

Diverse repetitive element RNA expression defines epigenetic and immunologic features of colon cancer

Niyati Desai,^{1,2} Dipti Sajed,² Kshitij S. Arora,^{1,2,3} Alexander Solovyov,⁴ Mihir Rajurkar,¹ Jacob R. Bledsoe,² Srinjoy Sil,¹ Ramzi Amri,³ Eric Tai,¹ Olivia C. MacKenzie,¹ Mari Mino-Kenudson,² Martin J. Aryee,^{1,2} Cristina R. Ferrone,³ David L. Berger,³ Miguel N. Rivera,^{1,2} Benjamin D. Greenbaum,⁴ Vikram Deshpande,² and David T. Ting^{1,5}

¹Massachusetts General Hospital Cancer Center, ²Department of Pathology, and ³Department of Surgery, Harvard Medical School, Charlestown, Massachusetts, USA. ⁴Tisch Cancer Institute, Department of Medicine, Hematology and Medical Oncology, Department of Oncological Sciences, and Department of Pathology, Icahn School of Medicine at Mount Sinai, New York, New York, USA. ⁵Department of Medicine, Harvard Medical School, Charlestown, Massachusetts, USA.

There is tremendous excitement for the potential of epigenetic therapies in cancer, but the ability to predict and monitor response to these drugs remains elusive. This is in part due to the inability to differentiate the direct cytotoxic and the immunomodulatory effects of these drugs. The DNA-hypomethylating agent 5-azacitidine (AZA) has shown these distinct effects in colon cancer and appears to be linked to the derepression of repeat RNAs. LINE and HERV are two of the largest classes of repeats in the genome, and despite many commonalities, we found that there is heterogeneity in behavior among repeat subtypes. Specifically, the LINE-1 and HERV-H subtypes detected by RNA sequencing and RNA in situ hybridization in colon cancers had distinct expression patterns, which suggested that these repeats are correlated to transcriptional programs marking different biological states. We found that low LINE-1 expression correlates with global DNA hypermethylation, wild-type *TP53* status, and responsiveness to AZA. HERV-H repeats were not concordant with LINE-1 expression but were found to be linked with differences in FOXP3⁺ Treg tumor infiltrates. Together, distinct repeat RNA expression patterns define new molecular classifications of colon cancer and provide biomarkers that better distinguish cytotoxic from immunomodulatory effects by epigenetic drugs.

Introduction

Although the general landscape of genetic and epigenetic features of colorectal cancer (CRC) initiation and progression have been well characterized (1, 2), there is inherent heterogeneity of these events that affect the behavior of individual tumor cells. Global DNA hypomethylation has been known to be a common feature across cancers associated with significant differences in cancer cell behavior (3). The long interspersed element, type 1 (LINE-1) is estimated to constitute 16%–20% of the human genome, which has led to the use of LINE-1 methylation status as a surrogate marker of global DNA methylation (4–6). The epigenetic drug 5-azacitidine (AZA) induces global DNA hypomethylation and is approved as a cytotoxic therapy in hematologic malignancies. However, response rates in solid tumors are limited, which may be due to the dose-limiting toxicities of AZA at effective concentrations for solid tumors (7) but could also be due to the lack of biomarkers to identify tumors that would be sensitive to epigenetic agents. Recently, two groups have shown that sustained low doses of AZA induced expression of human endogenous retroviruses (HERV), another major class of repetitive elements, which triggered the innate immune response (8, 9). This increased innate immune response led to suppression of cancer-initiating cell function (8) as well as induced susceptibility of a melanoma mouse model to immune checkpoint blockade (9). These data indicate that repeat expression in cancer can be used as a marker of global methylation status, and the presence of repeat expression may modulate the local immune tumor microenvironment. To define the landscape of repeat RNA expression in CRC, we have used a combination of RNA sequencing (RNA-seq) and RNA in situ hybridization (RNA-ISH) in cell line models and human primary tumors. We have found the high expression of LINE-1 and HERV-H repeats in CRC and that the pattern of their expression was distinct and associated with different histopathological, genetic, and clinical features.

Authorship note: N. Desai, D. Sajed, K.S. Arora, and A. Solovyov contributed equally to this work.

Conflict of interest: This work was partially supported by sponsored research with Affymetrix Inc. (N. Desai, K.S. Arora, M.N. Rivera, V. Deshpande, D.T. Ting).

Submitted: October 17, 2016

Accepted: December 19, 2016

Published: February 9, 2017

Reference information:

JCI Insight. 2017;2(3):e91078. <https://doi.org/10.1172/jci.insight.91078>.

Results

Repeat RNA expression is linked with global DNA methylation status. Given the high abundance of LINE-1 repeat sequences in the genome, we hypothesized that the expression of LINE-1 would be an accurate proxy of global DNA methylation (Figure 1A). To establish a baseline reference for DNA methylation status, HCT116, HCT8, and DLD1 cell lines were subjected to whole-genome bisulfite methyl sequencing. LINE-1 and HERV CpGs were highly correlated with each other ($R = 0.96$, $P < 0.0001$) and both had concordant increases in CpG methylation, which was highest in HCT116 cells, followed by HCT8 and then DLD1 cells (Figure 1B and Supplemental Table 1; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.91078DS1>). Treatment of these cell lines with AZA resulted in consistent and parallel CpG demethylation across all LINE-1 and HERV repeats (Figure 1B) as well as across the genome (Supplemental Figure 1). Given these features, we hypothesized that global LINE-1 RNA expression could be used as a surrogate marker of global DNA methylation. Application of RNA-ISH to a LINE-1-hypermethylated (HCT116) cell line and a LINE-1-hypomethylated (DLD1) cell line demonstrated an inverse relationship of LINE-1 RNA expression with methylation levels as predicted (Figure 1C and Supplemental Figure 2). We then analyzed 354 The Cancer Genome Atlas (TCGA) primary colon cancer tumors with matched CpG methylation and mRNA-seq data and found a concordant inverse correlation of median global CpG methylation with LINE-1 RNA expression (Figure 1D and Supplemental Figure 3; $R = -0.225$; $P = 1.9 \times 10^{-5}$). Altogether, LINE-1 RNA expression is linked with global methylation status and tracks with AZA-mediated demethylation in colon cancers.

HERV repeat subclasses have different coexpression patterns with LINE-1. Global LINE-1 and HERV repeat CpG methylation status was highly concordant in these cell lines, but we hypothesized that there may be meaningful differences in behavior among HERV subtypes. Indeed, analysis of the major HERV subtypes revealed that HERV-K CpG methylation was highly correlated to LINE-1 ($R = 0.97$), while HERV-H was less correlated ($R = 0.86$). To evaluate this relationship at the RNA level, we analyzed mRNA-seq data from 364 TCGA primary colon cancer tumors for HERV and LINE-1 expression. Consistent with cell line methylation data, there was a highly significant positive correlation of HERV-K and LINE-1 expression (Figure 1E; $R = 0.41$; $P = 7.2 \times 10^{-16}$), which was not seen between HERV-H and LINE-1 (Figure 1F; $R = 0.077$; $P = 0.14$). Notably, the HERV-K and HERV-H correlation was not as significant as the HERV-K and LINE-1 correlation, indicating that the expression of these HERV repeats is quite distinct (Supplemental Figure 4). Unsupervised clustering analysis of all HERV and LINE-1 subtypes demonstrated that, although a bulk of these repeats behaved similarly across samples, there were some repeats, like HERV-H, that did not follow the same pattern as the other HERV and LINE repeats across these TCGA colon cancer samples (Supplemental Figure 5). In summary, the expression of global HERV and LINE-1 elements appears to be linked to DNA methylation status, but specific subtypes of HERV elements (HERV-H) do not appear to be coregulated with the other HERV and LINE-1 elements. Together, this would suggest that not all repeats behave the same way and potentially reflect functional differences in cancer cells.

LINE-1 RNA-ISH defines subsets of primary colon cancer. To assess the applicability of the LINE-1 RNA-ISH to clinical specimens, a chromogenic version of the assay was applied to a set of 291 resected primary colorectal tumors. LINE-1 RNA-ISH signal was seen in nearly all samples (252 of 291; 87%), demonstrating the ability to detect these transcripts in standard formalin-fixed, paraffin-embedded material. Histological review revealed LINE-1 signal in tumor cells, normal adjacent tissue, and normal colon (Supplemental Figure 6), which is consistent with the known baseline expression of LINE-1 in normal tissues (10). Colon cancers were found to separate into staining patterns with levels of LINE-1 that were higher than those in normal adjacent tissue (LINE-1 HIGH) and those that were equal to normal adjacent tissue (LINE-1 LOW) (Figure 2A). LINE-1 status was scored in each sample and then used in correlative analysis with other clinical parameters (Table 1). LINE-1 HIGH tumors were inversely correlated with microsatellite instability (MSI⁺; Figure 2B; $P = 0.0001$) and BRAF mutation ($P = 0.041$) but positively correlated to earlier age at diagnosis (mean = 62 vs. 70 years; $P = 0.002$). There was a significant difference in location of primary tumor, with a higher proportion of LINE-1 LOW tumors on the right colon compared with the left colon (LINE-1 LOW: 25/12 vs. HIGH: 81/92 [right/left]; $P = 0.04$). There were no differences in nodal status between the two groups. Analysis of TCGA colon cancer mRNA-seq also revealed an inverse correlation of LINE-1 expression and MSI⁺ status consistent with RNA-ISH results (Supplemental Table 2).

HERV-H RNA-ISH correlates with intratumoral Treg infiltrates. We then evaluated if HERV-H and LINE-1 RNA-ISH would be different as we had seen in our cell line models and TCGA RNA-seq analysis. Indeed,

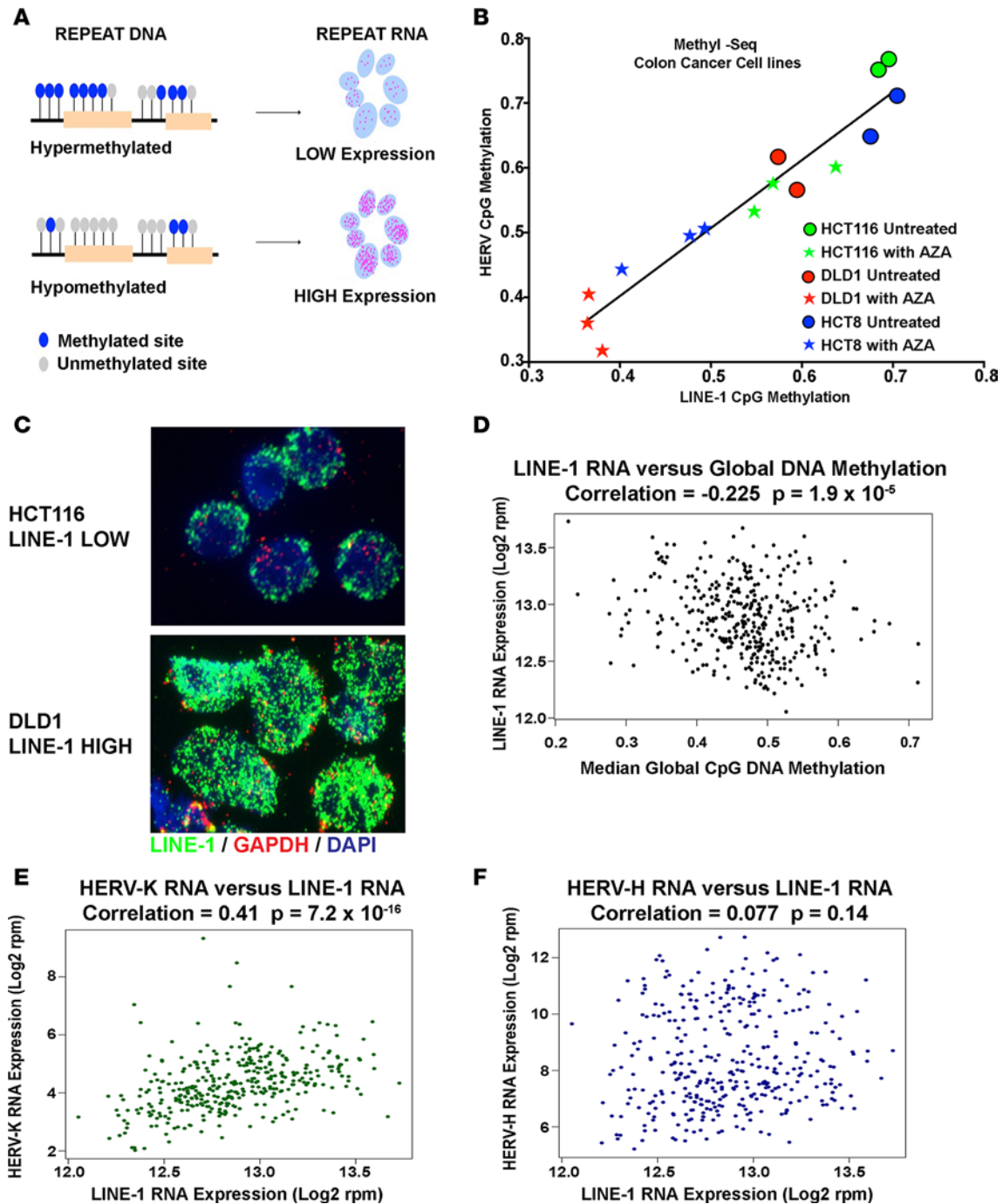


Figure 1. Global repeat RNA expression linked with CpG methylation, but distinct repeat classes had distinct behavior. (A) Schema of repeat DNA methylation linked to RNA expression visualized by RNA-ISH (red dots). (B) Total HERV versus LINE-1 CpG methylation fraction in colon cancer cell lines untreated and treated with AZA. (C) Representative images of LINE-1 LOW (HCT116) and LINE-1 HIGH (DLD1) cell lines stained by LINE-1 (green) and GAPDH (red) RNA-ISH. DAPI nuclear stain (blue). Original magnification, ×400. (D) Inverse relationship of LINE-1 RNA expression versus median global CpG methylation in TCGA colon cancer data set. (E) HERV-K versus LINE-1 RNA expression and (F) HERV-H versus LINE-1 RNA expression in TCGA colon cancer mRNA-seq data (log₂ read per million [rpm]). Statistical analysis was performed with Pearson correlation.

HERV-H and LINE-1 score concordance was only 54%, which suggested that HERV-H might correlate to different tumor biology. Given the linkage of HERV expression to differences in the immune microenvironment, we performed HERV-H RNA-ISH and immunohistochemistry (IHC) for CD8⁺ cytotoxic T cells and FOXP3⁺ Tregs, two of the most relevant T cell subpopulations for tumor immunology. In an initial test cohort, intratumoral FOXP3⁺ Tregs, but not CD8⁺ T cells, correlated with tumor cell HERV-H expression.

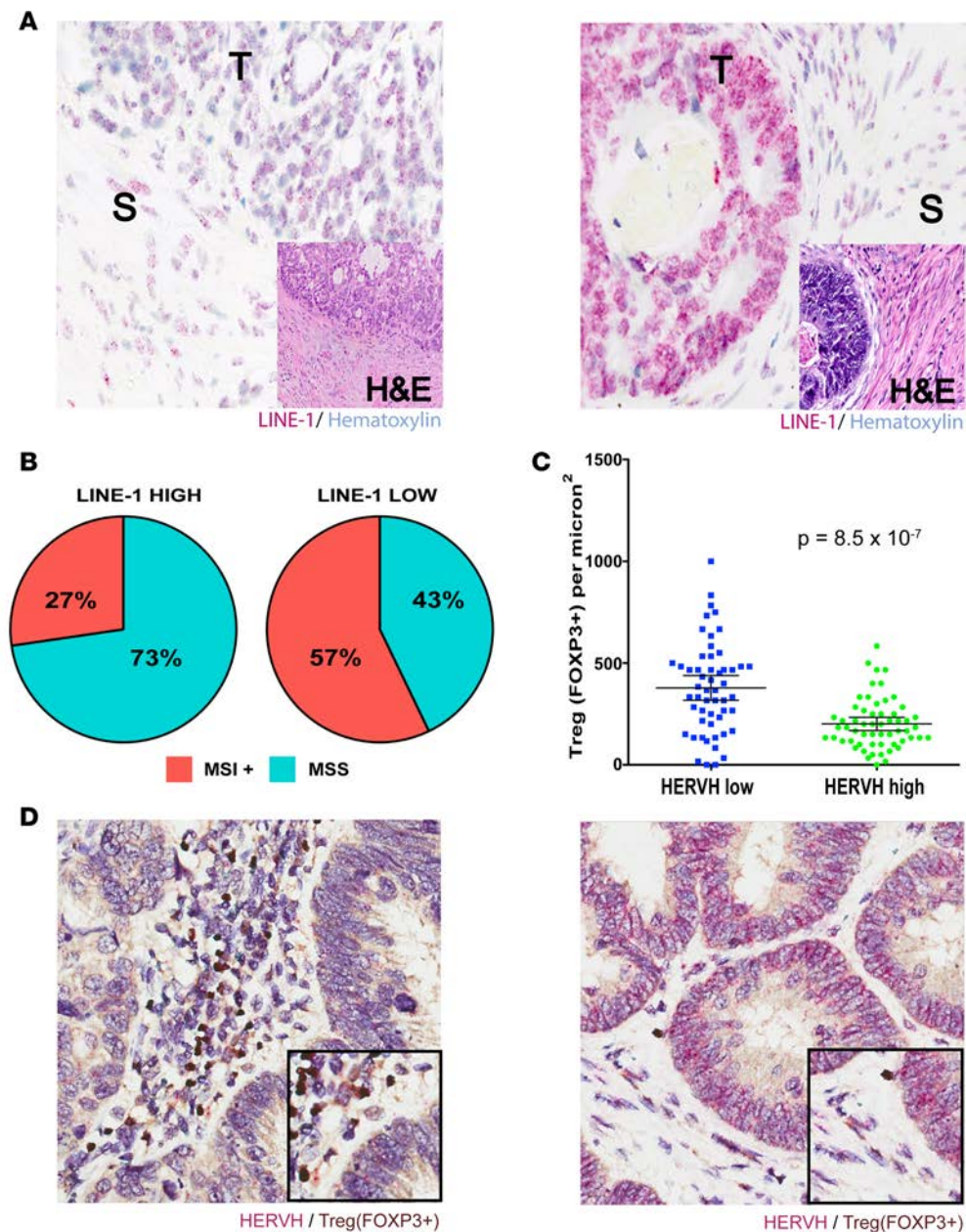


Figure 2. LINE-1 and HERV-H RNA-ISH of primary colon cancer tumors. (A) Image of LINE-1 RNA-ISH LOW (left) and HIGH (right) colon tumors ($n = 291$). LINE-1 RNA-ISH (red dots) scored based on relative tumor (T) expression compared with normal adjacent stroma (S). The inset shows H&E of sequential slide. (B) Breakdown of MSI⁺ and MSS colon cancers in LINE-1 LOW and LINE-1 HIGH populations. (C) Quantitation of intratumoral FOXP3⁺ Tregs in HERV-H LOW versus HIGH tumors ($n = 112$). Mean and error bars = 95% CI; statistical analysis was performed with unpaired 2-tailed t test. (D) Representative image of HERV-H RNA-ISH (red dots) LOW (left) and HIGH (right) colon tumors with costaining for FOXP3⁺ Tregs by IHC (brown cells). Original magnification, $\times 400$.

We then performed a combined HERV-H RNA-ISH with FOXP3 IHC on the same slide on a cohort of 112 colon primary tumors already stained and scored for LINE-1 (Figure 2C). While LINE-1 scores did not correlate with intratumoral FOXP3⁺ Treg density, there was a significantly higher density of intratumoral Tregs in HERV-H LOW compared with HIGH tumors (Figure 2D; mean 378 vs. 201 cells per mm²; $P = 8.5 \times 10^{-7}$). These findings confirm the distinct expression patterns between these repeats and the correlation of these markers to different clinical features. Interestingly, analysis of bulk TCGA colon cancer RNA-seq could not identify a clear relationship between Treg and HERV-H expression mainly attributable to limited reads for Treg-related genes. This is likely due to Tregs comprising a very low percentage of the tumor mass

and the high intratumoral regional variability of Treg infiltrates. This illustrates the significant advantages of RNA-ISH and IHC over bulk sample analyses to resolve the spatial heterogeneity of target RNAs and immune infiltrates in tumors.

LINE-1 RNA expression linked to TP53 status and predicts response to DNA-hypomethylating agents. To better understand the functional role of LINE-1 repeat expression in CRC, we evaluated a panel of 7 colon cancer cell lines for baseline LINE-1 expression using traditional quantitative reverse transcriptase PCR (qRT-PCR) and GAPDH as a reference gene (Figure 3A). Cell lines were classified as LINE-1 LOW (LoVo and HCT116), MEDIUM (HCT8 and HCT15), or HIGH (DLD1, LS123, and SW620). We found that microsatellite stable (MSS) cell lines (SW620 and LS123) had higher LINE-1 RNA levels compared with the MSI⁺ cell lines (LoVo, HCT116, HCT8, HCT15, and DLD1) consistent with our RNA-ISH results in primary tumors. Notably, quantitative LINE-1/GAPDH RNA-ISH by microscopy (Figure 3B) or flow cytometry (Figure 3C) provided improved resolution of LINE-1 LOW (ratio < 2), MEDIUM (ratio 2–3), and HIGH (ratio > 3) cell lines and an assessment of single-cell heterogeneity of LINE-1 expression not possible from bulk RNA analysis. Analysis of other molecular features annotated for these cell lines revealed wild-type *TP53* status as being associated with lower LINE-1 levels (Figure 3A). Suppression of *TP53* using lentiviral shRNA (Supplemental Figure 7) in a *TP53* wild-type cell line (HCT116) led to increased levels of LINE-1 RNA (Figure 3C), pointing to a role of *TP53* in the suppression of human LINE-1 repeats.

Given the linkage of LINE-1 expression with DNA methylation levels, we hypothesized that LINE-1 RNA levels could predict sensitivity to DNA-hypomethylating agents. We tested our CRC cell lines for AZA-mediated cell toxicity. In MSI⁺ CRC cell lines, we found that LINE-1 LOW (LoVo and HCT116) cell lines demonstrated higher sensitivity to AZA compared with LINE-1 MEDIUM (HCT8 and HCT15) and LINE-1 HIGH (DLD1) cell lines (Figure 3D). We found similar findings in the MSS cell lines; SW620 cells had lower LINE-1 expression and conjugated higher sensitivity to AZA compared with LS123 cells (Figure 3E). AZA-mediated cell toxicity was found to be in part due to induction of apoptosis, as confirmed by differential annexin V–FITC/propidium iodide staining in sensitive (HCT116) versus resistant (DLD1) cell lines (Figure 3F and Supplemental Figure 8). Together, this suggests that CRC LINE-1 RNA pretreatment levels can be used to identify tumors that are more susceptible to AZA-mediated cytotoxicity.

Discussion

We have demonstrated the use of repeat RNAs as markers of global DNA methylation and the innate immune response. Previous work using methylation pyrosequencing revealed an inverse correlation of LINE-1 hypomethylation with MSI⁺ and BRAF mutation, which is consistent with our RNA expression analysis (5, 6). Interestingly, LINE-1 extreme hypomethylated tumors were also found to correlate with younger age of onset, as seen in LINE-1 RNA high tumors (11). The concordant relationship of LINE-1 RNA expression in our study and work by others using LINE-1 methylation pyrosequencing demonstrates the ability of LINE-1 RNA-ISH to serve as a surrogate of global DNA methylation status with the simplicity of a stain that can be evaluated by standard microscopy. Importantly, the LINE-1 RNA-ISH can evaluate DNA methylation intratumoral heterogeneity, an epigenetic feature previously not assessable with sequencing-based technologies.

Both LINE-1 and HERV repeats can retrotranspose and result in mutagenesis (12, 13). However, repeat RNAs appear to have direct cellular effects aside from being substrates for retrotransposition. Recently, a variety of repeat RNAs have been shown to activate the innate immune response through the pattern recognition receptor pathways (8, 9, 14, 15). The expression of these repeats appears to be a combination of p53 function and DNA CpG methylation in model organisms (14, 16). Our results have defined these

Table 1. Clinical and molecular correlates with LINE-1 status

	LINE-1 LOW	LINE-1 HIGH	P value
MSI ⁺	24 of 42 (57%)	56 of 205 (27%)	0.0001
BRAF mutation	14 of 33 (42%)	30 of 123 (24%)	0.041
Mean age (yr)	70	62	0.002
Node +/-	20/22	107/99	0.58
Site (right/left)	25/12	81/92	0.04

Node, lymph node metastasis; +/-, node with or without metastasis.

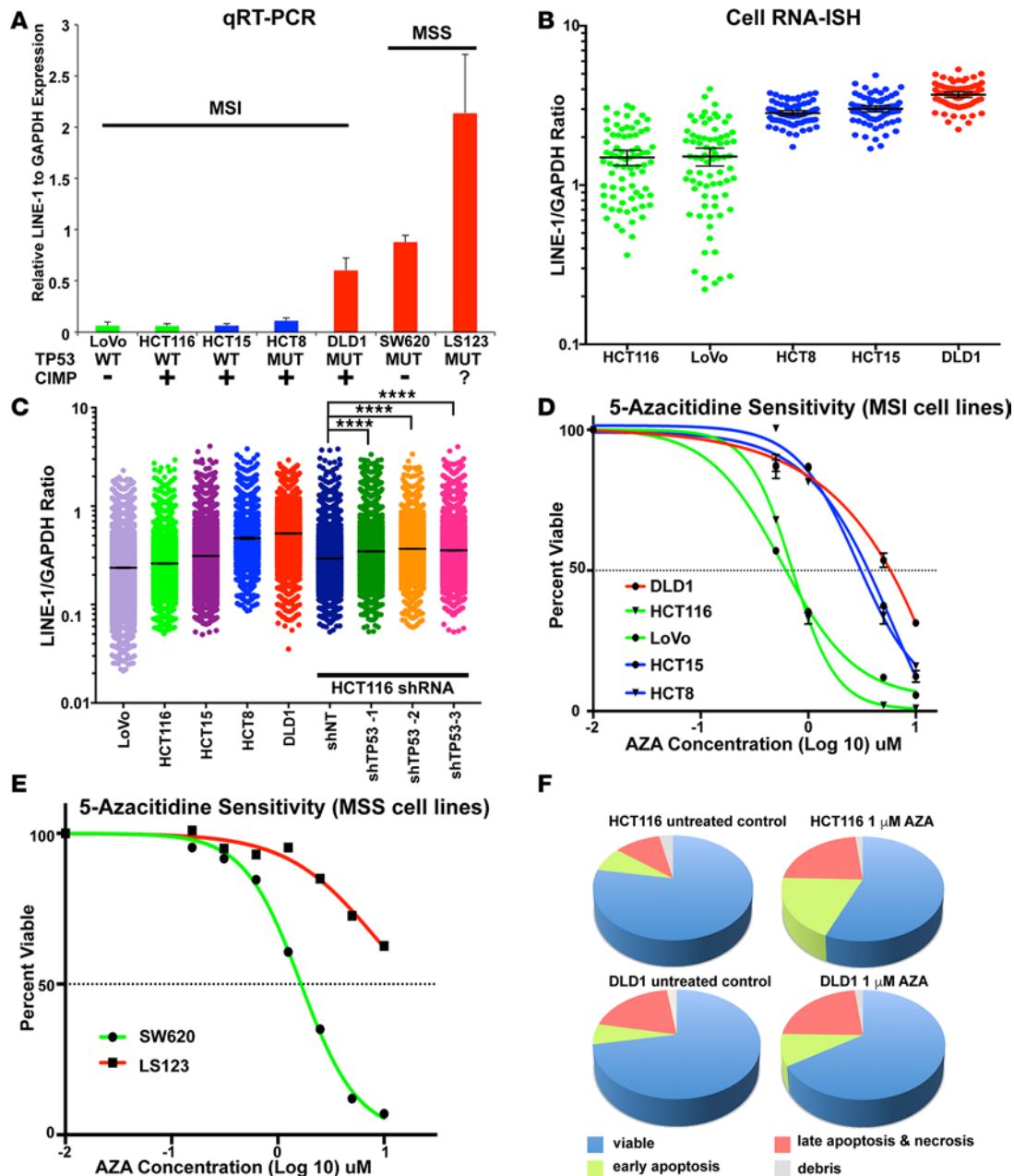


Figure 3. LINE-1 RNA expression linked to TP53 status and predictive of sensitivity to AZA. Relative LINE-1 to GAPDH expression in CRC lines determined by (A) quantitative reverse transcriptase PCR (qRT-PCR) ($n = 3$ technical replicates; bars = mean; error bars = SD), (B) cell RNA-ISH ($n = 75$), LINE-1 LOW (green), MEDIUM (blue), and HIGH (red) RNA expressers, and (C) flow RNA-ISH assays (minimum of 3,000 cells). TP53 wild-type versus mutant (MUT), and CpG island methylator phenotype (CIMP) (+, -, or unknown [?]) status is shown below cell line names. Suppression of TP53 in HCT116 cells with shRNA demonstrates increased LINE-1 levels (bar = mean; error bars = 95% CI; statistical analysis was performed with unpaired 2-tailed t test, **** $P < 0.0001$). (D-F) Colon cancer AZA sensitivity, as determined by CellTiter Glo viability assay in (D) MSI⁺ and (E) MSS CRC lines. Dotted line = IC₅₀. (F) Annexin V and propidium iodide cell apoptosis/necrosis assay of AZA-sensitive (HCT116) and -resistant (DLD1) cell line. All the experiments were repeated at least 3 times. IC₅₀ done in triplicate.

same patterns in CRC cell lines, identifying common mechanisms shared between species to restrain these mobile elements. In addition, we have shown that cell lines with lower basal repeat levels within a given subtype of CRC (MSI⁺ or MSS) are sensitive to AZA. This suggests that there are different repeat RNA set points between CRC subtypes and that increased repeat RNA load leads to increased cytotoxicity for cell lines with intact repeat-sensing mechanisms that trigger the innate immune response. In contrast, cell lines that have higher LINE-1 expression have developed accommodation mechanisms to deal with higher repeat RNA load, which merits further investigation. Last, we have demonstrated that repeat expression is

not generalizable and that certain repeats (HERV-H), but not others (LINE-1), correlate with differences in the composition of immune infiltrates. Our finding of HERV-H anticorrelation with FOXP3⁺ Tregs indicates that HERV-H HIGH tumors would have a more permissive environment for cytolytic activity and are potentially more “primed” for response to immune checkpoint inhibitors.

In summary, different repeat RNA expression patterns define new molecular classifications of colon cancers. The development of combined RNA-ISH and IHC in standard formalin-fixed, paraffin-embedded tissue provides the visual resolution to quantify tumor-specific expression of RNAs as they relate to different immune infiltrates, which is not easily identified in bulk RNA-seq analysis. This platform is readily translatable to the clinical realm and offers a paired diagnostic biomarker for epigenetic and immunomodulatory therapies.

Methods

For additional details, see the Supplemental Methods.

TCGA analysis. The results shown here are in part based upon data generated by the TCGA Research Network (<http://cancergenome.nih.gov/>), including 364 colorectal samples (COAD and READ). Quality filtered reads were mapped to the human genome (GENCODE, hg38 annotation; <https://www.encodegenes.org/>) and to Repbase elements (release 20; <http://www.girinst.org/rebase/index.html>) using the STAR aligner to identify coding and repeat RNAs, respectively.

Statistics. All correlations of RNA-seq or methylation data were calculated by using the Pearson correlation. Significance for contingency tables was performed with the χ^2 test. Ordinal data comparisons between groups were performed using an unpaired 2-tailed *t* test. Values of *P* < 0.05 were considered significant. Statistical analyses were performed using R, SPSS, or PRISM, as detailed in the Supplemental Methods.

Study approval. Human tumor tissues were obtained from the Massachusetts General Hospital under IRB-approved protocols 2012P000039 and 2014P000731. RNA-ISH was performed using the ViewRNA platform (Affymetrix Inc.).

Author contributions

ND, DS, KSA, AS, BDG, VD, and DTT conceived the study and design. ND, DS, KSA, AS, BDG, VD, and DTT developed methodology. ND, DS, KSA, MR, SS, ET, and OCM acquired data. ND, DS, KSA, AS, MR, SS, MJA, BDG, VD, and DTT analyzed and interpreted data. ND, DS, KSA, AS, BDG, VD, and DTT wrote, reviewed, and/or revised the manuscript. JRB, MMK, RA, CRF, DLB, and MNR provided administrative, technical, or material support. BDG, VD, and DTT supervised the study.

Acknowledgments

We are grateful to Alex Forrest-Hay, Manoj Gandhi, Frank Witney, and Emily Silva for their support. This work was supported by the Burroughs Wellcome Fund (to DTT and MNR), NIH grant K12CA087723-11A1 (to DTT), NIH grant P01CA087497-1 (to BDG), Stand Up to Cancer (to DTT and BDG), the Lustgarten Foundation (to DTT and BDG), the National Science Foundation (1545935 to DTT and BDG), the V Foundation (to AS and MR), and Affymetrix Inc. (to ND, KSA, MNR, DTT, and VD).

Address correspondence to: Benjamin Greenbaum, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, 1470 Madison Avenue, Box 1128, New York, New York 10029, USA. Phone: 212.824.8434; E-mail: benjamin.greenbaum@mssm.edu. Or to: Vikram Deshpande, Department of Pathology, Massachusetts General Hospital, 55 Fruit Street Warren 2, Boston, Massachusetts 02114-2696, USA. Phone: 617.726.7741; E-mail: vdeshpande@partners.org. Or to: David T. Ting, Massachusetts General Hospital Cancer Center, Building 149, Thirteenth Street, Room 6-618B, Charlestown, Massachusetts 02129, USA. Phone: 617.240.9402; E-mail: dting1@mg.harvard.org.

1. Vilar E, Gruber SB. Microsatellite instability in colorectal cancer—the stable evidence. *Nat Rev Clin Oncol*. 2010;7(3):153–162.
2. Lao VV, Grady WM. Epigenetics and colorectal cancer. *Nat Rev Gastroenterol Hepatol*. 2011;8(12):686–700.
3. Jones PA, Baylin SB. The epigenomics of cancer. *Cell*. 2007;128(4):683–692.
4. Matsuzaki K, Deng G, Tanaka H, Kakar S, Miura S, Kim YS. The relationship between global methylation level, loss of heterozygosity, and microsatellite instability in sporadic colorectal cancer. *Clin Cancer Res*. 2005;11(24 Pt 1):8564–8569.
5. Inamura K, et al. Tumor LINE-1 methylation level and microsatellite instability in relation to colorectal cancer prognosis. *J Natl Cancer Inst*. 2014;106(9):9.

6. Ogino S, et al. LINE-1 hypomethylation is inversely associated with microsatellite instability and CpG island methylator phenotype in colorectal cancer. *Int J Cancer*. 2008;122(12):2767–2773.
7. Ahuja N, Easwaran H, Baylin SB. Harnessing the potential of epigenetic therapy to target solid tumors. *J Clin Invest*. 2014;124(1):56–63.
8. Roulois D, et al. DNA-demethylating agents target colorectal cancer cells by inducing viral mimicry by endogenous transcripts. *Cell*. 2015;162(5):961–973.
9. Chiappinelli KB, et al. Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. *Cell*. 2015;162(5):974–986.
10. Ting DT, et al. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science*. 2011;331(6017):593–596.
11. Baba Y, et al. Epigenomic diversity of colorectal cancer indicated by LINE-1 methylation in a database of 869 tumors. *Mol Cancer*. 2010;9:125.
12. Lee E, et al. Landscape of somatic retrotransposition in human cancers. *Science*. 2012;337(6097):967–971.
13. Jern P, Coffin JM. Effects of retroviruses on host genome function. *Annu Rev Genet*. 2008;42:709–732.
14. Leonova KI, 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(1):E89–E98.
15. Tanne A, et al. Distinguishing the immunostimulatory properties of noncoding RNAs expressed in cancer cells. *Proc Natl Acad Sci U S A*. 2015;112(49):15154–15159.
16. Wylie A, et al. p53 genes function to restrain mobile elements. *Genes Dev*. 2016;30(1):64–77.