Molecular Therapy
Nucleic Acids
Original Article

Antisense transcription from lentiviral gene targeting linked to an integrated stress response in colorectal cancer cells

Taekyu Ha,[1] Michael DiPrima,[1] Vishal Koparde,[2,3] Parthav Jailwala,[2,3] Hidetaka Ohnuki,[1] Jing-Xin Feng,[1] Murali Palangat,[4] Daniel Larson,[4] and Giovanna Tosato[1]

[1]Laboratory of Cellular Oncology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, 37 Convent Drive, Bethesda, MD 20892, USA; [2]CCR Collaborative Bioinformatics Resource, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA; [3]Advanced Biomedical Computational Sciences, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Inc., Frederick, MD 21701, USA; [4]Laboratory of Receptor Biology and Gene Expression, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

Advances in gene therapy research have resulted in the successful development of new therapies for clinical use. Here, we explored a gene targeting approach to deplete ephrinB2 from colorectal cancer cells using an inducible lentiviral vector. EphrinB2, a transmembrane ephrin ligand, promotes colorectal cancer cell growth and viability and predicts poor patient survival when expressed at high levels in colorectal cancer tissues. We discovered that lentiviral vector integration and expression in the host DNA frequently drive divergent host gene transcription, generating antisense reads coupled with splicing events and generation of chimeric vector/host transcripts. Antisense transcription of host DNA was linked to development of an integrated stress response and cell death. Despite recent successes, off-target effects remain a concern in genetic medicine. Our results provide evidence that divergent gene transcription is a previously unrecognized off-target effect of lentiviral vector integration with built-in properties for regulation of gene expression.

INTRODUCTION

Remarkable advances in gene therapy research have resulted in the successful development of new therapies approved for clinical use.[1,2] Among these, chimeric antigen receptor (CAR) T cells engineered to target B cells expressing CD19 or B cell maturation antigen (BCMA) for the treatment of hematological malignancies rely on the *ex vivo* lentiviral gene transfer of patient T cells and re-infusion into the patient for adoptive immunotherapy.[3] A similar approach was successfully extended to the *ex vivo* lentiviral transduction of hematopoietic stem and progenitor cells (HSPCs) with a modified β-globin gene, enabling therapy for β-thalassemia.[4] Another promising lentiviral-based *ex vivo* application is for sickle cell disease, where the BCL11A repressor of fetal hemoglobin expression is knocked down in HSPCs.[1]

Recently, we and others have found that ephrinB2 and its Eph receptors control colorectal cancer cell growth and survival.[5,6] *EFNB2* (the gene coding for ephrinB2) expression is significantly higher in colorectal cancer than in normal colon, other normal tissues, and many other cancer types, and the probability of survival is significantly greater in colorectal cancer patients with low as opposed to high ephrinB2 in their cancer.[5] Ephrin ligands and Eph tyrosine kinase receptors comprise a family of transmembrane proteins that are pivotal regulators of cell function through cell-to-cell communication.[7,8]

Colorectal cancer is a heterogeneous disease of intestinal stem cells and the second leading cause of cancer-related death in the United States.[9,10] On the heels of recent successes in genetic medicine and the prospect of advances to overcome current challenges of genetic cancer targeting *in vivo*,[1,2] we explored a lentiviral-based gene therapy approach for the reversible depletion of ephrinB2 in colorectal cancer cells. The current study unveils the previously unrecognized occurrence of antisense gene transcription driven by lentiviral integration and expression in the host genome. Since antisense transcripts have built-in regulatory properties with important implications for gene regulation,[11–13] we speculate that the current discovery has general implications for current gene therapy approaches, cell infection by DNA-integrating viruses, and investigation of non-coding DNA.

RESULTS

Conditional *EFNB2* silencing in colorectal cancer cells

We engineered three short hairpin RNAs (shRNAs) for the conditional deletion of *EFNB2* (sh-7, sh-596, and sh-647) to target unique sequences of *EFNB2* mRNA from exon 2, exon 4, and the 3′ UTR. The target sequences were the same we had previously used successfully for the constitutive shRNA silencing of *EFNB2*.[5,14] The new conditional shRNAs were designed as Pol II-driven shRNA^{miRs} in the pTRIPZ

Received 3 February 2022; accepted 12 May 2022;
https://doi.org/10.1016/j.omtn.2022.05.029.

Correspondence: Giovanna Tosato, Laboratory of Cellular Oncology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, 37 Convent Drive, Bethesda, MD 20892, USA.
E-mail: tosatog@mail.nih.gov

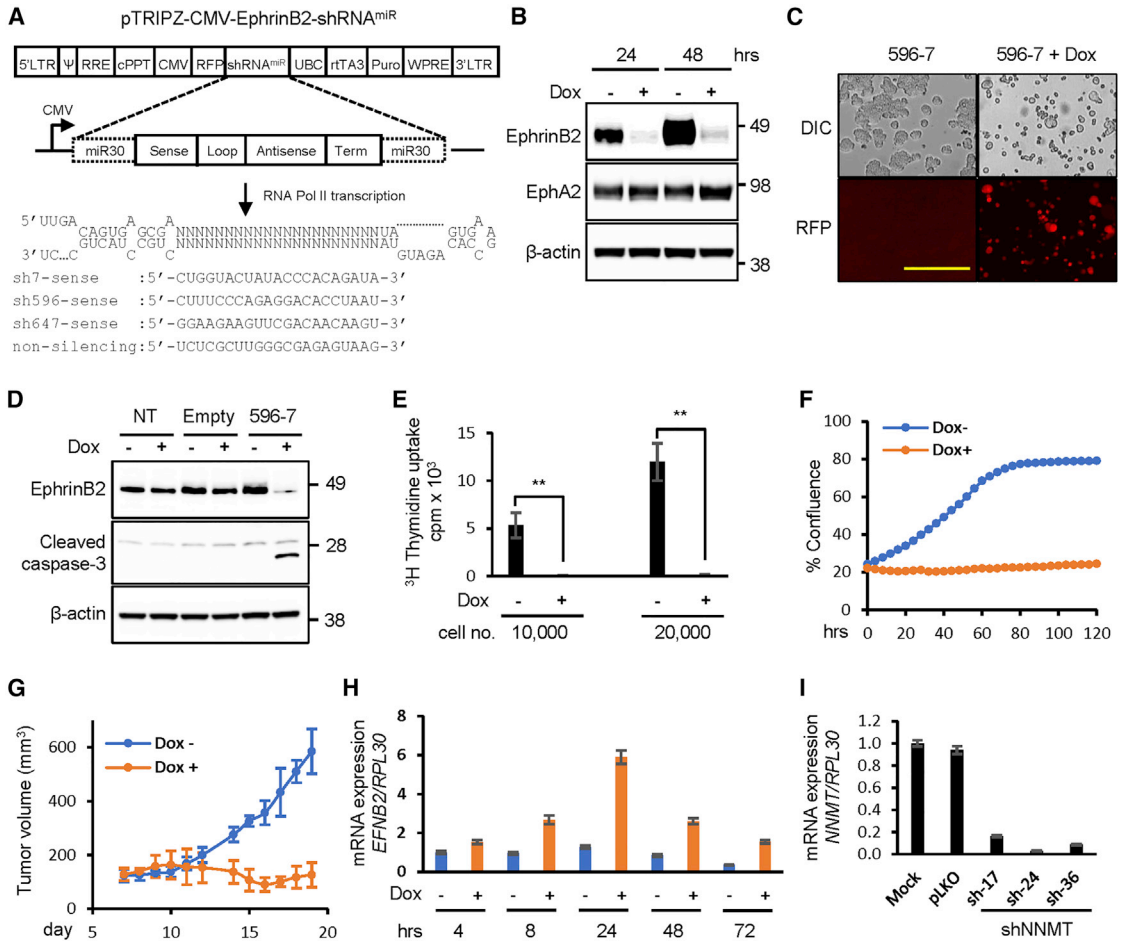


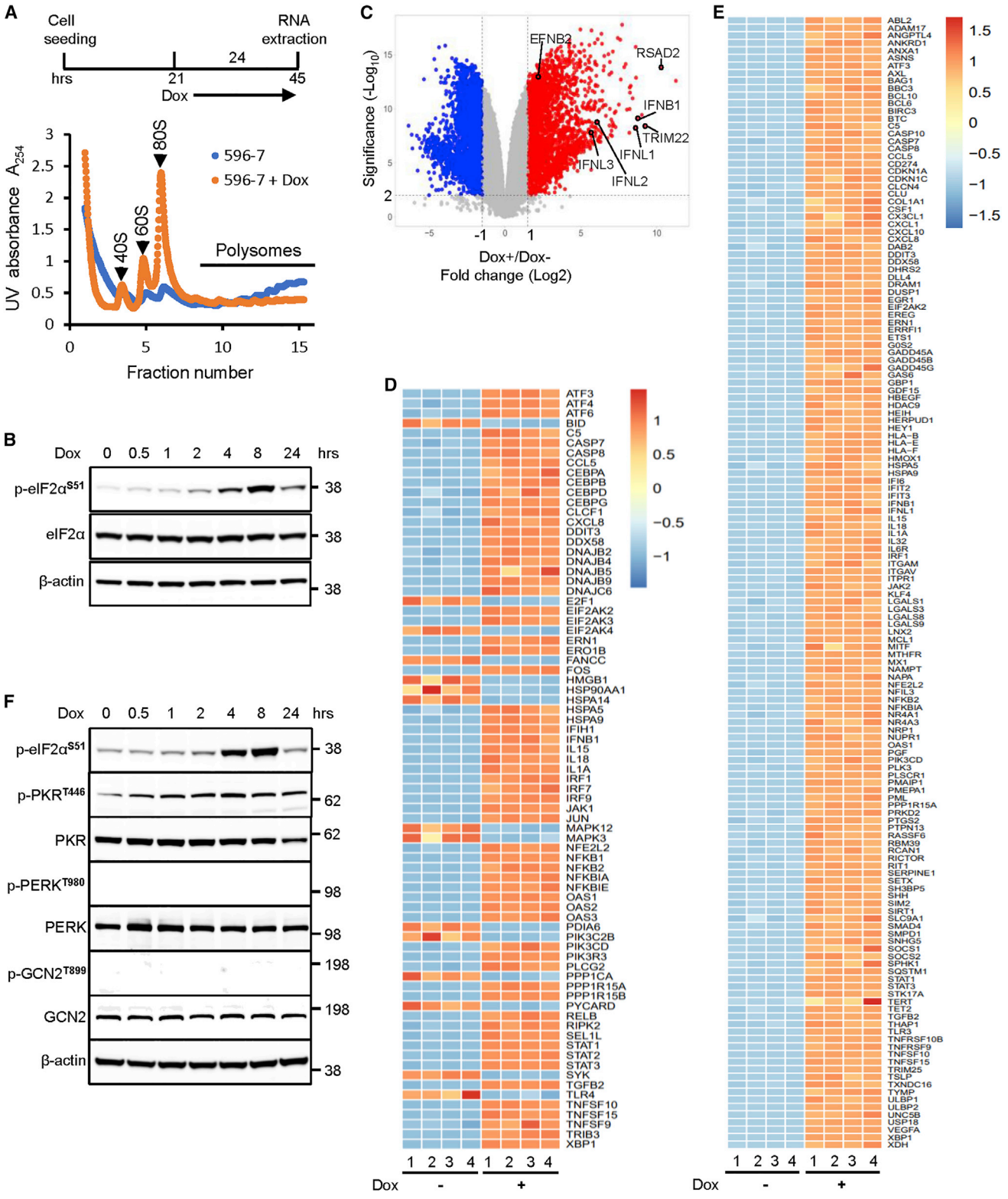
Figure 1. *EFNB2* knockdown in colorectal cancer cells

(A) pTRIPZ-CMV-shRNA^{miR} vector and DNA sequences for *EFNB2* targeting and non-silencing control. 5′ LTR, 5′ long terminal repeat; Ψ, Psi packaging sequence; RRE, Rev response element; cPPT, central polypurine tract; Puro, puromycin resistance gene; 3′ LTR, 3′ self-inactivating long terminal repeat; CMV, tetracycline-inducible minimal CMV promoter; RFP, TurboRFP reporter; shRNA^{miR}, microRNA (miR-30)-adapted shRNA; UBC, human ubiquitin C promoter; rtTA3, reverse tetracycline-transactivator 3; WPRE, woodchuck hepatitis posttranscriptional regulatory element. (B) EphrinB2 and EphA2 protein, representative of six immunoblots. (C) Cell death and RFP (red fluorescent protein) by differential interference contrast (DIC) and fluorescence; scale bar: 200 μm. (D) Cleaved caspase-3 in ephrinB2-depleted 596-7 clone, representative of three immunoblots. NT, non-silencing inducible vector; Empty, empty inducible vector. (E) Proliferation (mean ± SD, triplicate cultures; $^{**}p < 0.01$ by unpaired Student's t test), representative of three experiments. (F) Growth curves; clone 596-7 with or without Dox. (G) Tumor volume in NOD-SCID mice injected subcutaneously with clone 596-7. On day 7, groups of 10 mice were randomized to chow with or without Dox. (H) Relative *EFNB2* and (I) *NNMT* mRNA levels in clone 596-7 (mean ± SD of triplicate measurements; representative of three to five experiments). Clone 596-7 was superinfected with three constitutive *NNMT* shRNAs or control (I). See also Figure S1.

self-inactivating lentiviral vector containing a tetracycline-inducible cytomegalovirus (CMV) promoter, a puromycin resistance gene, and a red fluorescent protein (RFP) marker (Figure 1A). After infection of HT29 colorectal cancer cells, puromycin selection (1 μg/mL, 1 week), and doxycycline (Dox) induction (1 μg/mL, 48–96 h; virtually all cells were RFP$^+$), *EFNB2* mRNA and protein levels were variably reduced by all shRNAs compared with a non-targeting inducible control and the empty vector (Figures S1A and S1B). HT29 cell proliferation was also variably reduced by all these shRNAs (Figure S1C).

To select cells with the highest degree of ephrinB2 depletion, we cloned HT29 cells transduced with each of the three shRNAs; 169 clones were screened. Cell death was clearly observed in 2/74 sh-7 clones, 1/38 sh-596 clones, and 1/57 sh-647 clones after shRNA activation with Dox (24–72 h). Clone 596-7 transduced with sh-596 targeting the 3′ UTR was distinctive in displaying a virtually complete and selective depletion of ephrinB2 protein after Dox (Figure 1B), prominent cell death (Figure 1C), presence of the cell death marker protein cleaved caspase-3 (Figure 1D), markedly reduced proliferation (Figure 1E), and inability to reach confluency (Figure 1F).

Independent experiments reproducibly confirmed the depletion of ephrinB2 protein and the markedly reduced proliferation in clone 596-7 when the shRNA was expressed (Figures S1D and S1E). Also,



(legend on next page)

clone 596-7 did not give rise to tumors in immunodeficient NOD-SCID mice fed with Dox-containing chow (beginning 7 days after cell inoculation), but the same cells grew into tumors in control mice (no Dox-containing chow) (Figure 1G). Thus, clone 596-7 was selected for further analysis of the lethal effects of ephrinB2 depletion in colorectal carcinoma cells.

Surprisingly, Dox activation did not reduce *EFNB2* mRNA levels in clone 596-7, despite full complementarity of the antisense RNA to the *EFNB2* mRNA target sequence (3′ UTR CTTTCCCAGA GGACACCTAAT). Kinetic experiments showed that *EFNB2* mRNA levels increase beginning 4 h after Dox, peak at 24 h, and nearly normalize by 72 h (Figure 1H). In five independent experiments, *EFNB2* mRNA levels increased by 3.2- to 10.1-fold over control after Dox induction, and this increase was detected by primers for *EFNB2* products spanning exons 1/2 and 2 and the 3′ UTR (Figure S1F).

We examined whether the unexpected increase in *EFNB2* mRNA levels was attributable to defective mRNA degradation in clone 596-7. An intrinsic deficiency of the shRNA degradation pathway of Dicer, RNA-induced silencing complex (RISC), Argonaute (AGO), and DROSHA was unlikely to be responsible for the failure of clone 596-7 to degrade *EFNB2* mRNA, since expression in this clone of three shRNAs (#17, #24, and #36) targeting *NNMT* (coding for nicotinamide N-methyltransferase) efficiently reduced the specific target mRNA (Figure 1I), and shRNA #17 (sh-17) reduced NNMT protein levels (Figure S1G).

Mechanisms of ephrinB2 protein depletion and cell death in the presence of *EFNB2* shRNA

Since activation of the shRNA did not reduce *EFNB2* mRNA levels, the ephrinB2 protein reduction induced by activation of the shRNA could not be attributed to reduced transcription. To detect potential translational repression, we performed polysome profiling (Figure 2A). The results showed a reduction of polysome abundance in the Dox-treated 596-7 clone compared with the control, coupled with enrichment of monosome and small and large ribosomal subunits, indicative of reduced mRNA translation. We also examined the activation status of the translation initiation factor 2A (eIF2α), an essential regulator of mRNA translation and protein synthesis. Once phosphorylated at S51, eIF2α triggers a general reduction of mRNA translation. We found that eIF2α becomes phosphorylated at S51 in clone 596-7 after shRNA activation for 2–4 h (Figure 2B). Together, the results of polysome profiling and the time-dependent phosphorylation of eIF2α indicated that RNA translation was compromised in 596-7 cells after the shRNA was activated.

Phosphorylation (S51) of eIF2α is the core initiator of the integrated stress response (ISR), a complex regulatory network that helps cells adapt to stress or promotes cell death when stress surpasses the cell's adaptive responses.[15–17] Typically, the ISR is associated with a broad reduction in mRNA translation and activation of selective genes. These observations raised the possibility that the ISR was induced by Dox activation of the shRNA in clone 596-7, causing cell death. To evaluate the occurrence of the ISR, we compared the transcriptome of clone 596-7 incubated with or without Dox. Replicate samples of clone 596-7 incubated for 24 h with (n = 4) or without (n = 4) Dox were used as a source of RNA samples (n = 8). Principal-component analysis of annotated RNAs (Figure S2A) and volcano plots visualized substantial and significant change in RNA expression after Dox treatment (Figure 2C).

IFNB1 coding for interferon β1 (IFN-β1), *IFNL1* coding for interferon λ1 (IFN-λ1), and the interferon-induced *RSAD2*, *MX2*, and *TRIM22* genes were among the most significantly ($p < 0.001$; average counts >10) induced genes in clone 596-7 treated with Dox compared with untreated clone 596-7 (Figures 2C and S2B). The mRNAs encoding the transcriptional regulators ATF3,[18] CHOP (C/EBP-homologous protein encoded by *DDIT3*[19]), and ATF4,[20] which play critical roles in transcriptional reprogramming of the ISR,[17,21] were also significantly ($p < 0.005$) increased in clone 596-7 treated with Dox compared with the untreated clone 596-7 (Figure 2D). Expression of GADD34 (encoded by *PPP1CA*), the regulatory subunit of the phosphatase PP1, which is induced by p-eIF2α and de-phosphorylates eIF2α, providing feedback regulation of the ISR,[22] was significantly ($p < 0.005$) reduced in the Dox-treated clone 596-7 (Figure 2D).

We applied Ingenuity Pathway Analysis (IPA) to gain further insight into changes in gene expression induced by Dox in clone 596-7 (see Materials and methods for details). Consistent with the ISR, transcripts linked to signaling pathways for cell death and survival (Figure 2E); cell-cycle control/DNA biosynthesis (Figure S2C); interferon signaling (Figure S2D); metabolic pathways for glycolysis, gluconeogenesis, and fatty acid biosynthesis (Figure S2E); and cytokine-induced

Figure 2. Defective mRNA translation and altered gene expression after *EFNB2* silencing

(A) Experimental design. Linear sucrose gradient profile of polysomes isolated from clone 596-7 with or without Dox treatment. The positions of the small (S) and large (L) ribosomal subunits, monosomes, and polysomes are indicated; one of two experiments. (B) Time-dependent activation of phospho (p)-eIF2α (S51) in clone 596-7 treated with Dox; representative immunoblot of three experiments. (C) Magnitude (fold change) and statistical significance of changes in RNA expression induced by Dox in clone 596-7 over 24 h compared with uninduced 596-7 cells. RNA expression by total RNA-seq. Each dot in the volcano plot represents an annotated RNA; significantly induced RNAs are shown as red dots and significantly repressed RNAs are shown as blue dots. Some RNAs are identified as gene products. Vertical and horizontal dotted lines reflect cutoffs for the parameters fold change and significance. (D) Genes expressed at significantly different levels ($p < 0.005$) in control and Dox-treated 596-7 cells (n = 4/group) displayed by row-wide *Z* score (color bar) ordered alphabetically. The gene categories are inclusive of pathways "endoplasmic reticulum stress," "unfolded protein response," "role of PKR in interferon induction," and "pattern recognition receptors" (Ingenuity Pathway Analysis). (E) Differentially expressed ($p < 0.005$) "cell death and survival" genes (Ingenuity Pathway Analysis) in control and Dox-treated 596-7 cells (n = 4/group) displayed by row-wide *Z* score (color bar) ordered alphabetically. (F) Time-dependent PKR phosphorylation (T446) and eIF2α phosphorylation (S51) in 596-7 cells treated with Dox; PERK (T980) and GCN2 (T899) are not phosphorylated. Representative immunoblotting from three experiments. See also Figure S2.

Table 1. Characterization of shRNA integration in host DNA

Chromosome number[a]	Vector integration site on chromosome[b]	Gene annotation of vector integration site[c]	NCBI reference[d]	Gene start[e]	Gene end	Location[f]
14	83,814,307	FLRT2	NM_013231	85,530,144	85,654,428	1,716 kb away
13	95,131,748	ABCC4	NM_005845	95,019,835	95,301,475	intron
6	33,255,417	VPS52	NM_022553	33,250,272	33,272,047	intron
2	27,989,023	BRE	NM_199193	27,889,941	28,338,901	intron
15	73,062,530	NEO1	NM_001172623	73,051,710	73,305,205	intron
1	71,076,570	ZRANB2	NM_203350	71,063,291	71,081,289	intron
7	44,881,135	PURB	NM_033224	44,876,299	44,885,530	exon
7	15,681,655	MEOX2; LOC105375166	NM_005924	15,611,212	15,686,683	intron
5	73,643,317	ARHGEF	NM_001177693	73,626,158	73,941,993	intron
1	189,437,807	FAM5C	NM_199051	190,066,796	190,446,759	660 kb away
2	11,332,375	ROCK2	NM_004850	11,179,759	11,348,330	intron
12	132,023,419	EP400	NM_015409	131,949,942	132,080,460	intron
5	113,567,928	YTHDC2	NM_022828	113,513,694	113,595,285	intron
17	38,208,780	LOC440434	NR_036750	38,195,703	38,257,192	intron
4	68,935,607	UGT2A3	NM_024743	68,928,463	68,951,804	intron

[a]Chromosome in which the shRNA vector has integrated.
[b]Human genome 38 was used to map (bp) vector integration in the genome.
[c]Gene where the vector has integrated; when the integration occurred outside a gene, the closest gene is listed and its distance from the vector site of integration is noted.
[d]National Center for Biotechnology Information (NCBI) reference sequence accession number.
[e]Gene coordinates based on human genome 38 (Human hg38 chrX:15560138-15602945 UCSC Genome Browser v422).
[f]Vector integration relative to gene; when vector is outside gene boundaries the distance from the nearest gene is noted.

inflammation (Figure S2F) were significantly ($p < 0.005$) altered in Dox-treated clone 596-7 compared with the untreated control. Thus, induction of the *EFNB2* shRNA expression in clone 596-7 induces the ISR, broadly altering the transcriptome of clone 596-7.

eIF2α and the ISR are activated through the engagement of the sensor kinases heme-regulated inhibitor (HRI; gene *EIF2AK1*), double-stranded RNA-dependent protein kinase (PKR; gene *EIF2AK2*), PKR-like ER kinase (PERK; gene *EIF2AK3*), and general amino acid control nonderepressible 2 (GCN2; gene *EIF2AK3*), which contain distinct regulatory domains enabling them to respond to different stress stimuli. We found that PKR, activated by double-stranded (ds) RNA and other molecules resembling dsRNA,[23,24] becomes phosphorylated (T446) in clone 596-7 cells (Figure 2F), whereas the sensor kinases GCN2[25] and PERK[26] do not (Figure 2F). These results provide evidence that Dox activation induces the occurrence of PKR activation and ISR in clone 596-7.

Vector-driven antisense transcription in clone 596-7

We examined if vector integration and expression could explain off-target activation of PKR and ISR in clone 596-7 treated with Dox. First, we confirmed by long-range PCR, sequencing, alignment to the vector, and *de novo* assembly that the full-length provirus (5′ to 3′ LTR) and the *EFNB2*-targeting sequences were present in the host DNA (supplemental information). We then mapped vector integration in the host genome. Based on the identification of host/vector DNA junctions,[27] we identified 15 integration sites in 11 chromosomes: 12/15 located in introns, 1/15 located in an exon, and 2/15 mapping at distant sites (>65,000 bp) from the neighboring genes and lacking annotation (Table 1). Each of 10 subclones of the 596-7 clone displayed the same 15 integration sites, confirming that 596-7 cells are a clonal population.

There was no vector integration in the *EFNB2* locus. Several of the host genes where the vector integrated have the potential to globally regulate mRNA transcription and stability, particularly *ZRANB2*,[28] *PURB*,[29] *EP400*,[30] and *YTHDC2*.[31] However, expression of these genes changed by less than 2-fold after vector activation (Table S1). Protein levels of YTHDC2 and PURB changed modestly (Figure S2G). Therefore, a direct contribution of these genes to phenotypic changes in clone 596-7 was not supported by these observations.

Analysis of the host transcriptional landscape revealed the unexpected presence of antisense RNA gene reads (in antisense orientation to the host gene transcript, coming from the opposite DNA strand of host gene sense transcription), which originated at the shRNA vector site of integration in the host genome (Figures 3A, 3B, 4A, and 4B). Antisense reads (green tracks) and sense reads (red tracks) from genes YTHDC2 (Figure 3A), PURB (Figure 3B), VPS52 (Figure 4A), and ABCC4 (Figure 4B) are displayed in the upper section of each figure part. Results from the remaining ARHGEF28, ROCK2, NEO1, and ZRANB2 genes are displayed in Figures S3A–S3D. Thus, antisense reads were present at 8 of the 15 sites of vector integration (Table 1). No antisense reads were observed in the genes *FLRT2*,

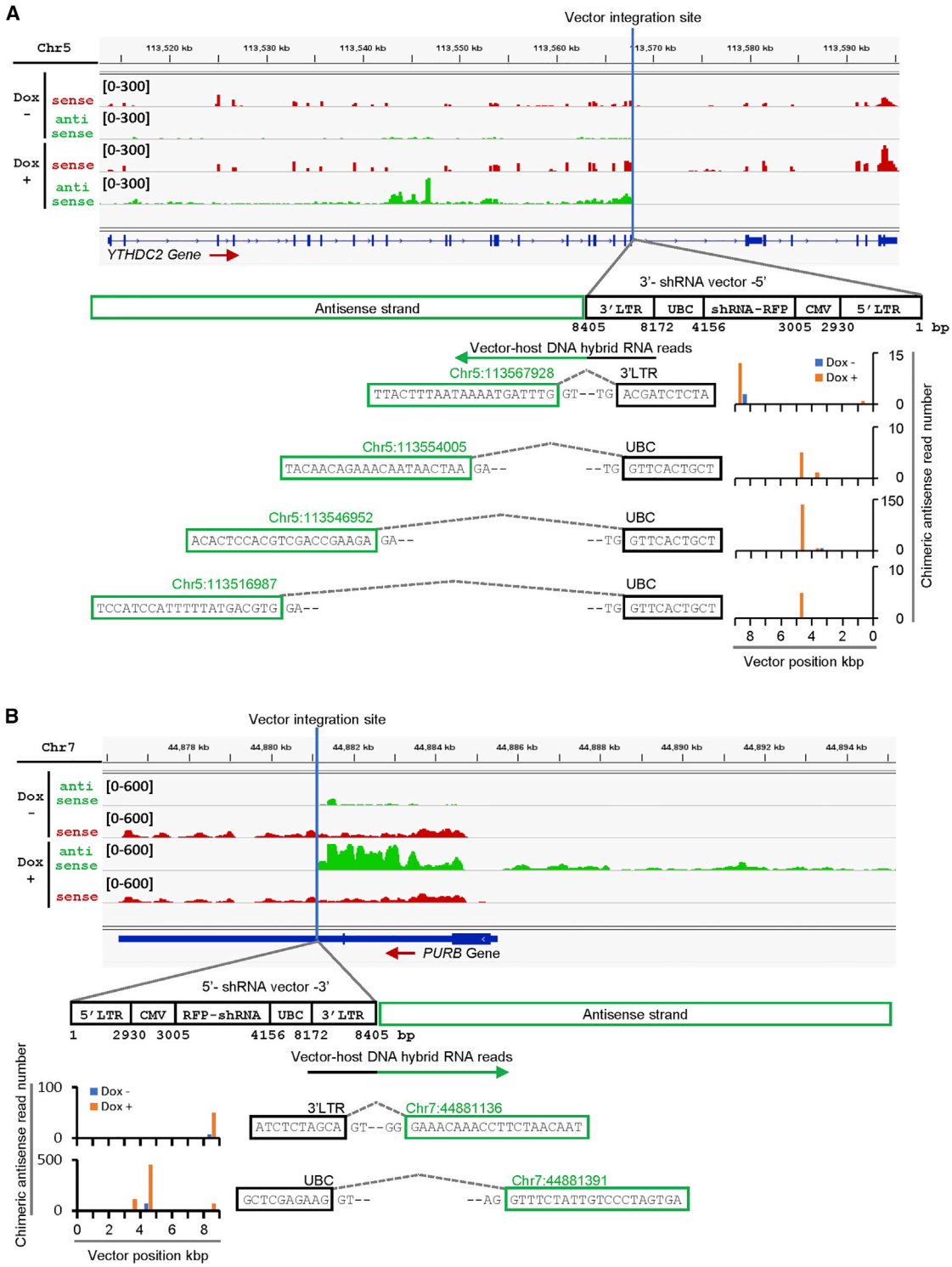


Figure 3. Antisense transcription from the *YTHDC2* and *PURB* genes and chimeric vector/host RNA reads

(A and B) Vector integration site in the host DNA and the junction between host RNA and vector RNA are marked by the blue vertical lines for the *YTHDC2* (A) and *PURB* (B) genes. The red arrow below the schematic gene structure indicates the sense (5′ to 3′) direction of gene transcription. Sense (red) and antisense (green) reads by stranded RNA-seq were derived from clone 596-7 treated or not treated with Dox. Each read is representative of four replicate samples. The read scale is indicated on the read panels.

(legend continued on next page)

FAM5C/BRINP2, LOC440434, *MEOX2*, *BRE/BABAM2*, *EP400*, and *UGT2A3* (Table 1). In all 8 cases in which they were present, the antisense gene reads were more abundant after clone 596-7 was treated with Dox, whereas the sense reads (red RNA-sequencing [RNA-seq] reads) from the same genes were quantitatively similar with or without Dox (Figures 3A, 3B, 4A, 4B, and S3A–S3D).

Since the antisense reads originated in all eight cases from the sites of shRNA integration and increased in number after the shRNA was induced by Dox, we focused on the vector (Figure S3E). Analysis of the data based on mapping to a hybrid reference (inclusive of the shRNA plus the Hg38 genome) revealed the presence of chimeric reads, i.e., reads containing vector-derived and host-derived RNA sequences. All chimeric reads at each of the eight integration sites contained vector reads transcribed in the 5′ to 3′ orientation (5′ LTR to 3′ LTR) and host gene reads in the antisense orientation (Figures 3A, 3B, 4A, 4B, and S3A–S3D). This suggested that the vector in the 5′ to 3′ orientation drives aberrant transcription of the host gene predominantly when activated by Dox. Focusing on the vector integration junctions, we captured vector-host chimeric RNA splicing events from cryptic splice donors within the vector to the host splice acceptor (Figures 3A, 3B, 4A, 4B, and S3A–S3D). In most cases, the chimeric reads at each DNA integration site showed the presence of two to four splice junction sites. Most splicing events followed the standard GT-AG mRNA processing rule for major class introns, whereas others were previously identified atypical splice sites[32] (Figures 3A, 3B, 4A, 4B, and S3A–S3D). Since the 5′ LTR and 3′ LTR vector regions function as transcriptional promoters, and the RFP, UBC, and woodchuck hepatitis posttranscriptional regulatory element (WPRE) sequences are 3′ of the CMV promoter (Figure S3E), these results indicate that the vector drives antisense host gene transcription and splicing, generating chimeric vector/host transcripts.

When all sense and antisense chimeric reads from 596-7 cells with or without Dox were combined, we confirmed that antisense chimeric reads are significantly more abundant than sense chimeric reads and that antisense reads are more abundant in Dox-treated cells compared with untreated cells. In addition, we observed that the antisense chimeric reads contain predominantly the same small set of vector reads mapped to a few vector loci (Figure 4C).

These results provide evidence that vector integration in clone 596-7 drives antisense transcription at most integration sites after the shRNA is induced by Dox.

Vector-induced antisense transcription in other colorectal cancer cell clones

To evaluate whether divergent transcription was unique to clone 596-7, we examined 9 additional clones among 38 that we derived from the parental HT29 cells infected in bulk with the sh-596 vector. The 9 clones (clones 1–4, 6, and 8–11) differed from clone 596-7 at the mRNA and protein levels of ephrinB2 targeted by shRNA-596, in the presence of the death cell marker cleaved caspase-3, and in the expression levels of *IFNB1*, *IFNL1*, and the shRNA (Figures 5A and 5B). Mapping the shRNA integration in the host DNA identified 13 to 30 integration sites in each of the clones (Figure 5C). Few integration sites were shared among the clones (Figure 5D).

Clones 596-1, 596-6, and 596-11, along with independent samples of clone 596-7, were selected for stranded RNA-seq based on their differences from one another and the originally evaluated clone 596-7. We found that all three clones display chimeric antisense reads originating from the shRNA integration sites in the host DNA, predominantly after Dox treatment (Figures 6A–6C: 6A, 596-6; 6B, 596-1; and 6C, 596-11; and Figures S4A, S4B, and S5). Each chimeric read had evidence of cryptic splicing from the shRNA splice donor to the host splice acceptor (Figures 6A–6C, S4A, S4B, and S5).

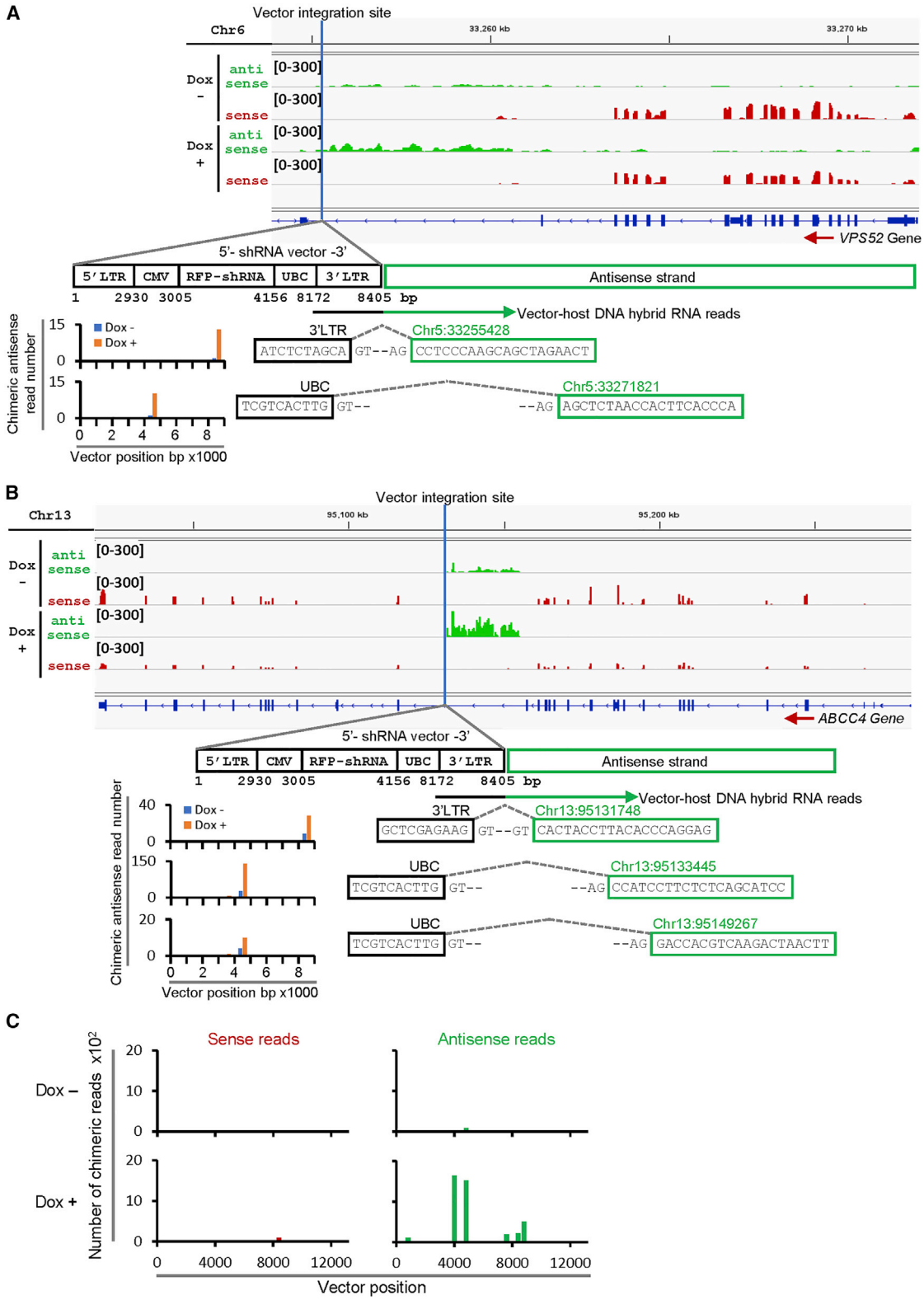
Overall, antisense reads were detected in 7/30 shRNA insertion sites of clone 596-1, in 6/17 insertion sites of clone 596-6, and in 6/18 insertion sites of clone 596-11. Clone 596-11 differed from the other clones, 596-1, 596-6, and 596-7, in showing a similar number of sense and antisense chimeric reads that were more numerous in Dox-treated cells compared with the untreated cells (Figures 7A–7D). In addition, the normalized antisense read number was lower in clones 596-1, 596-6, and 596-11 compared with clone 596-7 (results from repeat RNA-seq with independent sample) (Figure 7E). Overall, these results indicate that antisense transcription is a common event among HT29 cell clones infected with the shRNA.

To gain additional insight into the roles of vector integration, shRNA expression and the presence of antisense reads, we analyzed data from clones 596-1, 596-6, 596-7, and 596-11, which possess variable numbers of integrated vector copies, variable levels of shRNA expression, and variable numbers of antisense reads. Analysis of correlation using the non-parametric Spearman's test showed no direct correlation between vector integrated copy number and relative number of antisense reads (Figure 7F) but a moderately strong correlation between relative shRNA expression level and relative antisense read number ($r = 0.4$; Figure 7G). Overall, these results show that antisense transcription is a common event among HT29 cell clones infected with the shRNA and suggest that the abundance of antisense reads directly correlates with expression levels of the integrated vector, rather than the number of integrated copies of the vector.

DISCUSSION

Here, we report a previously unrecognized outcome of lentiviral DNA integration that holds important implications for lentivirus-based

The bar graphs reflect the quantification of chimeric antisense reads from cells treated with (orange bars) or without (blue bars) Dox. The results reflect the means of four samples. The vector-derived sequences within the chimeric reads are mapped to the vector sequence. A simplified schematic of the annotated shRNA vector is shown above the bar graphs joining to the host antisense DNA strand. The hybrid vector-host reads reveal splicing from the splice vector donor (3′ LTR and UBC) into the host splice acceptor site; the dotted lines indicate the shRNA-host splicing junctions.



(legend on next page)

gene targeting and more broadly for understanding regulation of the human genome. We found that some lentiviral integration events in the host human DNA drive divergent host gene transcription generating antisense reads. This divergent transcription is induced by the activated shRNA readthrough into the host DNA, generating chimeric vector/host transcripts. Since divergent transcripts have built-in properties for regulation of sense gene transcripts,[11–13] the current results uncover a previously unrecognized consequence of lentiviral gene integration in mammalian cells with potentially important functional implications. In the current study, the cell clone with the highest levels of antisense transcription among those investigated was unique in revealing a death phenotype emerging from an ISR, consistent with the idea that vector-induced antisense reads activate the sensor kinase PRK and ISR, perhaps through formation of antisense/sense RNA duplexes. Consistent with this possibility, dsRNA generated by synthesized sense and antisense SNORD113 RNA activated the sensor PKR kinase.[33]

In eukaryotes, gene transcription from each DNA strand initiates at non-overlapping sites, each marked by a separate transcription start site associated with a distinct RNA polymerase binding event. Promoter regions in prokaryotes and eukaryotes are bidirectional,[11,34] but despite this potential for bidirectional transcription, promoters are biased toward transcription of sense reads while suppressing antisense transcription through various mechanisms.[13,35–37] When generated, antisense non-coding RNAs may simply represent noise from open chromatin[13,38] and be rapidly degraded.[13,39] However, several functions have been suggested, including an evolutionary role by promoting *de novo* gene formation[11] and other functions.[40–44]

Integration into the host DNA is a key step in the life cycle of retroviruses and is essential for lentiviral vector function.[45] A drawback to retroviral vectors is that integration can subvert host gene transcription in a variety of ways.[46–49] As originally shown with ALV, a proportion of viral transcripts fail to be correctly polyadenylated at the correct LTR site, resulting in readthrough transcripts of adjacent host cell DNA.[48,50] This process can lead to viral activation of proto-oncogenes and cause cancer.[47,49] By integrating proximal to the LMO2 oncogene, a gene therapy vector derived from murine leukemia virus (MLV) conferred clonal growth advantage to T lymphocytes,[51] resulting in cases of human T cell leukemia.[52]

Antisense transcripts originating from the integrated pro-virus were uncovered in MLV-induced lymphomas[50] and in T cells naturally infected with HTLV,[53] but not in HIV-infected T cells, despite the presence of considerable aberrant host gene transcription.[54] It is possible that antisense transcription in HIV-infected T cells leads to cell death, preventing detection. To our knowledge, antisense transcription has not been previously reported with lentiviral vectors. In part, this may be attributable to the more recent introduction of RNA-seq platforms that track RNA strand derivation; to dilutional effects from using polyclonal, rather than clonal, cell populations; and to the use of an inducible rather than a constitutive system where cells with antisense reads would be potentially eliminated and not detected. However, the high frequency of divergent transcription we detected and diverse functions attributed to antisense non-coding RNAs[40–43] suggest a need for analysis of this off-target effect from integrating vectors.

In conclusion, we identify a previously unappreciated feature of lentiviral vector integration, namely, the occurrence of antisense transcription resulting in the generation of antisense reads, shedding new light on the off-target landscape surrounding lentiviral gene integration.

MATERIALS AND METHODS

Cells and culture

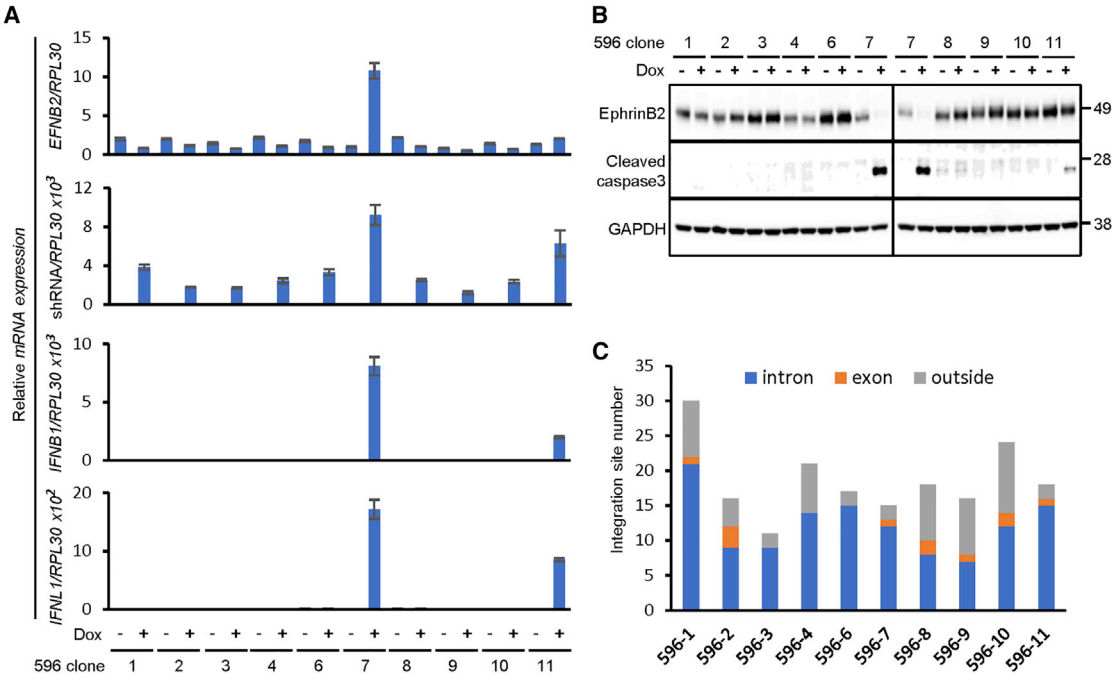
The HT29 human colorectal carcinoma cell line (ATCC; HTB-38) was grown in McCoy's 5A medium (Corning; 10-050-CV), with 10% fetal bovine serum (FBS) (Sigma-Aldrich; F2442) and 1% penicillin-streptomycin (Gibco; 15140-122). A cell clone from HT29 transduced in bulk with *EFNB2*-shRNA-596 lentivirus (clone 596-7) was selected with puromycin (Gibco; A11138-03; 3 μg/mL; 9 days) and maintained in tetracycline-free FBS (Corning; 35-075-CV) and 3 μg/mL puromycin. The monoclonality of clone 596-7 was verified by virus integration site mapping.[27]

pTRIPZ-*EFNB2*-Tet-on shRNA cloning, lentivirus production, and gene silencing

The human *EFNB2* shRNA^{miR} cassette, designed as a miR-30-based shRNA vector,[55] was cloned into the XhoI and EcoRI sites of the pTRIPZ vector (Thermo Scientific Open Biosystems; RHS4750). Lentiviruses were generated by transfecting 293T cells with Lipofectamine 2000 (Thermo Fisher, 11668), the shRNA-encoding plasmid (12 μg), the pMD2.G plasmid (Addgene; 12259; 4.8 μg), and the psPAX2 plasmid (Addgene; 12260; 9.6 μg). Lentivirus-containing supernatant was harvested (72 h), concentrated (ultracentrifugation at 19,500 rpm for 150 min), and stored at −80°C until use. Lentiviral

Figure 4. Antisense transcription from the *VPS52* and *ABCC4* genes and chimeric vector/host RNA reads

(A and B) Vector integration site in the host DNA and the junction between host RNA and vector RNA are marked by the blue vertical lines for the *VPS52* (A) and *ABCC4* (B) genes. The red arrow indicates the sense (5′ to 3′) direction of gene transcription. Sense (red) and antisense (green) reads from the *VPS52* (A) and *ABCC4* (B) genes were derived from clone 596-7 treated or not treated with Dox. Each read is representative of four replicate samples. The bar graphs in (A) and (B) reflect the quantification of chimeric antisense reads from cells treated (orange bars) or not (blue bars) with Dox. The results reflect the means of four samples. The vector-derived sequences within the chimeric reads are mapped to the vector sequence joining to the host antisense DNA strand. The hybrid vector-host reads reveal splicing from the splice vector donor (3′ LTR and UBC) into the host splice acceptor site; the dotted lines indicate the shRNA-host splicing junctions. (C) Quantification of all chimeric sense (host and vector-derived reads are in the sense orientation) and antisense (host-derived reads are in the antisense orientation and vector reads are in the sense orientation) reads coming from all 15 sites of vector integration in the host DNA. The vector component of the hybrid reads is mapped to the vector sequence.



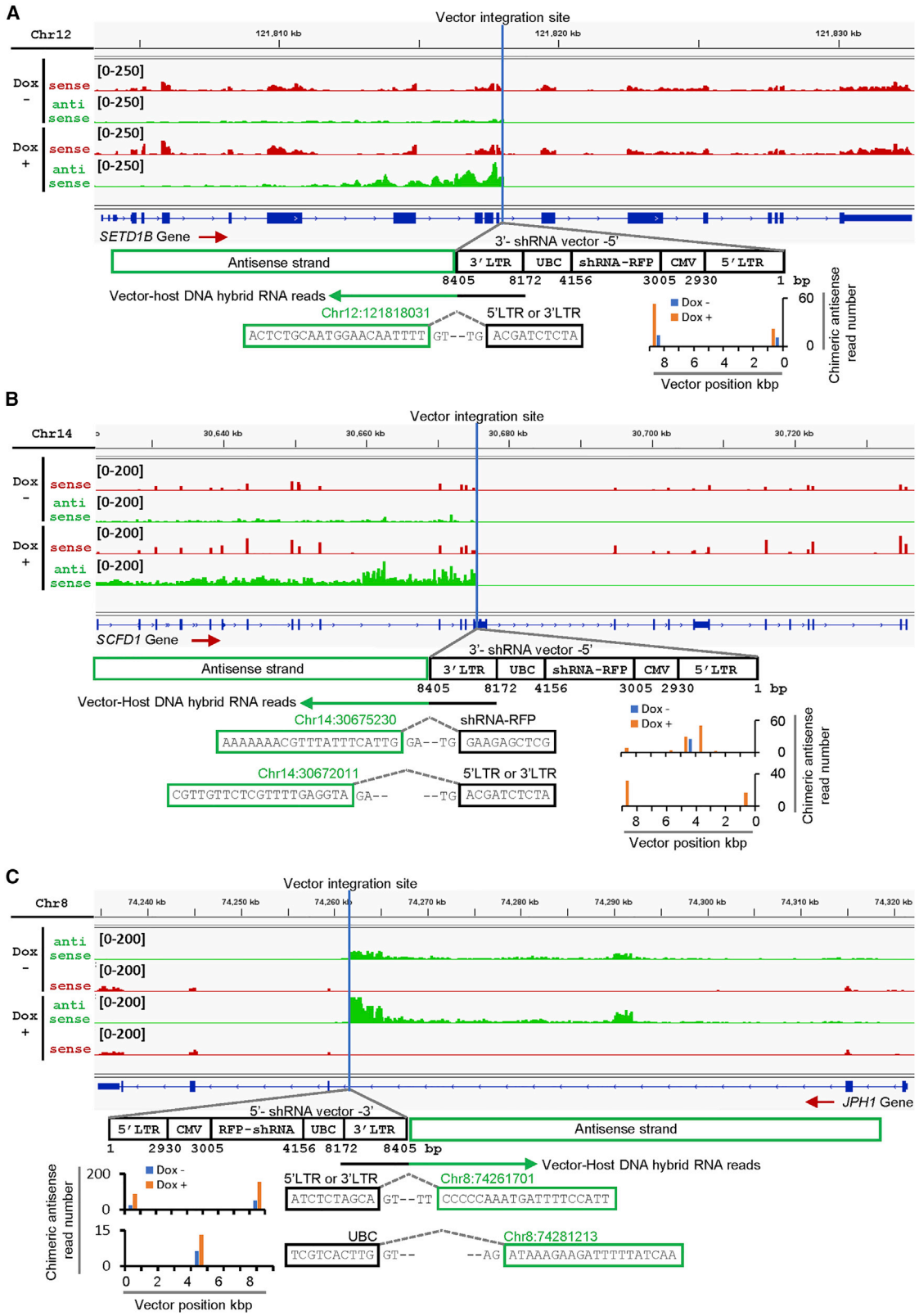
D

596 clone no.	Chromosome no.[a]	Vector Integration site[b]	Gene annotation[c]	NCBI reference[d]	Gene start[e]	Gene end	Location[f]
596-7	1	189437807	FAM5C	NM_199051	190097658	190478404	660kb away
596-5	1	190036918	FAM5C	NM_199051	190097658	190478404	61kb away
596-11	1	161853048	ATF6	NM_007348	161766320	161964070	Intron
596-11	1	161853053	ATF6	NM_007348	161766320	161964070	Intron
596-10	1	161814412	ATF6	NM_007348	161766320	161964070	Intron
596-7	2	11332375	ROCK2	NM_004850	11179759	11348330	Intron
596-6	2	11227158	ROCK2	NM_004850	11179759	11348330	Intron
596-1	3	17450074	TBC1D5	NM_001134381	17157162	18444817	Intron
596-6	3	17402188	TBC1D5	NM_001134381	17157162	18444817	Intron
596-10	8	106911472	ABRA	NM_139166	106759483	106770244	141kb away
596-8	8	106878913	ABRA	NM_139166	106759483	106770244	109kb away
596-4	11	129038432	ARHGAP32	NM_001142685	128965060	129279324	Intron
596-11	11	129163378	ARHGAP32	NM_001142685	128965060	129279324	Intron
596-4	17	82577414	FOXK2	NM_004514	82519713	82644662	intron
596-1	17	82572052	FOXK2	NM_004514	82519713	82644662	intron

[a] Chromosome where the shRNA vector has integrated.
[b] Human Genome 38 is used to map (bp) vector integration in the genome.
[c] Gene where the vector has integrated; when the integration occurred outside a gene, the closest gene is listed and its distance from the vector site of integration is noted.
[d] NCBI (National Center for Biotechnology Information) reference sequence accession number.
[e] Gene coordinates based on human genome 38 (Human hg38 chrX:15560138-15602945 UCSC Genome Browser v422)
[f] Vector integration relative to gene; when vector is outside gene boundaries the distance from the nearest gene is noted.

Figure 5. Characterization of HT29 cell clones infected with ephrinB2 shRNA 596

(A) Relative mRNA levels of *EFNB2*, *shRNA*, *IFNB1*, and *IFNL1* in each of the indicated clones infected with shRNA 596, with or without Dox, for 24 h. Values are normalized to mRNA levels in clone 7 without Dox. The results reflect the means of triplicate measurements. Error bars show the standard deviation. (B) Protein levels of ephrinB2, cleaved caspase-3, and GAPDH in HT29 clones infected with *EFNB2* shRNA 596, with or without treatment with Dox for 24 h. Immunoblotting results are shown. (C) Number and location of vector integration sites in each of the HT29 clones infected with *EFNB2* shRNA 596-7. (D) Shared DNA integration sites among 10 HT29 clones infected with *EFNB2* shRNA 596.



(legend on next page)

particles containing control shRNA (SHC002) or *NNMT* shRNA (TRCN0000315817, -24 and -36) (all from MilliporeSigma) were used to infect clone 596-7.

Cell growth and imaging

Cell proliferation was measured by [^{3}H]thymidine incorporation and Incucyte imaging (Essen BioSciences) as described.[14] Differential interference contrast (DIC) and red fluorescence images were done with an Olympus inverted microscope (Olympus; IX51).

Mouse xenografts

All mouse experiments were conducted in compliance with protocols approved by the National Cancer Institute's Institutional Animal Care and Use Committee (ACUC Bethesda, MD, USA). Groups of female NOD/SCID mice 8 to 11 weeks of age (Charles River Laboratories; strain code 394) were inoculated subcutaneously into the flank with human colorectal carcinoma tumor cells (10×10^6 cells/mouse). Conventional or Dox chow (Bio Serv, S3888) was provided 7 days after cell inoculation when a tumor was visible in all mice. Tumor volume (V) was calculated from caliper measurements as $V = D(d^2)/2$, where D and d are respectively the longest and shortest perpendicular dimensions.

Immunoblotting

After cell washing (1× PBS; Thermo Fisher, 10010-049), cell lysates were prepared in 1× RIPA buffer (Thermo Fisher, 89900) with 1× protease inhibitor cocktail (Thermo Fisher, 78430), 1× phosphatase inhibitor (Thermo Fisher, 1862495), and 5 mM EDTA. They were incubated (1 h at 4°C) and centrifuged (13,200 rpm, 12 min) and the protein concentration was measured (BCA method; Thermo Fisher, 23227). Then the samples were separated by SDS-PAGE (NuPage 4%–12% Bis-Tris gels; Thermo Fisher, NP0321) with 1× MOPS running buffer (Thermo Fisher, NP0001). After transfer to nitrocellulose, the membranes were blocked (1 h, room temperature) in 1× Tris-buffered saline (TBS) buffer (Quality Biological, 351-086-101) with 0.1% Tween 20 (Sigma, P1379) and 5% skim milk (Lab Scientific; M0841). After being washed (30 min in 1× TBS with 0.1% Tween 20, TBST), the membranes were incubated (18 h, 4°C) with primary antibody in 3% BSA in 1× TBST, washed (30 min, 1× TBST), and incubated (1 h, room temperature) with the appropriate horseradish peroxidase-linked secondary antibody in 1× TBST with 5% skim milk. A list of antibodies used is shown in Table S2. After being washed (1× TBST; 1 h, room temperature), the blots were developed using the ECL Prime (GE LifeSciences, RPN2232) reagent, and images were acquired with an LAS4000 (GE LifeSciences) or Amersham Imager 680 (GE LifeSciences) and the bands were quantified (ImageJ).

Real-time RT-PCR and RNA sequencing

Total RNA was extracted using the RNeasy kit (Qiagen; 74106); 1 μg RNA was used to synthesize cDNA with the QuantiTect reverse transcription kit (Qiagen; 205311). RNA integrity (RIN) was evaluated using the Bioanalyzer 2100 (Agilent Technologies). A list of primers is shown in Table S3. mRNA expression was measured (7900HT Fast Real-Time PCR system, Applied Biosystems) using the FastStart Universal SYBR Green Master Mix (ROX) (Roche, 04913914001). Relative mRNA expression was calculated using the ΔΔCt method (ΔCt = Ct[gene of interest] − Ct[housekeeping gene]; ΔΔCt = ΔCt[treated sample] − ΔCt [untreated sample]). Ct (cycle threshold) values were calculated by SDS 2.4.1 software (Applied Biosystems). For RNA-seq, RNA samples had a score higher than 9.9. A stranded total RNA library was prepared according to the Illumina library prep protocol (Illumina; document 1000000040499 v.00). Pair-end stranded RNA-seq was performed by the Sequencing Facility at the Frederick National Laboratory for Cancer Research using Illumina Nextseq (596-7, 596-7 + Dox) and NovaSeq 6000 SP (596-1, -6, -7, and -11). Differentially expressed genes were analyzed for pathway enrichment using IPA (https://digitalinsights.qiagen.com/). Normalized counts per million_trimmed mean of M values (CPM_TMM) counts (average of four samples, average counts >2; fold change [plus versus minus Dox > 3 or < −3] [fold change = Dox+/Dox−, when Dox+ is greater than Dox−; fold change = −(Dox−/Dox+), when Dox+ is no greater than Dox−]; fold change $p < 0.005$) were analyzed by IPA "core analysis." Gene lists from IPA canonical pathways were used for heatmaps.

Polysome profiling

Cultures of clone 596-7 (∼90% confluency) were incubated (10 min) in medium containing 100 μg/mL cycloheximide (BioVision; 1041), washed with PBS, trypsinized, washed with ice-cold PBS containing 100 μg/mL cycloheximide, and lysed in polysome lysis buffer (20 mM Tris [pH 7.2], 130 mM KCl, 15 mM $MgCl_2$, 0.5% [v/v] NP-40, 0.2 mg/mL heparin, protease inhibitors [Thermo Fisher], RNASin [Promega], 2.5 mM DTT, 0.5% deoxycholic acid, and 100 μg/mL cycloheximide). Cell lysates clarified by centrifugation (8,000*g*, 10 min, 4°C) were loaded onto a 10%–50% linear sucrose gradient (10 mM Tris [pH 7.2], 60 mM KCl, 15 mM $MgCl_2$, 1 mM DTT, 0.5% [v/v] NP-40, 0.1 mg/mL heparin) and centrifuged at 200,000*g*, 2 h at 4°C, in an SW41 Ti rotor (Beckman). After centrifugation, the contents were collected as fractions by a fractionator (BioComp Instruments) and optical density of the fractions was measured by UV profiling (254 nm), as described.[56]

Lentiviral vector integration

Lentiviral vector integration mapping was performed at the CCR Genomics Technology Laboratory (Frederick National Laboratory

Figure 6. Antisense reads detected in distinct HT29 cell clones

(A–C) Sense and antisense reads from the *SETD1B* gene in clone 596-6 (A), from the *SCFD1* gene in clone 596-1 (B), and from the *JPH1* gene in clone 596-11 (C). Vector integration site in the host DNA and the chimeric host-vector RNA junction are marked by the blue vertical lines. The direction of sense gene transcription is indicated by the red arrow below the schematic gene structure. The read scale is 0–200 for all *SCFD1* and *JPH1* reads and 0–250 for all *SETD1B* reads. Quantification of chimeric antisense reads is displayed in the bar graphs (orange bars reflect reads from cells treated with Dox and blue bars reflect reads from cells not treated with Dox). The results reflect the means of duplicate samples. The vector-derived sequences within the chimeric reads are mapped to the shRNA sequence. Splicing events from the vector splice donor (3′ LTR, 5′ LTR, or RFP) into the host splice acceptor site are identified; the dotted lines indicate the shRNA-host splice junctions. See also Figure S4.

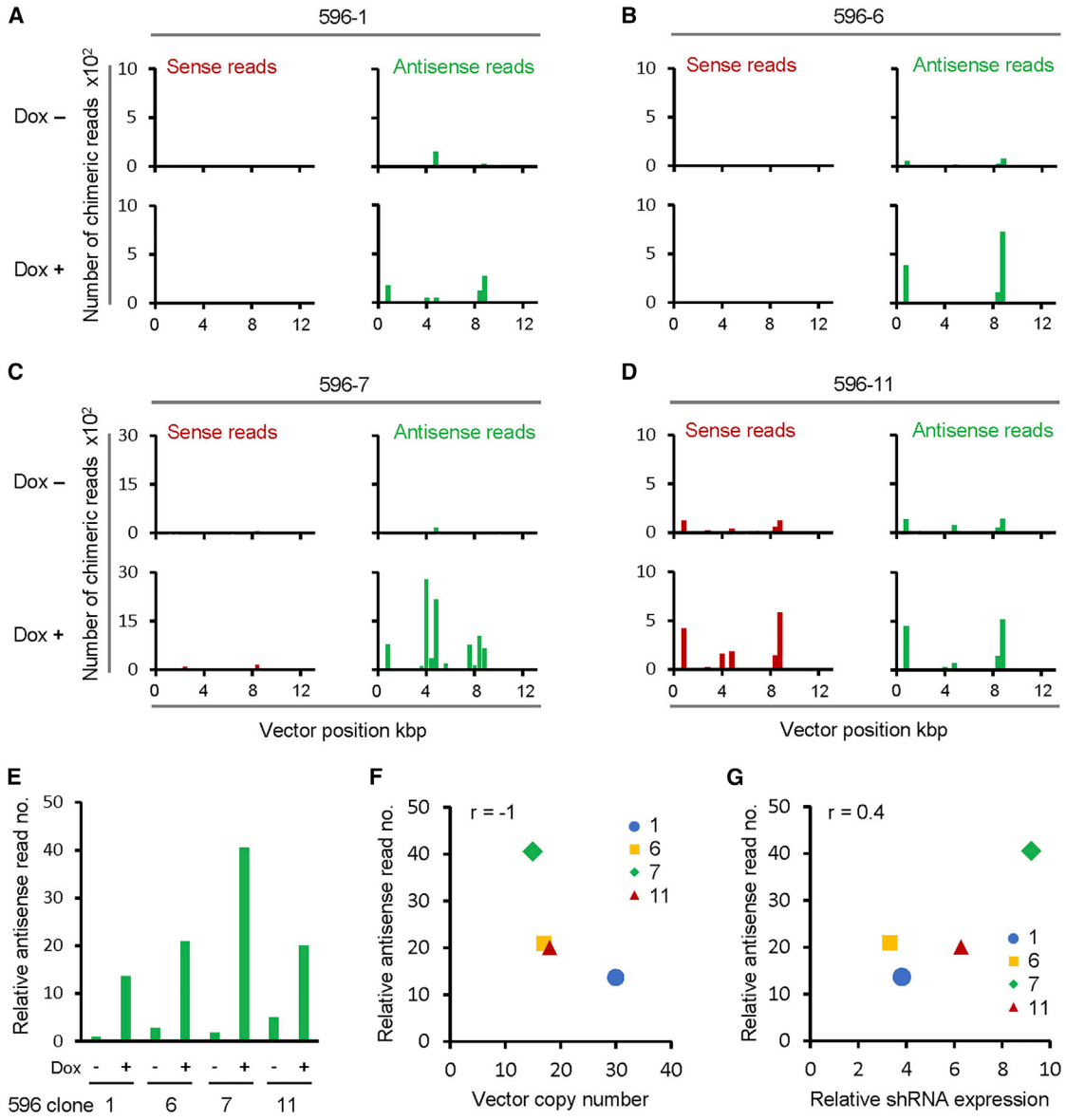


Figure 7. Antisense gene transcription in HT29 cell clones

(A–D) Quantification of all chimeric sense and antisense reads from all sites of vector integration in clones 596-1 (A), 596-6 (B), 596-7 (C), and 596-11 (D) with or without Dox. The vector hybrid reads are mapped to the vector sequence. (E) Quantification of all chimeric antisense reads from all DNA sites of vector integration in each of the clones 596-1, 596-6, 596-7, and 596-11 with or without Dox. The results reflect the means of two separate sequencing results and are displayed as normalized counts/10^6 reads. (F) Correlation between integrated vector copy number and relative antisense read number (antisense reads/10^6 reads) in clones 596-1, 596-6, 596-7, and 596-11. (G) Correlation between relative shRNA expression (shRNA/RLP30 × 10^3) and relative antisense read number (antisense reads/10^6 reads) in clones 596-1, 596-6, 596-7, and 596-11. Correlation between two variables was calculated by Spearman's rank correlation coefficient.

for Cancer Research), as described.[27] Briefly, integration sites were amplified using a linker-mediated PCR protocol. Primers for the 3′ and 5′ LTR sequences of the vector were used to selectively amplify the host/shRNA-vector DNA junctions. The 3′ LTR and 5′ LTR primer sequences are shown in Table S3. Libraries were generated from genomic DNA from shRNA-vector-infected clones 596-1, 596-6, and 596-11 using the NEBNext Ultra II FS DNA Library Prep Kit for Illumina. One microgram of DNA was randomly fragmented and a T-linker unique molecular identifier (UMI) sequence tag was ligated to the ends of the sheared DNA to distinguish independent DNAs with the same integration site. Paired-end sequencing was performed on an Illumina MiSeq and sequencing data were analyzed by a bioinformatics pipeline.[27]

Bioinformatics

Adapter-trimmed sequencing reads were aligned to host + shRNA (hybrid) genome using STAR aligner. RSEM was used to create a gene-level count matrix. Downstream analyses were then performed using custom scripting; details of the bioinformatics analyses are in the supplemental information.

Statistical analysis

Results are presented as means ± standard deviations. Unpaired two-tailed Student's t test was used for analysis of two groups with homoscedastic distribution. p values of less than 0.05 were considered statistically significant. The non-parametric Spearman's rank correlation coefficient was used to measure correlation between two variables; r values were used to estimate the strength of correlation.

DATA AVAILABILITY

The data associated with this study are presented in the paper or in the supplemental information. The sequencing read data generated in this study have been submitted to the GEO repository at NCBI (https://www.ncbi.nlm.nih.gov/geo/) under accession no. GSE186796. All codes used in this paper are available at https://github.com/kopardev/ccbr1060.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtn.2022.05.029.

ACKNOWLEDGMENTS

We are grateful to Drs. C. Buck and G. Starrett for constructive discussions; we thank Drs. J.M. Coffin, E. Freed, M. Hafner, J. Ziegelbauer, Y. Zhao, V. Chen, D. Lowy, and R. Yarchoan and all members of the laboratory and the animal facility personnel for helping with different aspects of this project. This work was supported by the intramural program of NCI/CCR and used the computational resources of the NIH High-Performance Computing Biowulf Cluster.

AUTHOR CONTRIBUTIONS

G.T. conceptualized and supervised the study; T.H. designed and executed most experiments; M.P. performed some experiments; V.K. and P.J. performed bioinformatic analysis; M.D., H.O., X.F., and D.L. provided intellectual input and interpreted data; G.T. wrote the manuscript with T.H. and it was edited by all authors. All authors have seen and approved the final manuscript prior to submission.

DECLARATION OF INTERESTS

All authors declare no competing interests.

REFERENCES

1. Mullard, A. (2020). Gene-editing pipeline takes off. Nat. Rev. Drug Discov. *19*, 367–372. https://doi.org/10.1038/d41573-020-00096-y.
2. Bulaklak, K., and Gersbach, C.A. (2020). The once and future gene therapy. Nat. Commun. *11*, 5820. https://doi.org/10.1038/s41467-020-19505-2.
3. Albinger, N., Hartmann, J., and Ullrich, E. (2021). Current status and perspective of CAR-T and CAR-NK cell therapy trials in Germany. Gene Ther. *28*, 513–527. https://doi.org/10.1038/s41434-021-00246-w.
4. Thompson, A.A., Walters, M.C., Kwiatkowski, J., Rasko, J.E.J., Ribeil, J.A., Hongeng, S., Magrin, E., Schiller, G.J., Payen, E., Semeraro, M., et al. (2018). Gene therapy in patients with Transfusion-dependent beta-Thalassemia. N. Engl. J. Med. *378*, 1479–1493. https://doi.org/10.1056/NEJMoa1705342.
5. DiPrima, M., Wang, D., Troster, A., Maric, D., Terrades-Garcia, N., Ha, T., Kwak, H., Sanchez-Martin, D., Kudlinzki, D., Schwalbe, H., and Tosato, G. (2019). Identification of Eph receptor signaling as a regulator of autophagy and a therapeutic target in colorectal carcinoma. Mol. Oncol. *13*, 2441–2459. https://doi.org/10.1002/1878-0261.12576.
6. Martini, G., Cardone, C., Vitiello, P.P., Belli, V., Napolitano, S., Troiani, T., Ciardiello, D., Della Corte, C.M., Morgillo, F., Matrone, N., et al. (2019). EPHA2 is a predictive Biomarker of resistance and a potential therapeutic target for improving antiepidermal growth factor receptor therapy in colorectal cancer. Mol. Cancer Ther. *18*, 845–855. https://doi.org/10.1158/1535-7163.MCT-18-0539.
7. Barquilla, A., and Pasquale, E.B. (2015). Eph receptors and ephrins: therapeutic opportunities. Annu. Rev. Pharmacol. Toxicol. *55*, 465–487. https://doi.org/10.1146/annurev-pharmtox-011112-140226.
8. Kania, A., and Klein, R. (2016). Mechanisms of ephrin-Eph signalling in development, physiology and disease. Nat. Rev. Mol. Cell Biol. *17*, 240–256. https://doi.org/10.1038/nrm.2015.16.
9. Keum, N., and Giovannucci, E. (2019). Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. Nat. Rev. Gastroenterol. Hepatol. *16*, 713–732. https://doi.org/10.1038/s41575-019-0189-8.
10. Siegel, R.L., Miller, K.D., Goding Sauer, A., Fedewa, S.A., Butterly, L.F., Anderson, J.C., Cercek, A., Smith, R.A., and Jemal, A. (2020). Colorectal cancer statistics, 2020. CA Cancer J. Clin. *70*, 145–164. https://doi.org/10.3322/caac.21601.
11. Wu, X., and Sharp, P.A. (2013). Divergent transcription: a driving force for new gene origination? Cell *155*, 990–996. https://doi.org/10.1016/j.cell.2013.10.048.
12. Churchman, L.S. (2017). Not just noise: genomics and genetics bring long noncoding RNAs into focus. Mol. Cell *65*, 1–2. https://doi.org/10.1016/j.molcel.2016.12.017.
13. Jin, Y., Eser, U., Struhl, K., and Churchman, L.S. (2017). The ground state and evolution of promoter region directionality. Cell *170*, 889–898.e10. https://doi.org/10.1016/j.cell.2017.07.006.
14. Salvucci, O., Ohnuki, H., Maric, D., Hou, X., Li, X., Yoon, S.O., Segarra, M., Eberhart, C.G., Acker-Palmer, A., and Tosato, G. (2015). EphrinB2 controls vessel pruning through STAT1-JNK3 signalling. Nat. Commun. *6*, 6576. https://doi.org/10.1038/ncomms7576.
15. Walter, P., and Ron, D. (2011). The unfolded protein response: from stress pathway to homeostatic regulation. Science *334*, 1081–1086. https://doi.org/10.1126/science.1209038.
16. Pakos-Zebrucka, K., Koryga, I., Mnich, K., Ljujic, M., Samali, A., and Gorman, A.M. (2016). The integrated stress response. EMBO Rep. *17*, 1374–1395. https://doi.org/10.15252/embr.201642195.
17. Costa-Mattioli, M., and Walter, P. (2020). The integrated stress response: from mechanism to disease. Science *368*, eaat5314. https://doi.org/10.1126/science.aat5314.
18. Jiang, H.Y., Wek, S.A., McGrath, B.C., Lu, D., Hai, T., Harding, H.P., Wang, X., Ron, D., Cavener, D.R., and Wek, R.C. (2004). Activating transcription factor 3 is integral to the eukaryotic initiation factor 2 kinase stress response. Mol. Cell. Biol. *24*, 1365–1377. https://doi.org/10.1128/MCB.24.3.1365-1377.2004.
19. Palam, L.R., Baird, T.D., and Wek, R.C. (2011). Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. J. Biol. Chem. *286*, 10939–10949. https://doi.org/10.1074/jbc.M110.216093.
20. Harding, H.P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M., and Ron, D. (2000). Regulated translation initiation controls stress-induced gene expression in mammalian cells. Mol. Cell *6*, 1099–1108. https://doi.org/10.1016/s1097-2765(00)00108-8.
21. Tabas, I., and Ron, D. (2011). Integrating the mechanisms of apoptosis induced by endoplasmic reticulum stress. Nat. Cell Biol. *13*, 184–190. https://doi.org/10.1038/ncb0311-184.
22. Lee, Y.Y., Cevallos, R.C., and Jan, E. (2009). An upstream open reading frame regulates translation of GADD34 during cellular stresses that induce eIF2α phosphorylation. J. Biol. Chem. *284*, 6661–6673. https://doi.org/10.1074/jbc.M806735200.

23. Dey, M., Cao, C., Dar, A.C., Tamura, T., Ozato, K., Sicheri, F., and Dever, T.E. (2005). Mechanistic link between PKR dimerization, autophosphorylation, and eIF2α substrate recognition. Cell *122*, 901–913. https://doi.org/10.1016/j.cell.2005.06.041.

24. Garcia, M.A., Meurs, E.F., and Esteban, M. (2007). The dsRNA protein kinase PKR: virus and cell control. Biochimie *89*, 799–811. https://doi.org/10.1016/j.biochi.2007.03.001.

25. Harding, H.P., Ordonez, A., Allen, F., Parts, L., Inglis, A.J., Williams, R.L., and Ron, D. (2019). The ribosomal P-stalk couples amino acid starvation to GCN2 activation in mammalian cells. Elife *8*. https://doi.org/10.7554/eLife.50149.

26. Kopp, M.C., Larburu, N., Durairaj, V., Adams, C.J., and Ali, M.M.U. (2019). UPR proteins IRE1 and PERK switch BiP from chaperone to ER stress sensor. Nat. Struct. Mol. Biol. *26*, 1053–1062. https://doi.org/10.1038/s41594-019-0324-9.

27. Wells, D.W., Guo, S., Shao, W., Bale, M.J., Coffin, J.M., Hughes, S.H., and Wu, X. (2020). An analytical pipeline for identifying and mapping the integration sites of HIV and other retroviruses. BMC Genomics *21*, 216. https://doi.org/10.1186/s12864-020-6647-4.

28. Plambeck, C.A., Kwan, A.H., Adams, D.J., Westman, B.J., van der Weyden, L., Medcalf, R.L., Morris, B.J., and Mackay, J.P. (2003). The structure of the zinc finger domain from human splicing factor ZNF265 fold. J. Biol. Chem. *278*, 22805–22811. https://doi.org/10.1074/jbc.M301896200.

29. Kelm, R.J., Jr., Elder, P.K., Strauch, A.R., and Getz, M.J. (1997). Sequence of cDNAs encoding components of vascular actin single-stranded DNA-binding factor 2 establish identity to purα and purβ. J. Biol. Chem. *272*, 26727–26733. https://doi.org/10.1074/jbc.272.42.26727.

30. Smith, J.A., White, E.A., Sowa, M.E., Powell, M.L.C., Ottinger, M., Harper, J.W., and Howley, P.M. (2010). Genome-wide siRNA screen identifies SMCX, EP400, and Brd4 as E2-dependent regulators of human papillomavirus oncogene expression. Proc. Natl. Acad. Sci. U S A *107*, 3752–3757. https://doi.org/10.1073/pnas.0914818107.

31. Tanabe, A., Konno, J., Tanikawa, K., and Sahara, H. (2014). Transcriptional machinery of TNF-alpha-inducible YTH domain containing 2 (YTHDC2) gene. Gene *535*, 24–32. https://doi.org/10.1016/j.gene.2013.11.005.

32. Sibley, C.R., Blazquez, L., and Ule, J. (2016). Lessons from non-canonical splicing. Nat. Rev. Genet. *17*, 407–421. https://doi.org/10.1038/nrg.2016.46.

33. Safran, S.A., Eckert, D.M., Leslie, E.A., and Bass, B.L. (2019). PKR activation by non-canonical ligands: a 5′-triphosphate requirement versus antisense contamination. RNA *25*, 1192–1201. https://doi.org/10.1261/rna.071910.119.

34. Seila, A.C., Core, L.J., Lis, J.T., and Sharp, P.A. (2009). Divergent transcription: a new feature of active promoters. Cell Cycle *8*, 2557–2564. https://doi.org/10.4161/cc.8.16.9305.

35. Mayer, A., di Iulio, J., Maleri, S., Eser, U., Vierstra, J., Reynolds, A., Sandstrom, R., Stamatoyannopoulos, J.A., and Churchman, L.S. (2015). Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. Cell *161*, 541–554. https://doi.org/10.1016/j.cell.2015.03.010.

36. Singh, S.S., Singh, N., Bonocora, R.P., Fitzgerald, D.M., Wade, J.T., and Grainger, D.C. (2014). Widespread suppression of intragenic transcription initiation by H-NS. Genes Dev. *28*, 214–219. https://doi.org/10.1101/gad.234336.113.

37. Marquardt, S., Escalante-Chong, R., Pho, N., Wang, J., Churchman, L.S., Springer, M., and Buratowski, S. (2014). A chromatin-based mechanism for limiting divergent noncoding transcription. Cell *157*, 1712–1723. https://doi.org/10.1016/j.cell.2014.04.036.

38. de Boer, C.G., van Bakel, H., Tsui, K., Li, J., Morris, Q.D., Nislow, C., Greenblatt, J.F., and Hughes, T.R. (2014). A unified model for yeast transcript definition. Genome Res. *24*, 154–166. https://doi.org/10.1101/gr.164327.113.

39. Malabat, C., Feuerbach, F., Ma, L., Saveanu, C., and Jacquier, A. (2015). Quality control of transcription start site selection by nonsense-mediated-mRNA decay. Elife *4*, e06722. https://doi.org/10.7554/eLife.06722.

40. Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. Science *322*, 1849–1851. https://doi.org/10.1126/science.1162253.

41. Faghihi, M.A., and Wahlestedt, C. (2009). Regulatory roles of natural antisense transcripts. Nat. Rev. Mol. Cell Biol. *10*, 637–643. https://doi.org/10.1038/nrm2738.

42. Scruggs, B.S., Gilchrist, D.A., Nechaev, S., Muse, G.W., Burkholder, A., Fargo, D.C., and Adelman, K. (2015). Bidirectional transcription arises from two distinct hubs of transcription factor binding and active chromatin. Mol. Cell *58*, 1101–1112. https://doi.org/10.1016/j.molcel.2015.04.006.

43. Warman, E.A., Forrest, D., Guest, T., Haycocks, J.J.R.J., Wade, J.T., and Grainger, D.C. (2021). Widespread divergent transcription from bacterial and archaeal promoters is a consequence of DNA-sequence symmetry. Nat. Microbiol. *6*, 746–756. https://doi.org/10.1038/s41564-021-00898-9.

44. Lenstra, T.L., Coulon, A., Chow, C.C., and Larson, D.R. (2015). Single-molecule imaging reveals a switch between Spurious and functional ncRNA transcription. Mol. Cell *60*, 597–610. https://doi.org/10.1016/j.molcel.2015.09.028.

45. Milone, M.C., and O'Doherty, U. (2018). Clinical use of lentiviral vectors. Leukemia *32*, 1529–1541. https://doi.org/10.1038/s41375-018-0106-0.

46. Coffin, J.M., Hughes, S.H., and Varmus, H.E. (1997). The interactions of retroviruses and their hosts. In Retroviruses, J.M. Coffin, S.H. Hughes, and H.E. Varmus, eds. (Cold Spring Harbor).

47. Nusse, R., and Varmus, H.E. (1982). Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. Cell *31*, 99–109. https://doi.org/10.1016/0092-8674(82)90409-3.

48. Swain, A., and Coffin, J.M. (1992). Mechanism of transduction by retroviruses. Science *255*, 841–845. https://doi.org/10.1126/science.1371365.

49. Uren, A.G., Kool, J., Berns, A., and van Lohuizen, M. (2005). Retroviral insertional mutagenesis: past, present and future. Oncogene *24*, 7656–7672. https://doi.org/10.1038/sj.onc.1209043.

50. Rasmussen, M.H., Ballarin-Gonzalez, B., Liu, J., Lassen, L.B., Fuchtbauer, A., Fuchtbauer, E.M., Nielsen, A.L., and Pedersen, F.S. (2010). Antisense transcription in gammaretroviruses as a mechanism of insertional activation of host genes. J. Virol. *84*, 3780–3788. https://doi.org/10.1128/JVI.02088-09.

51. Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E., et al. (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science *302*, 415–419. https://doi.org/10.1126/science.1088547.

52. Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., et al. (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J. Clin. Invest. *118*, 3132–3142. https://doi.org/10.1172/JCI35700.

53. Melamed, A., Yaguchi, H., Miura, M., Witkover, A., Fitzgerald, T.W., Birney, E., and Bangham, C.R. (2018). The human leukemia virus HTLV-1 alters the structure and transcription of host chromatin in cis. Elife *7*, e36245. https://doi.org/10.7554/eLife.36245.

54. Liu, R., Yeh, Y.H.J., Varabyou, A., Collora, J.A., Sherrill-Mix, S., Talbot, C.C., Jr., Mehta, S., Albrecht, K., Hao, H., Zhang, H., et al. (2020). Single-cell transcriptional landscapes reveal HIV-1-driven aberrant host gene transcription as a potential therapeutic target. Sci. Transl. Med. *12*, eaaz0802. https://doi.org/10.1126/scitranslmed.aaz0802.

55. Chang, K., Marran, K., Valentine, A., and Hannon, G.J. (2013). Creating an miR30-based shRNA vector. Cold Spring Harb. Protoc. *2013*. pdb.prot075853-635. https://doi.org/10.1101/pdb.prot075853.

56. Palangat, M., Anastasakis, D.G., Fei, D.L., Lindblad, K.E., Bradley, R., Hourigan, C.S., Hafner, M., and Larson, D.R. (2019). The splicing factor U2AF1 contributes to cancer progression through a noncanonical role in translation regulation. Genes Dev. *33*, 482–497. https://doi.org/10.1101/gad.319590.118.