

Cell fitness screens reveal a conflict between LINE-1 retrotransposition and DNA replication

Daniel Ardeljan^{1,2,3*}, Jared P. Steranka¹, Chunhong Liu¹, Zhi Li⁴, Martin S. Taylor⁵, Lindsay M. Payer¹, Mikhail Gorbounov¹, Jacob S. Sarnecki⁶, Vikram Deshpande⁵, Ralph H. Hruban¹, Jef D. Boeke⁴, David Fenyö⁴, Pei-Hsun Wu^{7,8}, Agata Smogorzewska⁹, Andrew J. Holland¹⁰ and Kathleen H. Burns^{1,2,11*}

LINE-1 retrotransposon overexpression is a hallmark of human cancers. We identified a colorectal cancer wherein a fast-growing tumor subclone downregulated LINE-1, prompting us to examine how LINE-1 expression affects cell growth. We find that nontransformed cells undergo a TP53-dependent growth arrest and activate interferon signaling in response to LINE-1. TP53 inhibition allows LINE-1⁺ cells to grow, and genome-wide-knockout screens show that these cells require replication-coupled DNA-repair pathways, replication-stress signaling and replication-fork restart factors. Our findings demonstrate that LINE-1 expression creates specific molecular vulnerabilities and reveal a retrotransposition–replication conflict that may be an important determinant of cancer growth.

Long interspersed element 1 (LINE-1) is the only functional, protein-coding retrotransposon in humans. LINE-1 is transcribed as a bicistronic RNA that encodes an RNA-binding protein, open reading frame 1 protein (ORF1p), and an endonuclease (EN) and reverse transcriptase (RT), ORF2p (ref. 1–3). Retrotransposition—the ‘copy-and-paste’ mechanism wherein an ‘active’ or ‘hot’ LINE-1 generates de novo insertions of itself—is a mutagenic process that cells limit by suppressing LINE-1 transcription via DNA methylation^{4,5} and other mechanisms.

Many studies have focused on host factors that alter retrotransposition efficiency or on the functional effects of acquired LINE-1 insertions; fewer have focused on cellular effects of LINE-1 expression^{6–10}. LINE-1 is known to be toxic, but the mechanisms underlying its toxicity are unclear. ORF2p appears to incite DNA double-strand breaks (DSBs) in some systems⁸, although it is thought to function as a single-strand nickase in retrotransposition¹¹. Despite its toxicity, LINE-1 promoter hypomethylation and protein expression are hallmarks of human cancers^{12,13}, and retrotransposition is commonplace in these diseases^{14–26}. This paradox reflects a lack of understanding surrounding LINE-1 toxicity and how malignant cells tolerate LINE-1 expression.

Here, we describe a case of colon cancer with an aggressive tumor subclone that shut down LINE-1 expression concurrent with its accelerated growth. This prompted us to explore how LINE-1 affects cell fitness. We find that LINE-1 triggers a tumor protein p53 (TP53)-mediated G1 arrest and an interferon response in nontransformed cells. In TP53-deficient cells, we conducted a knockout screen to identify genes that affect the fitness of LINE-1⁺ cells. These studies show that LINE-1⁺ cells rely on replication-coupled DNA-repair pathways, replication-stress

signaling responses and replication-fork restart factors for growth. We find that LINE-1 expression activates the Fanconi anemia pathway, induces markers of replication stress and sensitizes cells to mitomycin C (MMC). Accordingly, we propose a model for LINE-1 toxicity wherein LINE-1 retrotransposition conflicts with DNA replication.

Results

Heterogeneous LINE-1 expression in colon cancer. We assessed 22 colorectal cancers for ORF1p expression by immunohistochemistry. All were positive, with varied ORF1p staining intensity; immunoreactivity was limited to cancerous epithelium and not found in adjacent normal tissue (Fig. 1a)¹². One tumor showed dichotomous ORF1p expression, containing a well-differentiated LINE-1⁺ sector and an adjacent, poorly differentiated (CDX2^{dim}), LINE-1[−] sector (Fig. 1b). A metastatic site of disease closely resembled the former. To evaluate whether these two tumor regions were clonally related or independently derived, we genotyped driver point mutations and somatically acquired LINE-1 insertions to create a phylogenetic map (Fig. 1c and Extended Data Fig. 1). We found that the LINE-1⁺ and LINE-1[−] parts of the primary tumor both share a BRAF^{V600E} mutation, as well as numerous somatically acquired LINE-1 insertions incurred before retrotransposition ceased in the LINE-1[−] component (Extended Data Fig. 1c). The LINE-1[−] clone has a markedly increased proliferation index (Fig. 1d). Thus, the LINE-1[−] section derives from a LINE-1⁺ lineage, and loss of LINE-1 expression is associated with an enhanced growth rate.

The p53–p21 pathway restricts growth of LINE-1⁺ cells. To identify growth determinants of LINE-1⁺ cells, we developed an

¹Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ²McKusick-Nathans Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ³Medical Scientist Training Program, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁴Institute for Systems Genetics and Department of Biochemistry and Molecular Pharmacology, NYU Langone Health, New York City, NY, USA. ⁵Department of Pathology, Massachusetts General Hospital, Boston, MA, USA. ⁶Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD, USA. ⁷Johns Hopkins Physical Sciences Oncology Center, Johns Hopkins University, Baltimore, MD, USA. ⁸Institute for NanoBiotechnology, Johns Hopkins University, Baltimore, MD, USA. ⁹Laboratory of Genome Maintenance, The Rockefeller University, New York City, NY, USA. ¹⁰Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ¹¹Sydney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD, USA. *e-mail: ardeljan@jhmi.edu; kburns@jhmi.edu

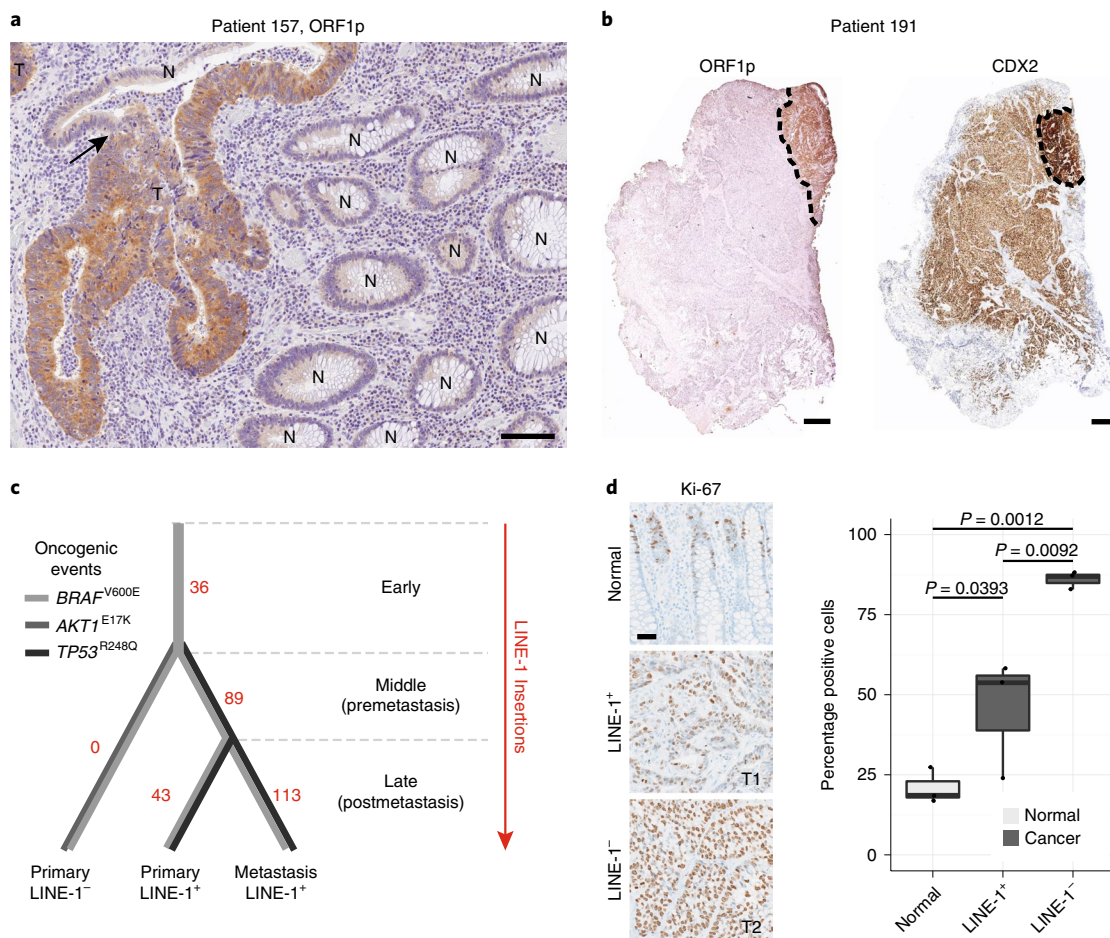


Fig. 1 | Heterogeneous LINE-1 expression in colon cancer. **a**, ORF1p immunohistochemistry (IHC) stain of formalin-fixed paraffin-embedded (FFPE) colon cancer tissue. LINE-1 immunostaining is seen in tumor (T) and not in normal colonic epithelium (N). The arrow indicates a transition from normal to tumor tissue within a gland. Scale bar, 50 μm . **b**, IHC stain of FFPE colon cancer tissue from patient 191. Left: low magnification of ORF1p intensely positive and negative tumor sectors. Right: low magnification of CDX2, a colon epithelium marker. LINE-1⁺ cells express higher CDX2 and form glands, whereas LINE-1⁻ cells express lower CDX2 and do not form glands. Scale bars, 500 μm . **c**, Phylogenetic tree of the tumor subclones in patient 191, based on transposon insertion sequencing and known tumor-driver alleles. The number of de novo LINE insertions is indicated in red along the line edges. Using Sanger sequencing, we genotyped known tumor-driver alleles and found an *AKT1*^{E17K} mutation in the CDX2^{dim} cells and a *TP53*^{R248Q} mutation in CDX2^{high} cells (both primary and metastatic sites). All tumor specimens possessed a *BRAF*^{V600E} allele regardless of LINE-1 expression status. The color of the lines indicates the presence or absence of known tumor-driver alleles. **d**, Ki-67 quantification of normal epithelium, LINE-1⁺ glandular cancer and LINE-1⁻ solid cancer in patient 191. The percentage of positive cells was calculated as the number of Ki-67⁺ nuclei divided by the total number of epithelial cell nuclei. Three independent high-powered fields were counted per tissue morphology, and results were compared using ANOVA and two-sided *t*-tests. Scale bar, 100 μm .

ectopic expression system in telomerase-immortalized retinal pigment epithelium-1 (RPE) cells, genetically stable diploid cells with intact p53 and DNA-damage responses (Fig. 2a,b). LINE-1 expression markedly inhibited RPE clonogenic growth by 98.2% compared with enhanced green fluorescent protein (eGFP) control (Fig. 2c). *TP53* loss-of-function mutations clinically correlate with LINE-1 activity^{12,25,27}, so we compared clonogenic growth of RPE cells expressing LINE-1 or eGFP (measured as number of LINE-1 cells per 100 eGFP colonies) with and without *TP53* knockdown (Fig. 2d and Extended Data Fig. 2a). *TP53* knockdown rescued LINE-1⁺ cell clonogenicity by 42.3-fold, but did not fully restore to LINE-1⁺ cells the clonogenic potential of eGFP-expressing cells. To test whether *TP53* function affects retrotransposition efficiency in this system, we used a reporter assay to compare LINE-1 insertion frequencies in control and *TP53*-knockdown cells, but found no significant difference (Extended Data Fig. 2b). Thus, p53 restricts growth of these cells but not retrotransposition potential.

We next performed a genome-wide CRISPR knockout screen to identify knockouts that rescue growth of LINE-1⁺ cells (Fig. 2e and Methods). Single-guide RNAs (sgRNAs) targeting *TP53* were the only ones to significantly enhance cell fitness (Fig. 2f and Extended Data Fig. 2c). Guides targeting *CDKN1A* (p21), a *TP53*-dependent growth-arrest effector and retrotransposition suppressor²⁸, were enriched but did not reach genome-wide significance (Fig. 2f and Extended Data Fig. 2c). Guide RNAs targeting other genes downstream of *TP53* did not tolerate cells to LINE-1 expression. To validate these findings, we transduced two individual sgRNAs targeting *TP53* or *CDKN1A*, or non-targeting control (NTC), in RPE cells expressing Cas9, and found that each knockout rescued growth of LINE-1⁺ cells (Fig. 2g). These data demonstrate that LINE-1 expression causes a p53–p21-dependent growth arrest.

LINE-1 induces p53-mediated G1 arrest and an interferon response. To characterize this growth arrest further, we performed

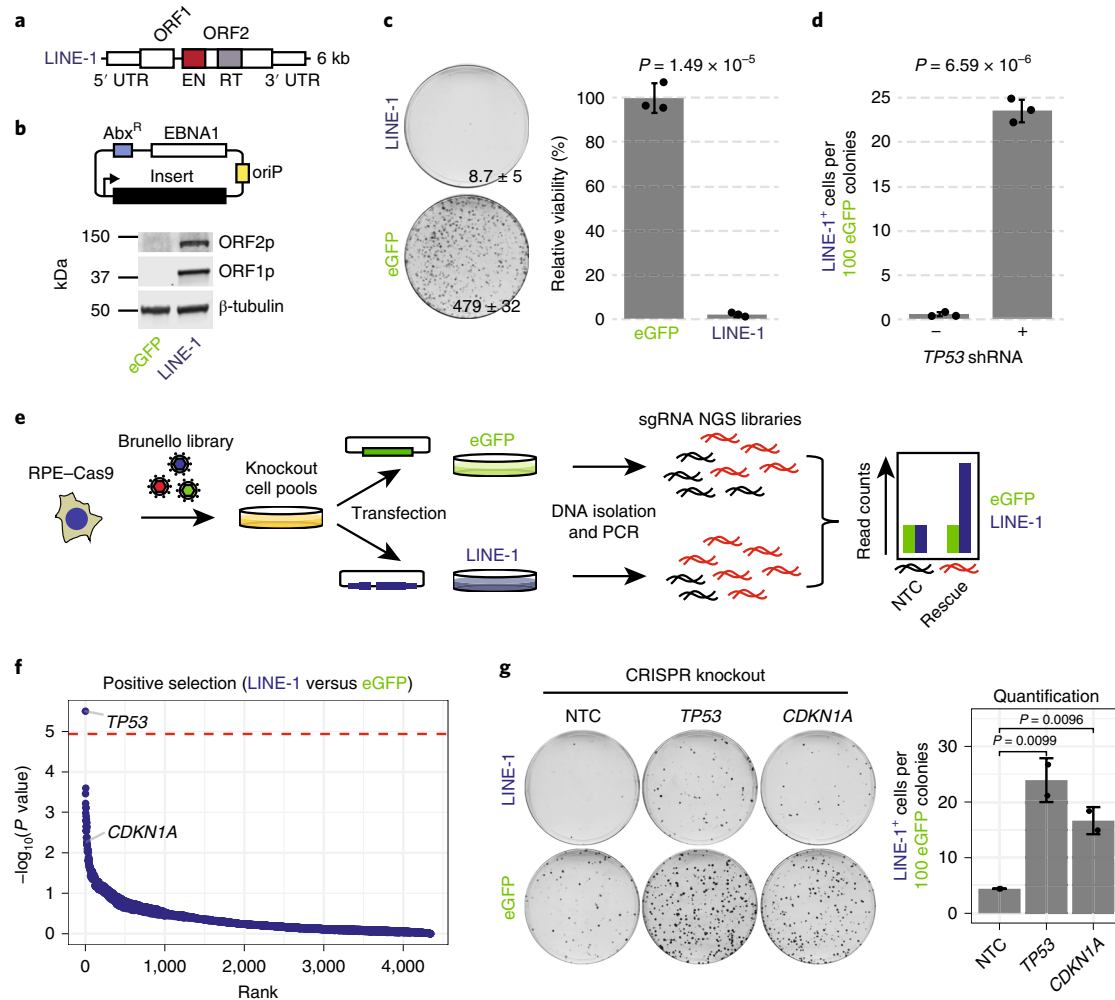


Fig. 2 | LINE-1 inhibits cell growth in RPE by activating the p53-p21 pathway. **a**, LINE-1 sequence. The 5' untranslated region (UTR) is a CpG-rich RNA polymerase II promoter. Open reading frame (ORF) 1 and ORF2 are separated by a 63-bp linker sequence. ORF2 has endonuclease (EN, red) and reverse transcriptase (RT, gray) domains. **b**, Top: episomal pCEP4 mammalian expression vector for eGFP (pDA083) or LINE-1 (pDA077). Abx^R, antibiotic selection marker; EBNA1, Epstein-Barr nuclear antigen 1; oriP, EBNA-1 replication origin. Bottom: western blot of ORF1p and ORF2p from RPE cells transfected with each plasmid. Uncropped blot is shown in Supplementary Data 1. **c**, Clonogenic assay (day 12). Cells are transfected with eGFP or LINE-1. Representative plates with number of colonies indicated \pm s.d. Right: quantification is normalized to eGFP-expressing cells set at 100%, from $n = 3$ independent experiments. P value calculated by two-sided unpaired t -test. **d**, Clonogenic assay (day 12). Cells are treated with lentivirus encoding *TP53* shRNA (+) or control vector (-). Data are presented as the number of LINE-1 colonies per 100 eGFP colonies \pm s.e.m.; $n = 3$ independent experiments. P value obtained by unpaired two-sided t -test. **e**, Positive-selection CRISPR-Cas9 knockout screen workflow using the Brunello CRISPR knockout library. RPE-Cas9, RPE cells constitutively expressing Cas9 protein. KO, knockout. NGS, next-generation sequencing. **f**, Screen enrichment rank versus significance values of gene knockouts that rescue growth of LINE-1⁺ cells. The dashed red line is the family-wise error rate (FWER)-adjusted genome-wide significance level. Low ranks indicate rescue of LINE-1⁺ cells. **g**, Compared with non-targeting control (NTC), CRISPR knockout of *TP53* or *CDKN1A* significantly rescues clonogenic growth of RPE. Representative plates with all data presented as the number of LINE-1 colonies per 100 eGFP colonies \pm s.e.m.; $n = 2$ biological replicates. P value was obtained by unpaired one-sided t -test.

RNA sequencing (RNA-seq) in RPE cells encoding a doxycycline-inducible (Tet-On) codon-optimized LINE-1 (ORFeus) or luciferase control. In total, 2,261 genes were differentially expressed by more than twofold and met Bonferroni-corrected significance (Fig. 3a). Gene set enrichment analysis revealed upregulation of the p53 pathway, and downregulation of cell-cycle progression genes (Fig. 3a, Extended Data Fig. 3a and Supplementary Table 1). Genes possessing p53 regulatory elements ('direct targets'), including *CDKN1A*, were upregulated in LINE-1⁺ cells ($P < 2.2 \times 10^{-16}$), and genes repressed via p21 ('indirect targets') were downregulated ($P < 2.2 \times 10^{-16}$) (Fig. 3b). We confirmed by flow cytometry that LINE-1⁺ cells accumulated in G1 in a LINE-1- and *TP53*-dependent manner (Extended Data Fig. 3b).

LINE-1 expression induces expression of the apoptotic effectors phorbol-12-myristate-13-acetate-induced protein 1 (*PMAIP1*, NOXA) and BCL2 binding component 3 (*BBC3*, PUMA), but does not activate caspase-3, as determined by western blot (data not shown); genes associated with the senescence associated secretory phenotype (SASP)²⁹ were not significantly upregulated (data not shown). These findings are consistent with LINE-1 inducing a p53-mediated G1 cell-cycle arrest.

Most (63.6%) of the gene sets upregulated by LINE-1 expression reflect interferon (IFN)-signaling (Fig. 3c and Supplementary Table 1) and IFN-stimulated genes (Extended Data Fig. 3c), consistent with prior reports³⁰⁻³⁴. This appears to be driven by IFN- β 1 (*IFNB1*) and the double-stranded RNA (dsRNA)-sensing pathway involving

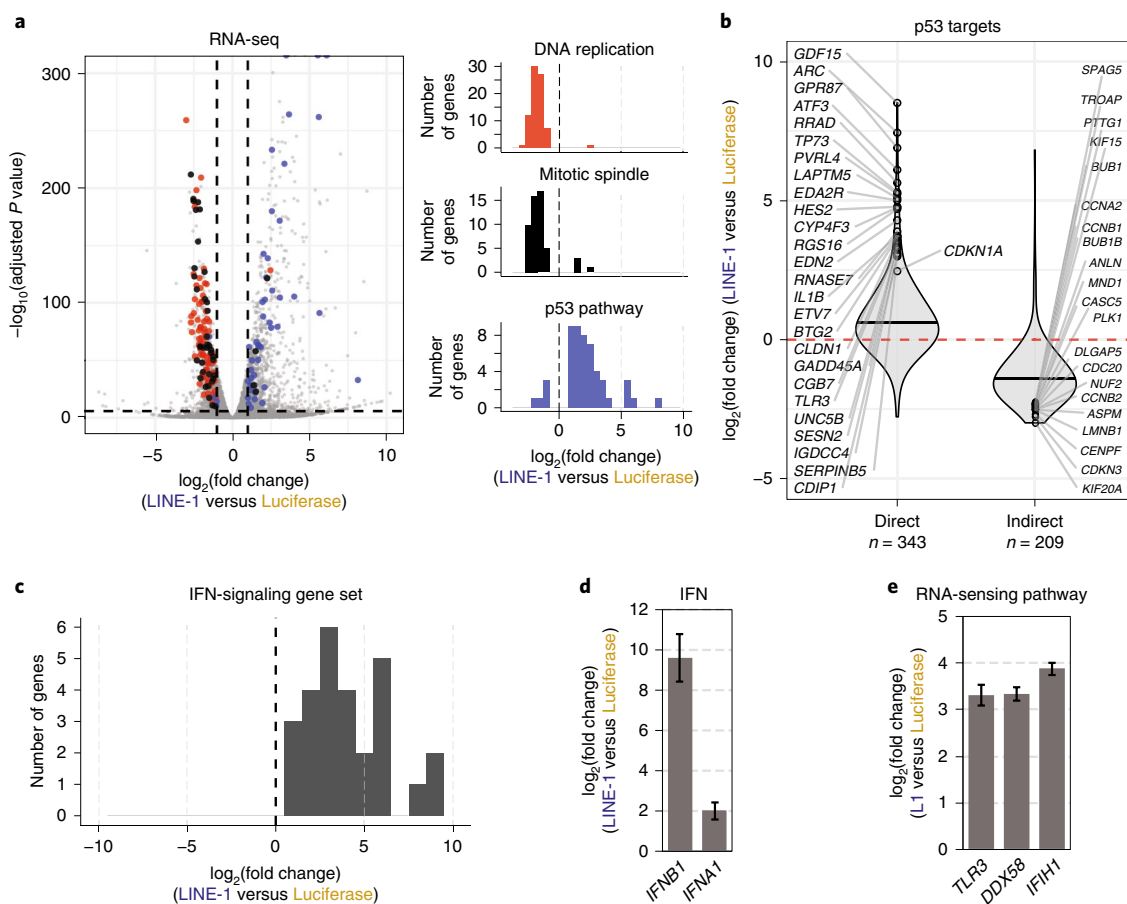


Fig. 3 | LINE-1 activates a p53 and IFN response. **a**, Left: volcano plot of differentially expressed genes. Vertical dashed lines indicate $\log_2(\text{fold change})$ of -1 or $+1$, and the horizontal dashed line indicates a FWER-controlled P value of 0.05 . Right: histograms of gene set enrichment analysis results. Gene set names are indicated above each plot. The number of genes is indicated on the y axis, and the x axis indicates differential expression bins. Individual genes comprising these datasets are highlighted in the volcano plot according to the colors of the bars in the histograms. Data are derived from $n=3$ independent replicates. **b**, Violin plots illustrating differential expression of p53 transcriptional targets. Direct and indirect target genes are curated from refs.^{71,72}. Horizontal bars mark median values. The number of genes in each group is indicated below the plot. **c**, Histogram of gene set enrichment results of IFN signaling genes. The number of genes is indicated on the y axis, and the x axis indicates differential expression. **d**, Relative fold change of *IFNB1* and *IFNA1* in LINE-1⁺ compared with luciferase⁺ cells, measured by RNA-seq. Error bars indicate s.e.m. **e**, RNA-seq analysis revealed upregulation of the RNA-sensing pathway involving *TLR3*, *RIG-I (DDX58)* and *MDA5 (IFIH1)* in LINE-1⁺ cells. Error bars indicate s.e.m.

Toll-like receptor 3 (*TLR3*), DExD/H-box helicase 58 (*DDX58*, or *RIG-I*) and interferon induced with helicase C domain 1 (*IFIH1*, or *MDA5*) (Fig. 3d,e). cGAS–STING is not expressed in these cells. LINE-1 also induces nuclear factor- κ B (NF- κ B)—an immune signaling transcription factor that can be activated by the RNA-sensing pathway³⁵—and NF- κ B transcriptional targets, including the proinflammatory cytokines interleukin-1 β (*IL-1B*) and *CXCL8* (Extended Data Fig. 3d). LINE-1 expression in *TP53*-knockdown cells similarly induces expression of *IFNB1* and interferon-inducible genes, including *TLR3*, *IFIT1* and *IFIT2* (Extended Data Fig. 3e), indicating the response is p53-independent. In contrast, addition of nucleoside reverse-transcriptase inhibitors known to act on LINE-1, zalcitabine (ddC) or didanosine (ddI)³⁶, attenuated the IFN response (Extended Data Fig. 3f). Thus, LINE-1 expression induces an IFN response that might contribute to its inhibitory effects on cell growth independent of p53.

Mapping LINE-1 fitness interactions in *TP53*-deficient cells. We next hypothesized that *TP53*-deficient (*TP53*^{KD}) LINE-1⁺ cells may rely on specific pathways to suppress LINE-1 toxicity. Their loss would be synthetic lethal with LINE-1 expression, and they would be potential therapeutic targets for LINE-1⁺ cancers.

To identify these pathways, we conducted a knockout screen in *TP53*^{KD} RPE cells expressing Cas9 protein (RPE–Cas9) with Tet-On transgenes encoding codon-optimized LINE-1 or luciferase (Fig. 4a). We generated knockout-cell pools in triplicate and expressed LINE-1 or luciferase for 27 d, sampling the populations for sgRNA representation every 4–5 d. Knockouts that become more highly represented in LINE-1⁺ cells relative to luciferase⁺ controls indicate a positive growth interaction, whereas those that are lost indicate a synthetic lethal interaction. Non-targeting control (NTC) sgRNAs were equally represented in LINE-1⁺ and luciferase⁺ cells (Extended Data Fig. 4a). *TP53* and *CDKN1A* knockouts exhibited null interactions in LINE-1⁺ and luciferase⁺ cells (Extended Data Fig. 4b), confirming that *TP53* knockdown was effective and that any p21 growth effects are p53-dependent. As expected, sgRNAs targeting essential genes were depleted from both LINE-1⁺ and luciferase⁺ populations (Extended Data Fig. 4c).

We found 1,390 gene knockouts with significant fitness interactions (Fig. 4b and Supplementary Table 2). Only 24 rescued LINE-1⁺ cell growth. Knockout of the *APC* tumor suppressor is among these (Extended Data Fig. 4d), which is notable since *TP53* and *APC* mutations frequently co-occur in colorectal cancer³⁷ and LINE-1 has mutated *APC* in colon cancers^{22,38}. *IFNAR1* (IFN receptor)

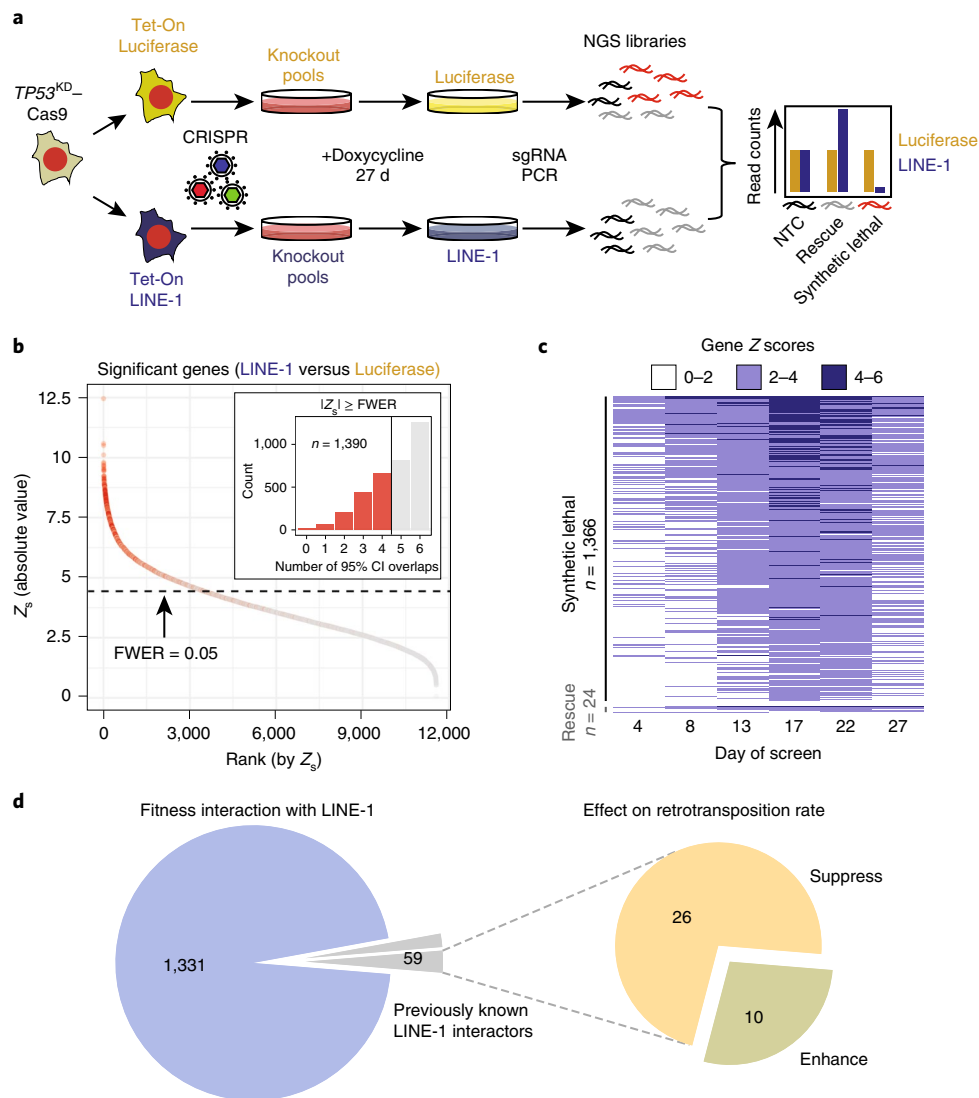


Fig. 4 | Mapping LINE-1 fitness interactions in TP53-deficient cells. a, $TP53^{KD}$ cells are RPE-Cas9 cells stably transduced with shRNA to knock down p53 and then engineered to express luciferase (pDA094) or codon-optimized LINE-1 (pDA095) in a doxycycline-inducible manner (Tet-On). Tet-On cells were transduced with the Brunello CRISPR knockout library at a multiplicity of infection of 0.3 and were puromycin-selected for 8 d before expression of LINE-1 or luciferase was induced for 27 d. Cell pools were sampled at 4- to 5-d intervals and analyzed for sgRNA representation with MAGeCK. Count data are normalized to reads that align to 1,000 built-in NTC sgRNAs (black). **b**, Genes shown as rank-ordered plot of Stauffer Z scores (Z_s) with a FWER of 0.05. Inset indicates the number of 95% confidence interval overlaps over all time points between LINE-1 and luciferase groups among gene knockouts that meet the FWER threshold (red) versus those that do not (gray). **c**, Heat map of 1,390 significant genes depicting the Z scores over time, ranked by Z_s . There are 1,366 synthetic lethal interactions and 24 rescue interactions. Most knockouts achieved detectable effects by 17–22 d into the screen, evidenced by increasing gene Z scores during these time points. **d**, Overlap of genes with LINE-1 fitness interactions observed in the present study with genes previously known to interact with LINE-1 proteins, physically or by modifying retrotransposition. Previously known LINE-1 interactors were identified by Liu et al.⁵; Goodier, Cheung & Kazazian³⁹; Taylor et al.⁴⁰; and Moldovan and Moran⁴¹.

knockout also enhanced cell growth (Extended Data Fig. 4e), highlighting that LINE-1-associated IFN activation suppresses cell growth independently of p53. In contrast, most genes identified in this screen ($n = 1,366$) demonstrate synthetic lethal interactions in LINE-1⁺ cells within 3 weeks of sustained expression (Fig. 4c).

We asked whether genes known to alter LINE-1 retrotransposition efficiency⁵ or that encode proteins that physically interact with ORF1p or ORF2p (ref. ^{39–42}) were enriched for fitness interactions (Fig. 4d and Supplementary Table 3). Of these 239 genes, 59 (24.7%) were identified in our fitness screen, compared with 12.0% (1,390/11,564) of all genes tested, a 2.05-fold enrichment ($\chi^2 = 8.4 \times 10^{-9}$). The majority, 58 of 59 (98.3%), demonstrated synthetic lethal interactions. Of these 59 genes, 10 enhance

retrotransposition, 26 suppress retrotransposition and 25 encode physical interactors. However, these 59 genes only account for 4.2% of genes identified in our study, indicating that most fitness interactors are distinct from host genes that regulate retrotransposition. We conclude that specific gene knockouts cause synthetic lethality in LINE-1⁺ cells. Relatively few knockouts act independently of p53 to enhance growth of LINE-1⁺ cells, and only a minor proportion of fitness interactors are known to influence retrotransposition.

We performed an overrepresentation analysis on all significant fitness interactors and found a 1.4-fold enrichment of genes encoding nuclear proteins ($\chi^2 = 6.61 \times 10^{-21}$; 50.1% of significant genes compared with 35.2% of genes in the library; Methods).

We found 41 gene ontology (GO) terms with a false-discovery rate (FDR) < 0.05 (Supplementary Table 4). The top enriched term was ‘mRNA processing’ (FDR = 2.29×10^{-10}); we also found terms related to maintenance of genome integrity, including ‘DNA repair’ (FDR = 4.47×10^{-7}) and ‘DNA replication’ (FDR = 0.01), and chromatin-related gene sets, including ‘histone modification’ (FDR = 3.07×10^{-8}) and ‘regulation of chromatin organization’ (FDR = 0.001).

HUSH complex loss increases LINE-1 transgene expression. Human silencing hub (HUSH) knockouts produced pronounced LINE-1 synthetic lethal interactions, which we validated by single-gene-knockout clonogenic growth studies (Extended Data Fig. 5a–c). HUSH is an epigenetic repressor complex that targets transgenic DNA sequences, including lentivirus insertions⁴³ and endogenous LINE-1 loci^{5,44}. Thus, we tested whether HUSH loss increases LINE-1 expression, either from endogenous LINE-1 loci or from the codon-optimized transgene. We did not detect ORF1p or ORF2p in no-doxycycline controls (Extended Data Fig. 5d), indicating that HUSH-mutant RPE cells do not upregulate endogenous LINE-1 proteins. In doxycycline-treated cells with the LINE-1 transgene, ORF1p, ORF2p and transgene messenger RNA expression increased with HUSH knockout (Extended Data Fig. 5e,f), and ORF2p protein level linearly correlated with transgene mRNA level (two- to fourfold increase, Extended Data Fig. 5g). ORF2p expression could be similarly increased in HUSH-intact cells transfected with Tet-On LINE-1 plasmid treated with higher doses of doxycycline (Extended Data Fig. 5h), and this is highly cytotoxic. We conclude that the synthetic lethal effect of HUSH mutants is caused by enhanced expression of the LINE-1 transgene. We note that high levels of ORF2p expression overwhelm the survival advantage conferred by *TP53* deficiency.

RNA-processing gene knockouts sensitize cells to LINE-1 expression. The GO term ‘mRNA processing’ encompasses 81 genes demonstrating fitness interactions in LINE-1⁺ cells; these genes are enriched for spliceosome components ($P = 2.24 \times 10^{-34}$), and knockouts of these are synthetic lethal in LINE-1⁺ cells (Extended Data Fig. 6a,b). We validated this effect by treating cells with the splicing inhibitor pladienolide B (PLA-B), which acts on the essential gene *SF3B1* (splicing factor 3b subunit 1), a component of the U2 small nuclear ribonucleoprotein (snRNP). At a PLA-B dose that reduced luciferase⁺ clonogenic growth by 6.8%, LINE-1⁺ cells grew 27.8% fewer colonies, a 4.1-fold increased sensitivity to PLA-B ($P = 0.044$, Extended Data Fig. 6c). We analyzed RNA-seq data from LINE-1⁺ RPE and did not observe alternatively spliced isoforms of the LINE-1 transgene (data not shown), indicating that these gene knockouts probably affect cell growth through an indirect mechanism, rather than by directly processing the LINE-1 RNA. Notably, cells subjected to DNA damage also are sensitized to loss of spliceosome components⁴⁵.

We found pronounced synthetic lethal interactions caused by knockouts of genes encoding the nuclear exosome-targeting (NEXT) complex, which degrades intronic RNAs and processed transcripts⁴⁶. Two of the three complex members demonstrate synthetic lethal interactions (RBM7 and ZCCHC8), whereas the third (SKIV2L2) is encoded by an essential gene (Extended Data Fig. 6d). Similarly, RNASEH2 knockout is synthetic lethal in LINE-1⁺ cells (Extended Data Fig. 6e). RNASEH2 facilitates retrotransposition by degrading LINE-1 RNA from RNA–DNA hybrids after reverse transcription occurs⁴⁷. Thus, when RNASEH2 is lost, this precludes LINE-1 retrotransposition and enhances toxicity.

Finally, we find that LINE-1⁺ cells require the dsRNA adenosine (A) to inosine (I) editing enzyme ADAR1 (Extended Data Fig. 6f), as do cancer cell lines with high expression of interferon-stimulated genes⁴⁸.

Fanconi anemia proteins suppress LINE-1 toxicity. DNA-repair genes that suppress LINE-1 toxicity were enriched for Fanconi anemia (FA)–BRCA1 pathway components ($P = 7.65 \times 10^{-13}$, Fig. 5a). The FA pathway is critical for resolving DNA interstrand crosslinks and transcriptional R-loops that interfere with progression of DNA replication⁴⁹. Knockout of the majority (83%) of the genes known to cause FA and several related genes⁵⁰ exhibited synthetic lethal interactions with LINE-1 (Fig. 5b and Extended Data Fig. 7a), including *BRCA1* (*FANCS*). We chose five genes to validate based on their functions in the pathway: *FANCM*, a helicase and branch translocase that has high affinity for stalled replication forks and RNA–DNA hybrids; *FANCA*, which is required for FA ‘core complex’ assembly; *FANCL*, the E3 ubiquitin ligase that activates the downstream effectors of the ‘ID Complex’; and ID complex members *FANCI* and *FANCD2*. We confirmed knockout efficacy by measuring MMC-induced *FANCD2* mono-ubiquitination (*FANCD2*-Ub) (Fig. 5c). MMC-induced *FANCD2*-Ub in NTC-treated cells, but not in the FA knockouts. These FA-deficient mutants were selectively sensitive to LINE-1 expression compared with NTCs (Fig. 5d), and displayed slight increases in chromatin-bound γ H2A.X, a marker of DNA damage, compared with NTC-treated LINE-1⁺ cells (1.1- to 1.7-fold, Extended Data Fig. 7b). Expression of native LINE-1 sequence is also synthetic lethal in *FANCD2*-knockout cells compared with NTC cells (Extended Data Fig. 7c).

On the basis of these data and reports that FA proteins suppress retrotransposition⁵, we hypothesized that the FA pathway is activated by LINE-1. To test this, we measured mono-ubiquitination of FA effector proteins *FANCD2* and *FANCI* and found 1.6- and 1.5-fold increases, respectively, with LINE-1 expression (Fig. 5e). Importantly, LINE-1 cytotoxicity has been previously reported to depend on endonuclease (EN) and reverse transcriptase (RT) activities^{8–10}, and we confirmed that expression of LINE-1 with inactivating EN and RT mutations is less toxic than wild-type (WT) LINE-1 (Extended Data Fig. 8). To dissect whether the enzymatic activities of LINE-1 are necessary for FA activation, we measured *FANCD2* monoubiquitination in HeLa cells expressing wild-type LINE-1 or mutants lacking EN activity and/or RT activity. Whereas wild-type LINE-1 increased *FANCD2*-Ub (2.6-fold), both EN- and RT-inactivating mutations (H230A and D702Y, respectively)^{2,51} did not (Fig. 5f). We next assessed FA activation by enumerating *FANCD2* nuclear foci. We expressed WT or RT mutant LINE-1 and quantified *FANCD2* nuclear foci in randomly imaged, EdU-labeled cells. Both hydroxyurea (HU) treatment and LINE-1 expression increased the number of *FANCD2* foci in S phase (EdU⁺) cells ($P = 1.7 \times 10^{-8}$ and 5.8×10^{-11} , respectively, Fig. 5g) but not those in G1/G2 (EdU⁻) phase (Extended Data Fig. 7d). The LINE-1 RT mutant did not induce *FANCD2* foci formation. Together, these data demonstrate that LINE-1 activates the FA complex and replication-coupled DNA repair. By contrast, LINE-1 EN and RT mutants do not have this effect, suggesting that the LINE-1 retrotransposition intermediate is crucial to the process.

To evaluate DNA damage associated with LINE-1 expression, we measured γ H2A.X and 53BP1 nuclear foci. We found that LINE-1⁺ cells have transient increases in numbers of γ H2A.X and 53BP1 foci as compared with control cells ($P = 3.4 \times 10^{-6}$ and 1.7×10^{-12} , respectively, Fig. 5h). These increases are detectable in S phase and resolve by G2, whereas doxorubicin-induced DNA damage foci continue to accumulate (data not shown). This pattern is more consistent with LINE-1-induced replication stress^{52,53} than with a large burden of persistent, dsDNA breaks.

Retrotransposition–replication conflict underpins LINE-1 toxicity. We next explored interactions between LINE-1 retrotransposition and DNA replication using our fitness screen data. Stalled replication forks activate signaling pathways involving ataxia telangiectasia and Rad3-related (ATR) and ATR-interacting protein

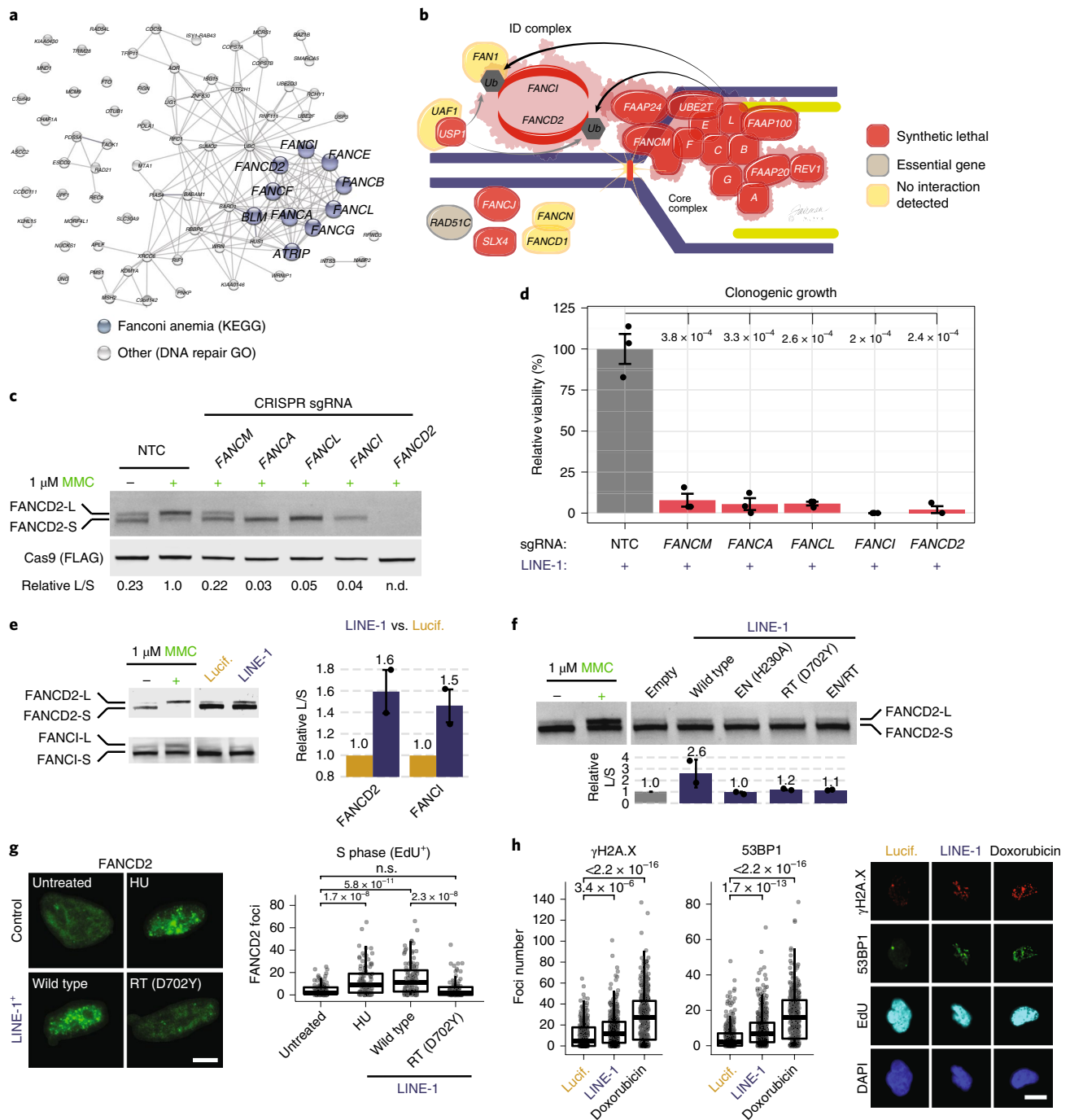


Fig. 5 | The Fanconi anemia pathway is essential in p53-deficient cells. **a**, A network of 75 DNA repair genes identified in the screen is enriched for Fanconi anemia genes (blue nodes). Edges indicate known physical interactions. **b**, Model of FA complexes responding to a DNA lesion (vertical line) encountered by a replication fork (blue line, genomic DNA; green line, nascent DNA). Genes are color-coded on the basis of the performance of their knockouts. Illustration generously provided by J. Fairman of the Department of Art as Applied to Medicine at Johns Hopkins University School of Medicine. **c**, Western blot of FANCD2 response to 24-h treatment with $1 \mu\text{g ml}^{-1}$ MMC. Cells are treated with FA member sgRNAs or NTC. FANCD2 mono-ubiquitination assessed as the ratio of FANCD2-L (long) to FANCD2-S (short) band intensities (relative L/S ratio) graphed relative to NTC, MMC-treated cells. n.d., not determined. **d**, Clonogenic growth assay of LINE-1⁺ TP53^{KD} cells with sgRNAs targeting the same genes as in **c**. $n = 3$ independent experiments. P value calculated with a one-sided t -test. **e**, Left: representative western blot of FANCD2 and FANCI following 72-h expression of LINE-1 or luciferase in RPE. MMC treatment reveals L (mono-ubiquitinated) and S (nonubiquitinated) protein bands. Right: quantification of data from $n = 2$ independent experiments \pm s.e.m.. **f**, Representative western blot of FANCD2 following 72-h expression of wild-type or mutant LINE-1 in HeLa cells. Quantification below of $n = 2$ independent experiments \pm s.e.m.. Effect of wild-type LINE-1 as assessed by ANOVA ($P = 0.0143$). **g**, Left: representative images of FANCD2 foci (green) in EdU⁺ nuclei. Scale bar, $6 \mu\text{m}$. Right: quantification of FANCD2 foci. Number of cells per group: untreated, $n = 134$; HU, $n = 105$; wild type, $n = 109$; RT (D702Y), $n = 101$. n.s., not significant. **h**, Left: $\gamma\text{H2A.X}$ and 53BP1 focus quantification in EdU⁺ TP53^{KD} cells. Number of cells per group: Lucif., $n = 326$; LINE-1, $n = 358$; doxorubicin, $n = 431$. Two-sided t -tests were used for statistical comparisons in **g** and **h**. Right: representative images of $\gamma\text{H2A.X}$ (red), 53BP1 (green), EdU (cyan) and DAPI (blue). Scale bar, $12 \mu\text{m}$. Uncropped blot images of **c**, **e** and **f** are shown in Supplementary Data 1.

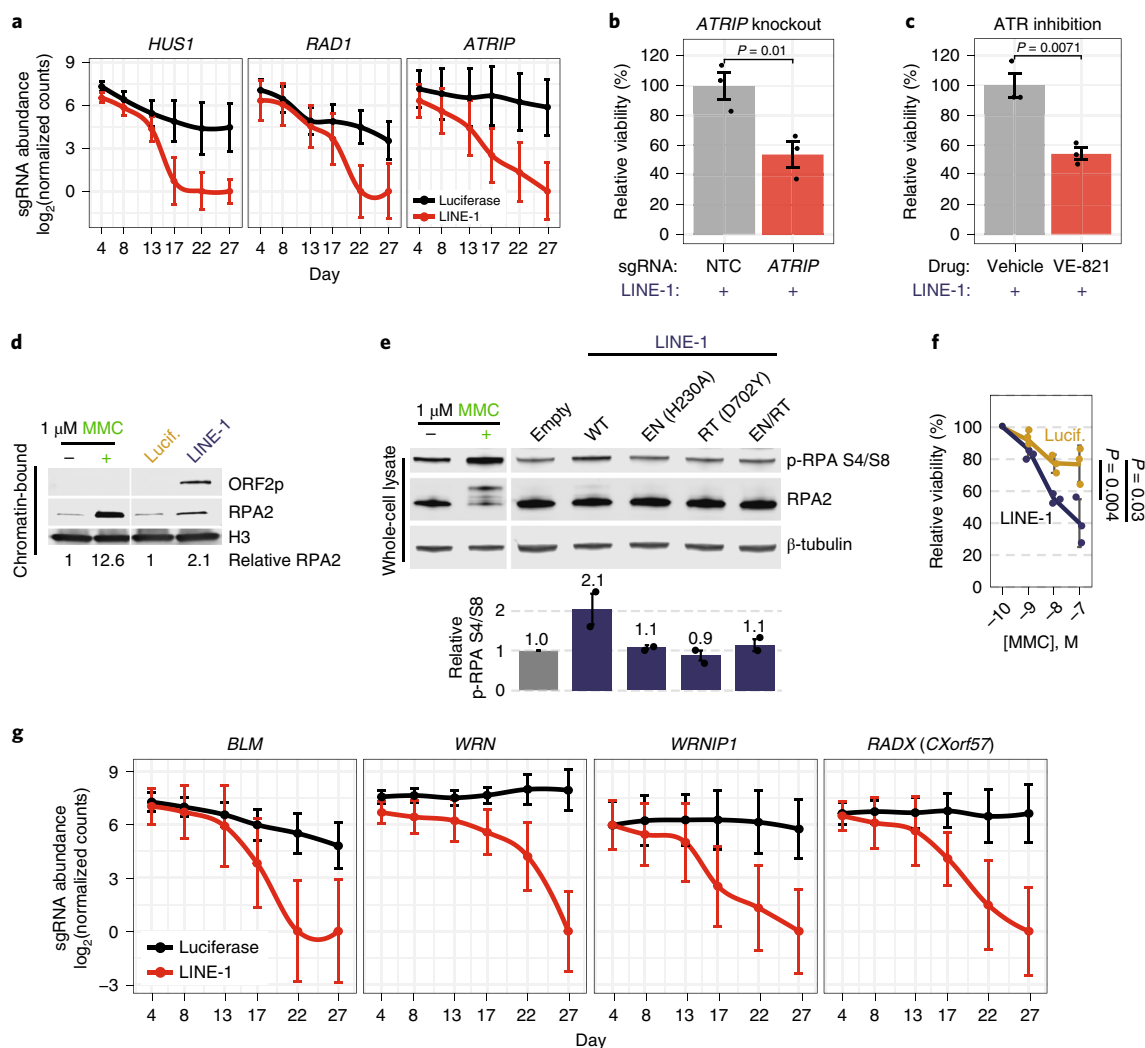


Fig. 6 | LINE-1 activity induces replication stress. **a**, Median count of sgRNAs targeting replication stress signaling genes *ATRIP* and the 9–1–1 complex (*HUS1* and *RAD1*) during the screen. Error bars indicate 95% confidence intervals. **b**, Clonogenic assay of LINE-1⁺ *TP53*^{KD} cells (induced with 1 μg ml⁻¹ doxycycline) with CRISPR knockout of *ATRIP* compared with NTC. Error bars indicate s.e.m.; *n* = 3 independent experiments. *P* value is calculated with an unpaired two-sided *t*-test. **c**, Clonogenic assay of LINE-1⁺ *TP53*^{KD} cells (induced with 1 μg ml⁻¹ doxycycline) with drug inhibition of ATR kinase by 1 μM VE-821 compared with vehicle (DMSO). Error bars indicate s.e.m.; *n* = 3 independent experiments. *P* value is calculated with an unpaired two-sided *t*-test. **d**, Western blot of RPA2 occupancy on chromatin induced by LINE-1 compared with luciferase control after 72 h of expression in RPE. Chromatin-bound protein lysates were used. We used 1 μM MMC as a control to verify that these cells respond to replication stress. **e**, Western blot of p-RPA S4/S8 after 72 h of WT or mutant LINE-1 expression in HeLa cells. Relative signal intensity for *n* = 2 independent experiments ± s.e.m. is quantified. 1 μM MMC was used as a replication stress control and produces RPA2 hyperphosphorylation and a gel shift in total RPA2. WT LINE-1 expression has this effect to a lesser degree. Statistical significance is assessed by ANOVA (*P* = 0.0007). **f**, MMC dose-response clonogenic assay of LINE-1⁺ cells or control. Molar concentration is indicated on the x axis. Data are plotted as the mean viability relative to 100 pM ± s.d.; *n* = 3 independent experiments. Two-sided *t*-tests were used to compare relative viability at each dose. **g**, Median count of sgRNAs targeting fork protection (*RADX*) and fork restart (*BLM*, *WRN*, *WRNIP1*) genes. Median values are depicted with 95% confidence intervals. Uncropped blot images of **d** and **e** are shown in Supplementary Data 1.

(*ATRIP*), as well as the tripartite *RAD9*, *HUS1*, *RAD1* (9–1–1) complex. *ATR* and *RAD9* are essential, but genes encoding all non-essential components of these complexes (*ATRIP*, *HUS1* and *RAD1*) are synthetic lethal LINE-1 interactors (Fig. 6a). We validated that *ATRIP*-knockout cells exhibited heightened sensitivity to LINE-1 expression (Fig. 6b); they also failed to sufficiently activate *FANCD2* upon MMC-induced DNA damage (data not shown). Similarly, *ATR* inhibition with the compound VE-821 sensitized cells to LINE-1 (Fig. 6c) at a dose that had no effect on viability in luciferase⁺ cells (data not shown). Thus, compromising replication stress signaling is synthetic lethal in LINE-1⁺ cells, potentially related to the role of *ATR*–*ATRIP* signaling in activating the FA pathway^{54,55}.

We next assayed for signs of replication-fork stall. Stalled replication forks accumulate single-stranded DNA (ssDNA) coated by replication protein A (RPA), a heterotrimer composed of RPA1, RPA2 and RPA3, to protect genomic DNA from nucleases⁵⁶. We isolated chromatin-bound protein fractions from cells treated with MMC or those expressing LINE-1 or luciferase and found that both MMC treatment and LINE-1 expression increased chromatin-bound RPA2 (Fig. 6d). These data show replication stress occurring in a LINE-1-dependent manner. We next asked whether LINE-1-associated replication stress depends on ORF2p enzymatic activity. We expressed wild-type or mutant LINE-1 from Tet-On plasmids in HeLa cells and measured p-RPA S4/S8, a phosphorylation modification placed on RPA2 during replication stress. Wild-type LINE-1 significantly

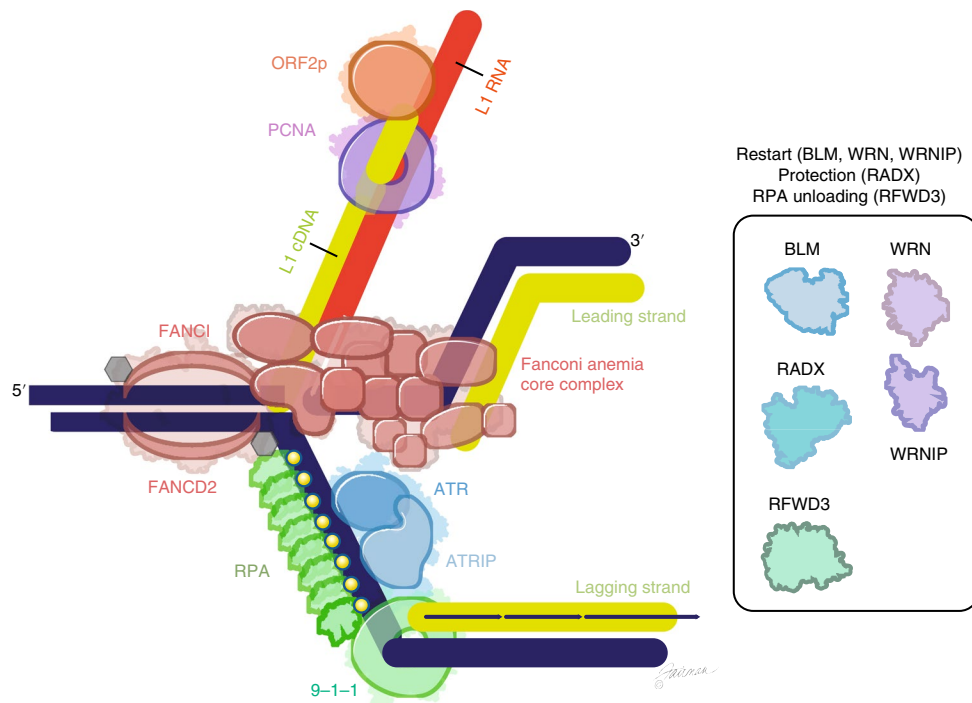


Fig. 7 | Model of LINE-1-induced replication stress. Collision of a replication fork, composed of genomic DNA (dark blue) and newly synthesized DNA (yellow), with a LINE-1 insertion intermediate—an RNA-DNA hybrid made of LINE-1 mRNA (red) and LINE-1 complementary DNA (yellow). The LINE-1 insertion intermediate is recognized by the Fanconi anemia pathway core complex and recruits and activates FANCD2 and FANCI, which are then mono-ubiquitinated. The stalled fork leads to an accumulation of RPA, which recruits ATR-ATRIP and the 9-1-1 complex, key replication-stress signaling proteins. These coordinate the cell response to the replication stress, including phosphorylation of RPA. Failure to resolve this collision reduces cell fitness. A similar conflict could occur upstream of the lagging strand as well.

induced phosphorylation by 2.1-fold ($P = 0.0007$), whereas EN- and RT-inactive mutants did not (Fig. 6e). These data indicate that ORF2p must nick DNA and reverse transcribe in order to induce replication stress, highlighting the importance of the retrotransposition intermediate in these events. Moreover, LINE-1⁺ cells were 1.9-fold more sensitive to MMC as compared with luciferase-expressing controls (Fig. 6f). Together, these data indicate that LINE-1 retrotransposition induces replication stress and sensitizes cells to compounds that increase demands on replication-coupled DNA repair.

Several key processes occur downstream of replication stress signaling, including: (1) fork reversal (that is, translocation of the replication fork away from the lesion and resection by nucleases including *ZRANB3*, *SMARCAL1* and *HLTF*); (2) fork protection from excess degradation by nucleases; and (3) fork restart⁵⁷. Fork-reversal genes do not score in our screen, whereas the fork protection factor *RADX* and proteins that are important for fork restart—including Bloom helicase (*BLM*), Werner helicase (*WRN*) and WRN interacting protein 1 (*WRNIP1*)—are LINE-1 synthetic-lethal interactors (Fig. 6g). Fork restart additionally requires the removal of RPA from the ssDNA. To this end, we note that knockout of *RFWD3*, an FA member whose E3 ubiquitin ligase activity regulates RPA unloading from chromatin⁵⁸, produces synthetic lethality (Extended Data Fig. 7a). These findings indicate that replication-fork protection and restart, but not reversal, are essential for LINE-1 cell growth.

Taken together, these data are consistent with a model wherein LINE-1 retrotransposition intermediates cause replication stress (Fig. 7). LINE-1⁺ cells rely on FA-mediated DNA repair, replication-stress signaling and fork-restart pathways for growth.

Discussion

LINE-1 expression slows cell growth, yet is a hallmark of many human cancers. Here, we used in vitro LINE-1 expression systems,

gene-expression profiling and CRISPR-Cas9 gene-knockout screening to characterize cellular responses to LINE-1 expression. We find that LINE-1 expression in nontransformed cells triggers p53-p21 mediated G1 arrest. Along with studies that place p53 as an upstream repressor of LINE-1 expression, our findings explain associations between LINE-1 expression and *TP53* loss in human cancers^{12,25,27}. Interestingly, although *TP53* loss promotes cell growth absent LINE-1 (ref. 59), we find LINE-1 enhances the relative growth advantage conferred by *TP53* mutation, raising the possibility that LINE-1 expression early in tumorigenesis may select for *TP53* mutations. This may be relevant in ovarian cancer, in which LINE-1 expression and fixation of *TP53* mutations appear to be essentially concordant events in serous tubal intraepithelial carcinoma (STIC) precursor lesions^{60,61}. Similarly, with implications for colon cancer development, we find LINE-1 enhances growth advantages conferred by *APC* mutation in p53-deficient cells. *APC* loss is an early event in these malignancies that can be antedated by LINE-1 expression and retrotransposition^{22,38}.

TP53 loss in turn tolerizes cells to LINE-1 expression. On the basis of a genome-wide CRISPR knockout screen, though, we find that LINE-1 expression confers specific molecular requirements for cell growth in a *TP53*-deficient background. LINE-1⁺ cells rely on RNA-processing machinery, including complexes that degrade RNA and spliceosome components. The former may directly act on retrotransposition intermediates⁴⁷. Compromised splicing may lead to the accumulation of dsRNA and exacerbate interferon responses to LINE-1 expression, or to an excess of transcriptional R-loops on chromatin that pose barriers to DNA replication⁶².

Most notably, our data indicate that retrotransposition conflicts with DNA replication. This model was suggested by the reliance of LINE-1⁺, p53-deficient cells on replication-coupled DNA-repair pathways mediated by the Fanconi anemia components.

All FA complex components show synthetic-lethal interactions with LINE-1 expression in our experimental system. Further, we demonstrate that the FA complex assembles in the S phase of the cell cycle in a manner that depends on LINE-1 enzymatic activity. In accordance with the importance of FA in reducing LINE-1 lesions, tumors that frequently express LINE-1 tend to amplify FA genes^{12,49}. Similarly, we find LINE-1⁺ cells have unique requirements for replication-stress signaling pathways (*ATRIP*, 9–1–1 complex components), replication-fork protection (*RADX*) and fork-restart factors (*BLM* and *WRN* helicases). We corroborate these genetic interactions biochemically by showing LINE-1 ORF2p enzymatic activities induce replication stress. Notably, both EN and RT activities are required to observe FA-pathway activation as well as replication-stress responses. Based on what is known about target-primed reverse transcription, this observation suggests that the branched LINE-1 insertion intermediate structures create physical blockades to replication-fork progression.

This model is further substantiated by independent, orthogonal observations in our field. In *in vitro* experimental systems, there is a predilection for *de novo* LINE-1 insertions to occur in S phase⁶³. Moreover, recent studies mapping LINE-1 insertion sites *in vitro*^{64,65} and *in vivo* in a wide variety of human cancers⁶⁶ indicate nonrandom distributions of insertions with respect to DNA-replication timing. Finally, FA and *BRCA1* inhibit LINE-1 retrotransposition, as has been shown by Liu *et al.*⁵, Mita *et al.*⁶⁷, and Moran and Garcia-Perez (personal communication). These findings indicate that retrotransposition is occurring in association with DNA replication, and that replication-coupled DNA-repair pathways are likely reducing retrotransposition intermediates. Loss of these repair pathways enhances both retrotransposition and LINE-1-associated toxicity.

We propose that the most crucial retrotransposition intermediates are found in unreplicated dsDNA positioned to collide with replication forks. It is possible that multiple intermediates form in each cell, and that most are normally reduced by FA repair or other mechanisms rather than resolved into new genomic insertions. Considering that LINE-1 is aberrantly expressed in half of human cancers¹² and many malignancies acquire between tens and thousands of somatic LINE-1 insertions^{14–21,23,26}, retrotransposition potentially represents an important source of endogenous replication stress and genomic instability in these malignancies.

Our findings underscore that limits on LINE-1 expression are required in order to preserve cell growth, and indeed we began our study after seeing a tumor subclone that lost LINE-1 expression and grew faster. Moreover, we provide the first evidence of unique molecular vulnerabilities in LINE-1⁺ cells, which has noteworthy implications for translational cancer research. From a therapeutic perspective, it is possible that LINE-1⁺ cancers will have characteristic drug sensitivities; for example, LINE-1 ORF2p expression and retrotransposition may prove a biomarker for tumors that respond to DNA-damaging agents, or inhibitors of ATR⁶⁸ or WRN helicase⁶⁹. We also demonstrate that LINE-1 promotes a type I IFN response, suggesting roles for LINE-1 in sensitivities to immunotherapies or ADAR inhibition^{48,70}. Experiments in disease-specific model systems that recapitulate chronic LINE-1 exposure are needed to address these possibilities.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41594-020-0372-1>.

Received: 8 February 2019; Accepted: 31 December 2019;
Published online: 10 February 2020

References

- Mathias, S. L., Scott, A. F., Kazazian, H. H. Jr., Boeke, J. D. & Gabriel, A. Reverse transcriptase encoded by a human transposable element. *Science* **254**, 1808–1810 (1991).
- Feng, Q., Moran, J. V., Kazazian, H. H. Jr. & Boeke, J. D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905–916 (1996).
- Hohjoh, H. & Singer, M. F. Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *EMBO J.* **15**, 630–639 (1996).
- Woodcock, D. M., Lawler, C. B., Linsenmeyer, M. E., Doherty, J. P. & Warren, W. D. Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. *J. Biol. Chem.* **272**, 7810–7816 (1997).
- Liu, N. *et al.* Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature* **553**, 228–232 (2018).
- Haoudi, A., Semmes, O. J., Mason, J. M. & Cannon, R. E. Retrotransposition-competent human LINE-1 induces apoptosis in cancer cells with intact p53. *J. Biomed. Biotechnol.* **2004**, 185–194 (2004).
- Belgnaoui, S. M., Gosden, R. G., Semmes, O. J. & Haoudi, A. Human LINE-1 retrotransposon induces DNA damage and apoptosis in cancer cells. *Cancer Cell Int.* **6**, 13 (2006).
- Gasior, S. L., Wakeman, T. P., Xu, B. & Deininger, P. L. The human LINE-1 retrotransposon creates DNA double-strand breaks. *J. Mol. Biol.* **357**, 1383–1393 (2006).
- Wallace, N. A., Belancio, V. P. & Deininger, P. L. L1 mobile element expression causes multiple types of toxicity. *Gene* **419**, 75–81 (2008).
- Kines, K. J. *et al.* The endonuclease domain of the LINE-1 ORF2 protein can tolerate multiple mutations. *Mob. DNA* **7**, 8 (2016).
- Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**, 595–605 (1993).
- Rodic, N. *et al.* Long interspersed element-1 protein expression is a hallmark of many human cancers. *Am. J. Pathol.* **184**, 1280–1286 (2014).
- Ardejan, D., Taylor, M. S., Ting, D. T. & Burns, K. H. The human long interspersed element-1 retrotransposon: an emerging biomarker of neoplasia. *Clin. Chem.* **63**, 816–822 (2017).
- Iskow, R. C. *et al.* Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell* **141**, 1253–1261 (2010).
- Lee, E. *et al.* Landscape of somatic retrotransposition in human cancers. *Science* **337**, 967–971 (2012).
- Shukla, R. *et al.* Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. *Cell* **153**, 101–111 (2013).
- Tubio, J. M. C. *et al.* Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science* **345**, 1251343 (2014).
- Rodic, N. *et al.* Retrotransposon insertions in the clonal evolution of pancreatic ductal adenocarcinoma. *Nat. Med.* **21**, 1060–1064 (2015).
- Ewing, A. D. *et al.* Widespread somatic L1 retrotransposition occurs early during gastrointestinal cancer evolution. *Genome Res.* **25**, 1536–1545 (2015).
- Doucet-O'Hare, T. T. *et al.* LINE-1 expression and retrotransposition in Barrett's esophagus and esophageal carcinoma. *Proc. Natl Acad. Sci. USA* **112**, E4894–E4900 (2015).
- Doucet-O'Hare, T. T. *et al.* Somatic acquired LINE-1 insertions in normal esophagus undergo clonal expansion in esophageal squamous cell carcinoma. *Hum. Mutat.* **37**, 942–954 (2016).
- Scott, E. C. *et al.* A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer. *Genome Res.* **26**, 745–755 (2016).
- Tang, Z. *et al.* Human transposon insertion profiling: Analysis, visualization and identification of somatic LINE-1 insertions in ovarian cancer. *Proc. Natl Acad. Sci. USA* **114**, E733–E740 (2017).
- Burns, K. H. Transposable elements in cancer. *Nat. Rev. Cancer* **17**, 415–424 (2017).
- Jung, H., Choi, J. K. & Lee, E. A. Immune signatures correlate with L1 retrotransposition in gastrointestinal cancers. *Genome Res.* **28**, 1136–1146 (2018).
- Schauer, S. N. *et al.* L1 retrotransposition is a common feature of mammalian hepatocarcinogenesis. *Genome Res.* **28**, 639–653 (2018).
- Wylie, A. *et al.* p53 genes function to restrain mobile elements. *Genes Dev.* **30**, 64–77 (2016).
- Kawano, K. *et al.* HIV-1 Vpr and p21 restrict LINE-1 mobility. *Nucleic Acids Res.* **46**, 8454–8470 (2018).
- Ruscetti, M. *et al.* NK cell-mediated cytotoxicity contributes to tumor control by a cytostatic drug combination. *Science* **362**, 1416–1422 (2018).
- Yu, Q. *et al.* Type I interferon controls propagation of long interspersed element-1. *J. Biol. Chem.* **290**, 10191–10199 (2015).
- Bregnard, C. *et al.* Upregulated LINE-1 activity in the fanconi anemia cancer susceptibility syndrome leads to spontaneous pro-inflammatory cytokine production. *EBioMedicine* **8**, 184–194 (2016).
- De Cecco, M. *et al.* L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature* **566**, 73–78 (2019).

33. Thomas, C. A. et al. Modeling of TREX1-dependent autoimmune disease using human stem cells highlights L1 accumulation as a source of neuroinflammation. *Cell Stem Cell* **21**, 319–331 (2017).
34. Simon, M. et al. LINE1 derepression in aged wild-type and SIRT6-deficient mice drives inflammation. *Cell Metab.* **29**, 871–885 (2019).
35. Pfeffer, L. M. The role of nuclear factor kappaB in the interferon response. *J. Interferon Cytokine Res.* **31**, 553–559 (2011).
36. Dai, L., Huang, Q. & Boeke, J. D. Effect of reverse transcriptase inhibitors on LINE-1 and Ty1 reverse transcriptase activities and on LINE-1 retrotransposition. *BMC Biochem.* **12**, 18 (2011).
37. Smith, G. et al. Mutations in APC, Kirsten-ras, and p53–alternative genetic pathways to colorectal cancer. *Proc. Natl Acad. Sci. USA* **99**, 9433–9438 (2002).
38. Miki, Y. et al. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res.* **52**, 643–645 (1992).
39. Goodier, J. L., Cheung, L. E. & Kazazian, H. H. Jr. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids Res.* **41**, 7401–7419 (2013).
40. Taylor, M. S. et al. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. *Cell* **155**, 1034–1048 (2013).
41. Moldovan, J. B. & Moran, J. V. The Zinc-finger antiviral protein ZAP inhibits LINE and Alu retrotransposition. *PLoS Genet.* **11**, e1005121 (2015).
42. Taylor, M. S. et al. Dissection of affinity captured LINE-1 macromolecular complexes. *Elife* **7**, e30094 (2018).
43. Tchasovnikarova, I. A. et al. GENE SILENCING. Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells. *Science* **348**, 1481–1485 (2015).
44. Robbez-Masson, L. et al. The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes. *Genome Res.* **28**, 836–845 (2018).
45. Adamson, B., Smogorzewska, A., Sigoillot, F. D., King, R. W. & Elledge, S. J. A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. *Nat. Cell Biol.* **14**, 318–328 (2012).
46. Lubas, M. et al. Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell* **43**, 624–637 (2011).
47. Benitez-Guijarro, M. et al. RNase H2, mutated in Aicardi-Goutieres syndrome, promotes LINE-1 retrotransposition. *EMBO J.* **37**, e98506 (2018).
48. Gannon, H. S. et al. Identification of ADAR1 adenosine deaminase dependency in a subset of cancer cells. *Nat. Commun.* **9**, 5450 (2018).
49. Nalepa, G. & Clapp, D. W. Fanconi anaemia and cancer: an intricate relationship. *Nat. Rev. Cancer* **18**, 168–185 (2018).
50. Richardson, C. D. et al. CRISPR–Cas9 genome editing in human cells occurs via the Fanconi anemia pathway. *Nat. Genet.* **50**, 1132–1139 (2018).
51. Moran, J. V. et al. High frequency retrotransposition in cultured mammalian cells. *Cell* **87**, 917–927 (1996).
52. Ward, I. M. & Chen, J. Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress. *J. Biol. Chem.* **276**, 47759–47762 (2001).
53. Her, J., Ray, C., Altshuler, J., Zheng, H. & Bunting, S. F. 53BP1 mediates ATR–Chk1 signaling and protects replication forks under conditions of replication stress. *Mol. Cell Biol.* **38**, e00472-17 (2018).
54. Shigechi, T. et al. ATR–ATRIP kinase complex triggers activation of the Fanconi anemia DNA repair pathway. *Cancer Res.* **72**, 1149–1156 (2012).
55. Cortez, D., Guntuku, S., Qin, J. & Elledge, S. J. ATR and ATRIP: partners in checkpoint signaling. *Science* **294**, 1713–1716 (2001).
56. Cimprich, K. A. & Cortez, D. ATR: an essential regulator of genome integrity. *Nat. Rev. Mol. Cell Biol.* **9**, 616–627 (2008).
57. Bhat, K. P. & Cortez, D. RPA and RAD51: fork reversal, fork protection, and genome stability. *Nat. Struct. Mol. Biol.* **25**, 446–453 (2018).
58. Feeney, L. et al. RPA-mediated recruitment of the E3 ligase RFWF3 is vital for interstrand crosslink repair and human health. *Mol. Cell* **66**, 610–621.e4 (2017).
59. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* **24**, 927–930 (2018).
60. Pisanic, T. R., 2nd et al. Long interspersed element 1 retrotransposons become deregulated during the development of ovarian cancer precursor lesions. *Am. J. Pathol.* **189**, 513–520 (2018).
61. Zhouchunyang, X. et al. Expression of L1 retrotransposon open reading frame protein 1 (L1ORF1p) in gynecologic cancers. *Hum. Pathol.* **92**, 39–47 (2019).
62. Hamperl, S. & Cimprich, K. A. Conflict resolution in the genome: how transcription and replication make it work. *Cell* **167**, 1455–1467 (2016).
63. Mita, P. et al. LINE-1 protein localization and functional dynamics during the cell cycle. *Elife* **7**, e30058 (2018).
64. Flasch, D. A. et al. Genome-wide de novo L1 retrotransposition connects endonuclease activity with replication. *Cell* **177**, 837–851 (2019).
65. Sultana, T. et al. The landscape of L1 retrotransposons in the human genome is shaped by pre-insertion sequence biases and post-insertion selection. *Mol. Cell* **74**, 555–570 (2019).
66. Rodriguez-Martin, B. et al. Pan-cancer analysis of whole genomes reveals driver rearrangements promoted by LINE-1 retrotransposition in human tumours. Preprint at *bioRxiv* <https://doi.org/10.1101/179705> (2018).
67. Mita, P. et al. BRCA1 mediated homologous recombination and S phase DNA repair pathways restrict LINE-1 retrotransposition in human cells. *Nat. Struct. Mol. Biol.* <https://doi.org/10.1038/s41594-020-0374-z> (2020).
68. Lecona, E. & Fernandez-Capetillo, O. Targeting A. T. R. in cancer. *Nat. Rev. Cancer* **18**, 586–595 (2018).
69. Chan, E. M. et al. WRN helicase is a synthetic lethal target in microsatellite unstable cancers. *Nature* **568**, 551–556 (2019).
70. Ishizuka, J. J. et al. Loss of ADAR1 in tumours overcomes resistance to immune checkpoint blockade. *Nature* **565**, 43–48 (2019).
71. Fischer, M. Census and evaluation of p53 target genes. *Oncogene* **36**, 3943–3956 (2017).
72. Fischer, M., Quaas, M., Steiner, L. & Engeland, K. The p53–p21–DREAM–CDE/CHR pathway regulates G2/M cell cycle genes. *Nucleic Acids Res.* **44**, 164–174 (2016).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020

Methods

Experimental model and subject details. *Cell lines.* We used Tet-On 3G HEK293 cells (ClonTech), Tet-On HEK293T (from J.D.B.⁴⁰), Tet-On 3G HeLa (ClonTech), HEK293FT (from A.J.H.), hTERT-RPE1^{puromycin} (from A.J.H.⁷³) and hTERT-RPE1^{puromycin}-Cas9 (from A.J.H.⁷³). RPE cells have been authenticated by STR profiling. Cells were grown in DMEM (293, HeLa) or DMEM/F12 with 1.5% sodium bicarbonate (RPE) with 10% tetracycline-free FBS (Takara Bio USA). Cells were cultured at 37 °C, 5% CO₂. Antibiotic selection was performed with puromycin (1 µg ml⁻¹), G418 (400 µg ml⁻¹), or blasticidin (10 µg ml⁻¹). Doxycycline was used at 1 µg ml⁻¹, unless otherwise stated. Cells were tested and were mycoplasma negative.

TP53^{KD} Generation. For shRNA growth experiments, wild-type TP53 RPE-Cas9 cells were transduced with pOT-p53-shRNA-TagRFP (ref. ⁷⁴) or pSicoR-mCh₂ empty, and then were transfected with LINE-1 or eGFP plasmids. To generate monoclonal knockout cells, RPE-Cas9 cells were transduced with pOT-p53-shRNA-TagRFP lentivirus and single red fluorescent protein (RFP⁺) cells were sorted by a FACS Aria into 96-well plates. Monoclonal cell lines were screened for p53 knockdown by western blot in cells treated with 200 ng ml⁻¹ doxorubicin.

Tet-On RPE generation. Wild-type TP53 (TP53^{WT}) or TP53^{KD} cells were transfected with Sleeping Beauty transposase plasmid (pCMV(CAT)T7-SB100) and a donor plasmid containing Tet-inducible codon-optimized LINE-1 (ORFeus) or Luciferase (pDA091, pDA093, pDA094, pDA095) following published guidelines⁷⁵. Cells were selected in G418 for 1 week, then sorted into 96-well plates by fluorescence. Monoclonal lines were screened for luciferase induction with the ONE-Glo assay (Promega) or ORF1p protein induction by western blot.

Method Details. Viability Assessments. Viability was determined by clonogenic growth or CellTiter-Glo assay (Promega). Wild-type RPE cells were assessed by clonogenic growth by transfecting 1×10^5 cells with 2 µg eGFP (pDA083) or 3 µg LINE-1 (pDA077) plasmid to achieve equimolar ratios. Cells were split to 10-cm growth dishes and selected with G418 24 h later. In Tet-On assays, 500 cells were plated and doxycycline was added to activate transgene expression. For MMC sensitivity experiments, cells were treated with 100 pM, 1 nM, 10 nM and 100 nM for 24 h on day 2 after plating. In VE-821 sensitivity, cells were treated with 1 µM drug or DMSO vehicle throughout the duration of the experiment. For assays in CRISPR-knockout cells, knockout-cell pools were generated by infecting TP53^{KD} Tet-On RPE cells with lentivirus encoding either non-targeting control or a gene-targeting guide and selecting with puromycin for 1 week (see Supplementary Table 6 for guide sequences). For all assays, after 10–14 d of LINE-1 or control expression, colonies were washed with PBS and fixed (6% glutaraldehyde, 0.5% crystal violet) for 10 min. Plates were rinsed in water and air-dried, and then imaged on a flatbed scanner. Colonies with >50 cells were counted.

CellTiter-Glo assays (Promega) were performed in HEK293T cells transfected with LINE-1 (pDA007), LINE-1 ORF2 H230A (pDA025), LINE-1 ORF2 D702Y (pDA034), LINE-1 ORF2 H230A/D702Y (pDA027) or empty vector (pDA019). There were 8,000 cells plated per well and treated with doxycycline (0–1000 ng ml⁻¹) for 72 h. CellTiter reagents were then added, and luminescence was measured using a Glomax Multi + Detection System (Promega).

CRISPR knockout screening. We used the Brunello GPP pooled CRISPR knockout library packaged into lentivirus for screening⁷⁶. The library comprises 76,441 guide RNAs targeting 19,114 genes, with 4 sgRNAs per gene. TP53^{WT} RPE cells expressing Cas9 protein (TP53^{WT}-Cas9) were transduced at 100-fold library representation at a multiplicity of infection (MOI) of 0.2, in duplicate. TP53^{KD} RPE cells expressing Cas9 protein (TP53^{KD}-Cas9) with LINE-1 or luciferase transgenes were transduced at 100-fold library representation at an MOI of 0.3, in triplicate. Knockout pools were puromycin-selected for 8 d. TP53^{WT}-Cas9 cells were transfected with LINE-1 (pDA077) or eGFP (pDA083) at 150-fold library representation and assayed for library representation at day 19. TP53^{KD}-Cas9 cells were started at 500-fold library representation and maintained at 200-fold representation during passages through day 27. For TP53^{KD}-Cas9 screens, cells were continuously doxycycline-treated and sampled every 4–5 d. Cells were lysed (50 mM Tris, 50 mM EDTA, 1% SDS, pH 8) and incubated with RNase A and proteinase K, and DNA was extracted by isopropanol precipitation. DNA concentrations were measured by Nanodrop. Library preparation was performed with a one-step PCR by Q5 Hot-start polymerase master mix (cat. no. M0494, New England Biolabs) (98 °C for 30 s; 24 cycles: 98 °C for 5 s, 68 °C for 30 s, 72 °C for 30 s; 72 °C for 2 min; hold at 10 °C). See Supplementary Table 6 for primer sequences. Bar-coded libraries were quantified using the NEB Library Quant Kit and mixed to obtain equal coverage, then sequenced with single-end 75-base reads on an Illumina NextSeq 500.

Samples were demultiplexed, and 20 bp CRISPR sgRNA sequences were aligned to the Brunello reference index using Bowtie⁷⁷, allowing no mismatches. We restricted our analysis to genes with fragments per kilobase of exon model per million reads mapped (FPKM) > 1 in RPE cells⁷⁸. Read count data were analyzed to quantify knockout cell proportions with MAGeCK software v0.5.6 or v0.5.7 (ref. ⁷⁹) with the following key parameters: –norm-method control, –additional-rra-parameters, –permutation 1000–min-percentage-goodsrna 0.6. Gene *P* values from MAGeCK

were converted into *Z* scores and combined by Stouffer's method ($Z_s = \sum_{i=1}^n Z_i / \sqrt{n}$),

where *i* is an individual time point, and *n* is the total number of time points in which a gene was identified. We filtered this list by limiting the number of overlapping 95% confidence intervals among timepoints to fewer than 5. Gene knockouts with differential fitness effects on LINE-1⁻¹ cells as compared with control were analyzed for overrepresentation of GO terms using Webgestalt⁴⁰. Individual GO categories were then analyzed in StringDB⁸¹ to generate network plots. To determine enrichment of genes encoding nuclear proteins, we used a Chi-square test following the null hypothesis that only 35.2% of genes should encode nuclear proteins on the basis of the genetic composition of the Brunello library. Analysis of HUSH complex genes was pursued on the basis of knowledge of the LINE-1 literature, as this complex was not annotated in gene sets at the time of this analysis.

RNA-seq analysis. LINE-1 or luciferase was induced for 3 d with 1 µg ml⁻¹ doxycycline, and RNA was collected with the Quick-RNA Microprep kit (Zymo). Libraries were prepared with the TruSeq stranded mRNA library preparation kit (Illumina). Paired-end 150-bp reads were obtained on an Illumina HiSeq4000. Demultiplexed libraries were aligned to hg38 using STAR v2.4.5. Quantification and differential expression analysis was performed using the HTseq and DESeq2 packages in R. For gene set enrichment analysis, we isolated genes with $|\log_2(\text{fold change})| > 1$ and *P*-adjusted < 1.8×10^{-6} and used GSEA software v2.0 from the Broad Institute against Hallmark, Biocarta, KEGG and Reactome genesets v6.2. We used $\log_2(\text{fold change})$ values to perform a preranked analysis. Direct and indirect target genes are curated from published reports^{71,72}. Cell-cycle phase genes were curated from CycleBase 3.0 (ref. ⁸²).

Western blots. Cells were lysed in RIPA buffer with protease/phosphatase inhibitor (cat. no. 5872, Cell Signaling Technology) or Laemmli Sample Buffer (cat. no. 1610747, Biorad) by sonication. PAGE was carried out with manufacturer-recommended buffers on 4–20% or 7.5% Mini TGX Gels (Biorad), NuPAGE 4–12% BisTris gels or NuPAGE 3–8% Tris-Acetate gels (Thermo). Semi-dry transfers were carried out for Biorad gels or NuPAGE BisTris gels at 2.5 A for 5–15 min using the Trans-Blot-Turbo (Biorad). Wet transfers were carried out for Tris-Acetate gels at 30 V overnight at 4 °C. All blocking was performed with Odyssey Blocking Buffer (Licor). Primary antibodies were incubated with membranes overnight at 4 °C, then infrared-conjugated secondary antibodies (Licor) were added at a 1:10,000 concentration and imaged on a Licor Odyssey Scanner. Quantifications were carried out using Image Studio v4.0. Blots were stripped with Reblot Plus Strong Solution (Millipore Sigma). A list of antibodies used can be found in the Supplementary Methods Key Reagents table.

Cloning. Plasmids used in this study are listed in Supplementary Table 5. Several are available at https://www.addgene.org/Kathleen_Burns/. The mammalian expression vector pCEP4 (Invitrogen) was modified to possess a second- or third-generation Tet-inducible promoter (ClonTech) by Gibson assembly. LINE-1 sequences were inserted into the vector backbone by Gibson assembly with PCR amplicons of endogenous LINE-1 sequence (LINE-1 RP) or ORFeus codon-optimized sequence⁸³. Control pCEP4 vectors encoded either eGFP or lacked expression inserts. LINE-1 point-mutant constructs were also created by amplification and Gibson assembly. For Sleeping-Beauty-integrated LINE-1, ORFeus codon-optimized LINE-1 was cloned into the donor vector pSBtet-RN or pSBtet-GN (ref. ⁸⁴) by Gibson assembly. Briefly, pSBtet-RN or GN was digested with SfiI and DraIII, gel-purified and assembled with PCR-amplified LINE-1 (primers SB-ORFeus-5 and SB-ORFeus-3 in Supplementary Table 6) using the HiFi 2× Assembly Master Mix (New England Biolabs).

Single-gene CRISPR knockout-cell generation. To validate screen hits, 20 bp CRISPR sgRNAs were cloned into the pLentiGuide-Puro vector digested with BstBI restriction enzyme as previously described⁸⁵, and the plasmids were packaged into lentivirus. We selected sgRNAs that were enriched in the screens. See Supplementary Table 6 for sgRNA sequences. Cells were incubated with lentiviral supernatants supplemented with 10 µg ml⁻¹ polybrene for 24 h, then selected with puromycin for 1 week, and used in downstream clonogenic assays and western blots.

Transfection. HEK293 and HeLa cells were transfected with Fugene HD reagent (Promega) following standard protocols. RPE cells were transfected using midi- or maxi-prepped plasmid DNA with Viafect reagent (Promega) at a DNA/Viafect ratio of 1/3.

Lentivirus packaging. HEK293FT cells were transfected with Fugene HD (Promega), following the manufacturer's recommendations. Insert vector was added to packaging plasmids pMD.G and psVAX2 at a ratio of 3/4/1 by mass. Medium was changed after 24 h and 48 h, and viral supernatants were collected and filtered through 0.45-µm filters. For screen libraries, complex lentivirus pools were packaged by a similar method by Applied Biological Materials.

Retrotransposition reporter assay. We used an eGFP reporter assay to measure retrotransposition⁸⁶. We transfected 2×10^5 RPE cells with 2 µg LINE-1 reporter plasmids (MT525, JM11) or 2 µg eGFP plasmid and selected with 1 µg ml⁻¹

puromycin for 12 d. Cells were trypsinized and resuspended in cytometry buffer (HBSS, no phenol red, 1% FBS, 1 mM EDTA) at a concentration of $\sim 1 \times 10^6$ cells per ml, then analyzed on a BD Accuri C6 Flow Cytometer. Singlets were gated on SSC-A/SSC-H and FSC-A/FSC-H, then eGFP thresholds were set such that untransfected cells showed 0.1% eGFP⁺ cells. We normalized the percentage of eGFP⁺ cells in experimental groups to the percentage of eGFP⁺ in eGFP-transfected controls.

Nucleoside reverse transcriptase inhibitor treatments for qRT-PCR. We plated 250,000 Tet-On TP53^{KD} cells expressing luciferase or LINE-1 on T25 flasks with 1 ng ml⁻¹ doxycycline added, which were then treated with 5 μ M zalcitabine (ddC) or 5 μ M didanosine (ddI) for 72 h. Cells were lysed, and RNA was extracted using Quick-RNA MicroPrep kit (Zymo Research).

qRT-PCR. cDNA was generated using the iScript kit (Biorad) following RNA extraction using the Quick-RNA Microprep kit (Zymo). Primers were designed using Primer3 and tested against cDNA to ensure single bands were generated in the PCR. Real-time PCR was performed for 40 cycles (98°C for 15 s, 60°C for 30 s) using SSOAdvanced 2x Master mix (Biorad) on the MyIQ cyclor (Biorad). Fold change expression was determined by the 2^{- $\Delta\Delta$ Ct} method. See Supplementary Table 6 for primer sequences.

Immunofluorescence Imaging. HEK293T cells were transfected with doxycycline-inducible LINE-1 plasmid (pDA055) and stably selected with hygromycin for 2 weeks. We plated 5,000 cells in a black 96-well, glass-bottom plate (Corning, cat. no. 3603), treated with doxycycline (0–5,000 ng ml⁻¹, 24 h), fixed (3% paraformaldehyde, 10 min), permeabilized (0.5% Triton X-100/PBS-glycine, 3 min) and blocked (1% BSA/PBS-glycine, 30 min). Cells were incubated with anti-ORF1p (1:500 dilution, Millipore Sigma) and anti-FLAG (1:500 dilution, Sigma) primary antibodies; the Hoechst 33342 (1:50 dilution, Sigma) DNA marker; and HCS CellMask deep red cytoplasmic stain (1:20,000 dilution, Invitrogen). After brief washing in TBST, cells were incubated with anti-rabbit Alexa Fluor 488 (1:200 dilution, Invitrogen) and anti-mouse Alexa Fluor 568 (1:200 dilution, Invitrogen) secondary antibodies. Imaging was performed with a TE300 epifluorescent microscope (Nikon) with a motorized stage and excitation/emission filters (Prior). Images acquired with a DS-QiMc camera at low magnification (20x Plan Fluor lens; 0.285 μ m per pixel, Nikon) using Nikon Elements software (Nikon). Twenty-five images were acquired per sample in a 5 \times 5 grid (1.88 mm²). Images were analyzed using a custom MATLAB software⁸⁷ to segment single cells using the HCS CellMask stain and nuclei using Hoechst 33342. Accurate cell segmentation was manually verified to create a subset of 100 single cells in which ORF1p and ORF2p signal strengths were measured as the total intensity within each segmented cell for each fluorescence channel.

Nuclear foci quantification. We used either Tet-On TP53^{KD} cells expressing luciferase or LINE-1 or Tet-On 3G HeLa cells transfected with doxycycline-inducible LINE-1 plasmids (pDA007, pDA025, pDA027, pDA033, pDA019), and stably selected with puromycin for 1–2 weeks. Positive controls were treated with either 6 mM hydroxyurea for 4 h or 200 ng ml⁻¹ doxorubicin for 2 h. We plated 100,000 cells on cover slips, and treated them with 1,000 ng ml⁻¹ doxycycline for 72 h. EdU was added for 2 h, and cells were pre-treated with 0.5% Triton X-100 for 5 min, fixed with 3.7% paraformaldehyde for 10 min, then permeabilized with 0.5% NP-40 for 10 min. EdU Click-iT reaction (ThermoFisher) was performed following manufacturer's instructions. Slides were blocked (1% BSA/PBS-Glycine, 30 min) and incubated with polyclonal rabbit FANCD2 (1:1,000, Novus Biologicals), rabbit 53BP1 (1:500, Novus Biologicals) or mouse γ H2A.X (1:1,000, Millipore) for 1 h at room temperature, and then anti-rabbit Alexa Fluor 488 for FANCD2 (1:200, ThermoFisher) and anti-rabbit Alexa Fluor 488 (1:2,000, ThermoFisher) and anti-mouse Alexa Fluor 555 (1:2,000, ThermoFisher) for 53BP1 and γ H2A.X, respectively. Slides were imaged at low magnification with the same equipment as described above with key methodological differences. Randomly selected nuclei (>200 per sample) were imaged at high magnification. Foci were quantified using a previously published method in MATLAB (ref. ⁸⁸). We categorized cells as S phase (EdU⁺) or G1/G2 phase (EdU⁻), and excluded cells with low DNA content (dying cells). We compared foci counts using unpaired two-sided *t*-tests.

Transposon insertion sequencing and PCR validations. Tissues for transposon insertion sequencing (TIP-seq) were acquired as flash-frozen de-identified surgical specimens. Small sections of each frozen tissue sample were isolated and TIP-seq was performed as previously described^{18,23,89}. Briefly, 10 μ g of DNA was digested with *AseI*, *BspHI*, *BstYI*, *HindIII*, *NcoI*, or *PstI* (New England Biolabs). Vectors matching the sticky ends were ligated, and touchdown PCR was run with an L1PA1-specific primer (5'-AGATATACCTAATGCTAGATGACACA-3') and ExTaq HS polymerase (Takara Bio). We combined six PCR reactions for each sample, and purified the DNA for sequencing library preparation, shearing amplicons to an average size of 300 bp. We then performed end-repair, dA-tailing and index-specific adapter ligation steps according to Illumina's TruSeq DNA Sample Prep v4 kit protocol (Illumina). Using 2% Size-Select E-gels (Life Technologies), we size-selected our adapter-ligated DNA at approximately 450 bp

before performing a final PCR amplification. After purifying the PCR-amplified libraries, we submitted them for quality control and Illumina HiSeq4000 150-bp paired-end sequencing at the NYU Genome Technology Center. Insertions were called using TIPseqHunterV2 (ref. ²³) after alignments to reference genome hg19. We validated insertions by designing PCR primers with Primer3 and amplifying the insertions. We performed genotyping PCR reactions using 1 ng input DNA of both flash-frozen surgical specimens and DNA obtained from FFPE tissue using the QIAamp DNA FFPE Tissue Kit (Qiagen).

Quantification and statistical analysis. In CRISPR knockout screens and RNA-seq analyses, statistical testing was included in the software packages (MAGECK, DESeq2, WebGestalt, GSEA, StringDB). For all other analyses, appropriate statistical tests were performed using R, which is indicated in figure legends. Tests were typically unpaired and included both one- and two-sided *t*-tests or ANOVA, depending on the a priori hypothesis.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

MAGECK-normalized sgRNA read counts from CRISPR knockout screens and RNA-seq counts and differential expression values have been deposited in the GEO database under accession number GSE119999. Source data for Figs. 2b, 5c,e,f and 6d,e are available online. Requests for resources and reagents should be directed to and will be fulfilled by K.H.B.. Select plasmids created in the Burns Lab can be accessed at Addgene (https://www.addgene.org/Kathleen_Burns/).

References

- Lambrus, B. G. et al. A USP28–53BP1–p53–p21 signaling axis arrests growth after centrosome loss or prolonged mitosis. *J. Cell Biol.* **214**, 143–153 (2016).
- Lambrus, B. G. et al. p53 protects against genome instability following centriole duplication failure. *J. Cell Biol.* **210**, 63–77 (2015).
- Grabundzija, I. et al. Comparative analysis of transposable element vector systems in human cells. *Mol. Ther.* **18**, 1200–1209 (2010).
- Doench, J. G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR–Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
- Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
- Hart, T. et al. High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* **163**, 1515–1526 (2015).
- Li, W. et al. MAGECK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).
- Wang, J., Vasaike, S., Shi, Z., Greer, M. & Zhang, B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* **45**, W130–W137 (2017).
- Szklarczyk, D. et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).
- Santos, A., Wernersson, R. & Jensen, L. J. Cyclebase 3.0: a multi-organism database on cell-cycle regulation and phenotypes. *Nucleic Acids Res.* **43**, D1140–D1144 (2015).
- An, W. et al. Characterization of a synthetic human LINE-1 retrotransposon ORF1p-Hs. *Mob. DNA* **2**, 2 (2011).
- Kowarz, E., Loscher, D. & Marschalek, R. Optimized Sleeping Beauty transposons rapidly generate stable transgenic cell lines. *Biotechnol. J.* **10**, 647–653 (2015).
- Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
- Ostertag, E. M., Prak, E. T., DeBerardinis, R. J., Moran, J. V. & Kazazian, H. H. Jr. Determination of L1 retrotransposition kinetics in cultured cells. *Nucleic Acids Res.* **28**, 1418–1423 (2000).
- Wu, P. H. et al. Evolution of cellular morpho-phenotypes in cancer metastasis. *Sci. Rep.* **5**, 18437 (2015).
- Wu, P. H., Hung, S. H., Ren, T., Shih, I. M. & Tseng, Y. Cell cycle-dependent alteration in NAC1 nuclear body dynamics and morphology. *Phys. Biol.* **8**, 015005 (2011).
- Steranka, J. P. et al. Transposon insertion profiling by sequencing (TIPseq) for mapping LINE-1 insertions in the human genome. *Mob. DNA* **10**, 8 (2019).

Acknowledgements

Human Brunello CRISPR knockout pooled library was a gift from D. Root and J. Doench (Addgene no. 73178). pSBtet-RN and pSBtet-GN were gifts from E. Kowarz (Addgene plasmid no. 60501 and 60503). pCMV(CAT)T7-SB100 was a gift from Z. Izsvak (Addgene plasmid no. 34879). JM111 was a gift from H. Kazazian. pSicoR-mCh_{empty} was a gift from M. Ramalho-Santos (Addgene no. 219070). LentiGuide-Puro was a gift from F. Zhang (Addgene no. 52963). We thank J. Gucwa at the Sidney Kimmel Flow

Cytometry Core and the staff of the NYU Genome Technology center. We thank J. S. Bader for his statistical expertise. We thank B. A. Bari, R. M. Hughes, B. Vogelstein, J. V. Moran and H. H. Kazazian for helpful discussion and review of the manuscript. We thank J. Fairman of the Department of Art as Applied to Medicine at Johns Hopkins University School of Medicine for illustrations. This study was funded by F30CA221175 (D.A.), P50GM107632 (K.H.B., J.D.B., D.F.), U54CA210173 (P.W.) and the Sol Goldman Pancreatic Research Center (K.H.B., R.H.H.).

Author contributions

D.A. and K.H.B. conceptualized this work and wrote the manuscript. Experiments were performed by D.A., J.P.S., C.L., P.-H.W., J.S.S., M.G. and Z.L. All data were primarily analyzed by D.A. Key resources were provided by A.J.H. (RPE cells), A.S. (FANCI mAb), and M.S.T. and V.D. (deidentified colorectal cancer samples). All authors participated in manuscript revisions. Funding was acquired by D.A., J.D.B. and K.H.B.

Competing interests

The authors declare no competing interests.

Additional information

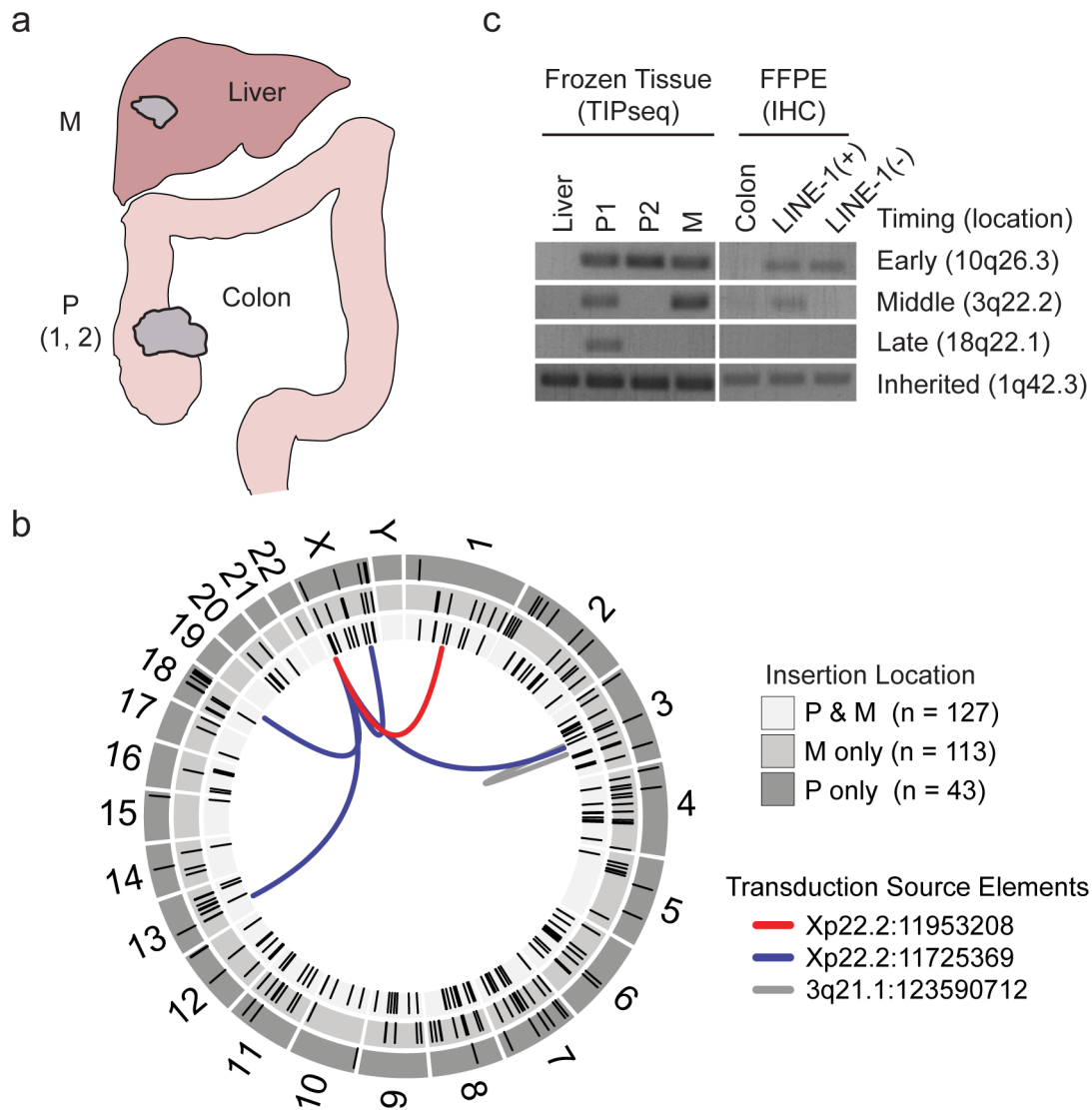
Extended data is available for this paper at <https://doi.org/10.1038/s41594-020-0372-1>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41594-020-0372-1>.

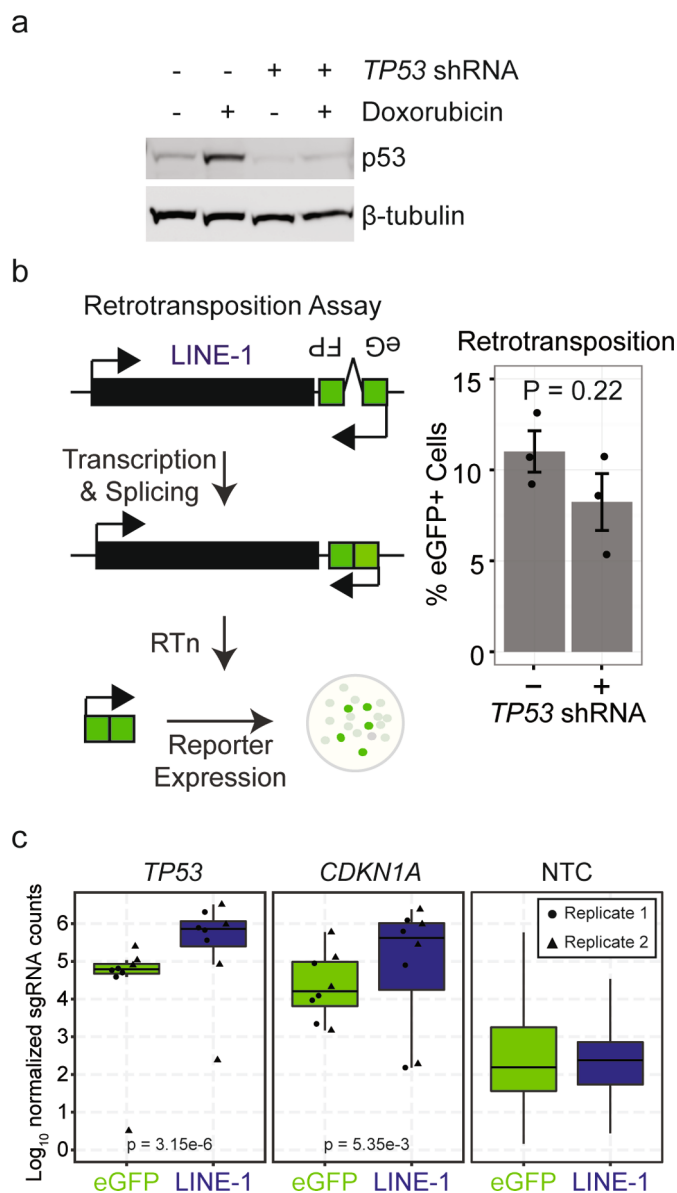
Correspondence and requests for materials should be addressed to D.A. or K.H.B.

Peer review information Beth Moorefield was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

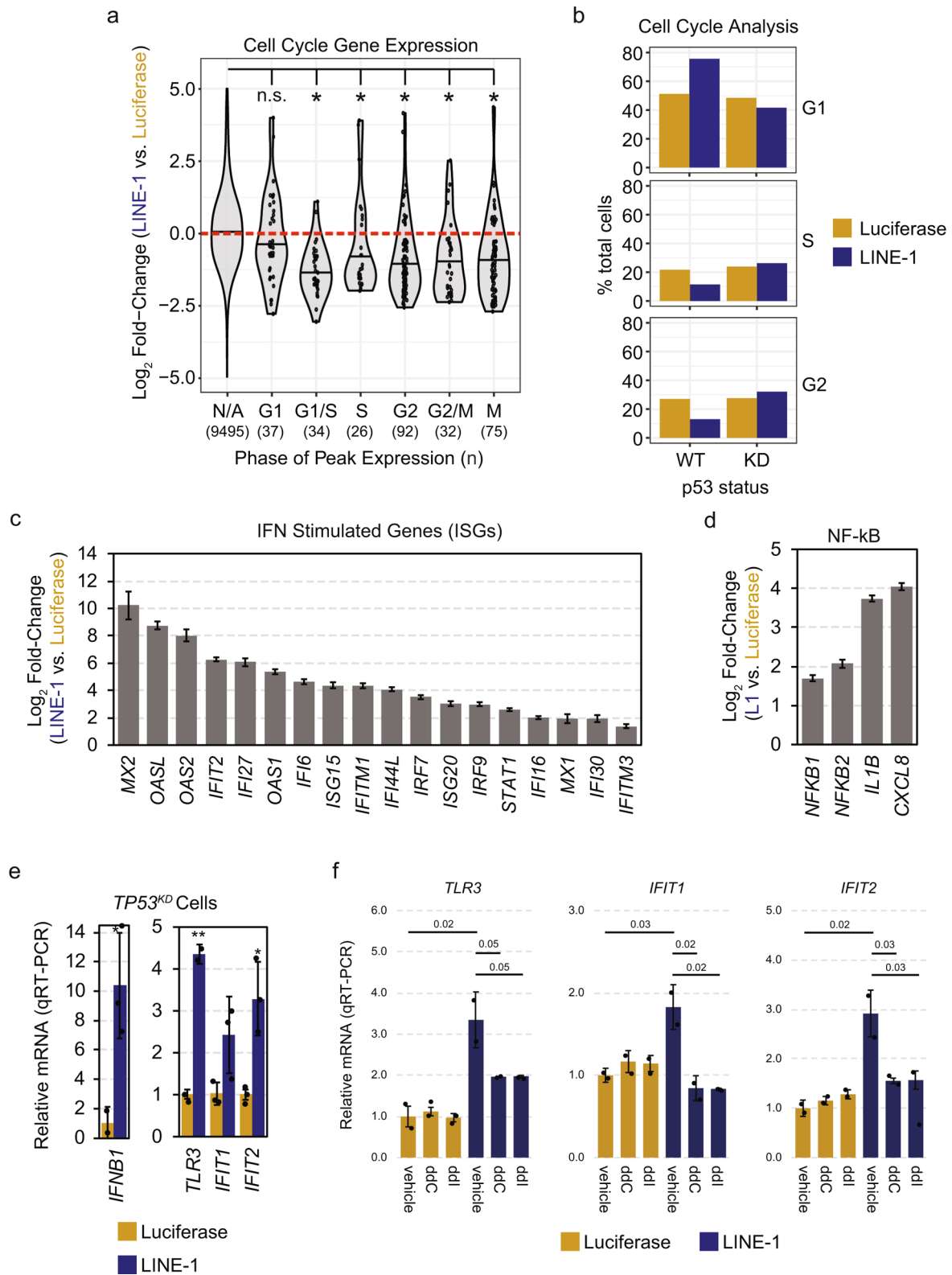
Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 1 | LINE-1 heterogeneity in colon cancer. (a) Tissues collected for transposon insertion profiling by sequencing (TIP-seq) mapping of tumor-specific LINE insertions. Fresh frozen tissue was collected from two sites in the primary tumor (P1, P2) in the colon and one site in the metastatic tumor (M) in the liver. Normal tissue was collected from the liver. The liver metastasis exhibited ORF1p immunoreactivity as well (data not shown). (b) Circos plot detailing TIP-seq results and whether insertions were found in the primary (P only), metastasis (M only) or in both (P & M). In the validation process, we identified 11 3' transduction events, 6 of which mapped to two LINE-1 sequences on Xp22.2 and one on 3q21.1 that are known to be highly active tumor alleles. As expected, the majority of this tumor's *de novo* insertions were intronic or intergenic and not near known tumor suppressors or oncogenes. (c) We genotyped the insertions using hemi-specific PCR in genomic DNA obtained from dissected histology slides and compared to the allele's presence in bulk frozen tissue used for TIP-seq. In all samples, we detected an inherited LINE-1 on 1q42.3, indicating that our PCR conditions were sufficient to genotype LINE-1 alleles. An early *de novo* insertion on 10q26.3 was found in all frozen tissue samples (primary and metastasis) and both CDX2^{high} and CDX2^{dim} slide-dissected samples. An insertion on 3q22.2 is present in the primary tumor subclonally and in the metastasis and therefore occurred before metastasis but after dedifferentiation of the CDX2^{dim} clone. An insertion on 18q22.1 occurred late, after metastasis to the liver had occurred, since it was found in the primary CDX2^{high} clone and not in the metastasis.

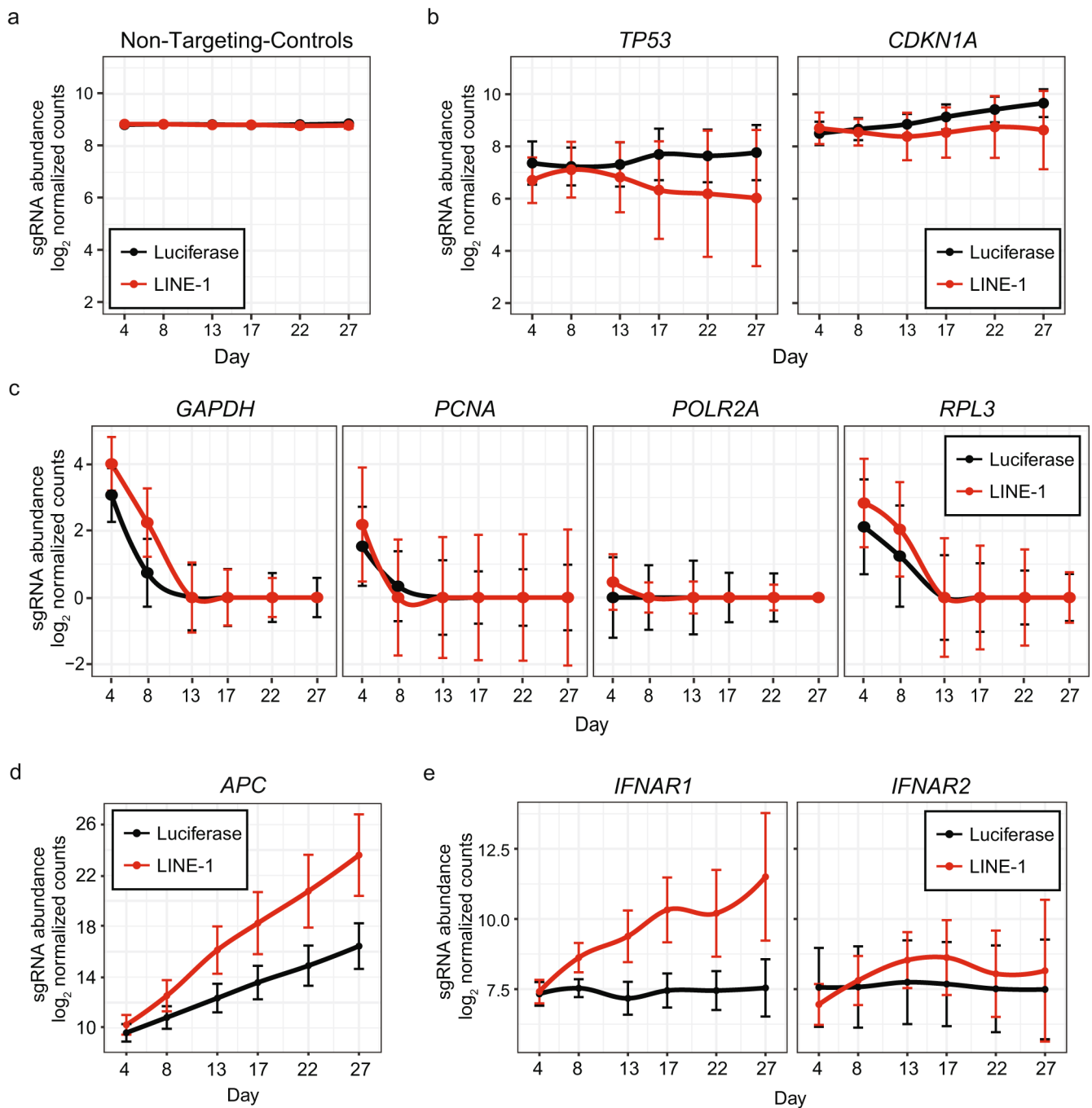


Extended Data Fig. 2 | *TP53* effects on LINE-1⁺ cell growth and retrotransposition. (a) Demonstration of effective *TP53* knockdown. RPE cells were treated with *TP53* shRNA lentivirus (pDA079) or control lentivirus (pDA081). The Western blot shows the p53 response to treatment with the DNA intercalator doxorubicin (200 ng ml⁻¹ for 24 h). (b) Left, the retrotransposition reporter assay. LINE-1 is expressed from a plasmid with an antisense eGFP in the 3'UTR that is interrupted by a sense intron. During transcription, the intron is spliced, reconstituting the coding potential of the eGFP reporter. The eGFP reporter carries with it a CMV promoter and is inserted into the genome by LINE-1. Expression of eGFP from the genome allows for fluorescence-based quantification of retrotransposition rate by flow cytometry. Right, reporter assay performed in RPE with *TP53* knockdown or control \pm s.e.m., $n = 3$ independent experiments. P value was calculated by two-sided *t*-test. (c) Normalized median read counts of sgRNAs targeting *TP53* and *CDKN1A* in cells expressing either LINE-1 (navy blue) or eGFP (green) control compared to non-targeting-controls (NTC). Individual sgRNAs are indicated by circles or triangles. Results from two biological replicates are depicted.

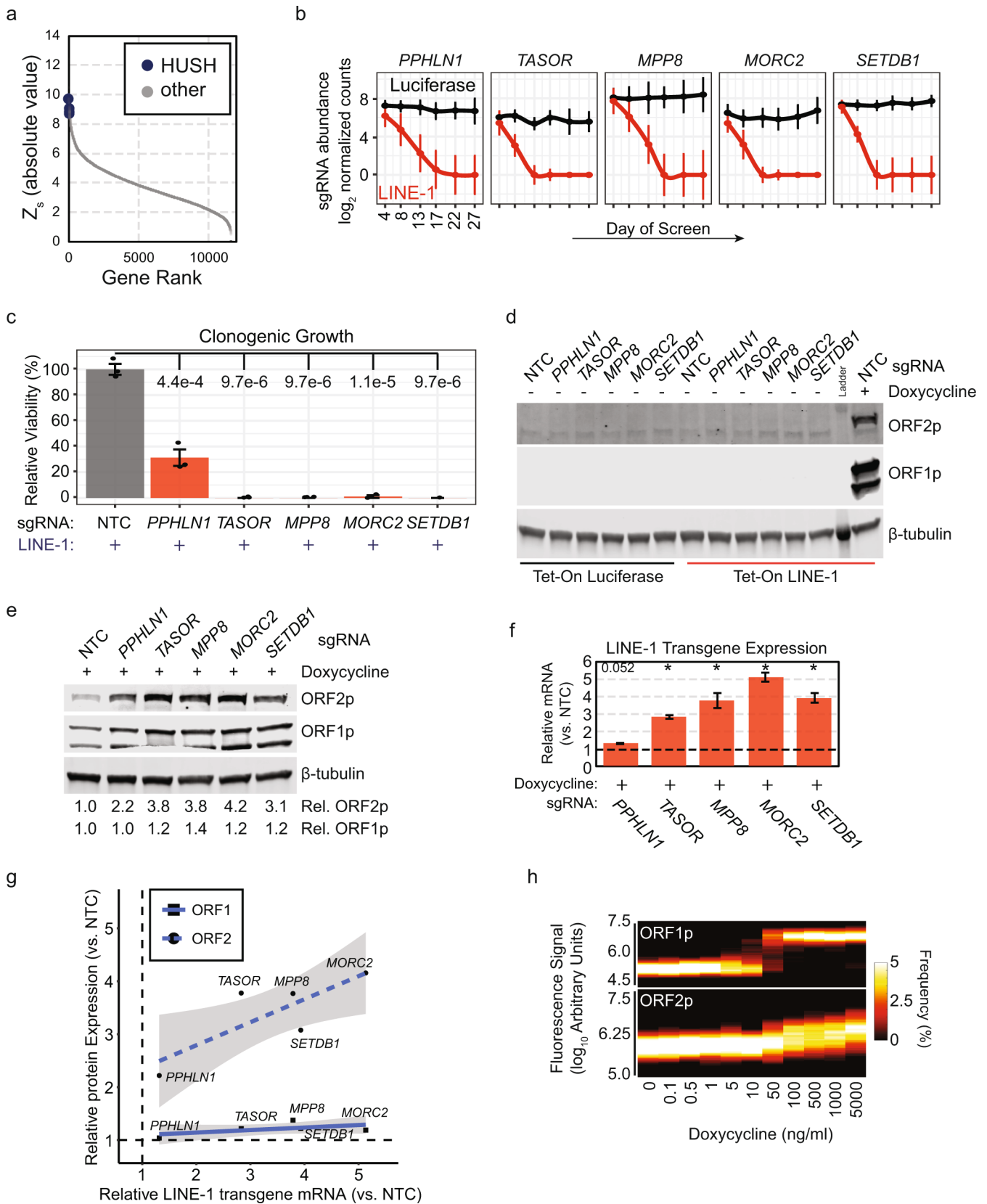


Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | LINE-1 RNAseq analysis. (a) Genes regulated by cell cycle were curated from CycleBase v3.0⁷² and differential expression values were plotted. S, G2, and M phase genes were significantly downregulated in LINE-1⁺ cells. Unpaired two-sided *t*-tests were used for statistical testing. N/A = not applicable. **p*-values vs. N/A: G1 = not significant (n.s.), G1/S = 1.7e-9, S = 1.5e-2, G2 = 2.1e-13, G2/M = 5.2e-6, M = 3.4e-10. **(b)** Flow cytometry was used to assess cell cycle by quantifying DNA content using a PI DNA stain in Tet-On LINE-1 or Tet-On luciferase cells induced with 1 μg ml⁻¹ doxycycline for 48 h. LINE-1⁺ cells with wild-type (WT) p53 accumulated in G1 phase (2n DNA copy number), whereas *TP53*^{KD} resulted in more even cell cycle proportions. These data are from one experiment. **(c)** Relative fold-change of interferon-stimulated genes in LINE-1 compared to luciferase-expressing cells measured by RNAseq. Error bars indicate s.e.m. **(d)** RNAseq analysis revealed upregulation of NF-κB and several target genes in LINE-1⁺ cells. Error bars indicate s.e.m. **(e)** Differential expression of *IFNB1* (right) and interferon-stimulated genes (left) in p53-knockdown cells expressing LINE-1 or luciferase for 72 h. Measured by qRT-PCR. Error bars indicate s.d., n = 3 biological replicates. * *p* < 0.05, ** *p* < 0.001. **(f)** Differential expression of *TLR3*, *IFIT1*, and *IFIT2* with the addition of 5 μM zalcitabine (ddC) or 5 μM didanosine (ddI) in p53-knockdown cells expressing LINE-1 or luciferase for 72 h. Measured by qRT-PCR, n = 3 independent experiments. *P* values indicated within the plots.

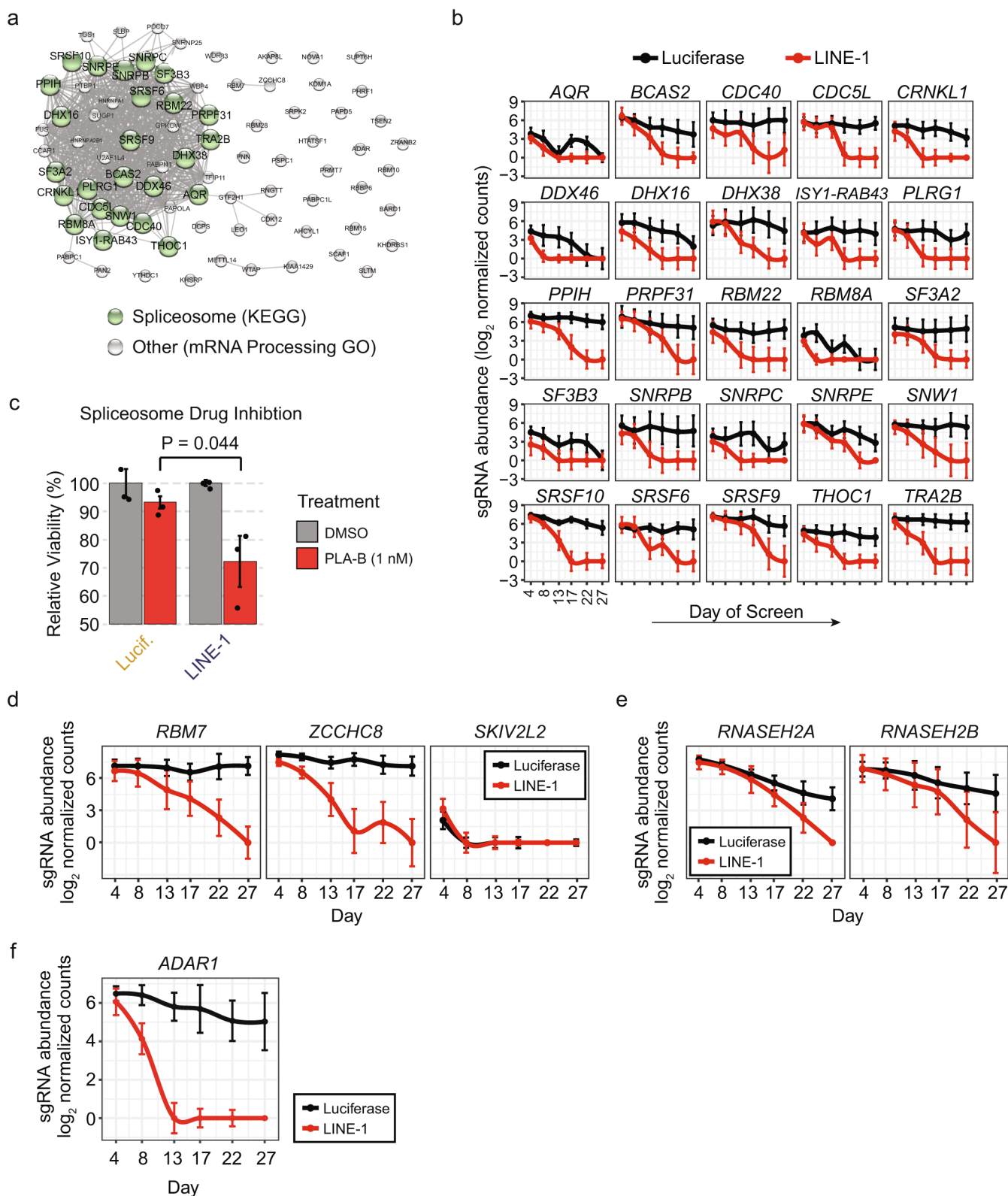


Extended Data Fig. 4 | *TP53*-Knockdown Screen Supplement. (a) Behavior of non-targeting-control sgRNAs in the screen over time. Data points indicate the median sgRNA count per replicate and error bars the 95% confidence interval. (b) Behavior of *TP53*- and *CDKN1A*-targeting sgRNAs. Median values are depicted with 95% Confidence Intervals. There is no appreciable change in *TP53* sgRNA representation between LINE-1⁺ and luciferase control cells, indicating loss of p53 function due to the shRNA. *CDKN1A* sgRNAs do not differ between groups as well, suggesting that *CDKN1A* effects are contingent on p53 function. (c) Examples of essential gene knockouts that deplete from both LINE-1⁺ and luciferase⁺ cells. Median values are depicted with 95% Confidence Intervals. (d) Knockout of *APC* provides a growth advantage to LINE-1⁺ cells. Median values are depicted with 95% Confidence Intervals. (e) Knockout of the interferon alpha and beta receptor subunit 1 (*IFNAR1*) but not subunit 2 (*IFNAR2*) provides a growth advantage in LINE-1⁺ cells. Median values are depicted with 95% Confidence Intervals.

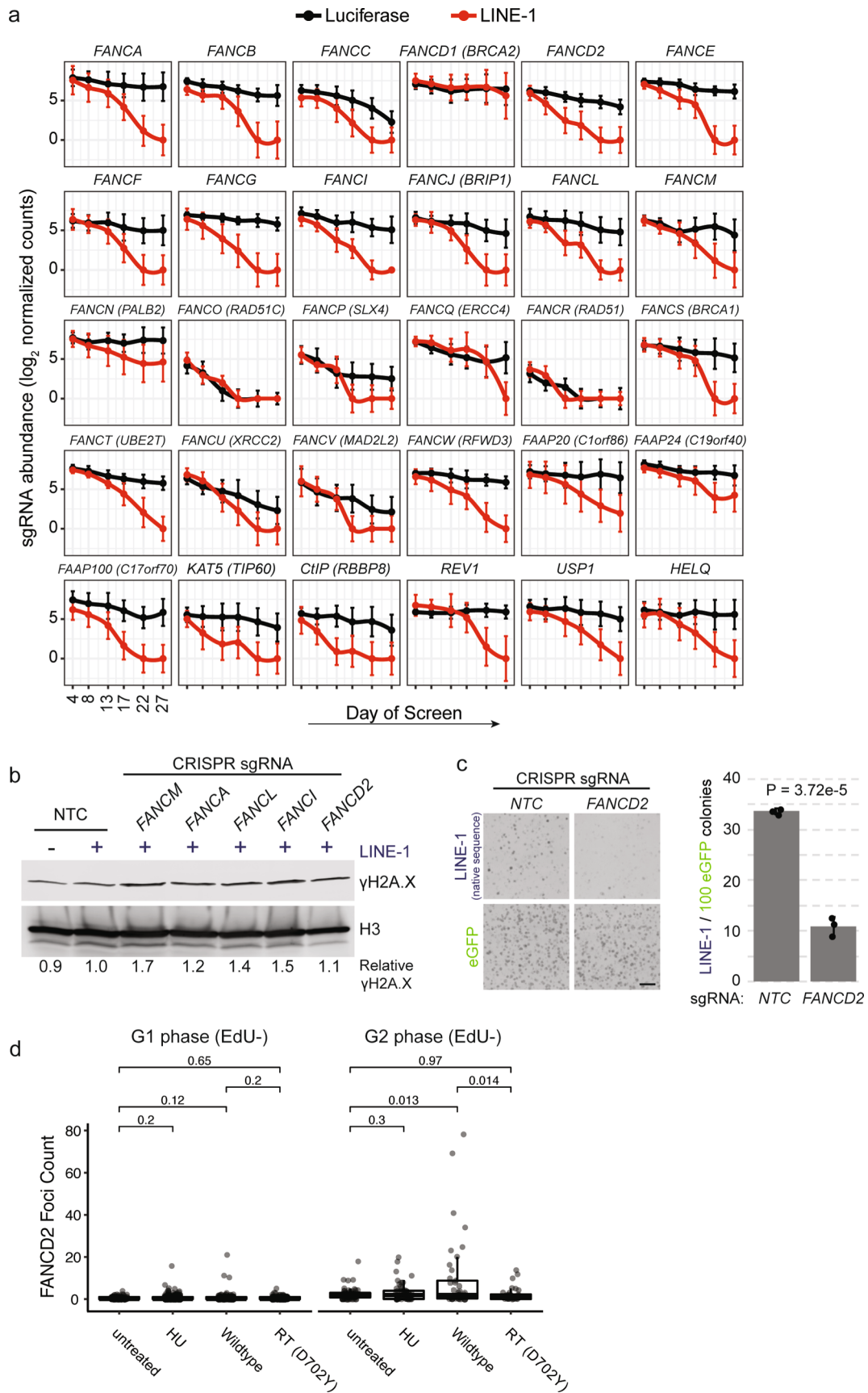


Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | HUSH knockout is synthetic lethal due to derepression of the LINE-1 transgene. (a) Gene screen ranks by Z_s scores. HUSH genes are in blue. **(b)** HUSH complex sgRNA performance during the screen. All knockouts drop out early from LINE-1⁺ cells (red) and do not affect growth of luciferase⁺ cells (black). Median values are depicted with 95% Confidence Intervals. **(c)** 12 d clonogenic growth assay in cells expressing LINE-1 (doxycycline-induced) with targeted knockouts of HUSH components compared to non-targeting-control (NTC). $n = 3$ independent experiments. Error bars indicate \pm s.e.m. P values calculated by one-sided t-test. **(d)** Western blot comparing ORF1p and ORF2p expression in HUSH knockout cells or non-target-controls (NTC) that have not been treated with doxycycline compared to NTC with 24 h of $1 \mu\text{g ml}^{-1}$ doxycycline treatment. ORF1p and ORF2p expression are only detected in NTC-treated cells with doxycycline added to the culture media. The double banding pattern for ORF1p is consistently seen with codon-optimized LINE-1. **(e)** Western blot comparing ORF1p and ORF2p expression 24 h after $1 \mu\text{g ml}^{-1}$ doxycycline treatment in HUSH knockouts compared to NTC. The ORF2p antibody cannot distinguish between endogenous or transgenic LINE-1 expression. **(f)** qRT-PCR analysis of LINE-1 transgene expression in HUSH knockouts compared to NTC (induced with $1 \mu\text{g ml}^{-1}$ doxycycline). Because the LINE-1 transgene is codon-optimized, qRT-PCR is specific for the transgene and does not amplify endogenous LINE-1 sequences. $*p < 0.001$. **(g)** Linear regression plot of LINE-1 transgene expression and ORF1p and ORF2p expression in HUSH knockouts compared to NTC. Shaded area indicates 95% confidence interval for regression line. Both ORF1p and ORF2p increase in expression with higher transgene mRNA expression, although the increase in ORF1p is minimal compared to that observed with ORF2p. **(h)** Heatmap of immunofluorescence imaging depicting the proportion of cells expressing ORF1p and ORF2p at different levels in HEK293T cells expressing Tet-On LINE-1 (pDA055) at increasing doses of doxycycline.

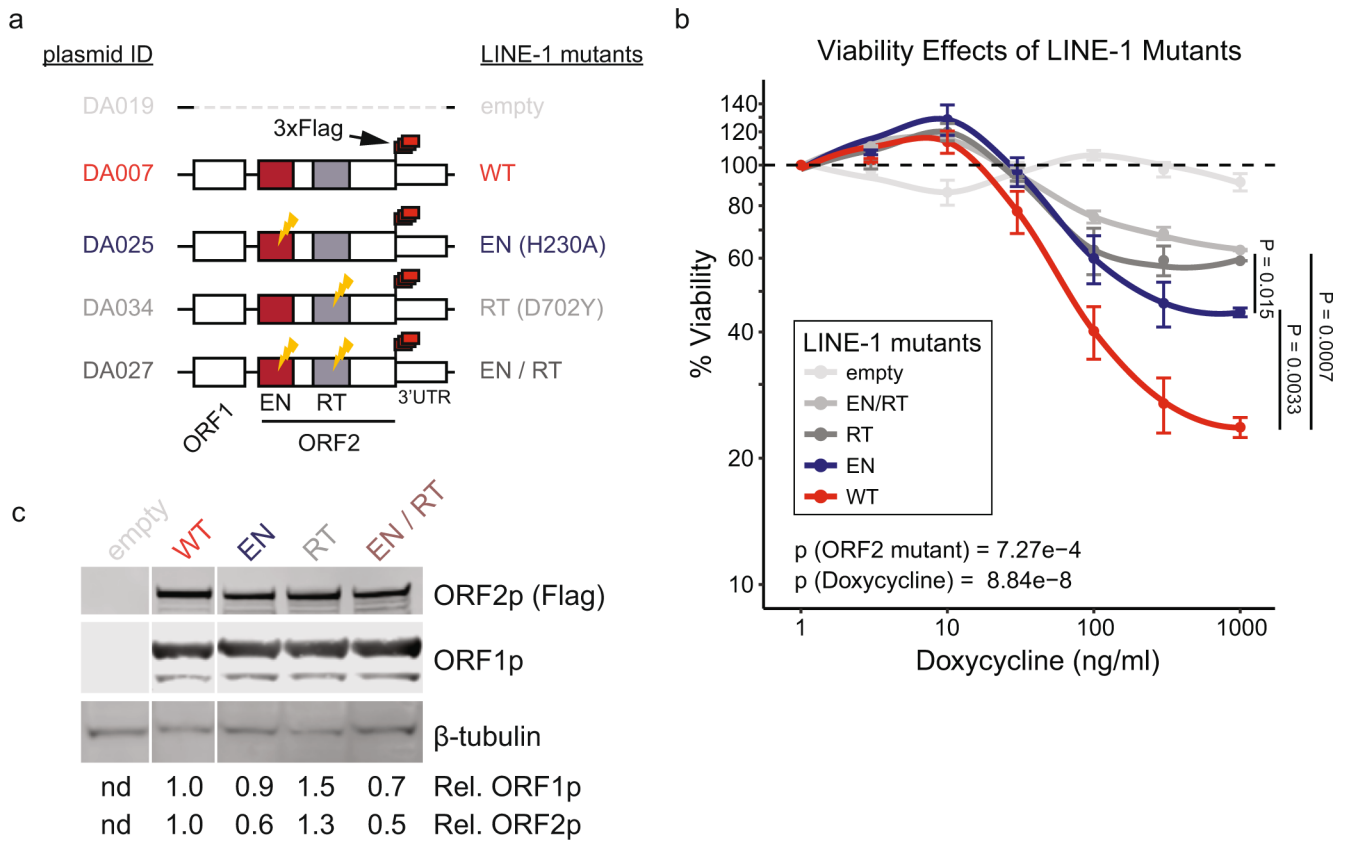


Extended Data Fig. 6 | RNA processing gene knockouts sensitize cells to LINE-1. (a) StringDB network plot of the 81 mRNA processing genes identified by this screen. Edges indicate known protein-protein interactions. This network is enriched for spliceosome machinery (green nodes). (b) Screen behavior of significant genes belonging to the spliceosome KEGG GO term. Median sgRNA counts are depicted with 95% Confidence Intervals. (c) Clonogenic assay (12 d) comparing growth of luciferase⁺ and LINE-1⁺ cells (induced with 1 μg ml⁻¹ doxycycline) treated with 1 nM pladienolide B (PLA-B) or vehicle (DMSO). n = 3 independent experiments. Error bars indicate s.e.m. P value calculated by unpaired one-sided t-test. (d) Behavior of nuclear exosome complex genes in the screen. Median values are depicted with 95% Confidence Intervals. (e) Behavior of RNASEH2 component sgRNAs in the screen. Median values are depicted with 95% Confidence Intervals. (f) Behavior of ADAR1 sgRNAs in the screen. Median values are depicted with 95% Confidence Intervals.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | The Fanconi Anemia Pathway is required for growth of LINE-1⁺ cells. **(a)** Behavior of sgRNAs targeting Fanconi Anemia pathway genes in the screen. Median values are depicted with 95% Confidence Intervals. **(b)** Western blot of DNA damage marker γ H2A.X in chromatin-bound protein fractions of LINE-1⁺ cells with or without perturbations to the FA pathway. H3 was used as loading control. γ H2A.X levels were quantified and graphed relative to NTC-treated, LINE-1⁺ cells. **(c)** Clonogenic assay (10 d). *TP53*^{KD} cells constitutively expressing Cas9 are treated with lentivirus encoding non-targeting-control (NTC) or *FANCD2* sgRNA and then transfected with eGFP (pDA083) or the native LINE-1 sequence L1RP (pDA077). Left, representative images of colonies. Scale bar = 1 cm. Right, data are presented as the rate of LINE-1 per 100 eGFP colonies \pm s.d. to control for transfection efficiency across samples, $n = 3$ independent experiments. P value obtained by unpaired two-sided *t*-test. **(d)** Quantification of FANCD2 foci in G1 and G2 phase (EdU-) HeLa cells. Number of cells per group: G1 untreated ($n = 104$), G1 HU ($n = 352$), G1 wildtype LINE-1 ($n = 186$), G1 RT (D702Y) ($n = 138$), G2 untreated ($n = 60$), G2 HU ($n = 58$), G2 wildtype LINE-1 ($n = 42$), G2 RT (D702Y) ($n = 32$). Two-sided *t*-tests were used for statistical comparisons. HU = hydroxyurea. RT = reverse transcriptase. ns = not significant.



Extended Data Fig. 8 | Viability assays with LINE-1 mutants. (a) Tet-On constructs for wild-type and mutant LINE-1 expression. **(b)** Viability of HEK293T cells after 4 days expressing wild-type or a mutant at increasing doxycycline doses. A multivariate ANOVA (Viability ~ ORF2 * doxycycline) was performed in R to calculate p values for ORF2 mutant status and doxycycline dose. Tests of viability differences among ORF2 mutants were further performed using two-sided *t*-tests at the 1000 ng ml⁻¹ doxycycline dose. N = 6 replicates per doxycycline dose. **(c)** Western blot of ORF1p and ORF2p 24 hours after inducing protein expression with 1000 ng ml⁻¹ doxycycline.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted <i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

| | |
|-----------------|--|
| Data collection | No software was used to collect data. |
| Data analysis | Software to analyze RNAseq data include STAR v2.4.5, R (HTseq and DESeq2 packages), and GSEA software v2.0 from the Broad Institute. Screen data were analyzed using MAGeCK, StringDB, and Webgestalt. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

MAGeCK-normalized sgRNA read counts from CRISPR KO screens and RNAseq counts and differential expression values are included in the GEO database under accession number GSE119999. Source data for 2, 5, and 6 are provided.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|-----------------|--|
| Sample size | For screens, sample size was determined based on the design as a positive selection or synthetic lethal screen. Higher library representation (500x starting compared to 100x for positive selection) and increased KO library replicate (triplicate as opposed to duplicate) numbers were chosen for synthetic lethal screens, as is described in the literature describing factors to consider in powering of genome-wide knockout screens. For all other assays, experiments were performed in duplicate, triplicate, or with higher replicate numbers, which is indicated throughout the text. |
| Data exclusions | Data were not excluded. |
| Replication | To follow up on screen hits, we used a variety of assays, including generation of individual (as opposed to pooled) CRISPR knockout cell lines and we confirmed knockout using functional assays (i.e. MMC response of FANCD2-Ub in Fanconi Anemia knockout cells). |
| Randomization | During screen growth periods, plates were held in a 37C 5% CO2 incubator and cells were sampled every 4-5 days. It took approximately 8 hours to perform the required cell culture procedures on such days. Plates were split up randomly into 4 batches and thus rotated throughout the incubator over the course of the entire experiment. Otherwise, randomization was not an issue in the experiments described in this manuscript. |
| Blinding | The most relevant assay wherein blinding was necessary was in quantifying clonogenic growth. To address this, the investigator who set up the experiment and fixed and stained the colonies (D.A.) printed out de-identified images of the plates and asked a blinded investigator (J.P.S., C.L.) to count the colony numbers, in order to verify that the counts were reflective of the true data and not the investigators' bias. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| n/a | Involved in the study |
|-------------------------------------|---|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Antibodies |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |

Methods

| n/a | Involved in the study |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Antibodies

| | |
|-----------------|---|
| Antibodies used | <p>anti-mouse Alexa Fluor 568, ThermoFisher, cat# A-11004 anti-rabbit Alexa Fluor 488, ThermoFisher, cat# A-11034 IRDye 680RD Goat anti-mouse IgG, LI-COR, cat# 925-68070 IRDye 680RD Goat anti-rabbit IgG, LI-COR, cat# 925-68071 IRDye 800CW goat anti-rabbit IgG, LI-COR, cat# 925-32211 IRDye 800CW goat anti-mouse IgG, LI-COR, cat# 925-32210 mouse anti-flag, Sigma, F1804 mouse anti-human FANCD2 Clone F117, Santa Cruz Mouse anti-human ORF1p Clone 4H1, Millipore Sigma, MABC1152 mouse anti-human p53 Clone DO-1, CalBiochem, cat# OP43 mouse anti-human yH2A.X, Clone JBW301, EMD Millipore, cat# 05-636 mouse anti-RPA2, Clone 9H8, Abcam, Cat# ab2175 rabbit anti p-RPA S4/S8, Bethyl, cat # A300-245A rabbit anti-H3, abcam cat# 1791 rabbit anti-human b-tubulin clone 9F3, Cell Signalling Technology, cat# 2128 rabbit anti-human FANCD2, Novus Bio, cat# NB100-182</p> |
|-----------------|---|

rabbit anti-human FANCI clone 589, from Agata Smogorzewska lab
 rabbit anti-human ORF2p clone MT49, from Burns Lab
 rabbit anti 53BP1, Novus Biologicals, cat# NB100-904

Validation

All commercially available antibodies have been previously validated. The FANCI clone 589 antibody has been validated by the Smogorzewska lab and detects a band at the proper size that is responsive to positive control MMC. The ORF2p MT49 antibody produced by the Burns lab has been extensively validated by multiple conventional assays and is described in more detail here: [biorxiv.org/content/10.1101/744425v1](https://www.biorxiv.org/content/10.1101/744425v1).

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

Tet-On 3G HEK293 Cells (ClonTech)
 Tet-On HEK293T LD cells (JD Boeke Lab)
 Tet-On 3G HeLa (ClonTech)
 HEK293FT (AJ Holland Lab)
 hTERT-RPE-1, puromycin sensitive [RPE] (AJ Holland lab)
 RPE-Cas9 (AJ Holland lab)
 RPE with p53 knockdown - produced in this study
 RPE-Cas9 with p53 knockdown - produced in this study
 RPE-Cas9, p53 knockdown, Tet-On LINE-1 (ORFeus) - produced in this study
 RPE-Cas9, p53 knockdown, Tet-On Luciferase - produced in this study
 RPE, Tet-On LINE-1 (ORFeus) - produced in this study
 RPE, Tet-On Luciferase - produced in this study

Authentication

RPE cells obtained from AJ Holland underwent STR profiling and were confirmed as RPE. The remaining cell lines were either purchased commercially or not tested.

Mycoplasma contamination

Cells were periodically tested with the MycoAlert Lonza assay and confirmed negative.

Commonly misidentified lines
 (See [ICLAC](#) register)

Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

The flow data presented do not include traditional flow plots. The data were collected in one dimension to assess retrotransposition efficiency using a field-specific reporter assay described in extended data figure 2. Single cells are first selected based on FSC/SSC parameters. Then cells are analyzed for GFP (FITC). First, a GFP-negative control is gated such that 0.1% are considered GFP+. Next, a GFP-positive control is analyzed to ensure that >90% of the cells fall into the GFP+ gate. Finally, the experimental groups are analyzed to assess GFP-content. We also assessed cell cycle stage based on DNA content.

Instrument

BD Accuri C6

Software

BD Accuri C6 software

Cell population abundance

Retrotransposition events occur in up to 20% of cells, making these an abundant target.

Gating strategy

n/a

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

