PA2 element (Fig. 2D; Supplementary Fig. S3B) that was also seen in chromatin immunoprecipitation sequencing (ChIP-seq) analysis indicating this particular L1 repeat subfamily is directly regulated by p53. The major active human LINE-1 retrotransposon (L1HS) was also expressed at significantly higher levels in p53-Mut cell lines (Fig. 2D), though we did not see differential enrichment from p53-IP sequencing. To use an alternative method to quantify HSATII repeat expression given the potential heterogeneity of expression within each cell line, we used RNA *in situ* hybridization (RNA-ISH) and utilized quantitative digital image analysis of multiple areas of each cell line (Fig. 2E). The HSATII repeat had significantly higher expression in the p53-Mut cell lines (Fig. 2E; Supplementary Fig. S4A and S4B). Dysregulation of SAT and L1 repeats was previously shown to be highly expressed in a broad set of cancers (3, 6, 26), and the higher expression of HSATII and L1PA2 paired with loss of p53 binding in these cell lines is consistent with a direct suppression of these repeats by p53. In addition, the active L1HS retrotransposon was significantly higher in p53-Mut cell lines, but there was no indication of changes in enrichment in p53-IP-seq analysis. This suggests L1HS derepression is due to secondary effects of a p53 mutation. Although we cannot rule out potential p53 mutation gain-of-function effects (27) on repeat expression, the loss of DNA binding to repeat sequences in p53-Mut cell lines would point toward a loss-of-function effect of p53 on repeat expression including L1HS. Overall, these findings in colorectal cancer are consistent with other models demonstrating a direct role of p53 in regulating repeat RNAs (24, 28), which would indicate p53 mutation is a potential genetic marker of colorectal cancer tumors sensitive to 3TC.

Phase II Clinical Trial of 3TC in Metastatic Colorectal Cancer

These preclinical data provided evidence to support the initiation of a single-arm phase II clinical trial (NCT03144804) of 3TC in patients with p53-Mut metastatic colorectal cancer. Eligibility included patients with *TP53*-mutant refractory metastatic colorectal cancer with progression on or intolerance to 5-fluorouracil (5FU), oxaliplatin, and irinotecan, and anti-EGFR therapy if RAS WT. Eligible patients were age ≥18 years, with histologically or cytologically documented, advanced (metastatic and/or unresectable) disease that was incurable and had progressed on at least two lines of systemic therapy (Fig. 3A; Supplementary Tables S1–S3). See Methods for details of inclusion and exclusion criteria. Primary endpoint of the study was to describe overall response rate with the null hypothesis that the response rate is ≤1% versus the alternative

Figure 1. RT inhibitor 3TC has preclinical efficacy in colorectal cancer. **A,** Schema illustrating experimental procedure for measuring cell migration. **B–E,** Colon cancer cell line migration data with representative images and quantification of area covered by stained cells in (**B** and **C**) p53-Mut and (**D** and **E**) p53-WT colon cancer cell lines treated with DMSO or 3TC across 8-μm pore Transwell after 24–72 hours following fixation and staining with crystal violet. Statistical significance calculated by Student two-tailed *t* test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. **F–K,** Response to 3TC vs. DMSO control in soft-agar colony assay in (**F**–**H**) p53-Mut cell lines SW620 and DLD1 and (**I**–**K**) p53-WT cell lines HCT116 and HCT8. **F** and **I,** Representative images of soft-agar colonies of SW620 and HCT116 (left) with image quantification markup (right). Scale bar, 2 mm (left), 1 mm (right). Quantification of colony size by digital image analysis shown with violin plot with median and interquartile range in (**G** and **H**) p53-Mut cells and (**J** and **K**) p53-WT cells treated with DMSO or 3TC at 5 and 100 μmol/L. Statistical significance calculated by Student two-tailed *t* test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. **L** and **M,** p53-Mut (SW620) and p53-WT (HCT116) colorectal cancer xenograft tumors treated with 3TC vs. PBS (control) treatment. Luciferase-expressing (**L**) SW620 or (**M**) HCT116 cells were subcutaneously implanted in immunocompromised Nu/Nu mice and grown for 2 weeks, after which mice were treated with PBS or 3TC at 50 mg/kg administered by intraperitoneal injection 3 times a week. Tumor luminescence was measured using IVIS imaging every 5 days. Graph represents relative luminescence units (RLU) normalized to day 0. Significance determined by two-way ANOVA test: ****, $P < 0.0001$.



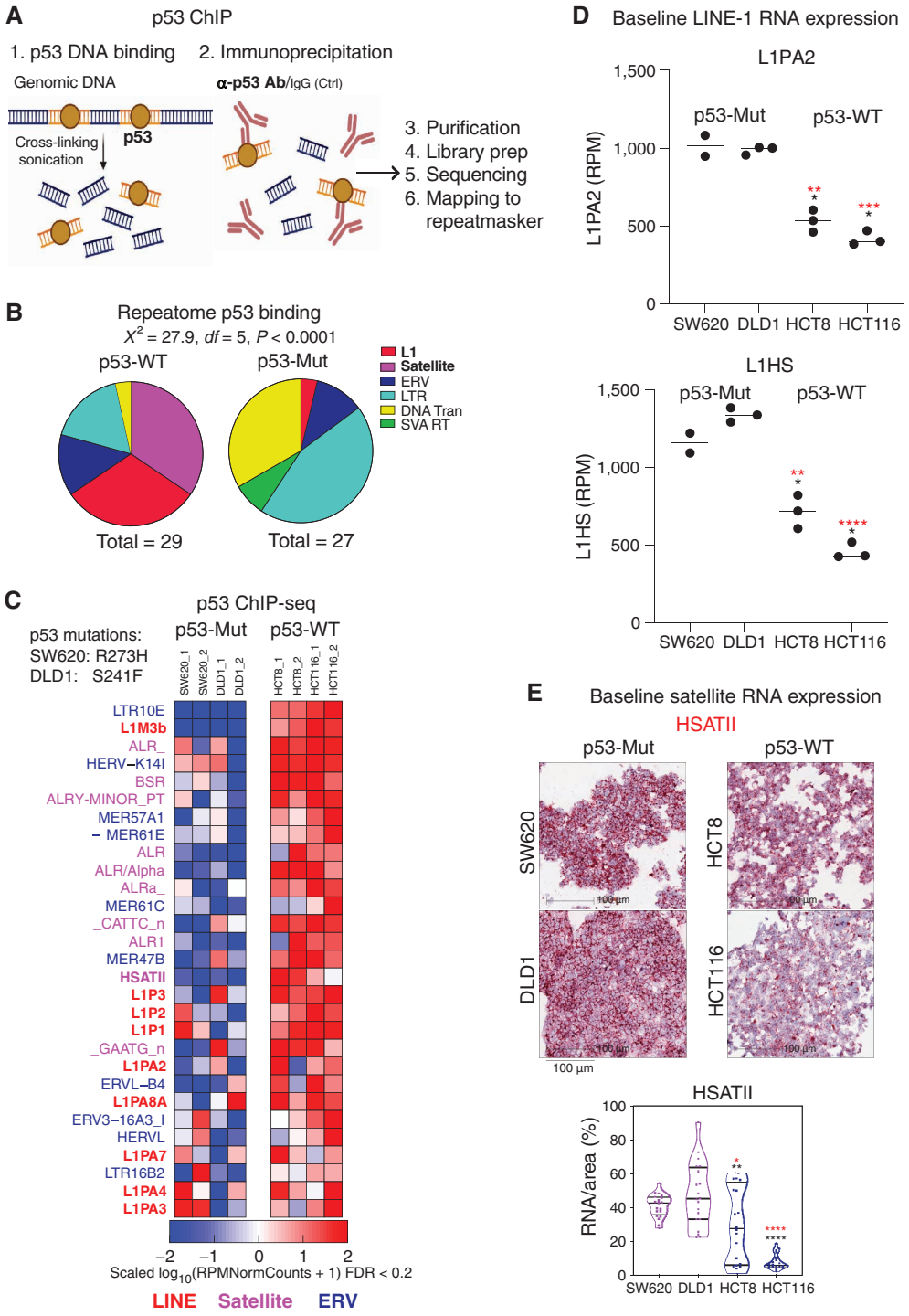


Figure 2. p53 is a direct repressor of repeat RNA expression. **A,** Schematic demonstrating p53-ChIP-seq conducted on colorectal cancer tumorspheres to identify repeatome p53 binding. **B,** Distribution of significantly enriched repeats (FDR < 0.2) from p53 ChIP-seq in p53-Mut and p53-WT cell lines (SAT, satellite; L1, LINE-1; ERV, endogenous retrovirus). p53-WT cells show significantly enriched p53 binding to repeat elements compared with p53-Mut cells. Statistical significance is calculated by Chi-squared test. **C,** Differential expression of repetitive genomic elements measured by RNA-seq of p53-Mut (SW620, DLD1) vs. p53-WT (HCT116, HCT8) tumorspheres grown for 14 days represented as volcano plot [*y*-axis $-\log_{10}(P)$ and *x*-axis $\log_2$(fold difference)]. Highlighted are satellite (SAT), LINE, and ERV repeats. Biological triplicates of RNA-seq were used for each cell line. **D,** Baseline levels of LINE element RNA L1PA2 (top) and L1HS (bottom) in p53-Mut and p53-WT cell lines as measured by RNA-seq. Significance determined by Student two-tailed *t* test with Welch correction calculated pairwise individually for HCT8 and HCT116 cells compared with SW620 and DLD1: *, $P < 0.05$; **, $P < 0.01$; ****, $P < 0.0001$. Red stars indicate significant difference compared with SW620. Black stars indicate significant difference compared with DLD1. **E,** HSATII RNA expression in p53-Mut (SW620, DLD1) and p53-WT (HCT8, HCT116) tumorspheres grown for 14 days, measured by RNA-ISH. Signal is quantified as percentage of HSATII RNA-positive signal per area in tumorspheres across 20 fields. Significance determined by Student two-tailed *t* test with Welch correction calculated pairwise individually for HCT8 and HCT116 cells compared with SW620 and DLD1: *, $P < 0.05$; **, $P < 0.01$; ****, $P < 0.0001$. Red stars indicate significant difference compared with SW620. Black stars indicate significant difference compared with DLD1.

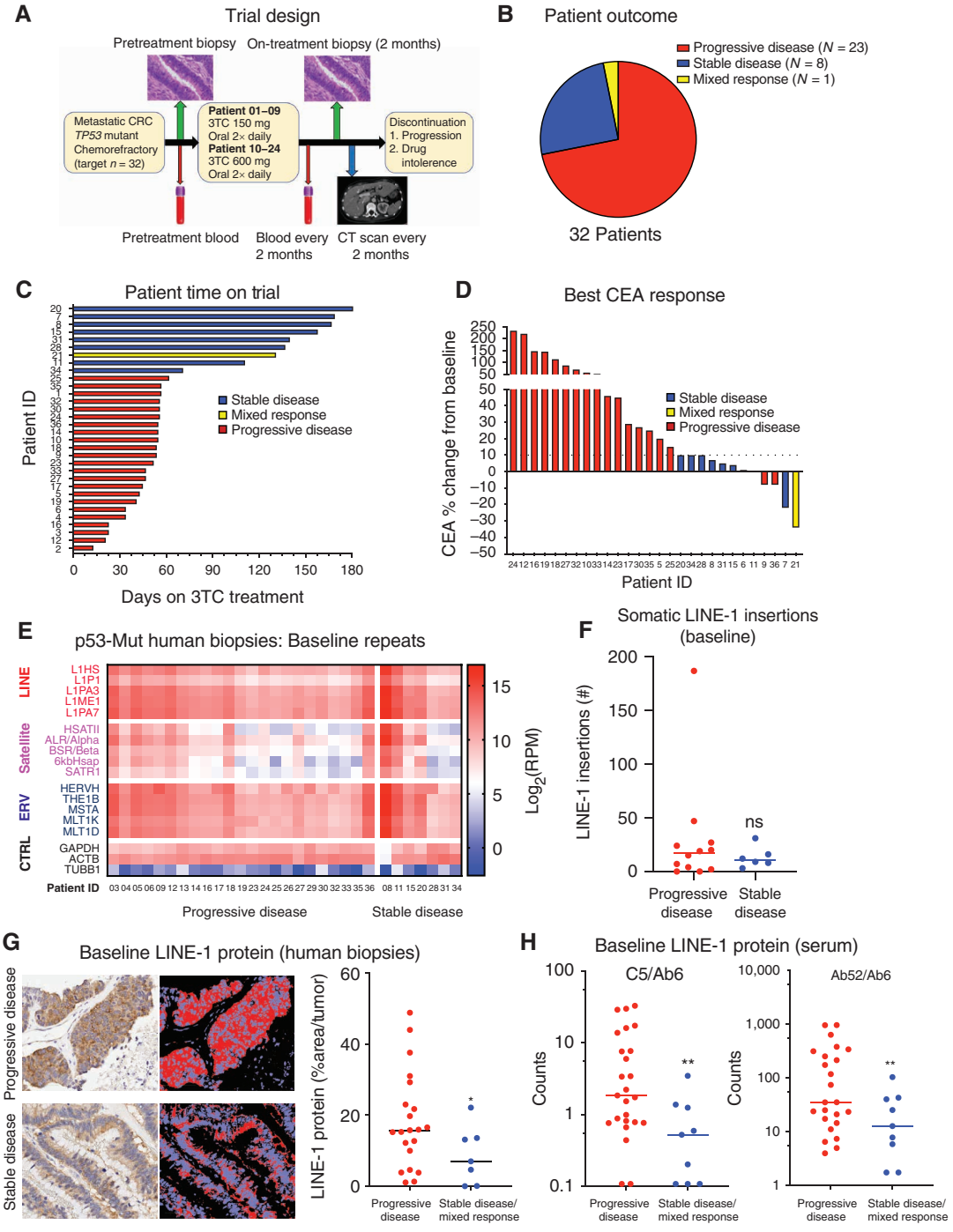


Figure 3. Phase II clinical trial of lamivudine (3TC) for colorectal cancer. **A,** Schematic of trial of 3TC single agent in p53-Mut colorectal cancer (CRC) with correlative blood, biopsy, and staging scans. **B,** Representation of patient response to 3TC. Eight of 32 patients achieved stable disease (SD) after treatment, and 1 patient achieved mixed response. **C,** Swimmer plot of time on 3TC treatment (*x*-axis days) for 32 patients (*y*-axis patient ID). **D,** Best serum CEA response in patients on the clinical trial. Patients with SD (blue: 7, 8, 11, 15, 20, 28, 31, 34) had unchanged or decreased serum CEA levels, whereas most patients with progression (PD; red) had increased CEA. Pt 21 had a mixed response (yellow) with the largest CEA response but had a new metastasis at restaging scans and for grouped with SD for comparative analysis. **E,** RNA-seq repeat RNA expression of SAT, L1, and ERV and housekeeping genes in pretreatment biopsy specimens from patients on the 3TC trial. Expression shown in $Log_2(RPM + 1)$. Pretreatment biopsies have high expression repeats, without significant differences in baseline repeat expression between patients with SD and PD. **F,** Somatic L1 genomic DNA insertions in pretreatment biopsies. No significant difference is seen in genomic L1 insertions between patients with SD and PD. **G,** L1 ORF1 protein (L1 ORF1p) expression in pretreatment biopsy specimens from patients on the 3TC trial. Representative image of L1 ORF1p IHC (left) in SD and PD specimen with image analysis markup (right). Quantification shown for each pretreatment biopsy in 28 patients. Patients with SD have significantly lower baseline L1 ORF1p. Significance determined by Student two-tailed *t* test with Welch correction: *, $P < 0.05$. **H,** Serum pretreatment L1 ORF1p levels in patients with PD and SD measured with two different antibodies (C5/Ab6 and Ab52/Ab6). Patients with SD have significantly lower baseline serum L1 ORF1p. Significance determined by Student two-tailed *t* test with Welch correction: **, $P < 0.01$.

that response rate is ≥10%. Secondary endpoint included median progression-free survival (PFS), overall disease control rate, and median overall survival (OS). The first 9 patients received 150 mg orally twice daily for 28-day cycles, the maximum FDA-approved dose of 3TC for HIV. After indication of safety in the first 9 patients, an institutional review board (IRB) amendment was approved allowing increased dosing to 600 mg orally twice daily for 28-day cycles, which was the highest dose evaluated in patients with HIV patients in phase I trials. Tumor assessments were performed every 8 weeks until documented disease progression by RECIST criteria or drug intolerance. The median age was 59 years (range, 27–83) with 18 males and 14 females. The median number of prior therapies was 3. The study did not meet the primary endpoint of response. However, remarkably, stable disease (SD) was seen in 8 of 32 (25%; Fig. 3B) patients (Pts 7, 8, 11, 15, 20, 28, 31, and 34) on single-agent 3TC with a median PFS of 149 days (Fig. 3C), with Pt 20 remaining on therapy for 230 days. Notably, 1 patient had mixed response with some reduction in tumor size of target lesions with a concordant 34% drop of the colon cancer serum marker carcinoembryonic antigen (CEA) but had developed new metastases (Pt 21). This patient had the largest maximal CEA response of all patients on trial and continued with 3TC treatment despite new metastases and remained on drug for 131 days given partial disease stability and limited adverse effects. The best CEA response was relatively unchanged (0%–10%) or decreased from baseline in all patients with SD (Fig. 3D). There were no grade 5 adverse events, and the majority of adverse events for patients on trial were grade 3 (18/32; 56.3%), which were mostly not attributable to treatment (Supplementary Tables S4 and S5). There were very few treatment-related grade 3 adverse events, which included anemia (1/32; 3.1%) and diarrhea (2/32; 6.3%). Altogether, these preclinical and clinical data supported the promising single-agent activity of 3TC with minimal adverse effects as a novel class of agents for chemorefractory metastatic colorectal cancer.

We then evaluated our pretreatment biopsies from the clinical trial to determine the patterns of repeat RNA, L1 protein expression, and repeat genomic retrotransposition in our patients. Pretreatment biopsies were obtained on 30 of 32 patients who were processed for RNA-seq, 28 of 32 were evaluable for L1 ORF1 protein (L1 ORF1p) IHC (Supplementary Fig. S4C and S4D), and 19 of 32 for whole-genome sequencing (WGS) with paired germline from peripheral blood mononuclear cells. RNA-seq demonstrated high expression of multiple repeat RNA classes across all patients, as expected given all patients on trial had a p53 mutation in their tumor (Fig. 3E). We looked for differentially expressed repeat RNA from patients with stable disease and Pt 21 mixed response (SD) compared with progressive disease (PD) on treatment, but we did not identify either any consistent repeat RNA difference in pretreatment biopsies or any clear differences in expression of coding genes after multiple hypothesis correction (Supplementary Fig. S5A and S5B and Supplementary Data S3 and S4). Analysis of tumor mutation with clinical targeted gene panels did not reveal any particular mutations enriched in patients with SD compared with PD (Supplementary Table S3). Analysis of LINE-1 retrotransposition using the xTea platform (29) on WGS of paired baseline tumor and peripheral blood germline DNA samples also did not reveal any significant differences in transposition frequency between SD and PD tumors (Fig. 3F; Supplementary Fig. S5C and S5D). However, L1 ORF1p levels determined by quantitative analysis of IHC noted significantly lower L1 levels in patients with SD compared with PD (Fig. 3G; *t* test $P < 0.05$), suggestive of patients with PD having sufficient L1 RT activity to overcome 3TC at the doses given in this trial. Given L1 ORF1p immunostaining was done on a biopsy of a single lesion, we wanted to see if L1 ORF1p could be found systemically in our patients using a highly sensitive microbead immunoassay (Simoa) technology to quantify circulating L1 ORF1p levels (30). Using two independent L1 ORF1p antibodies, we found significantly lower levels of plasma L1 ORF1p in our patients with SD compared with PD (Fig. 3H). This may indicate differences in disease burden or may suggest single-agent 3TC may not have the potency to block tumors with higher RT capacity. In summary, disease stability on 3TC is associated with lower baseline levels of L1 ORF1 protein, but not expression of specific repeat RNAs or somatic retrotransposition that points to 3TC effects on the repeat RNA lifecycle as a proximal event before genomic insertion.

3TC Decreases Cytoplasmic DNA Levels and Increases RNA:DNA Hybrids

To better understand the broader effects of 3TC on colorectal cancer tumors, we evaluated all available paired pretreatment and on-treatment biopsies from the trial. Analysis of paired total RNA-seq ($n = 13$) found a consistent pattern of significantly decreased expression of most repeats (Fig. 4A–C). The same pattern of repeat downregulation was seen in both tumorspheres and xenografts treated with 3TC (Supplementary Figs. S6A–S6F; S7A–S7E; Supplementary Data S5). Although we did not find individual coding genes that were statistically different between pre- and on-treatment biopsies, gene ontology (GO) analysis noted significant upregulation of interferon response genes, as has been reported to be associated with repeat RNA (7, 8, 10, 11, 13) and DNA (22, 23) sensing (Fig. 4D; Supplementary Table S6). We found no changes in retrotransposition between pre- and on-treatment biopsies (Supplementary Fig. S7). Altogether, these findings indicate that repeat RNA levels decrease in patients and our cell line models from 3TC treatment with no detectable changes in LINE-1 retrotransposition. Given the decreased repeat RNA seen, we hypothesized that these changes might represent alterations in complementary DNA (cDNA) or RNA:DNA hybrids generated from RT, as PCR-based sequencers cannot differentiate native cDNA and RNA species from RNA-seq preparation using exogenous RT. To evaluate the potential effects of 3TC on RT products, we analyzed cytoplasmic DNA and RNA:DNA hybrids in our colorectal cancer model system. We performed quantitative immunofluorescence (IF) for cytoplasmic double-stranded DNA (dsDNA; Supplementary Fig. S8A) using an anti-dsDNA antibody as previously reported (17) and RNA:DNA hybrids using the S9.6 antibody (Supplementary Fig. S8B). Notably, we found significant reduction in cytoplasmic dsDNA in SW620 and DLD1 cells treated with 3TC compared with DMSO control (Fig. 4E; *t* test $P < 0.0001$), but not in HCT116 or HCT8 (Supplementary Fig. S9A). RNA:DNA hybrid quantification showed a

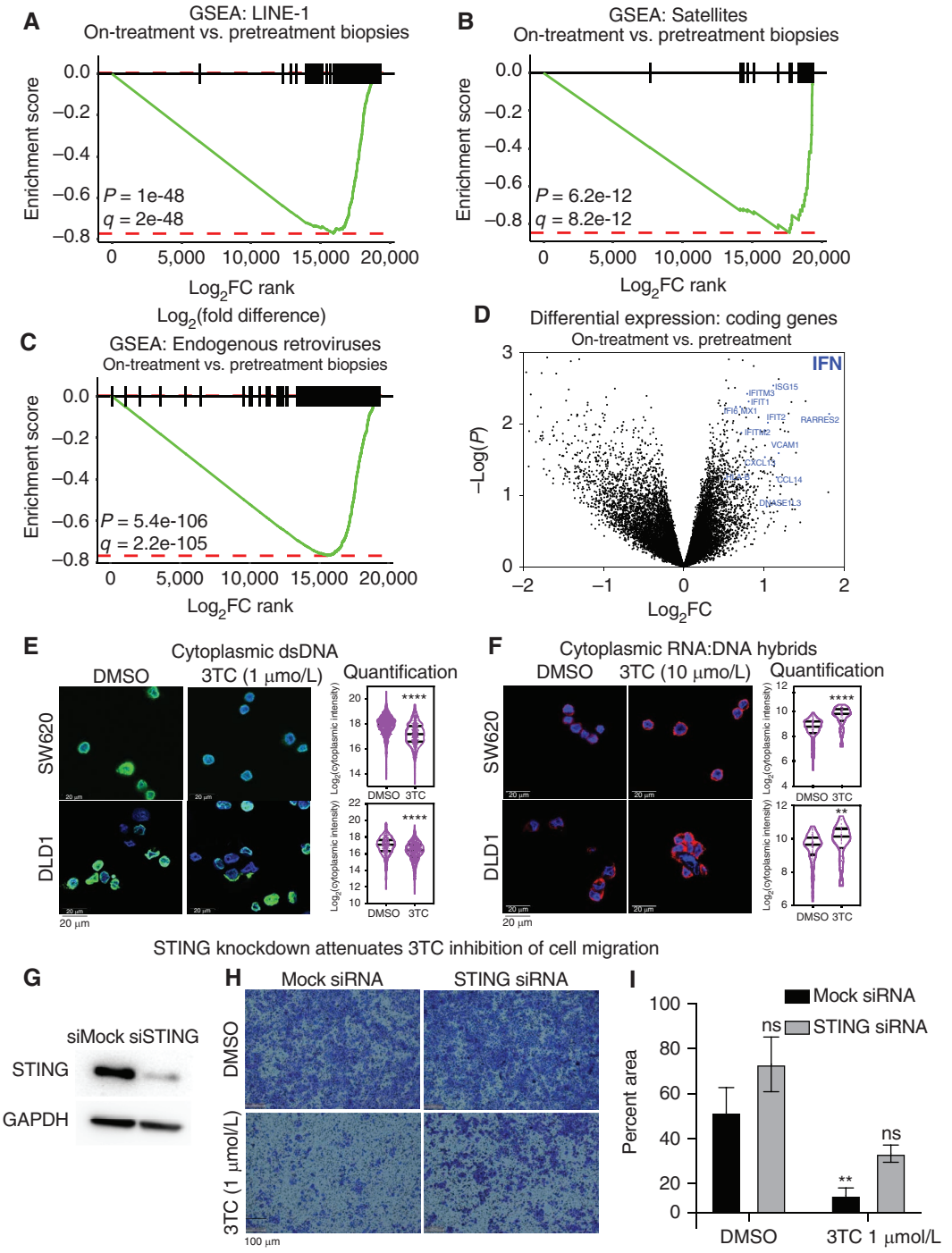


Figure 4. RT inhibition leads to decreased cytoplasmic DNA and increased RNA:DNA hybrid intermediates. **A–C,** GSEA comparing 3TC patient pretreatment and on-treatment biopsy RNA-seq (**A**) L1, (**B**) satellite, and (**C**) ERV repeats. Repeats were ranked by $\log_2$(fold change, on-treatment vs. pretreatment tumor sample) for GSEA. Significant decrease in L1, satellite, and ERV repeat levels after treatment with 3TC shown in patient biopsies with *P* value and corrected q value shown (see Supplementary Methods). **D,** Volcano plot representing differential expression (Log_2FC) vs. significance (–Log *P*) of coding genes in tumor biopsies before treatment and after treatment with 3TC. Interferon response genes (blue) are significantly upregulated after treatment with 3TC. **E,** Immunofluorescence (IF) of cytoplasmic dsDNA in p53-Mut tumorspheres treated with DMSO or 3TC (1 μmol/L) for 7 days. Quantification of cytoplasmic IF signal indicates significant decrease in cytoplasmic dsDNA in response to 3TC. Significance determined by two-tailed *t* test with Welch correction: ****, $P < 0.0001$. **F,** IF of cytoplasmic RNA:DNA hybrids in p53-Mut tumorspheres treated with DMSO or 3TC (10 μmol/L) for 7 days. Quantification of cytoplasmic IF signal indicates significant increase in cytoplasmic RNA:DNA in response to 3TC. Significance determined by two-tailed *t* test with Welch correction: ****, $P < 0.0001$. **G–I,** Rescue of effects of 3TC on cell migration by STING knockdown. **G,** Western blot analysis confirms knockdown of STING (*TMEM173*) protein using pooled siRNA compared with nontargeting control in DLD1 cells. **H,** Representative images of DLD1 cell migration with siRNA against STING (right) or Mock (left) and treated with DMSO or 3TC. **I,** Quantification of DLD1-migratedcells in siRNA STING and 3TC experiments in **H**. Statistical significance determined by Student two-tailed *t* test as compared with nontargeting control siRNA treated with DMSO: **, $P < 0.01$.

significant increase with 3TC treatment in both p53-Mut (SW620; DLD1) and p53-WT (HCT116; HCT8) cells treated with 3TC compared with DMSO control (Fig. 4F; Supplementary Fig. S9B). Given the alterations in cytoplasmic RT products from 3TC, we evaluated the potential impact of STING, a known receptor for cytoplasmic DNA (17) and RNA:DNA hybrids (31), in our cell line models. Interestingly, siRNA-mediated reduction in STING partially rescued the antimigratory effects of 3TC in the DLD1 cell line (Fig. 4G–I). This indicated that 3TC reduces cytoplasmic dsDNA and increases RNA:DNA hybrids that block migratory function in p53-Mut colorectal cancer cell lines partially through STING signaling. To evaluate potential other mechanisms for antimigratory effects of 3TC, we analyzed coding genes differentially expressed between colorectal cancer tumorspheres treated with 3TC compared with DMSO (Supplementary Fig. S10A and Data S6). This revealed a number of genes downregulated (*t* test FDR < 0.15), and we noted *S100A4* as a gene that is well known to be a negative prognostic marker in colorectal cancer (32–34) and demonstrated functional effects on metastatic potential in preclinical models (35–37). Suppression of *S100A4* using siRNA noted reduced migration of p53-Mut colorectal cancer cell lines (DLD1 and HCT15) relative to p53-WT cell lines (HCT116 and HCT8) when compared with mock siRNA controls (Supplementary Fig. S10B and S10C). Altogether, these findings support a combination of 3TC effects on cytosolic nucleic acid species and *S100A4* expression that reduce migratory function in colorectal cancer cell lines.

Combination Repeatome Targeting Induces Cancer Cell Death through Necroptosis

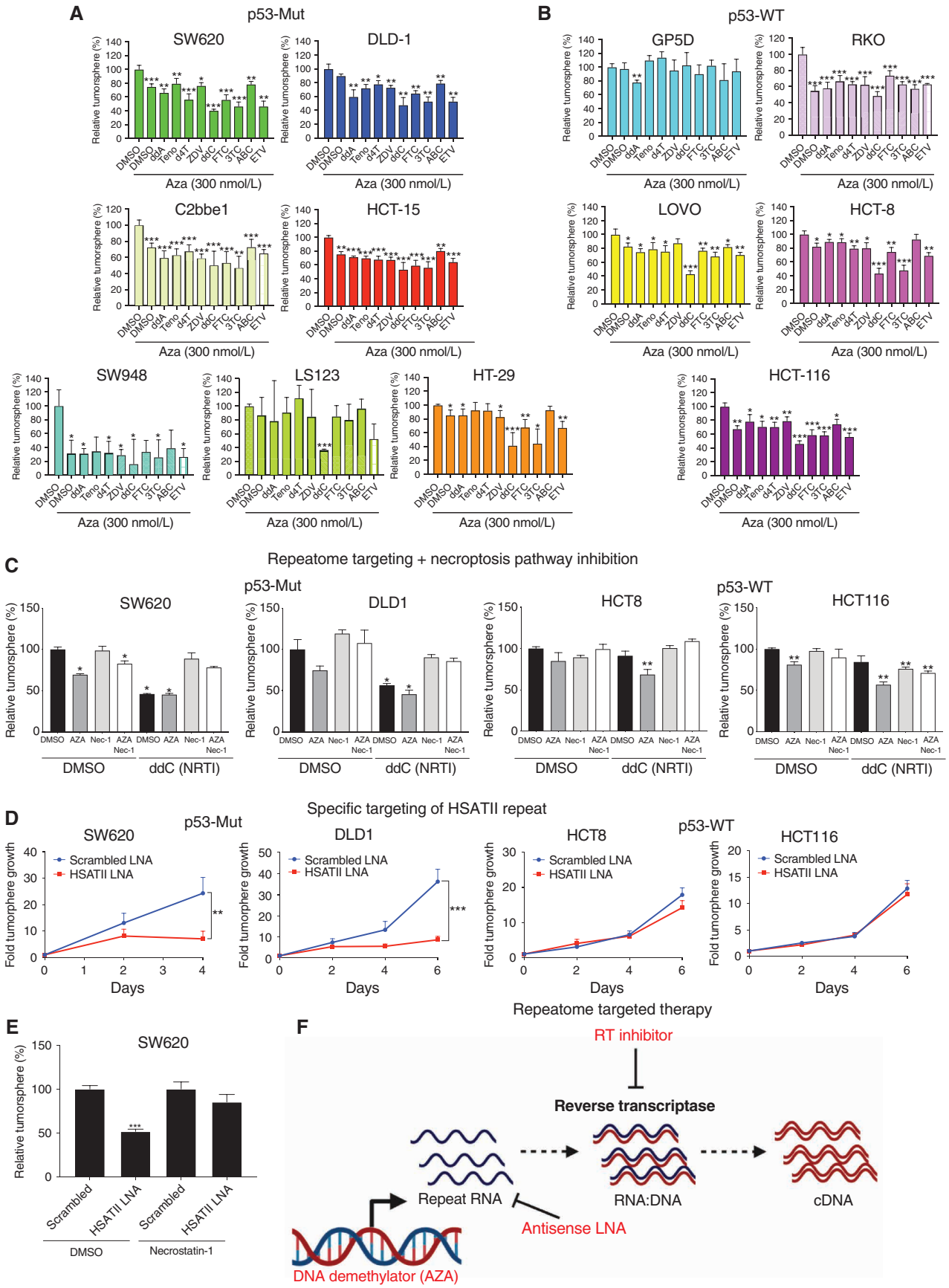
Given the effect of 3TC on the cytosolic pool of nucleic acid species, we explored the possibility of combining NRTIs with the DNA hypomethylating agent 5-azacitidine (AZA), which has been shown to derepress a wide range of repeat RNAs and activate an interferon response (3, 7, 8). We evaluated a broad panel of 9 NRTIs and combined with low-dose AZA (300 nmol/L) to achieve epigenetic repeatome effects, which resulted in significant reduction in tumorsphere growth across both p53-Mut and p53-WT colorectal cancer cell lines (Fig. 5A and B). The efficacy of AZA in our cell lines is consistent with analogous work in mouse cancer cell lines demonstrating significant induction of repeat species triggering and IFN response linked with p53 status (7). Evaluation of all NRTIs revealed that ddC had the most consistent cytotoxic effect across our colorectal cancer cell lines, and in general the C analogues (FTC, 3TC, and ddC) and G analogue [entecavir (ETV)] had the most toxicity in our colorectal cancer cell lines. This indicates that not all NRTIs have the same functional effects, which may be driven by differences in CpG sequence motif biases in repeat RNAs (13). We next investigated the mechanism of cancer cell toxicity induced by repeat dysregulation and focused on necroptosis given recent work indicating that mutant p53 primes epithelial cells to necroptosis (38). Inhibition of the necroptosis effector RIPK1 by the small-molecule inhibitor necrostatin-1 was sufficient to rescue multiple colorectal cancer cell lines from ddC-, AZA-, and combination ddC + AZA–mediated toxicity (Fig. 5C).

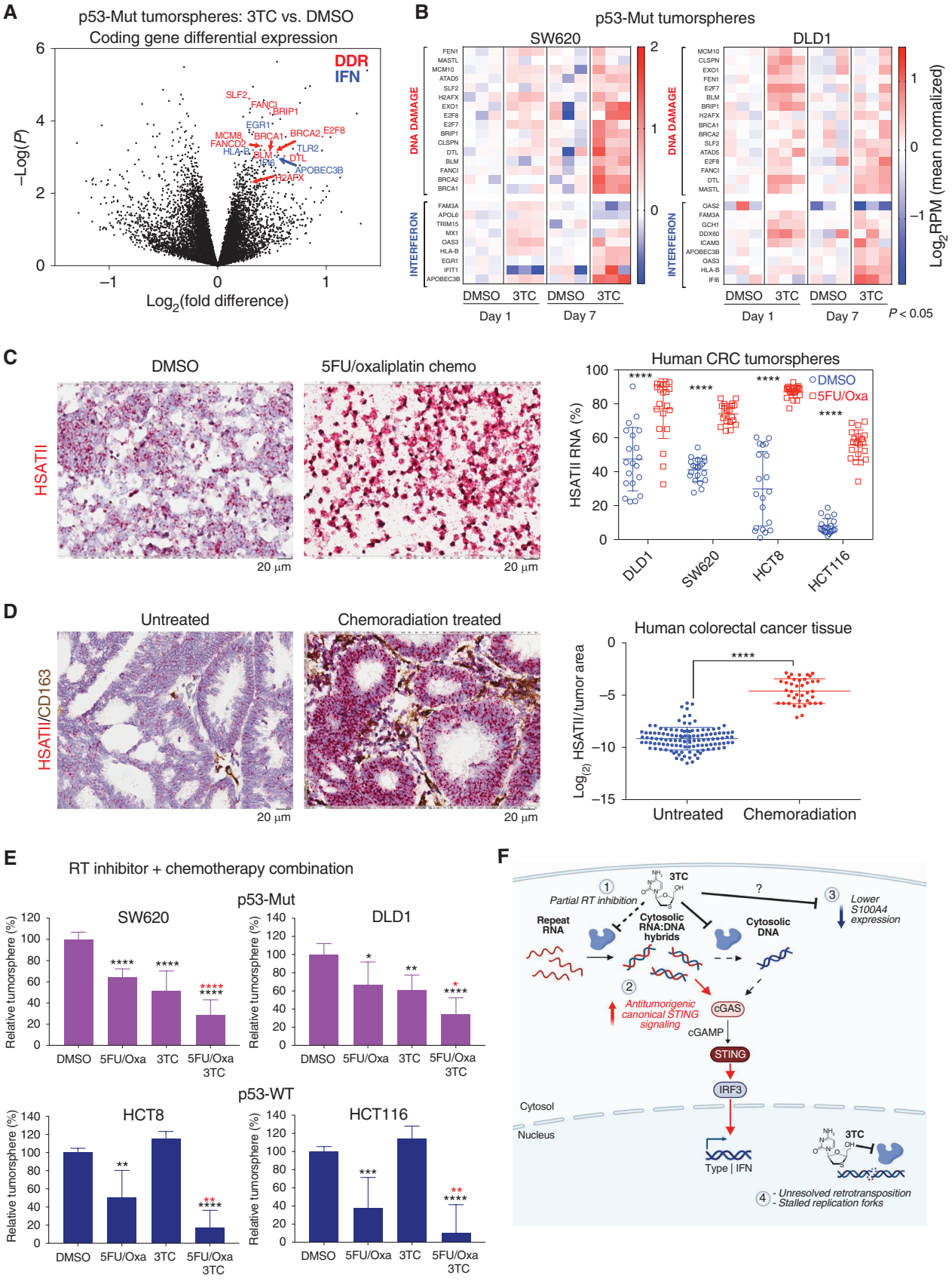
We next wanted to determine if the effects seen with ddC and AZA were linked specifically to repeat mediated effects. We chose to focus on the HSATII satellite repeat given the consistent loss of HSATII binding by mutant p53 (Fig. 2C), the associated higher HSATII RNA expression (Fig. 2E), and our previous success in targeting HSATII with locked nucleic acids (LNA) in colorectal cancer cell lines (16). We modified these original HSATII LNAs with phosphorothioate to enhance cellular uptake and LNA stability to maximize cancer cell toxicity (see Methods). We found specific inhibition of p53-Mut (SW620, DLD1) compared with p53-WT (HCT8 and HCT116) cell line growth with HSATII LNAs compared with scrambled LNA controls (Fig. 5D). Necrostatin was able to rescue cell toxicity in the SW620 cell line, which supports necroptosis as the mechanism of cytotoxicity induced with HSATII modulation (Fig. 5E). Collectively, these data support a model of repeat mediated IFN response and necroptotic cell death induced by agents that disrupt genomic repeat epigenetic suppression, directly target the repeat RNA, or inhibit repeat RT (Fig. 5F).

RT Inhibition Results in Increased DNA Damage

Emerging data have recently demonstrated a relationship between genomic instability with satellite RNA expression (39–41) and LINE-1 activity (42, 43), but the effects of blocking repeat RNA reverse transcription in the setting of high retrotransposon activity have not been fully characterized. To understand the acute effects of 3TC, we analyzed total RNA-seq data from p53-Mut colorectal cancer cell lines treated for 1 and 7 days. We noted significant upregulation (*t* test FDR < 0.15) of innate immunity and interferon response genes, including *EGR1*, *IFI6*, *HLA-B*, *TLR2*, and *APOBEC3B* (Fig. 6A), consistent with our findings in paired patient biopsies (Fig. 4D). Unbiased GO analysis of upregulated genes (*t* test FDR < 0.15) was highly enriched for gene signatures involved with DNA damage response, including the genes *BRCA1*, *BRCA2*, *BRIP1*, *FANCI*, and *FANCD2* (Fig. 6A; Supplementary Table S7 and Supplementary Data S6). Western blot analysis for gamma-H2AX confirmed increased double-strand DNA breaks in cell lines treated with 3TC (Supplementary Fig. S11A). Downregulated genes were enriched for genes involved with amino acid metabolism and endoplasmic reticulum stress including multiple aminoacyl-tRNA synthetase

Figure 5. Combination repeatome targeting induces cancer cell death through necroptosis. **A** and **B,** Efficacy of combination of 9 RT inhibitors (5 μmol/L) with DNA demethylating agent AZA at 300 nmol/L in (**A**) p53-Mut and (**B**) p53-WT colorectal cancer tumorspheres measured by CellTiter-Glo viability assay after 7 days of treatment. **C,** Necroptosis pathway inhibition by treatment with RIPK1 inhibitor necrostatin-1 (Nec-1) at 10 μmol/L attenuates effect of NRTI ddC (5 μmol/L) and DNA demethylating agent AZA (300 nmol/L) as measured by CellTiter-Glo assay after 7 days of treatment. Significance determined by Student two-tailed *t* test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. **D,** Treatment with antisense LNA targeting HSATII in p53-Mut cell lines (SW620 and DLD1) and p53-WT (HCT8 and HCT116) cell lines grown as tumorspheres compared with scrambled LNA, as measured by CellTiter-Glo assay. Significance determined by two-way ANOVA test: **, $P < 0.01$. **E,** Necroptosis pathway inhibition with Nec-1 (10 μmol/L) rescues effect of anti-HSATII LNA in SW620 tumorspheres. Significance determined by Student two-tailed *t* test: ***, $P < 0.001$. **F,** Schematic showing multiple avenues for repeatome targeting in cancer.





genes, which implies effects on translational machinery (Supplementary Table S8 and Supplementary Data S6). Temporal analysis of post-3TC exposure demonstrated DNA damage and interferon/innate immune genes starting at day 1 and increasing expression at day 7 posttreatment in the p53-Mut cell lines (Fig. 6B). As LINE-1 RT can act at DNA damage breakpoints (15, 42, 43), the accumulation of DNA damage and interferon signatures in cell lines and patient biopsies suggests a model of 3TC-mediated disruption of the retroviral-like life cycle of repeats leading to RNA:DNA hybrid accumulation that enhances the ability of the innate immune system to detect their presence and leads to unresolved DNA damage. Given the enhanced DNA damage response signatures in colorectal cancer cell lines, we wanted to determine the relationship of repeat RNA expression in the setting of DNA-damaging agents. We treated colorectal cancer cell lines with the standard combination chemotherapy 5FU/oxaliplatin (5FU/Oxa) for 14 days and detected marked elevation of HSATII RNA by RNA-ISH (Fig. 6C; Supplementary Fig. S11B). We then applied HSATII RNA-ISH to 160 human primary colorectal cancer tumors that were untreated or pretreated (neoadjuvant) with chemoradiation before resection, which demonstrated significant enrichment of HSATII repeat RNAs in tumors that received cytotoxic therapy (Fig. 6D). To determine if this had therapeutic implications, we treated colorectal cancer lines with 5FU/Oxa ± 3TC, which showed significantly increased cytotoxicity in all 4 cell lines with the combination compared with 5FU/Oxa alone (Fig. 6E). These combination approaches we present boost repeat RNA expression and inhibit RT activity that leads to accumulation of immunostimulatory RNA:DNA hybrids. Overall, these data support a model of RT inhibition leading to enhanced DNA damage due to unresolved, chain terminated intermediates of retrotransposition or secondary DNA-damaging effects from alterations in the balance of cytosolic RT products (Fig. 6F).

DISCUSSION

The combination of preclinical and clinical results provides a path for the evaluation of NRTIs as a new class of anticancer therapeutics. We have shown 3TC is able to decrease migration of mutant *TP53* colorectal cancer cell lines partially through a STING-dependent nucleic acid mechanism and reduction in *S100A4* levels. Reduction of cytoplasmic dsDNA species and increased RNA:DNA hybrids from 3TC appears to trigger STING signaling and induce DNA damage. These combined effects of 3TC (Fig. 6F) had antitumorigenic effects in mouse models and potential benefit to patients in a clinical trial. This unexpected RNA:DNA elevation with 3TC suggests incomplete RT inhibition with potential reengagement of RT for multiple rounds of synthesis or delayed chain termination after initial RT of the polyA tail; however, these mechanisms require further elucidation. Cytoplasmic L1 DNA has been found elevated in mouse models of senescence, and, notably, 3TC has been shown to reduce L1 cytoplasmic DNA in senescence model systems (22, 23). Although the effects we have observed could be from L1 repeat species, we acknowledge that 3TC effects on RNA:DNA hybrids and cytoplasmic DNA species likely affect a broad range of RNA species that can be processed by L1 (44). The associated accumulation of immunogenic RNA:DNA species drives an interferon response potentially through cGAS–STING (31), DDX41 (45), or other nucleic acid cytoplasmic sensor. We demonstrate in our colorectal cancer cell lines that STING is likely involved in RNA:DNA detection that attenuates migratory capability. Notably, others have shown that increased cytoplasmic DNA levels are linked with genomic instability and increased metastatic potential in cancers through noncanonical chronic STING signaling (17). Our findings are complementary to this work by demonstrating the importance of the nuances in STING signaling on cancer cell phenotype and the use of 3TC and other NRTIs providing a therapeutic modality to decrease the protumorigenic effects of cytoplasmic DNA.

Cytosolic nucleic acid species have been shown to trigger IFN response and adaptive immunity across mouse model and human tumors (6–13) with recent papers demonstrating a relationship to immune-checkpoint response (46–48). Similarly, we found increased IFN response from 3TC affecting the repeat life cycle, which was also seen in recent work demonstrating the association of IFN response gene activation of normal colon samples from non-HIV patients on NRTIs for antiviral prevention (49). These findings may explain in part the apparent decreased incidence of breast, prostate, and colorectal cancers in patients living with HIV who are on stable anti-HIV regimens that include NRTIs (50). Finally, we speculate that incomplete RT creates DNA lesions that are difficult to repair or prevents the retrotransposition-mediated repair of other forms of DNA damage. Altogether, our preclinical and proof-of-concept clinical trial provides the foundation to evaluate combination RT inhibitors to obtain more potent effects of disrupting the cancer repeatome, a strategy that clearly changed the course of HIV treatment (51, 52). The lack of significant dose-limiting toxicity of 3TC in

Figure 6. RT inhibition induces DNA damage in colon cancer. **A,** Volcano plot of coding gene RNAs with 3TC (5 μmol/L) or DMSO treatment for 7 days in p53-Mut cell lines (DLD1 and SW620; $n = 3$ for 2 independent experiments per cell line). Highlighted are DNA damage response (DDR) and interferon/innate response (IFN) genes. **B,** Tumorsphere RNA-seq expression heatmap of p53-Mut colorectal cancer cell lines for DDR and IFN genes at days 1 and 7 after 3TC vs. DMSO treatment. Expression is Log2RPM normalized to mean of DMSO per time point and cell line. Genes represented have significantly increased expression in response to 3TC at either day 1, or day 7, or both ($P < 0.05$ Student two-tailed t test). **C,** HSATII expression in HCT8 tumorspheres grown in the presence of DMSO (control) or 50 μmol/L 5FU + 1.25 μmol/L oxaliplatin (5FU/Oxa) for 2 weeks. Scale bar, 20 μm. HSATII RNA-ISH in DLD1, SW620, HCT8, and HCT116 tumorspheres treated with DMSO or 5FU/Oxa for 2 weeks. Plots represent HSATII staining as a percentage of tumor area across 20 fields. Significance determined by two-tailed t test with Welch correction: ****, $P < 0.0001$. CRC, colorectal cancer. **D,** HSATII RNA-ISH on human colorectal cancer tumors from untreated patients (left) and patients who received neoadjuvant chemoradiation (right). Scale bar, 20 μm. HSATII RNA levels quantified as a percentage of the tumor area. Significance determined by the Student two-tailed t test with Welch correction: ****, $P < 0.0001$. **E,** Cell viability of colorectal cancer tumorspheres treated with either DMSO or 5 μmol/L 3TC for 7 days in the presence of 5FU/Oxa. Significance determined by the Student two-tailed t test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$. Black stars indicate statistical significance compared with DMSO alone. Red stars indicate statistical significance compared with 5FU/Oxa. **F,** Schematic of multiple effects of 3TC on colorectal cancer including alterations in cytosolic nucleic acids (1 and 2), S100A4 expression (3), and DNA damage (4).

patients with metastatic colorectal cancer affords the use of 3TC or related compounds in combination with both existing and novel cancer therapies in future clinical trials to augment the efficacy of these drugs as we have shown with 5FU/Oxa and AZA in colorectal cancer cell line models. Interestingly, a recent paper has also shown that reactivation of p53 with a novel MDM2 inhibitors can also induce HERV expression through epigenetic targets that triggers an antitumoral immune response (53). These intricate relationships between tumor suppressors, epigenetics, and repeat RNA expression will need to be further elucidated to fully understand the impacts of these drugs alone and in combination. The inherent massive derepression of repeat RNAs in cancer provides a rich opportunity to target this genomic "dark matter" as a therapeutic strategy in human disease.

METHODS

Cell Culture

Cell lines were originally obtained from ATCC. Cell lines were cultured in either RPMI-1640 (SW620, HCT116, HCT8, DLD1, HCT15, and LOVO), DMEM (GP5D), or DMEM/F12 (SW948, C2bbe1, and RKO) media supplemented with 10% FBS + 1% penicillin/streptomycin (Gibco/Life Technologies). Cell lines were maintained in adherent 2-D culture on standard tissue culture plates (Corning). For subculturing, media were aspirated, and cells were washed in cell culture grade PBS (pH 7.4, Thermo Fisher). Cells were passaged using 0.25% Trypsin-EDTA (Thermo Fisher) for 5 minutes, and reaction was inhibited using media with serum. Cells were centrifuged and resuspended in media for subculture. For 3-D culture (tumorspheres), serum-free media (RPMI, DMEM, or DMEM/F12) was supplemented with 20 μL/mL B27 + 20 ng/mL EGF + 20 ng/mL bFGF (Thermo Fisher). Cells were plated in ultralow attachment plates (Corning). All cell lines were routinely tested for *Mycoplasma* on a monthly basis, or earlier if there were indications of slower cell growth or changes in morphology. All cell lines were used within 2 months of thawing for experiments. Cell line authentication was not performed on all cell lines. All cell lines will be provided upon request with appropriate Material Transfer Agreement approvals.

Drug Treatment

NRTIs lamivudine (3TC), zalcitbine (ddC), emtricitabine (FTC), entecavir (ETV), abacavir (ABC), ddA, tenofovir (Teno), zidovudine (ZDV), and stavudine (d4T), and DNA demethylating agent AZA were purchased from Selleck Chemicals. Drugs were reconstituted in tissue culture grade DMSO (Sigma). The RIPK1/necroptosis inhibitor necrostatin-1 was obtained from Sigma-Aldrich and reconstituted in DMSO. Chemotherapy drug 5FU was obtained from Invivogen. Oxaliplatin was obtained from Selleck Chemicals.

Migration Assays

Colon cancer cells were preincubated with or without 3TC into ultralow attachment plates for 24 hours. Then, cells were dissociated to single-cell suspensions and seeded onto 8-μm pore Transwell membranes (CELLTREAT Scientific Products; cat. #230639) in serum-free media at 50,000 to 100,000 cells/well in a 24-well plate containing complete growth media (10% FBS) in the bottom chamber. 3TC was treated both upper chamber and bottom chamber during migration. After 24 to 96 hours at 37°C, Transwell chambers were fixed with methanol for 10 minutes and stained with 0.1% of crystal violet (Sigma-Aldrich; cat. #V5265), and cells on top of the chamber were stripped off with cotton swabs. Images were acquired, and the total crystal violet–stained area was quantitated using HALO software.

Soft-Agar Assays

Cells were seeded at a density of 5,000 cells per well in a 6-well plate of 0.3% agarose (Seaplaque, Lonza) in RPMI containing 10% (vol/vol) FBS in the presence of DMSO or lamivudine (3TC). After 14 days of growth cells were stained in 0.005% crystal violet solution in 4% PFA/PBS solution for 30 minutes followed by 3× washes in water. Crystal violet–stained colonies were imaged using a dissection microscope and quantified using HALO software.

Cell Viability Assays

Cell viability assays were conducted on both adherent cell (2-D) colorectal cancer cell lines, as well as cells grown as tumorspheres (3-D) in response to drug treatment. Cells were plated at a concentration of 5,000 cells per well of 96-well plates (Corning). Cells were treated for the indicated concentrations as described for the indicated number of days, and cell viability was measured using the CellTiter-Glo reagent (Promega) according to the manufacturer's recommendations, and luminescence was measured using the Spectramax i3X (Molecular Devices) plate reader. Percent cell/tumorsphere viability was represented as luminescence units normalized to DMSO control.

siRNA Knockdown

STING. Colorectal cancer cells were transfected with pooled siRNA targeting STING (Horizon Discovery, TMEM173 SMARTpool L-024333-00-0005) or Nontargeting siRNA (Horizon Discovery, D-001810-01-05) with Lipofectamine 2000 at a concentration of 100 nmol/L. Knockdown was assessed 48 hours posttransfection using Immunoblotting for STING (TMEM173).

S100A4. Transfection with siRNA was performed with Lipofectamine 2000 reagent (Thermo Fisher Scientific; cat. #11668019) according to the manufacturer's instructions. 100 nmol/L siRNA targeting human *S100A4* ON-TARGET plus SMARTpool (GE Healthcare Dharmacon), mixture of 4 siRNAs targeting the following sequences: 5′-GAUGUGUAACGAAUUCUUU-3′, 5′-CCACAAGUACUCGGGCAAA-3′, 5′-GUGACAAGUUCAAGCUCAA-3′, 5′-GAAAACUCCUCUGAUGUGG-3′, or ON-TARGETplus Nontargeting siRNA as negative control (GE Healthcare Dharmacon) were used. Knockdown efficiency was assessed by qRT-PCR, and migration assay was performed for 48 hours.

Quantitative RT (qRT-PCR)

RNA was isolated using the miRNeasy Mini Kit (Qiagen; cat. #217004) according to the manufacturer's instructions including on-column DNase treatment (Qiagen; cat. #79254). The 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) according to the manufacturer's instructions. Primers used include the following: GAPDH Forward 5′-ACATCATCCCTGCCTCTACT-3′, Reverse 5′-TCCACCACTGACACGTTG-3′; S100A4 Forward 5′-CAGAACTAAAGGAGCTGCTGACC-3′, Reverse 5′-CTTGGAAGTCCACCTCGTTGTC-3′. Reactions were run on a QuantStudio 3 (Applied Biosystems) thermocycler. The level of gene expression was calculated by the $2^{-\Delta\Delta CT}$ method and normalized to the C_t value for GAPDH.

Antisense Locked Nucleic Acid Treatment

Colorectal cancer cells were transfected with antisense locked nucleic acids (Exiquon, scrambled LNA: GATTCCATTCGATGAT, anti-HSATII LNA: +A*+T*+g*+G*A*A*T*C*A*T*C*A*T*+C*+G*+A*+A) with Lipofectamine 2000 at a concentration of 500 nmol/L. Cell viability assay was conducted 24 hours after LNA treatment.

Animal Studies

Mouse xenograft studies were performed according to an animal protocol approved by the MGH Institutional Animal Care and Use Committee (IACUC) under protocol 2014N000321. SW620 or HCT116 cell lines stably transduced with luciferase (1×10^6) were injected in the flank of 6-week-old female nude mice (Charles River Laboratories) at a 1:1 ratio with Matrigel (Sigma-Aldrich). Tumors were allowed to grow for 2 weeks after implantation. At 2 weeks before beginning drug treatment, mice were intraperitoneally injected with luciferin (Promega) to visualize luciferase expression. Tumors were visualized through Luciferase Luminescence measurement with the IVIS Spectrum Imaging System (PerkinElmer). Mice were randomized and divided into treatment arms: control (PBS) and lamivudine (3TC). 3TC was solubilized in PBS and administered intraperitoneally at a dosage of 50 mg/kg 3 times a week. In the control group, PBS was administered intraperitoneally. Mice were treated for 20 days, and tumor volume was measured every 5 days through IVIS imaging. At 20 days, mice were euthanized according to IACUC guidelines. Animals did not show any sign of systemic toxicity upon drug administration. Tumor growth curves were calculated using relative luminescence units normalized to day 0 of treatment. After euthanasia, tumors were dissected and used for RNA isolation using TRIzol or standard formalin-fixed, paraffin-embedded (FFPE) processing.

Clinical Trial

This is a single-arm phase II clinical trial of 3TC in patients who have progressed on systemic therapy for metastatic colorectal cancer with *TP53* mutations (NCT03144804). The trial was reviewed and approved by the Dana-Farber Harvard Cancer Center (DFHCC) IRB (Protocol 17-044). Eligibility included patients with *TP53*-mutant refractory metastatic colorectal cancer with progression on or intolerance to 5FU, oxaliplatin, and irinotecan and anti-EGFR therapy if RAS WT. Eligible patients were age ≥18 years, with histologically or cytologically documented, advanced (metastatic and/or unresectable) disease that was incurable and had disease progressed on at least two lines of systemic therapy. Patients were required to provide an evaluable tissue sample for biomarker analysis from a tumor lesion not previously irradiated; had radiologically measurable disease per RECIST version 1.1 as assessed by independent central radiologic review; had an Eastern Cooperative Oncology Group performance status of 0, 1, or 2; and had adequate organ function. Key exclusion criteria included uncontrolled intercurrent illness, HIV-positive patients on combination antiretroviral therapy, as well as patients with hepatitis B. The study was designed as a two-stage disease with target accrual of 20 patients in stage I and a total of 36 patients with continuation beyond stage I allowed if at least one patient demonstrated a response. Primary endpoint of the study was to describe overall response rate with the null hypothesis that the response rate is ≤1% versus the alternative that response rate is ≥10%. Secondary endpoint included median PFS, overall disease control rate, and median OS. The first 9 patients received a dose of 150 mg orally twice daily for 28-day cycles, the maximum FDA-approved dose of 3TC for HIV. The subsequent patients received 600 mg orally twice daily for 28-day cycles. Tumor assessments were performed every 8 weeks until documented disease progression by RECIST criteria or drug intolerance. CEA was obtained as part of clinical care.

RNA-seq

RNA was isolated using TRIzol reagent (Thermo Fisher) or the AllPrep DNA/RNA/protein mini kit (Qiagen; #80004) according to the manufacturer's recommendations. RNA-seq was performed using the two different sequencing library construction methods. For cell lines and xenografts, we performed Illumina TruSeq stranded RNA-seq Ribo-zero TM Gold (Illumina RS-122-2301) library construction per protocol for coding gene transcriptional profiling. Given known issues with repeat RNA quantification with standard RNA-seq protocols, we utilized the Clontech/Takara SMARTer Stranded Total RNA-seq PICO v2 kit (Clontech/Takara 634414) for repeatome profiling. Given limited sample material in biopsies, only the Clontech/Takara SMARTer Stranded Total RNA-seq PICO v2 kit (Clontech/Takara 634414) was performed on samples. All libraries then had 75-bp paired-end sequencing on the NextSeq500 using a 150-cycle high-output kit.

RNA-seq Computational Analysis

Illumina reads were mapped to the human or mouse genomes (or their union for xenograft samples), build 38 using STAR aligner (54). Aligned reads were assigned to genes using the featureCounts function of Rsubread package with the external Gencode annotations (55, 56). Counts for repeat elements were obtained using RepeatMasker annotation (https://www.repeatmasker.org). This produced the raw read counts for each gene and repeat element. Gene expression in terms of $\log_2$ CPM (counts per million reads) was computed and normalized across samples using the trimmed mean of M-values method (TMM), as implemented in the calcNormFactors function of edgeR package (57, 58).

Differential Expression. For comparisons with a small number (typically two or three) of samples per phenotype, we used DESeq2 (59), which employs moderated estimate of dispersion useful for a small sample size. For comparisons with a larger (typically ten or more) sample size, we computed the normalized log-transformed CPM values and then performed the *t* test between the conditions. Where applicable, we subtracted the mean within each group of samples (e.g., each cell line is a separate group when comparing the treated and untreated samples using multiple cell lines). This was done using a linear model approach.

Gene set enrichment analysis (GSEA) was performed using repeats of various classes as gene sets (Satellite repeats, LINEs, SINEs, and ERVs). We used the preranked GSEA as implemented in fgsea R package (60). We performed two types of analysis ranking the genes and repeat elements by either the fold change or the *t*-statistic from the differential expression analysis.

WGS

DNA and RNA were extracted from samples using the AllPrep DNA/RNA/protein mini kit (Qiagen; #80004). DNA was quantified fluorometrically using the Quant-iT Picogreen dsDNA assay kit (Invitrogen; #P7589) and 100 ng DNA input was used for library generation. The Nextera DNA library prep kit (Illumina; #20020188) was used in combination with the Nextera DNA CD Indexes (Illumina; #20018708) for library preparation. In brief, DNA samples were fragmented using Illumina transposon-based technology. After tagmentation, right-sided size selection was carried out using AMPure-XP beads (Beckman Coulter; #A63881; 1:0.4, followed by 1:1.1) to remove insufficiently tagmented gDNA. Unique 5′ and 3′ index adapters were added to each sample by PCR using five cycles of amplification. Amplified and barcoded libraries were captured during AMPure-XP bead clean-up (1:0.8). Libraries were quantified using the KAPA-library quantification kit for Illumina platform (Roche; #07960140001). Library size was evaluated with a High-Sensitivity DNA kit (Agilent; #5067-4626) on an Agilent 2100 Bioanalyzer. Samples were pooled and sequenced on the NovaSeq 6000 platform.

WGS Analysis

For each patient, both the case and control samples were sequenced with Illumina paired-end reads in length 150 bp. BWA-mem (61) was used to align the reads to the human reference genome hg38. GATK (62) was used to mark duplicates and recalibrate base quality scores. Then, SAMtools (63) was used to sort and index the alignment files.

In addition, NGSCheckMate (64) was used to make sure the case–control samples were well matched.

With the preprocessed case–control bams, we ran the xTea (29) "–case_ctrl" mode on each pair of bams to identify somatic L1, Alu, and SVA insertions. We used the xTea default parameters, where the cutoffs are automatically adjusted based on the average read depth.

Measurement of Circulating ORF1p with Single Molecule Arrays (Simoa)

Affinity reagents for the Simoa assays were obtained from John LaCava (capture nanobody) and Abcam (ab246317 and ab246320). Recombinant ORF1p was obtained from Dr. Martin Taylor and Dr. Kathleen Burns. All buffers, beads, and consumables used for Simoa reagent preparation and assay running were purchased from Quanterix Corporation.

Preparation of antibody-coated paramagnetic beads. 7 × 108 carboxylated 2.7-μm paramagnetic beads were washed three times with 400 μL Bead Wash Buffer, followed by two times with 400 μL Bead Conjugation Buffer, before being resuspended in 390 μL of cold Bead Conjugation Buffer. Ten microliters of freshly dissolved 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC; Thermo Fisher Scientific) in cold Bead Conjugation Buffer (10 mg/mL) was then added to the beads, which were shaken at 4°C for 30 minutes. After activation with EDC, the beads were washed once with 400 μL cold Bead Conjugation Buffer, resuspended in 400 μL of the capture antibody (0.25 mg/mL) or nanobody (0.025 mg/mL), and shaken at 4°C for two hours for conjugation to the nanobody. The conjugated beads were then washed two times with 400 μL Bead Wash Buffer, resuspended in 400 μL Bead Blocking Buffer, and shaken at room temperature for 30 minutes. After blocking, the beads were washed once each with 400 μL Bead Wash Buffer and 400 μL Bead Diluent, before resuspending in 400 μL Bead Diluent for storage at 4°C. The conjugated beads were counted using a Beckman Coulter Z1 Particle Counter.

Preparation of Biotinylated Detector Antibody. Detector antibody (Abcam, ab246317) was biotinylated by adding 80-fold molar excess of sulfo-NHS-LC-LC-biotin (Thermo Fisher Scientific) and incubating for 30 minutes at room temperature. Excess biotin was removed using an Amicon filter (50 kDa, MilliporeSigma), using Biotinylation Reaction Buffer (Quanterix) as the buffer and centrifuging five times at 14,000 × *g* for five minutes. The filter was then inverted and centrifuged at 1,000 × *g* for two minutes to recover the purified antibody. The filter was then washed with 50 μL Biotinylation Reaction Buffer and centrifuged again at 1,000 × *g* for two minutes.

L1ORF1p Simoa Assays. Simoa assays were performed on an HD-X Analyzer (Quanterix Corp.). Beads, detector antibodies, streptavidin-beta-galactosidase (SβG), diluted plasma samples, and necessary consumables were loaded onto the HD-X Analyzer according to the manufacturer's instructions. Conjugated beads were diluted in Bead Diluent, detector antibody was diluted in Homebrew Detector and Sample Diluent to 0.3 μg/mL, and SβG was diluted in SβG Diluent to 150 pmol/L. Plasma samples were diluted 4-fold in Homebrew Detector and Sample Diluent with added protease inhibitor (Halt Protease Inhibitor Cocktail, Thermo Fisher Scientific). For each assay, 100 μL diluted plasma sample was incubated with 250,000 conjugated beads and 250,000 unconjugated helper beads for 15 minutes, followed by washing and incubation with 100 μL biotinylated detector antibody (0.3 μg/mL) for five minutes. The beads were then washed, incubated with 100 μL SβG (150 pmol/L) for five minutes, and washed again before resuspending in 25 μL of the enzyme substrate, resorufin-β-D-galactopyranoside (RGP), and loaded into a 216,000 microwell array, in which each femtoliter-sized microwell can fit at most one bead. The wells were sealed with oil, with each well containing an enzyme-labeled bead subsequently generating a detectable fluorescent signal for counting of fluorescent "on" and "off" wells. The average enzyme per bead (AEB) was calculated by the HD-X Analyzer software, and calibration curves were fit with a four-parameter logistic regression. Limits of detection were determined as three standard deviations above the blank AEB.

GSEA

Coding genes were ranked according to $\log_2$(fold change), and enriched gene sets were obtained using the GSEA of the preranked gene list (65, 66).

TP53 Chromatin Immunoprecipitation

HCT116, HCT8, DLD1, and SW620 cells were cross-linked with formaldehyde (1% final) for 10 minutes at room temperature. Cross-linked cells were quenched with glycine (125 mmol/L final) for 5 minutes, followed by two washes in cold PBS. Nuclei were then isolated from 20 million cells as previously described (67), and chromatin was sheared to 250-bp average size using a Covaris S220. Immunoprecipitations were performed using 1,000 μg of sheared chromatin lysate and 5 μg of p53 antibody (p53 D07; Santa Cruz, sc-47698) preconjugated to protein G beads (Invitrogen). ChIP reactions were incubated for 16 hours at 4°C with rotation and then washed four times in wash buffer [50 mmol/L Hepes-HCl (pH 8), 100 mmol/L NaCl, 1 mmol/L EDTA, 0.5 mmol/L EGTA, 0.1% sodium deoxycholate, and 0.5% N-laurylsarcosine], followed by one wash in ChIP final wash buffer (1 × tris-EDTA (TE) buffer and 50 mmol/L NaCl). Immunoprecipitated DNA was eluted from washed beads, reverse cross-linked overnight, purified, and used to construct libraries. Sequencing for ChIP experiments were prepared using NEBNext Ultra reagents (New England Biolabs). All ChIP samples and input were double-end–sequenced on an Illumina NextSeq 550.

ChIP-seq Computational Analysis

Raw Illumina reads were quality filtered as follows. First 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. Reads that are shorter than 40 bases after trimming were discarded as well. If at least one of the reads in the pair failed the quality check and had to be discarded, then the mate was discarded as well. Quality filtered reads were mapped to the human genome (HG38) using Bowtie2 aligner (68). First the reads are mapped to the genome and then the duplicate reads are removed using Picard MarkDuplicates tool. The mapped BAM files containing the reads were used to call peaks using MACS peak calling algorithm (69). The files were also used to get counts of reads for each repeat region annotated using repeatmasker and BEDTools coverage tool (70). These counts for each repeat region were then further analyzed in R to create plots and derive inference.

Histology/Cytology

Tumorsphere Cell Block Preparation for RNA-ISH/IHC. Tumorspheres were cultured in ultralow attachment cell culture flasks (Corning) in 3-D culture media. For cell block preparation, tumorspheres were transferred to 15 mL Falcon tubes with media and allowed to settle at the bottom of the tube. Media were removed, and tumorspheres were washed 3× in PBS. Tumorspheres were resuspended in human plasma (Sigma-Aldrich) and reconstituted in MilliQ water. Cell pellet/plasma mixture was coagulated with addition of Bovine Thrombin reconstituted in MilliQ water/BSA (Sigma-Aldrich). Coagulated cell pellets were fixed for 3 hours in 10% formalin/PBS. Pellets were processed using standard FFPE. For mice xenograft tumors, tissue was fixed overnight in 10% formalin/PBS, and processed using standard FFPE.

Cytospin Preparation for Immunofluorescence. Tumorspheres were transferred grown in ultralow attachment flasks in 3-D media were allowed to settle in 15-mL falcon tubes. Media were removed and washed 3× in PBS. Tumorspheres were dissociated in trypsin for 5 minutes at 37°C followed by neutralization with media containing 10% FBS. Cells were centrifuged at 150 × *g* and resuspended in PBS. Cells were counted and diluted to a concentration of 500,000/1 mL of PBS. The cell solution (300 μL) was added to a cytology funnel attached to a slide and centrifuged at 350 rpm for 5 minutes in the Cytospin 4 centrifuge (Thermo Fisher).

Human Tissue. For lamivudine clinical trial, human biopsy tissue was fixed in 10% formalin/PBS and processed using standard FFPE. Standard FFPE processed tissue for untreated and chemo/radiation-treated human colorectal cancer tissue was obtained from the MGH Pathology Tissue Bank. All slides were cut at 5-μm thickness.

RNA-ISH and Protein IHC

Automated RNA-ISH assay was performed using the ViewRNA eZ-L Detection 1-Plex Kit from Affymetrix (cat. #19500) on the BondRx 6.0 platform (Leica Biosystems Inc.). Sections (5 μm) 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. The Bond RX user selectable protocol was as follows: The RNA unmasking conditions for the tissue consisted of a 10-minute incubation at 90°C in Bond Epitope Retrieval Solution 1 (Leica Biosystems). This was followed by incubation with Proteinase K with 1:1,000 dilution (Leica Biosystems; cat. #AR9551). Proteinase K incubation on human TMAs and biopsies was done for 10 minutes; and cell blocks for 5 minutes at 40°C. This was followed by HSATII-Type1 (cat. #VA1-10874) probe hybridization with 1:20 dilution for 3 hours at 40°C. After the run, the slides were rinsed with water and allowed to air dry for 1 hour at room temperature. The slides were mounted using Micromount (Leica Biosystems; cat. #3801731) and visualized using a standard brightfield microscope. "Dot like" red color hybridization signals in the tumor cell nuclei and cytoplasm were defined as HSATII-positive signals and quantified using Visiopharm software.

HSATII (ACD 512018) and LINE-1 (ACD 565098) were also detected using RNAscope 2.5 LS Reagent Kit-Brown from Advanced Cell Diagnostics (ACD; cat. #322100) on the BondRx platform. 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.

For IHC, FFPE sections were deparaffinized by baking them for 1 hour at 60°C. IHC staining was done on the BondRx using the Bond Polymer Refine Detection kit (cat. #DS9800). Antigen retrieval was carried out for 10 minutes for cell blocks, 15 minutes for xenografts, and 20 minutes for human tissue using Bond Epitope Retrieval Solution 2 (Leica Biosystems). LINE-1 Orf1p antibody was used at a concentration of 1:3,000 (EMD Millipore, MABC1152).

Immunofluorescence

Cells were washed with PBS, resuspended to 50,000 cells/mL, and fixed in 4% paraformaldehyde for 15 minutes at room temperature. Cell pellets were washed 3 times in PBS and then cytospun using a volume of 300 μL/spot. Cytospun cells were treated with 0.02% saponin for 5 minutes at room temperature. Permeabilization of cells using 1% Triton for 4 minutes at room temperature. Cells were washed 3 × 10 minutes with standard IF Wash Buffer. Samples blocked in 50 μL blocking buffer for 1 hour at room temperature. Blocking buffer was then aspirated, and 50 μL primary Ab was added. Samples were incubated overnight at 4°C. Cells were then washed for 3 × 10 minutes with IF Wash Buffer and 50 μL secondary antibody (Goat Anti-Mouse IgG H&L Alexa Fluor 488 (ab150113) used at 1:1,000) was added and incubated for 1 hour at room temperature in the dark. Samples washed for 3 × 10 minutes with IF Wash Buffer. The samples were mounted with Vectashield and coverslipped with clear nail polish at the edges.

IF Wash Buffer
0.5% NP40
0.3% w/v sodium azide
In 1× PBS made in MilliQ H_2O

IF Block Buffer
10% BSA in IF Wash Buffer

Antibodies
dsDNA (AB27156): mouse mAb; 0.910 mg/mL. 1:1,000 dilution
S9.6 (Kerafast): 1:400 dilution

Image Analysis and Quantification

Cell Migration Assay. The migration assay plates stained with crystal violet were imaged in brightfield with a 4× objective on an inverted tissue culture microscope. Migration was quantified as an area percentage using color deconvolution in the HALO software IHC area quantification module.

Soft-Agar Assay (Brightfield). Soft-agar plates stained with crystal violet were imaged in brightfield with a 1× objective on an Olympus MVX10 macro zoom microscope. The size of each colony and colony density for each treatment condition was quantified using the HALO software IHC module. Colonies were identified based on contrast, intensity, and size of the color-deconvolved blue stain. Colony size and colony number were represented in figures.

IHC/RNA-ISH (Brightfield). Slides were scanned at 40× resolution using Aperio Scanscope slide scanner (Leica Biosystems). The human rectal carcinoma tissue sections and tumorsphere sections stained with HSATII RNA-ISH, LINE-1 RNA-ISH, or LINE-1 IHC were imaged with a Leica Aperio Scanscope CSO (Leica Biosystems). Stain quantification was performed with the digital image analysis software VIS (Visiopharm). Tumor regions of the human tissue images were annotated by a pathologist to mark the regions for analysis. Twenty 100 μm × 100 μm square regions of interest (ROI) were manually annotated throughout each tumorsphere image to mark the regions for analysis. The next step of the analysis was the training of classifiers to extract and separate the hematoxylin, red or brown, and background regions. Representative stain regions were manually selected to constitute the training set. The classification method was chosen to be Bayesian or K-means clustering, depending on visual validation of the resulting classification. The classifiers were trained using multiple image features, including color deconvolution of the individual stains, spatial mean, and median filters to remove noise and enhance punctate signals, and the Visiopharm Polynomial Blobs filter to further enhance punctate signals for ISH quantification. The red and hematoxylin classifier used the following color features: the green band of the RGB color space, red-blue contrast, and red chromaticity. The brown ISH or IHC and hematoxylin classifier used the following color features: the green and blue bands of the RGB color space, the saturation band of the IHS color model, green chromaticity, and a custom-defined hematoxylin and DAB color model. After classification, postprocessing steps were performed to finetune the separation of the hematoxylin, red or brown, and background regions. Postprocessing steps included hole-filling and minimum thresholds for size and signal intensity. The following metrics were outputted by the analysis algorithm for each image: total red or

brown area, total hematoxylin area, red or brown area percentage of the total stain area, and mean signal intensity of red or brown (per image for human tissue and per ROI for tumorspheres). The red or brown area percentage of the total stain area was calculated as follows: 100 × red or brown area/(red or brown area + hematoxylin area).

Immunofluorescence (Fluorescent). Slides stained with dsDNA or RNA:DNA hybrid (S9.6) antibody (Kerafast, ENH001) were imaged in DAPI and AF488 with a Zeiss LSM710 confocal microscope (MGH Cancer Center/Molecular Pathology Confocal Core). dsDNA slides were imaged with a 40× oil immersion objective, and RNA/DNA hybrid slides were imaged with a 63× oil immersion objective. AF488 signal per cell was quantified using digital image analysis, which was performed with the HALO software platform (Indica Labs) Fluorescence Module. Noncellular material and cells with staining artifacts were manually excluded from analysis, leaving about 800 cells per treatment condition to be included for analysis. The first step in analysis was segmentation of nuclei based on contrast, intensity, and size of the DAPI stain. Next, a distance from the nucleus was chosen to define the cytoplasmic region in the absence of neighboring cells. Average AF488 signal per cell was quantified in three compartments: nucleus, cytoplasm, and the entire cell. Cytoplasmic signal was represented in figures.

Immunoblotting

For Western blot, cells were lysed in Laemlli buffer and run on a standard denaturing SDS-PAGE gel (Thermo Fisher) followed by transfer to PVDF membrane (primed in 10% methanol). Blocking was conducted for 1 hour at room temperature with 5% BSA/PBS-T followed by overnight incubation at 4°C with primary antibody for phospho-histone-h2a-x-ser139 (Cell Signaling Technology, 9718S), STING (Cell Signaling Technology, 13647), or GAPDH (Cell Signaling Technology, 2118S), followed by 3× washes with PBS-T and incubation with anti-rabbit secondary antibody diluted 1:1,000 in 5% BSA/PBS-T. Signal was detected using Supersignal WestPico (Thermo Fisher).

Data and Materials Availability

All RNA-seq and WGS data from clinical biopsy specimens have been uploaded to NCBI SRA accession number phs002833.v1.p1. All RNA-seq and ChIP-seq from cell lines and xenografts have been uploaded to NCBI GEO accession number GSE199097. All software for WGS, RNA-seq, ChIP-seq, and digital image data analysis is described in the Methods and all software will be provided upon request.

Authors' Disclosures

A.R. Parikh reports personal fees from C2I Genomics, Guardant, Checkmate, Natera, Eli Lilly, Pfizer, Roche, Inivata, and Bio Fidelity, and other support from Puretech, PMV Pharma, Plexxicon, Takeda, Bristol Myers Squibb, Novartis, Mirati, Genentech, and Daiichi Sankyo outside the submitted work. A. Solovyov reports grants from the NIH during the conduct of the study; personal fees from ROME Therapeutics outside the submitted work; and a patent for targeting repeat RNAs and their use as novel biomarkers pending. M.S. Taylor reports personal fees from ROME Therapeutics outside the submitted work. C. Wu reports a patent for ultrasensitive assays for detection of ORF1p in biofluids pending. R.Y. Ebright reports personal fees from Nextech Invest outside the submitted work. E.E. Van Seventer reports personal fees and other support from Blueprint Medicines outside the submitted work. N. Desai reports other support from ACD during the conduct of the study, as well as other support from ACD outside the submitted work. M. Ligorio reports a patent for treatments pending unrelated to the topic of this article. L. Goyal serves as a consultant or advisory board member for Alentis Therapeutics AG, AstraZeneca, Black Diamond, Basilea, Exelixis, H3 Biomedicine, Incyte Corporation, QED Therapeutics, Servier, Sirtex Medical Ltd., and Taiho Oncology; reports grant/research support (to institution) from Adaptimmune, Bayer, Eisai, Merck, Macrogenics, Genentech, Novartis, Incyte, Eli Lilly, Loxo Oncology, Relay Therapeutics, QED Therapeutics, Taiho Oncology, Leap Therapeutics, Bristol Meyers Squibb, Nucana, and Servier; and is on the Data Safety Monitoring Committee for AstraZeneca. D.P. Ryan reports personal fees and other support from MPM, personal fees from Boehringer Ingelheim, UpToDate, and McGraw Hill, grants from SU2C, and other support from Exact Sciences during the conduct of the study. R.B. Corcoran reports personal fees from AbbVie, Elicio, FOG Pharma, Guardant Health, Navire, Qiagen, Syndax, Taiho, and Zikani Therapeutics, other support from Avidity Biosciences, grants and personal fees from Asana Biosciences and Pfizer, personal fees and other support from C4 Therapeutics, Cogent Biosciences, Kinnate Biopharma, Nested Therapeutics, Interline Therapeutics, nRichDx, Remix Therapeutics, Revolution Medicines, and Theonys, and grants from Novartis and Lilly outside the submitted work. M.N. Rivera reports nonfinancial support from ACD during the conduct of the study, as well as nonfinancial support from Merck Serono outside the submitted work. M.J. Aryee reports personal fees from SeQure Dx outside the submitted work. T.S. Hong reports personal fees from Merck, Novocure, Boston Scientific, Inviata, Merck, GlaxoSmithKline, PanTher Therapeutics, and the Lustgarten Foundation, and grants from Taiho, AstraZeneca, Bristol Myers Squibb, GlaxoSmithKline, IntraOp, and Ipsen outside the submitted work. D.R. Walt reports personal fees from Quanterix Corporation outside the submitted work; multiple patents issued, licensed, and with royalties paid from Quanterix Corporation; and is a founder, equity holder, and director of Quanterix Corporation. D.R. Walt's interests were reviewed and are managed by Brigham and Women's Hospital and Mass General Brigham in accordance with their conflict of interest policies. K.H. Burns reports personal fees and other support from ROME Therapeutics, Transposon Therapeutics, and Oncolinea Pharmaceuticals outside the submitted work, as well as a patent for LINE-1–encoded ORF2p for cancer therapeutics licensed to Axcelius Holding Company, LLC and a patent for detection of ORF1p in biofluids pending. B.D. Greenbaum reports grants from Merck and Bristol Myers Squibb, and personal fees from ROME Therapeutics outside the submitted work, as well as a patent for WO-2019157087-A1 issued. D.T. Ting reports grants from ACD/Biotechne, the NIH/NCI, the National Science Foundation, the Robert L. Fine Cancer Research Foundation, the SU2C–Lustgarten Foundation 2015-002, Affymetrix, SU2C AACR-PS17, and the Burroughs Wellcome Fund during the conduct of the study; personal fees and other support from ROME Therapeutics, personal fees from Tekla Capital, Ikena Oncology, NanoString Technologies, Pfizer, and Foundation Medicine Inc., and other support from PanTher Therapeutics and TellBio, Inc. outside the submitted work; and a patent for biomarkers of cancer issued and licensed to ROME Therapeutics, a patent for targeting human satellite II (HSATII) issued and licensed to ROME Therapeutics, and a patent for repeat RNA as biomarkers of tumor immune response pending and licensed to ROME Therapeutics. No disclosures were reported by the other authors.

Authors' Contributions

M. Rajurkar: Conceptualization, data curation, formal analysis, investigation, visualization, methodology, writing–original draft, writing–review and editing. **A.R. Parikh:** Conceptualization, data curation, formal analysis, supervision, investigation, writing–review and editing. **A. Solovyov:** Software, formal analysis, investigation, visualization, methodology, writing–review and editing. **E. You:** Formal analysis, investigation, visualization, writing–review and editing. **A.S. Kulkarni:** Formal analysis, investigation. **C. Chu:** Formal

analysis, investigation, visualization, methodology. **K.H. Xu:** Data curation, formal analysis, investigation, visualization, methodology. **C. Jaicks:** Resources, data curation, formal analysis, investigation, visualization, methodology, writing–review and editing. **M.S. Taylor:** Resources, formal analysis, investigation, visualization, methodology, writing–review and editing. **C. Wu:** Resources, formal analysis, investigation, visualization. **K.A. Alexander:** Formal analysis, investigation, visualization. **C.R. Good:** Formal analysis, investigation, visualization. **A. Szabolcs:** Data curation, investigation, methodology. **S. Gerstberger:** Data curation. **A.V. Tran:** Software, formal analysis, methodology. **N. Xu:** Investigation. **R.Y. Ebright:** Investigation. **E.E. Van Seventer:** Resources, data curation. **K.D. Vo:** Resources, investigation. **E.C. Tai:** Resources, writing–review and editing. **C. Lu:** Resources, formal analysis, investigation. **J. Joseph-Chazan:** Resources, investigation. **M.J. Raabe:** Resources, formal analysis, investigation, methodology. **L.T. Nieman:** Resources. **N. Desai:** Resources, formal analysis, investigation, methodology. **K.S. Arora:** Resources, formal analysis, investigation. **M. Ligorio:** Formal analysis, investigation. **V. Thapar:** Resources, data curation, supervision. **L. Cohen:** Resources, supervision, funding acquisition, methodology. **P.M. Garden:** Resources, supervision, funding acquisition. **Y. Senussi:** Resources, supervision, funding acquisition, methodology, writing–review and editing. **H. Zheng:** Resources, supervision. **J.N. Allen:** Conceptualization, resources, formal analysis, supervision, funding acquisition, writing–original draft, project administration, writing–review and editing. **L.S. Blaszkowsky:** Conceptualization, resources, data curation, formal analysis, supervision, funding acquisition, investigation, visualization, methodology, writing–original draft, project administration, writing–review and editing. **J.W. Clark:** Resources, formal analysis, investigation, visualization, methodology, writing–review and editing. **L. Goyal:** Resources, data curation, investigation, methodology. **J.Y. Wo:** Resources, data curation. **D.P. Ryan:** Resources, software, formal analysis, methodology. **R.B. Corcoran:** Resources, investigation. **V. Deshpande:** Resources, formal analysis, investigation. **M.N. Rivera:** Resources, data curation, formal analysis, investigation. **M.J. Aryee:** Formal analysis, investigation. **T.S. Hong:** Resources, data curation, formal analysis, supervision, investigation. **S.L. Berger:** Resources, supervision, funding acquisition, investigation. **D.R. Walt:** Resources, formal analysis, supervision, funding acquisition, investigation, methodology. **K.H. Burns:** Resources, supervision, funding acquisition, investigation, methodology, writing–review and editing. **P.J. Park:** Resources, supervision, investigation. **B.D. Greenbaum:** Conceptualization, resources, formal analysis, supervision, funding acquisition, methodology, writing–original draft, project administration, writing–review and editing. **D.T. Ting:** Conceptualization, resources, data curation, formal analysis, supervision, funding acquisition, investigation, visualization, methodology, writing–original draft, project administration, writing–review and editing.

Acknowledgments

We are grateful to Laura Libby for mouse colony care, Emily Silva, and Danielle Bestoso for administrative support. We appreciate productive discussions with Arnold Levine and Junne Kamihara on this work. This study was supported by NIH grants R01CA240924 (D.T. Ting and B.D. Greenbaum), U01CA228963 (B.D. Greenbaum and D.T. Ting), R01GM130680 (K.H. Burns), R01CA240816 (K.H. Burns), T32GM007753 (R.Y. Ebright), 1F30CA232407-01 (R.Y. Ebright); Gateway for Cancer Research G-17-1000 (D.T. Ting and A.R. Parikh); National Science Foundation NSF PHY-1549535 (D.T. Ting and B.D. Greenbaum); Burroughs Wellcome Fund 1010968.01 (D.T. Ting); V Foundation for Cancer Research D2015-034 (M. Rajurkar); SU2C AACR-PS-17 (D.T. Ting and S.L. Berger); SU2C-Lustgarten Foundation 2015-002 (D.T. Ting and B.D. Greenbaum); Affymetrix, Inc. (D.T. Ting, K.S. Arora, N. Desai, M.N. Rivera, and V. Deshpande); ACD-Biotechne (D.T. Ting, A.S. Kulkarni, M.N. Rivera, and V. Deshpande); Robert L. Fine Cancer Research Foundation (D.T. Ting); and The Pershing Square Sohn Prize–Mark Foundation Fellowship (B.D. Greenbaum).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 17, 2021; revised January 27, 2022; accepted March 8, 2022; published first March 23, 2022.

REFERENCES

1. Probst AV, Okamoto I, Casanova M, El Marjou F, Le Baccon P, Almouzni G. A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. Dev Cell 2010;19:625–38.
2. Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, et al. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. Genes Dev 2009;23:1303–12.
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;331:593–6.
4. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell 2015;160:48–61.
5. 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;52:320–30.
6. 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:e91078.
7. 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. Proc Natl Acad Sci U S A 2013;110:E89–98.
8. 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;162:961–73.
9. Guler GD, Tindell CA, Pitti R, Wilson C, Nichols K, KaiWai Cheung T, et al. Repression of stress-induced LINE-1 expression protects cancer cell subpopulations from lethal drug exposure. Cancer Cell 2017;32:221–37.
10. 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;162:974–86.
11. 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;174:549–63.
12. 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;595:309–14.
13. 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. Proc Natl Acad Sci U S A 2015;112:15154–9.
14. Lee E, Iskow R, Yang L, Gokcumen O, Haseley P, Luquette LJ 3rd, et al. Landscape of somatic retrotransposition in human cancers. Science 2012;337:967–71.
15. Rodriguez-Martin B, Alvarez EG, Baez-Ortega A, Zamora J, Supek F, Demeulemeester J, et al. Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. Nat Genet 2020;52:306–19.

16. 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. Proc Natl Acad Sci U S A 2015;112:15148–53.
17. Bakhoum SF, Ngo B, Laughney AM, Cavallo JA, Murphy CJ, Ly P, et al. Chromosomal instability drives metastasis through a cytosolic DNA response. Nature 2018;553:467–72.
18. Kwon J, Bakhoum SF. The cytosolic DNA-sensing cGAS-STING pathway in cancer. Cancer Discov 2020;10:26–39.
19. Dai L, Huang Q, Boeke JD. Effect of reverse transcriptase inhibitors on LINE-1 and Ty1 reverse transcriptase activities and on LINE-1 retrotransposition. BMC Biochem 2011;12:18.
20. Contreras-Galindo RA, Dube D, Fujinaga K, Kaplan MH, Markovitz DM. Susceptibility of human endogenous retrovirus type-K to reverse transcriptase inhibitors. J Virol 2017;91:e01309–17.
21. Wood JG, Jones BC, Jiang N, Chang C, Hosier S, Wickremesinghe P, et al. Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in Drosophila. Proc Natl Acad Sci U S A 2016;113:11277–82.
22. De Cecco M, Ito T, Petrashen AP, Elias AE, Skvir NJ, Criscione SW, et al. L1 drives IFN in senescent cells and promotes age-associated inflammation. Nature 2019;566:73–8.
23. Simon M, Van Meter M, Ablaeva J, Ke Z, Gonzalez RS, Taguchi T, et al. LINE1 derepression in aged wild-type and SIRT6-deficient mice drives inflammation. Cell Metab 2019;29:871–85.
24. Wylie A, Jones AE, D'Brot A, Lu WJ, Kurtz P, Moran JV, et al. p53 genes function to restrain mobile elements. Genes Dev 2016;30:64–77.
25. Jung H, Choi JK, Lee EA. Immune signatures correlate with L1 retrotransposition in gastrointestinal cancers. Genome Res 2018;28:1136–46.
26. 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;23:512–21.
27. Zhu J, Sammons MA, Donahue G, Dou Z, Vedadi M, Getlik M, et al. Gain-of-function p53 mutants co-opt chromatin pathways to drive cancer growth. Nature 2015;525:206–11.
28. Levine AJ, Ting DT, Greenbaum BD. P53 and the defenses against genome instability caused by transposons and repetitive elements. BioEssays 2016;38:508–13.
29. Chu C, Borges-Monroy R, Viswanadham VV, Lee S, Li H, Lee EA, et al. Comprehensive identification of transposable element insertions using multiple sequencing technologies. Nat Commun 2021;12:3836.
30. Cohen L, Cui N, Cai Y, Garden PM, Li X, Weitz DA, et al. Single molecule protein detection with attomolar sensitivity using droplet digital enzyme-linked immunosorbent assay. ACS Nano 2020;14:9491–501.
31. Mankan AK, Schmidt T, Chauhan D, Goldeck M, Höning K, Gaidt M, et al. Cytosolic RNA:DNA hybrids activate the cGAS-STING axis. EMBO J 2014;33:2937–46.
32. Boye K, Jacob H, Frikstad KA, Nesland JM, Maelandsmo GM, Dahl O, et al. Prognostic significance of S100A4 expression in stage II and III colorectal cancer: results from a population-based series and a randomized phase III study on adjuvant chemotherapy. Cancer Med 2016;5:1840–9.
33. Boye K, Nesland JM, Sandstad B, Haugland Haugen M, Mælandsmo GM, Flatmark K. EMMPRIN is associated with S100A4 and predicts patient outcome in colorectal cancer. Br J Cancer 2012;107:667–74.
34. Kang YG, Jung CK, Lee A, Kang WK, Oh ST, Kang CS. Prognostic significance of S100A4 mRNA and protein expression in colorectal cancer. J Surg Oncol 2012;105:119–24.
35. Sack U, Walther W, Scudiero D, Selby M, Aumann J, Lemos C, et al. S100A4-induced cell motility and metastasis is restricted by the Wnt/β-catenin pathway inhibitor calcimycin in colon cancer cells. Mol Biol Cell 2011;22:3344–54.
36. Wang H, Shi J, Luo Y, Liao Q, Niu Y, Zhang F, et al. LIM and SH3 protein 1 induces TGFβ-mediated epithelial-mesenchymal transition in human colorectal cancer by regulating S100A4 expression. Clin Cancer Res 2014;20:5835–47.
37. Zuo Z, Zhang P, Lin F, Shang W, Bi R, Lu F, et al. Interplay between Trx-1 and S100P promotes colorectal cancer cell epithelial-mesenchymal transition by up-regulating S100A4 through AKT activation. J Cell Mol Med 2018;22:2430–41.
38. Watanabe H, Ishibashi K, Mano H, Kitamoto S, Sato N, Hoshiba K, et al. Mutant p53-expressing cells undergo necroptosis via cell competition with the neighboring normal epithelial cells. Cell Rep 2018;23:3721–9.
39. Zhu Q, Hoong N, Aslanian A, Hara T, Benner C, Heinz S, et al. Heterochromatin-encoded satellite RNAs induce breast cancer. Mol Cell 2018;70:842–53.
40. Zhu Q, Pao GM, Huynh AM, Suh H, Tonnu N, Nederlof PM, et al. BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. Nature 2011;477:179–84.
41. Kishikawa T, Otsuka M, Yoshikawa T, Ohno M, Ijichi H, Koike K. Satellite RNAs promote pancreatic oncogenic processes via the dysfunction of YBX1. Nat Commun 2016;7:1–12.
42. Ardeljan D, Steranka JP, Liu C, Li Z, Taylor MS, Payer LM, et al. Cell fitness screens reveal a conflict between LINE-1 retrotransposition and DNA replication. Nat Struct Mol Biol 2020;27:168–78.
43. Mita P, Sun X, Fenyo D, Kahler DJ, Li D, Agmon N, et al. BRCA1 and S phase DNA repair pathways restrict LINE-1 retrotransposition in human cells. Nat Struct Mol Biol 2020;27:179–91.
44. Taylor MS, Altukhov I, Molloy KR, Mita P, Jiang H, Adney EM, et al. Dissection of affinity captured LINE-1 macromolecular complexes. Elife 2018;7:e30094.
45. Stavrou S, Aguilera AN, Blouch K, Ross SR. DDX41 recognizes RNA/DNA retroviral reverse transcripts and is critical for *in vivo* control of murine leukemia virus infection. mBio 2018;9:e00923–18.
46. Parikh AR, Szabolcs A, Allen JN, Clark JW, Wo JY, Raabe M, et al. Radiation therapy enhances immunotherapy response in microsatellite stable colorectal and pancreatic adenocarcinoma in a phase II trial. Nat Cancer 2021;2:1124–35.
47. Smith CC, Beckermann KE, Bortone DS, De Cubas AA, Bixby LM, Lee SJ, et al. Endogenous retroviral signatures predict immunotherapy response in clear cell renal cell carcinoma. J Clin Invest 2018;128:4804–20.
48. 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:e121522.
49. Hughes SM, Levy CN, Calienes FL, Stekler JD, Pandey U, Vojtech L, et al. Treatment with commonly used antiretroviral drugs induces a type I/III interferon signature in the gut in the absence of HIV infection. Cell Rep Med 2020;1:100096.
50. Coghill AE, Engels EA, Schymura MJ, Mahale P, Shiels MS. Risk of breast, prostate, and colorectal cancer diagnoses among HIV-infected individuals in the United States. J Natl Cancer Inst 2018;110:959–66.
51. Ho DD. Time to hit HIV, early and hard. N Engl J Med 1995;333:450–1.
52. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science 1996;271:1582–6.
53. Zhou X, Singh M, Sanz Santos G, Guerlavais V, Carvajal LA, Aivado M, et al. Pharmacological activation of p53 triggers viral mimicry response thereby abolishing tumor immune evasion and promoting anti-tumor immunity. Cancer Discov 2021;candisc.1741.2020.
54. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2012;29:15–21.
55. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res 2013;41:e108.
56. Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, et al. GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res 2019;47:D766–D73.
57. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139–40.

58. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 2010; 11:R25.
59. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014; 15:550.
60. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. bioRxiv 2021:060012.
61. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754–60.
62. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics 2013;43:11.10.1–33.
63. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25:2078–9.
64. Lee S, Lee S, Ouellette S, Park WY, Lee EA, Park PJ. NGSCheckMate: software for validating sample identity in next-generation sequencing studies within and across data types. Nucleic Acids Res 2017; 45:e103.
65. 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. Proc Natl Acad Sci U S A 2005;102:15545–50.
66. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003;34:267–73.
67. Shah PP, Donahue G, Otte GL, Capell BC, Nelson DM, Cao K, et al. Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. Genes Dev 2013;27:1787–99.
68. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–9.
69. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008;9:R137.
70. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841–2.