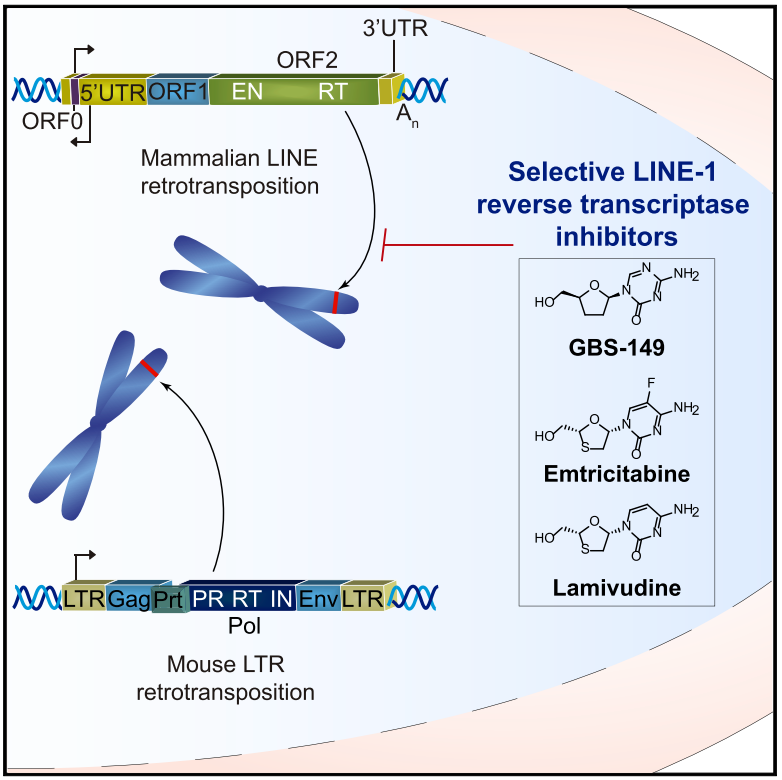
Article

Cell Chemical Biology

Synthesis and Characterization of Specific Reverse Transcriptase Inhibitors for Mammalian LINE-1 Retrotransposons

Graphical Abstract



Highlights

- Characterization of 33 compounds has identified specific inhibitors of LINE-1s
- GBS-149, FTC, and 3TC are non-toxic and mammalian-specific LINE-1 inhibitors
- GBS-149 is a new and potent RTi of active human LINE-1s

Authors

Guillermo Banuelos-Sanchez, Laura Sanchez, Maria Benitez-Guijarro, ..., Francisco Franco-Montalban, Juan A. Tamayo, Jose L. Garcia-Perez

Correspondence

jtamayo@ugr.es (J.A.T.), jose.garcia-perez@igmm.ed.ac.uk (J.L.G.-P.)

In Brief

The mobilization of active LINE-1 retrotransposons continues to impact the human genome. LINE-1s move using a copy-and-paste mechanism that relies on reverse transcriptase activity. The identification of compounds that selectively inhibit the reverse transcriptase activity of mammalian LINE-1s will allow studying the impact of their mobilization.

Banuelos-Sanchez et al., 2019, Cell Chemical Biology *26*, 1095–1109
August 15, 2019
https://doi.org/10.1016/j.chembiol.2019.04.010

CellPress

Cell Chemical Biology
Article

Synthesis and Characterization of Specific Reverse Transcriptase Inhibitors for Mammalian LINE-1 Retrotransposons

Guillermo Banuelos-Sanchez,[1] Laura Sanchez,[2] Maria Benitez-Guijarro,[2] Valentin Sanchez-Carnerero,[2] Carmen Salvador-Palomeque,[2,5] Pablo Tristan-Ramos,[2,3] Meriem Benkaddour-Boumzaouad,[2] Santiago Morell,[2,6] Jose L. Garcia-Puche,[2] Sara R. Heras,[2,3] Francisco Franco-Montalban,[1] Juan A. Tamayo,[1,*] and Jose L. Garcia-Perez[2,4,7,*]

[1]Department of Medicinal and Organic Chemistry, Faculty of Pharmacy, University of Granada, Campus de Cartuja, s/n 18071 Granada, Spain

[2]GENYO. Centro de Genómica e Investigación Oncológica: Pfizer-Universidad de Granada-Junta de Andalucía, Avenida de la Ilustración 114, PTS Granada, 18016 Granada, Spain

[3]Department of Biochemistry and Molecular Biology II, Faculty of Pharmacy, University of Granada, Campus Universitario de Cartuja, 18071 Granada, Spain

[4]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine (IGMM), University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, UK

[5]Present address: Mater Research Institute - University of Queensland, TRI Building, Woolloongabba, Australia

[6]Present address: Department of Genetics, University of Cambridge, UK

[7]Lead Contact

*Correspondence: jtamayo@ugr.es (J.A.T.), jose.garcia-perez@igmm.ed.ac.uk (J.L.G.-P.)

https://doi.org/10.1016/j.chembiol.2019.04.010

SUMMARY

Retrotransposons are a type of transposable element (TE) that have amplified to astonishing numbers in mammalian genomes, comprising more than a third of the human and mouse genomes. Long interspersed element class 1 (LINE-1 or L1) retrotransposons are abundant and currently active retroelements in the human and mouse genomes. Similarly, long terminal repeat (LTR)-containing retrotransposons are abundant in both genomes, although only active in mice. LTR- and LINE-1-retroelements use different mechanisms for retrotransposition, although both involve the reverse transcription of an intermediate retroelement-derived RNA. Retrotransposon activity continues to effect the germline and somatic genomes, generating interindividual variability over evolution and potentially influencing cancer and brain physiology, respectively. However, relatively little is known about the functional consequences of retrotransposition. In this study, we have synthesized and characterized reverse transcriptase inhibitors specific for mammalian LINE-1 retrotransposons, which might help deciphering the functional impact of retrotransposition *in vivo*.

INTRODUCTION

Most vertebrate genomes contain a large fraction of transposable element (TE)-derived sequences (Ivancevic et al., 2018), and in most mammalian genomes retrotransposons have successfully been amplified to high numbers over evolution (Ivancevic et al., 2018; Lander et al., 2001; Waterston et al., 2002). Indeed, at least a third of the human and mouse genomes are made of retrotransposons, which include long terminal repeat (LTR)-containing retrotransposons and non-LTR retrotransposons (Lander et al., 2001; Richardson et al., 2015; Waterston et al., 2002). Retrotransposons are a type of TE that move (i.e., retrotranspose) using a replicative copy-and-paste mechanism that involves reverse transcription of an intermediate retroelement-derived RNA (Boeke and Chapman, 1991; Richardson et al., 2015). LTR retrotransposons or endogenous retroviruses (ERVs) are similar to retroviruses in their structure (resemble the proviral integrated form) and mode of reverse transcription, and comprise 8% and 10% of the human and mouse genome, respectively (Lander et al., 2001; Richardson et al., 2015; Waterston et al., 2002). Because of the lack of a functional envelope gene, ERVs are considered obligated intracellular retrotransposons (Mager and Stoye, 2015). However, evolutionary analyses have demonstrated that germline reinfection, rather than retrotransposition, is the main insertion pathway exploited by ERVs, which further suggest that a small pool of ERVs might retain infection activity in the germline (Belshaw et al., 2004; Kim et al., 1994). In humans, all ERVs (i.e., HERVs) have been inactivated by the accumulation of mutations over evolution, and although some HERVs from the K subfamily (HERV-K) are polymorphic in the human population (Hughes and Coffin, 2004), they are thought to be currently immobile in the human genome (Richardson et al., 2015). However, several ERVs remain active in the mouse genome, including intracisternal A-particle (IAP), endogenous type D murine (MusD), and early transposon retroelements (reviewed in Mager and Stoye, 2015). Indeed, active mouse LTR retrotransposons are often drivers of mutations in mice (Mager and Stoye, 2015), and autonomous IAP and MusD elements have been characterized in the mouse genome

(Dewannieux et al., 2004; Ribet et al., 2004). On the other hand, ~17% and ~19% of the human and mouse genome is made of long interspersed element class 1 (LINE-1 or L1) non-LTR retrotransposons (Lander et al., 2001; Waterston et al., 2002), and active copies have been characterized in both genomes (DeBerardinis et al., 1998; Goodier et al., 2001; Kazazian et al., 1988; Naas et al., 1998; Sassaman et al., 1997). Although only one LINE-1 subfamily is currently active in the human genome (L1Hs elements), at least three active subfamilies of LINE-1 elements currently coexist in the mouse genome (termed L1Md-A, L1Md-G_F, and L1Md-T_F) (reviewed in Mager and Stoye, 2015; Richardson et al., 2015). Despite their genomic abundance, and similar to ERVs, most LINE-1 copies are no longer active because of mutation accumulation over evolution and/or because of 5′ truncation during LINE-1 insertion (Beck et al., 2010; Brouha et al., 2003; Grimaldi et al., 1984; Lander et al., 2001). However, ~100 and up to ~3,000 LINE-1 copies retain retrotransposition capability in the human and mouse genome, respectively (Beck et al., 2010; Brouha et al., 2003; DeBerardinis et al., 1998; Goodier et al., 2001; Naas et al., 1998). Active LINE-1s are also responsible for the mobilization of non-autonomous active short interspersed elements (SINEs, also classified as non-LTR-retrotransposons), such as Alu and SVA in humans, and B1, B2, and B4 in rodents (Dewannieux et al., 2003; Dewannieux and Heidmann, 2005; Lander et al., 2001; Waterston et al., 2002). SINE mobilization has been very prolific over evolution, and millions of insertions have accumulated in the human and mouse genomes, comprising >11% and >5% of their genomes, respectively (Lander et al., 2001; Mager and Stoye, 2015; Richardson et al., 2015; Waterston et al., 2002).

The ongoing activity of LTR and non-LTR retrotransposons continues to impact the mammalian genome, and in mice and humans both types of elements have generated sporadic disease-causing germline mutations (Mager and Stoye, 2015; Richardson et al., 2015). Furthermore, it is becoming increasingly evident that LINE-1 retrotransposons might participate in human interferonopathies, such as Aicardi-Goutieres syndrome (AGS) (Benitez-Guijarro et al., 2018; Stetson et al., 2008; Thomas et al., 2017). Unexpectedly, active LINE-1s can also affect selected somatic tissues, including cancer and brain cells (reviewed in Burns, 2017; Garcia-Perez et al., 2016; Richardson et al., 2015; Singer et al., 2010; Thomas et al., 2012). Indeed, it has now been very well documented that most epithelial-derived human tumors are characterized for accumulating new LINE-1 insertions (reviewed in Burns, 2017; Scott and Devine, 2017). Data on mouse is limited, although mouse LINE-1 retrotransposition has been reported in a mouse model of hepatic carcinogenesis (Schauer et al., 2018), and IAP ERVs can generate new insertions in mouse radiation-induced myeloid leukemia tumors (Takabatake et al., 2008). In the brain, it has been demonstrated that human and mouse LINE-1s can retrotranspose in neuronal progenitor cells and mature non-dividing neurons *in vitro* and *in vivo* (Coufal et al., 2009; Erwin et al., 2016; Evrony et al., 2012; Macia et al., 2017; Muotri et al., 2005; Upton et al., 2015). Although there is currently no evidence of mouse ERV activity in the brain, a recent report demonstrated that *gypsy* LTR retrotransposons could retrotranspose in the *Drosophila* brain (Li et al., 2013); thus, additional studies are required to determine if ERVs are expressed and active in the mouse brain, or in other somatic cells. In summary, we know relatively little about the functional impact of retrotransposition in heritable disorders, mendelian interferonopathies, cancer, brain physiology, or even brain disorders, and additional research is needed to reveal the potential contribution of retrotransposition, if any, to these biological processes.

Inhibiting ongoing retrotransposition represents a simple and effective manner to infer the role that the mobilization of a given retroelement might manifest in any biological process, using a loss-of-function approach. Although during retrotransposition the mechanism of LTR and non-LTR reverse transcription is markedly different, both rely on an encoded reverse transcriptase (RT) activity. Previous studies have demonstrated that certain RT inhibitors (RTi) used to treat HIV/AIDS with a nucleoside structure can also inhibit retrotransposon mobilization in a dose-dependent manner (Contreras-Galindo et al., 2017; Dai et al., 2011; Jones et al., 2008). In fact, cellular phosphorylated nucleoside analogs are known to inhibit different types of retroviral RTs through competitive inhibition with the native dNTP for RT binding and chain termination of elongating cDNAs; thus, their inhibitory efficiency is influenced by their cellular uptake, efficiency of phosphorylation by cellular kinases, efficiency of the analog to bind the RT, and incorporation efficiency into elongating cDNAs among other factors (reviewed in Arts and Wainberg, 1996). The simplicity of their use and their wide activity spectrum has resulted in many studies using RTis to analyze the role of retrotransposition on mouse oocyte attrition (Malki et al., 2014), on human tumor growth in mouse xenografts (Sciamanna et al., 2005), on cellular aging (De Cecco et al., 2019; Simon et al., 2019), or even in mouse memory formation (Bachiller et al., 2017). However, conflicting results have been found when using RTis to ameliorate the phenotype of AGS mice models, and further research with more specific RTis would help to determine the therapeutic potential of RTis to treat AGS (Achleitner et al., 2017; Beck-Engeser et al., 2011). Recently, RTis have been used in a pilot clinical study involving a small number of AGS patients, revealing a significant amelioration of their interferon score (Rice et al., 2018). However, these studies had a number of significant biases, as relatively nothing is known about the specificity of RTis on mammalian retrotransposons, together with potential secondary effects of RTis such as intrinsic anti-inflammatory activity (Fowler et al., 2014), and toxicity effects due to off-target effects on cellular DNA polymerases (Arts and Wainberg, 1996), etc.

To better understand the specificity of nucleoside analog RTis, and aiming to identify a specific and potent inhibitor of active human LINE-1 retrotransposons, in this study we have analyzed the specificity and inhibitory strength of 33 nucleoside analogs (commercially available and synthesized *de novo*) on active human and mouse LINE-1s and on active mouse ERVs. Notably, we have identified three non-toxic and selective LINE-1 RTis that exhibit no activity against mouse LTR retrotransposons. Furthermore, an RTi synthesized and characterized in this study, GBS-149, is on average 18-fold more active at inhibiting human LINE-1s than mouse LINE-1 retrotransposons. In summary, we have generated robust and novel tools to study the impact of LINE-1 retrotransposition in arguably any cellular context.

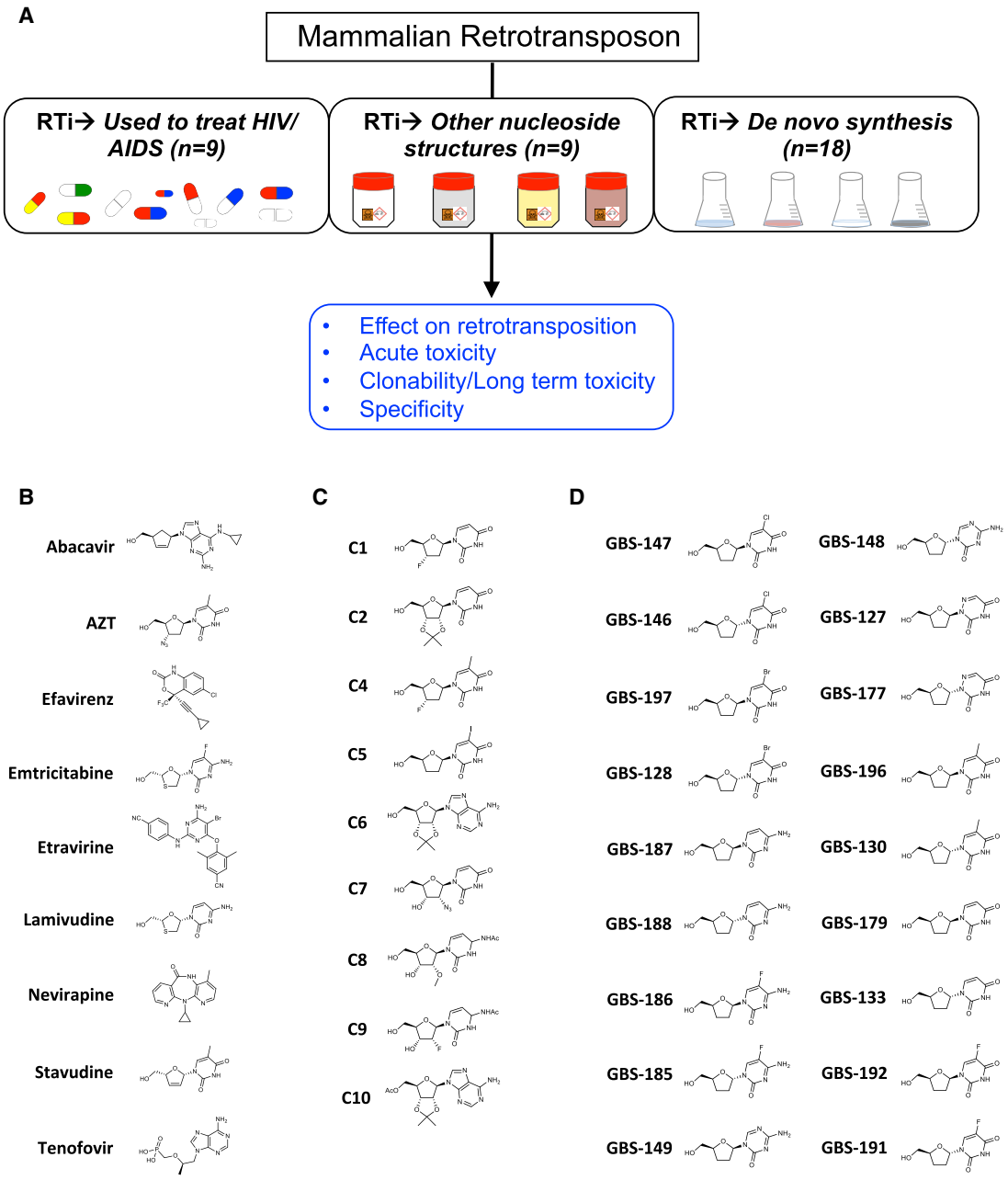


Figure 1. Rationale of the Study and Structures of Drug Tested

(A) Scheme of the rationale of the study. Further details are provided in the main text.

(B–D) Chemical structure of drugs tested in this study. (B) The chemical structure of drugs used to treat AIDS that were tested in this study. (C and D) The chemical structure of commercially available and synthesized nucleoside analogs tested in this study, respectively (see Table 1).

RESULTS

Rationale of the Study and Structures of Nucleoside Analogs Tested

To identify a specific and potent RTi for human LINE-1 retrotransposons, we analyzed the effect that 33 nucleoside analog structures exert on the retrotransposition rate of human and mouse LINE-1s and on mouse ERVs, together with short-term/acute and long-term/clonability toxicity controls (Figure 1A). As controls, we also tested the effect of three non-nucleoside analog HIV RTis (efavirenz, etravirine, and nevirapine, Figure 1B). To test the effect of these drugs on retrotransposons, we took advantage of previously established cell-based retrotransposition assays for LINE-1s and ERVs (Figure S1) (Boeke et al., 1985; Freeman et al., 1994; Heidmann and Heidmann, 1991; Moran et al., 1996). As retrotransposition occurs through an RNA intermediate, LINE-1 and ERV retrotransposition assays are based on the same principle. In brief, the presence of an engineered intron within a

Table 1. HIV/AIDS RTis and Commercially Available Nucleoside Analogues Used in This Study

Name	CAS	Provider	Solvent
Abacavir	136470-78-5	Hospital	Water
AZT	30516-87-1	Hospital	Water
Efavirenz	154598-52-4	Hospital	DMSO
Emtricitabine	143491-57-0	Hospital	Water
Etravirine	269055-15-4	Hospital	DMSO
Lamivudine	134678-17-4	Hospital	Water
Nevirapine	129618-40-2	Hospital	DMSO
Stavudine	3056-17-5	Hospital	Water
Tenofovir disoproxil	147127-20-6	Hospital	Water
2′,3′-Dideoxy-3′-fluorouridine (C1)	41107-56-6	Aldrich	Water
2',3'-Isopropylideneuridine (C2)	362-43-6	Aldrich	Water
2',3′-Dideoxy-3′-fluorothymidine (C4)	25526-93-6	Aldrich	Water
2′,3′-Dideoxy-5-iodouridine (C5)	105784-83-6	Aldrich	Water
2',3'-O-isopropylideneadenosine (C6)	362-75-4	Aldrich	Water
2'-Azido-2'-deoxyuridine (C7)	26929-65-7	Aldrich	Water
N4-Acetyl-2'-O-methylcytidine (C8)	113886-71-8	Carbosynth	Water
N4-Acetyl-2'-deoxy-2'-fluorocytidine (C9)	159414-97-8	Carbosynth	Water
5'-O-Acetyl-2',3'-O-isopropylideneadenosine (C10)	15888-38-7	Carbosynth	DMSO

reporter gene ensures that reporter gene expression can only occur after a round of retrotransposition (Figures S1A–S1C). The effect of each nucleoside analog was tested at two concentrations, 5 and 25 μM, using HeLa cells because LINE-1 and ERV retrotransposition occur at high frequencies in this cervix cancer-derived cell line (Moran et al., 1996; Ribet et al., 2004). As a reporter gene, we used the neomycin phosphotransferase gene (NEO), which activates resistance to G418 upon retrotransposition (Figures S1A–S1C). Notably, the NEO-based retrotransposition assay is quantitative, and the resulting number of drug-resistant foci provides a readout of retrotransposition activity. To control for long-term toxicity/clonability, HeLa cells were transfected with a plasmid expressing the NEO gene (pU6ineo, Figure S1D), while short-term/acute toxicity was tested using an MTT colorimetric assay (Figure 1A). Figures 1B–1D show the structure of each nucleoside analog tested, including drugs used to treat HIV/AIDS (Figure 1B: abacabir, AZT, emtricitabine [FTC]; lamivudine [3TC], stavudine [d4T], and tenofovir; the non-nucleoside analogs efavirenz, etravirine, and nevirapine were used as internal controls), commercially available nucleoside analogs (Figure 1C: C1, C2, C4, C5, C6, C7, C8, C9, and C10, see the STAR Methods) and the 18 nucleoside analogs that were synthesized and purified in this study (Figure 1D: GBS-147, 146, 197, 128, 187, 188, 186, 185, 149, 148, 127, 177, 196, 130, 179, 133, 192, and 191).

Effect of Drugs Used to Treat AIDS on Mammalian Retrotransposons

Here, we tested the effect that RTis used to treat AIDS exert on the retrotransposition rate of an active human LINE-1 (L1.3 [Sassaman et al., 1997]) tagged with the NEO-based retrotransposition indicator cassette (Figures 2A and 2B). These assays revealed that several of the drugs used to treat AIDS are potent inhibitors of human LINE-1 retrotransposition (Figure 2A), consistent with two previous reports (Dai et al., 2011; Jones et al., 2008). However, MTT control assays revealed that etravirine, and to a lesser extent AZT and efavirenz, are mildly toxic to HeLa cells in a dose-dependent manner (Figure 2E). Consistently, similar data were observed in long-term toxicity/clonability controls, confirming that etravirine, and to a lesser extent AZT and efavirenz, are toxic to HeLa cells in a dose-dependent manner (Figures S2C and S2D). Thus, we used the toxicity values induced by RTis to quantify the overall effect of HIV RTis on human L1.3 retrotransposition. These analyses revealed that abacavir, AZT, emtricitabine, lamivudine, stavudine, and tenofovir inhibit human L1 retrotransposition in a dose-dependent manner (Figures 2A and 2B). Among the RTis tested, emtricitabine and lamivudine were the more potent non-toxic inhibitors, reducing human L1.3 retrotransposition by more than 3- and 6-fold when tested at 25 μM (Figure 2B). Notably, tenofovir reduced retrotransposition of human L1.3 >75-fold, although it showed mild acute toxicity in HeLa cells (Figure 2E). As expected, no effect on human L1 retrotransposition was detected in cells treated with non-nucleoside analog RTis, with the exception of nevirapine, which led to a ~30% reduction in L1 retrotransposition, but in a dose-independent manner (Figure 2B). Consistently, a previous study found that nevirapine led to a minor reduction on human L1 retrotransposition (Dai et al., 2011).

To test the specificity of RTis used to treat HIV/AIDS on human LINE-1 retrotransposons, we next tested their inhibitory effect on the retrotransposition rate of two mouse LTR retrotransposons (MusD, Figures 2C and S2A; IAP, Figures 2D and S2B), also using HeLa cells. Taking into account their short- and long-term toxicity (Figures 2E, S2C, and S2D), our data revealed that AZT and tenofovir are also potent inhibitors of both mouse LTR retrotransposons, while abacavir and nevirapine mildly inhibit the retrotransposition frequency of both mouse ERVs, all in a dose-dependent manner (Figures 2C, 2D, S2A, and S2B). Furthermore, we observed that nevirapine and stavudine lead to statistically significant but modest reductions in MusD retrotransposition levels (~30% reduction at 25 μM, Figures 2D and S2A).

In summary, our data demonstrate that AZT and tenofovir are potent but not selective RTis, inhibiting human LINE-1 and mouse ERV retrotransposition, but inducing short- and long-term toxicity in HeLa cells. However, emtricitabine and lamivudine are non-toxic-specific inhibitors of human LINE-1 retrotransposition, as was stavudine, although the latter is mildly toxic to HeLa cells.

Effect of Commercially Available Nucleoside Analogs on Mammalian Retrotransposons

Next, we tested the inhibitory retrotransposition potential of nine commercially available nucleoside analogs (Figure 1C). MTT assays revealed that C2, C8, and C9 were mildly toxic to HeLa cells (Figure S3A), while long-term/clonability assays revealed that C4

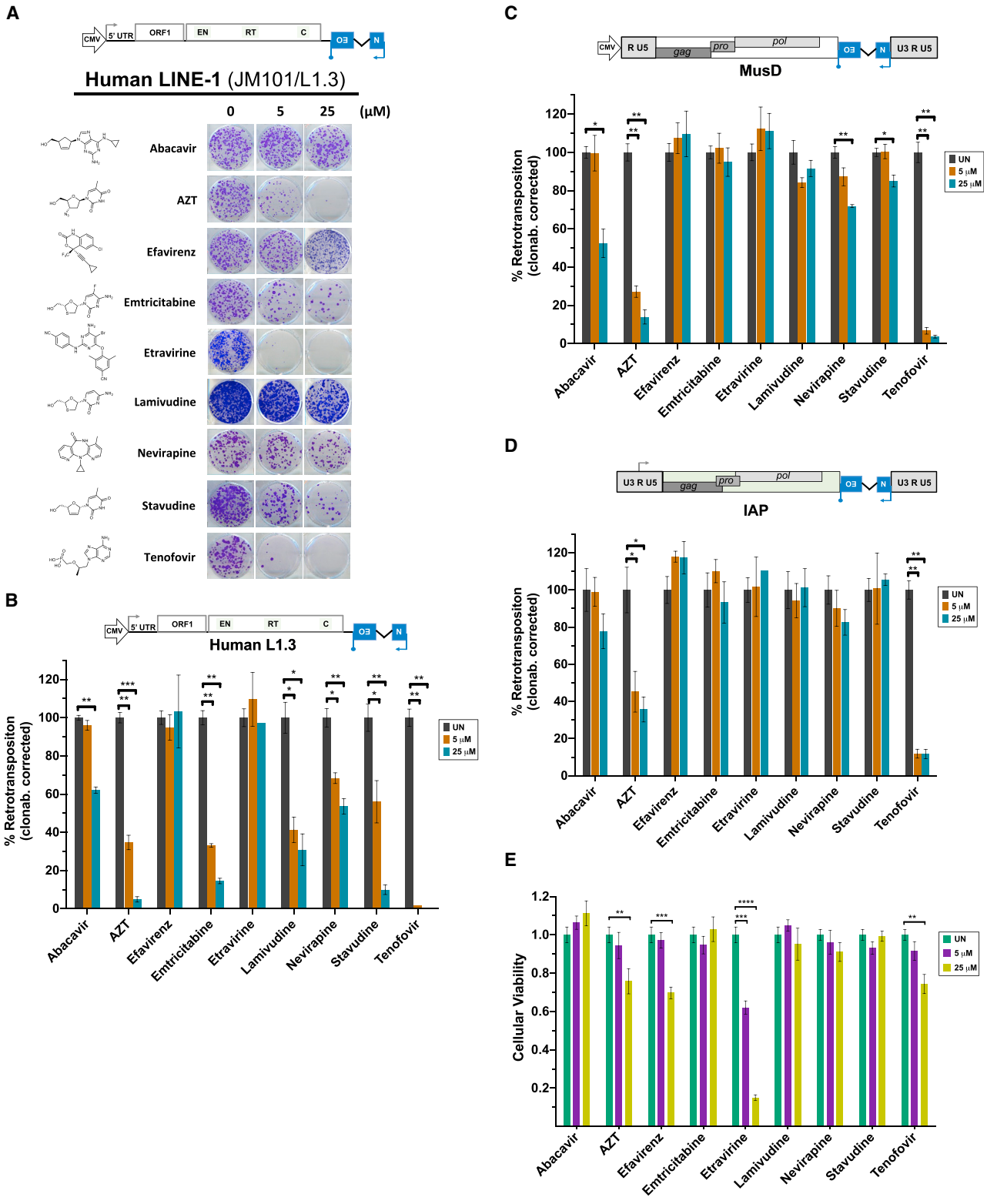


Figure 2. Effect of Drugs Used to Treat AIDS on Mammalian Retrotransposons

(A and B) Effect on a human LINE-1 retrotransposon (L1.3). (A) Representative results of LINE-1 retrotransposition assays conducted in the presence of the indicated HIV RT inhibitor (tested at 5 and 25 μM) using HeLa cells (2×10^4 cell/well). (B) The quantification of the above assays in triplicate (including SD; black bars, no drug; orange bars, 5 μM; green bars, 25 μM). Retrotransposition values of untreated cells were assigned 100% for comparisons. A cartoon of the L1 construct used (plasmid JM101/L1.3) is shown at the top of (A and B). *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

(legend continued on next page)

was very toxic to HeLa cells, as was C1, although to a much lesser degree (Figures S3D and S3E). When retrotransposition assays were quantified and normalized taking into consideration the toxicity of RTis, we observed that C4 inhibited the retrotransposition rate of human L1.3, mouse IAP, and mouse MusD, in a dose-dependent manner (Figures 3, S3B, and S3C). In addition, C9 mildly inhibited the retrotransposition rate of human L1.3 (Figures 3A and 3B).

In summary, our data demonstrate that C4 is a mild and not selective RTi that can inhibit human L1s and mouse LTR retrotransposons in HeLa cells. However, C4 induced significant long-term toxicity to HeLa cells, which could limit its applications.

Effect of Synthesized Nucleoside Analogs on Mammalian Retrotransposons

The data presented above revealed that RTis with a very similar chemical structure could have markedly different effects on the retrotransposition rate of human LINE-1 and mouse ERV retrotransposons. Aiming to increase the strength and specificity of nucleoside analog RTis for human LINE-1 retrotransposons, we next modified the structure of active 2′,3′-dideoxynucleotide RTis by including different substituents. Thus, we next synthesized and purified 18 nucleoside analogs, all pyrimidine analogs (Figure 1D), and we then tested their effect on mammalian retrotransposition rates using HeLa cells. Control MTT assays revealed that several of the synthesized nucleoside analogs induced significant short-term toxicity (Figure S4A), including GBS-177, -196, -179, -133, -192, -191, -187, and -186. However, clonability controls revealed that GBS-133, -192, -191, -187, and -186 also induced significant long-term toxicity (Figure S5B). When the effect of the 18 nucleoside analogs was tested on human LINE-1 retrotransposition, we observed that, while GBS-196, -191, and -186 mildly reduced retrotransposition in a dose-dependent manner, GBS-149 reduced L1.3 retrotransposition by more than 30-fold when tested at 25 μM (Figures 4A and S4B). Notably, GBS-149 had no effect on mouse LTR retrotransposition (neither MusD, Figures 4B and S4C; nor IAP, Figure S5A), while GBS-179 and -192 mildly reduced MusD LTR retrotransposition (Figures 4B and S4C). Interestingly, no significant reduction on IAP retrotransposition was detected with any of the 18 nucleoside analogs tested (Figure S5A).

In summary, our data revealed that GBS-149 is a new potent, non-toxic, and specific human LINE-1 RTi.

GBS-149, Emtricitabine, and Lamivudine Dose-Response Assays

The above experiments demonstrated that, of the 33 nucleoside analogs tested, GBS-149, emtricitabine, and lamivudine are potent and non-toxic inhibitors of human LINE-1 retrotransposition, with no inhibitory effect on mouse LTR retrotransposition when tested at low concentrations (5 and 25 μM, Figures 2 and 4). Thus, we selected these three nucleoside analogs for further experimentation.

To confirm the inhibitory potential of GBS-149, emtricitabine, and lamivudine on active LINE-1 retrotransposons, we conducted retrotransposition assays using: (1) a different retrotransposition indicator reporter gene and a different human LINE-1, (2) lower RTi concentrations, and (3) additional cell lines. First, we evaluated the potential toxicity of GBS-149, emtricitabine, and lamivudine on human PA-1 cells, an ovarian teratocarcinoma-derived pluripotent cell line that naturally overexpresses endogenous LINE-1 ribonucleoprotein particles (L1-RNPs) (Garcia-Perez et al., 2010; Zeuthen et al., 1980). MTT short-term toxicity assays revealed that none of the three RTis tested induced significant toxicity to PA-1 or HeLa cells when tested at low concentrations (5 and 25 μM, Figure S6B [PA-1], 5B [HeLa]). Consistently, none of the three RTis tested induced significant toxicity to PA-1 cells when tested at higher concentrations (50 and 150 μM, Figure S6C). Additional controls revealed that none of the RTis tested induced changes in endogenous L1-ORF1p or p53 expression (Figures 5 and S6A; Data S1; STAR Methods). Similar data were observed in cultured HeLa cells (data not shown). Thus, these data suggest that the three RTis selected, GBS-149, emtricitabine and lamivudine, are not toxic to cultured cells expressing high (PA-1) or low (HeLa) levels of endogenous L1-RNPs.

Next, we conducted retrotransposition assays using the blasticidin S deaminase gene as a retrotransposition indicator reporter gene (BLAST, *mblastI*-tagged constructs, Figure S1E) (Goodier et al., 2007; Morrish et al., 2002). The BLAST retrotransposition assay is similar to the NEO-based assay, but the kinetics of retrotransposition and selection is different due to blasticidin S inducing rapid cell death (Izumi et al., 1991). Thus, we tested the effect of a range of GBS-149, emtricitabine, and lamivudine concentrations on the retrotransposition rate of an active human LINE-1 tagged with the *mblastI* cassette (construct JJ101/L1.3) using HeLa cells. Notably, we confirmed robust inhibition of human LINE-1 retrotransposition with the three RTis tested (Figure 5A), even at low concentrations (2.5 μM). We confirmed these results using an active mouse LINE-1 from the L1Md-T_F class (L1SM; Han and Boeke, 2004) tagged with the same *mblastI* cassette (Goodier et al., 2007; Morrish et al., 2002) (Figures S6D–S6F). Thus, these data confirm that GBS-149, emtricitabine, and lamivudine are potent, non-toxic, and LINE-1-specific RTis, at least on HeLa cells.

We next tested the inhibitory potential of GBS-149, emtricitabine, and lamivudine on LINE-1 retrotransposition using human HEK293T cells. However, in HEK293T cells we used a human LINE-1 retrotransposition assay that activates expression of firefly luciferase upon retrotransposition (*mFlucI*, Figure S1F) (Xie et al., 2011). This assay is not based on antibiotic selection/generation

(C and D) Effect on mouse LTR-retrotransposons. (C) (MusD) and (D) (IAP) show the effect of the indicated HIV RT inhibitor on mouse LTR-retrotransposition using HeLa cells when tested at 5 μM (orange bars) and 25 μM (green bars) (black bars, no drug). In (C) and (D), 2×10^4 and 1×10^4 HeLa cells were plated per well, respectively, and a cartoon of each construct used is also shown. *$p \leq 0.05$, **$p \leq 0.01$.

(E) MTT assays in the presence of HIV inhibitors. The graph shows the results of MTT cell metabolic activity assays conducted on HeLa cells in the presence of 5 μM (purple bars) and 25 μM (yellow bars) of the indicated HIV inhibitor (triplicate). Colorimetric values of untreated cells (green bars) were used to normalize and were arbitrarily designated as 1. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

See also Figure S2.

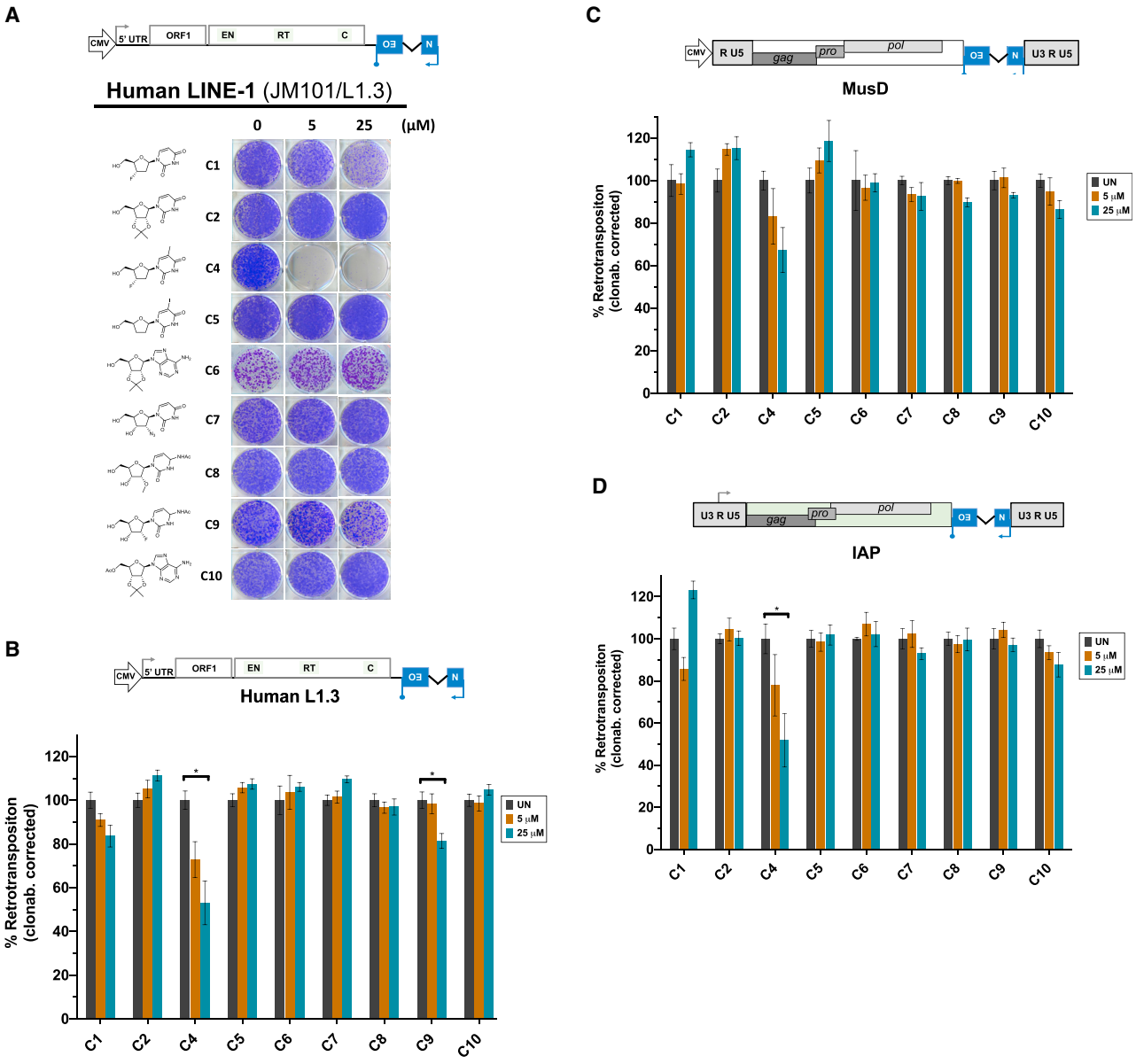


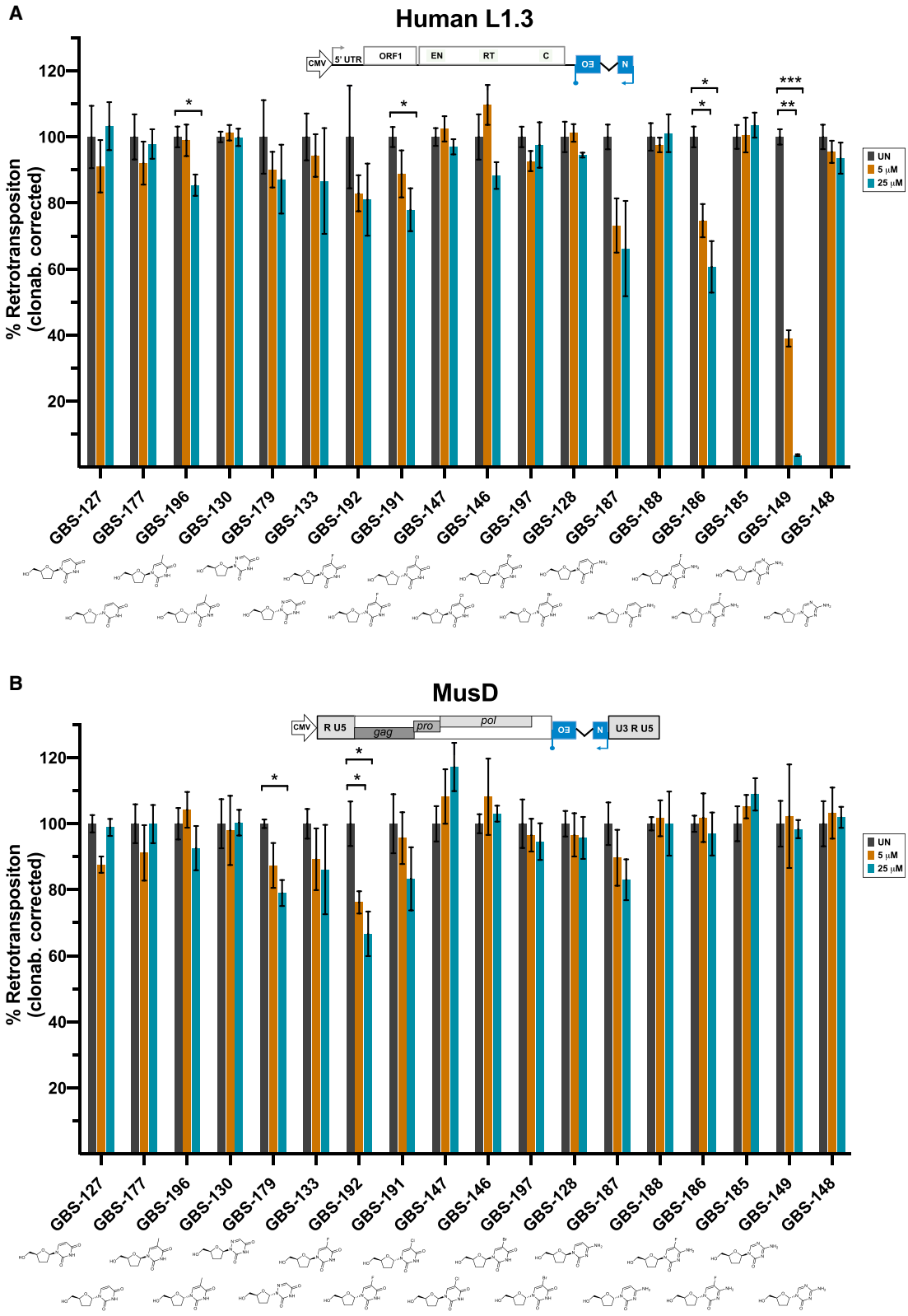
Figure 3. Effect of Commercially Available Nucleoside Analogs on Mammalian Retrotransposons

(A and B) Effect on a human LINE-1 retrotransposon (L1.3). (A) Representative results of LINE-1 retrotransposition assays conducted in the presence of the indicated amount of each nucleoside analog using HeLa cells (triplicate, 2×10^4 cells per well). (B) Quantification of assays shown in (A); black bars, no drug; orange bars, 5 μM; green bars, 25 μM. A cartoon of the L1 construct used (plasmid JM101/L1.3) is shown at the top of (A and B) *$p \leq 0.05$.

(C and D) Effect on mouse LTR-retrotransposons. (C) (MusD) and D (IAP) show the effect of the indicated nucleoside analog on mouse LTR-retrotransposition using HeLa cells (triplicate, 2×10^4 [MusD] and 1×10^4 [IAP] HeLa cells per well were transfected and selected). A cartoon of each construct used is also shown. Black bars, no drug; orange bars, 5 μM; green bars, 25 μM; (D). *$p \leq 0.05$.

See also Figure S3.

of drug-resistant colonies, and might avoid confounding effects generated by changes in the growth characteristics of RTi-treated cells. In addition, in this assay we used a different LINE-1, $L1_{RP}$, an active human element that inserted in the retinitis pigmentosa-2 gene and generated a human disorder (Kimberland et al., 1999). Using this assay, we confirmed significant inhibition of human LINE-1 retrotransposition on HEK293T cells in a dose-dependent manner (Figures S6G–S6I). Notably, we observed that GBS-149 is the most potent inhibitor of human LINE-1 retrotransposition in HEK293T, followed by lamivudine and emtricitabine. Finally, additional quality controls revealed that several RTis tested, with (GBS-149, emtricitabine, lamivudine, tenofovir, and AZT) or without (C5) inhibitory effect on human LINE-1 retrotransposition, did not affect the overall stability of transfected-engineered LINE-1 DNA constructs or expression of L1-derived RNAs from transfected plasmids (Figures S7A–S7D).



(legend on next page)

Specificity of GBS-149, Emtricitabine, and Lamivudine on Mammalian Retrotransposons

Overall, our data demonstrated that GBS-149, emtricitabine, and lamivudine are potent inhibitors of human LINE-1 retrotransposition, with no effect on mouse LTR retrotransposition rates. To confirm the lack of activity on mouse LTR retrotransposons, we next evaluated the activity of GBS-149, emtricitabine, and lamivudine at higher concentrations. Toxicity control assays revealed that none of the three RTis tested induced significant short- or long-term toxicity in HeLa cells when tested at low (5 and 25 μM, Figures 5B and S8A) or high concentrations (50 and 150 μM, Figures S8C and S8D). Notably, GBS-149, emtricitabine, and lamivudine failed to inhibit mouse LTR retrotransposition (MusD and IAP) when tested at low (5 and 25 μM, Figures 6A and 6B) or high concentrations (50 and 150 μM, Figure S8B). Parallel assays confirmed robust human L1 retrotransposition inhibition with GBS-149, emtricitabine, and lamivudine at low concentrations (5 and 25 μM, Figures 6A and 6B). Thus, these data suggest that GBS-149, emtricitabine, and lamivudine are RTis specific to LINE-1 retrotransposons.

However, to further analyze the specificity of GBS-149, emtricitabine, and lamivudine, we next tested their retrotransposition inhibitory potential on three active mouse LINE-1s from the L1Md-A, -T_F, and -G_F subfamilies (DeBerardinis et al., 1998; Goodier et al., 2001; Naas et al., 1998). To do that, we used *mneoI*-tagged active mouse LINE-1s on retrotransposition assays using HeLa cells. These assays revealed that GBS-149, emtricitabine, and lamivudine could also inhibit the retrotransposition of mouse L1Md-A, -T_F, and -G_F retrotransposons, but not to the same extent in comparison with human LINE-1 retrotransposition inhibition (using L1.3, Figures 6A and 6B). On average, GBS-149, emtricitabine, and lamivudine reduced mouse LINE-1 retrotransposition by ~2-fold when tested at 25 μM (GBS-149: L1Md-A, 1.6-fold; L1Md-T_F 1.5-fold; L1Md-G_F 2.6-fold; emtricitabine: L1Md-A, 1.7-fold; L1Md-T_F 2.1-fold; L1Md-G_F 1.9-fold; lamivudine: L1Md-A, 2-fold; L1Md-T_F 2.2-fold; L1Md-G_F 1.8-fold, Figures 6A and 6B). Remarkably, GBS-149, emtricitabine, and lamivudine reduced human LINE-1 retrotransposition by 33-, 7-, and 3-fold, respectively (at 25 μM, Figures 6A and 6B). Thus, these data revealed that GBS-149 is the more potent and specific human LINE-1 RTi tested, followed by emtricitabine, and then by lamivudine, which seems to inhibit all tested LINE-1s to the same extent (Figures 6A and 6B).

DISCUSSION

There is increasing interest in deciphering the role of active LINE-1 retrotransposition in a number of biological processes, and inhibiting ongoing retrotransposition using RTis is a simple, straightforward, and effective manner to achieve this goal. The pandemic infection with HIV has resulted in the development of effective drugs to treat AIDS, and some of these drugs target the RT encoded by HIV. Non-nucleoside and nucleoside analog RTis are the two main classes of drugs targeting the RT of HIV, and nucleoside analogs can potentially interfere with other RTs due to their mechanism of action (Arts and Wainberg, 1996). Indeed, two previous studies have previously tested the effect that selected nucleoside analogs used to treat HIV/AIDS exert on the retrotransposition rate of human LINE-1 retrotransposons. Jones and colleagues demonstrated that stavudine, AZT (or zidovudine), tenofovir, and lamivudine all inhibited human LINE-1 retrotransposition using HeLa cells, while nevirapine, a non-nucleoside analog HIV RTi, showed no effect on L1 retrotransposition (Jones et al., 2008). Although no toxicity was evaluated in this study, and although a different human LINE-1 (LRE3) retrotransposon tagged with a different retrotransposition indicator reporter gene (EGFP) (Ostertag et al., 2000) were used, our data are very similar to those published by Jones et al. (2008). More recently, in an elegant study Dai et al. (2011) tested the inhibitory effect of several HIV RTis (including AZT, abacavir, lamivudine, stavudine, and efavirenz), but using recombinant purified RT proteins; the RT domain of an active human LINE-1 ($L1_{RP}$, expressed from a codon optimized construct), the RT domain of HIV (p66), and the RT-RNase H domain of an LTR retrotransposon from yeast, Ty1, were tested *in vitro*, and it was found that all nucleoside analogs inhibited the RT activity of these three retroelements, but with different strengths. As controls, non-nucleoside analogs HIV RTis were tested, and no activity was found to human LINE-1 RT activity. Furthermore, Dai et al. (2011) also examined their inhibitory potential on human (a codon optimized $L1_{RP}$ element was used) and mouse LINE-1s (using the same codon optimized mouse L1Md-T_F element used in this study) retrotransposons using HeLa cells and *mneoI*-tagged constructs. In general, our data fully agree with those included in the Dai et al. study, and we both found that nucleoside analogs used to treat AIDS also inhibit human and mouse L1 retrotransposition, with different potencies. Similarly, no inhibitory effect on mammalian L1 retrotransposition was detected with non-nucleoside RTis (efavirenz, etravirine, and nevirapine), with the exception of nevirapine, which mildly reduced human L1 retrotransposition, but in a dose-independent manner. In summary, our data are in agreement with these two previous studies, although here we: (1) increased the number of nucleoside analogs tested, (2) tested more mammalian retrotransposons, (3) tested more cell lines, and (4) we analyzed their short- and long-term induced toxicity to cultured cells. As a result, and after testing 33 nucleoside analogs, our study has identified several drugs that specifically inhibit human and mouse active LINE-1s, without interfering with mouse LTR-retrotransposition.

To date, the two studies mentioned above have been used as a reference in other studies analyzing the implication of endogenous LINE-1 retrotransposition in a myriad of biological processes, from a potential role in memory to oocyte attrition to cite a few (Bachiller et al., 2017; Malki et al., 2014). Many of these studies used the mouse as a model, and given its known efficiency, most also used AZT as an RTi, which affects both LTR and LINE-1 retrotransposition; similarly, studies using tenofovir

Figure 4. Effect of Synthesized Nucleoside Analogs on Mammalian Retrotransposons

(A and B) The effect of the indicated nucleoside analog on human LINE-1 (A) and mouse MusD (B) retrotransposition using HeLa cells (triplicate, 2×10^4 HeLa cells per well were transfected and selected). A cartoon of each construct used is also shown. Black bars, no drug; orange bars, 5 μM; green bars, 25 μM. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. See also Figures S4 and S5.

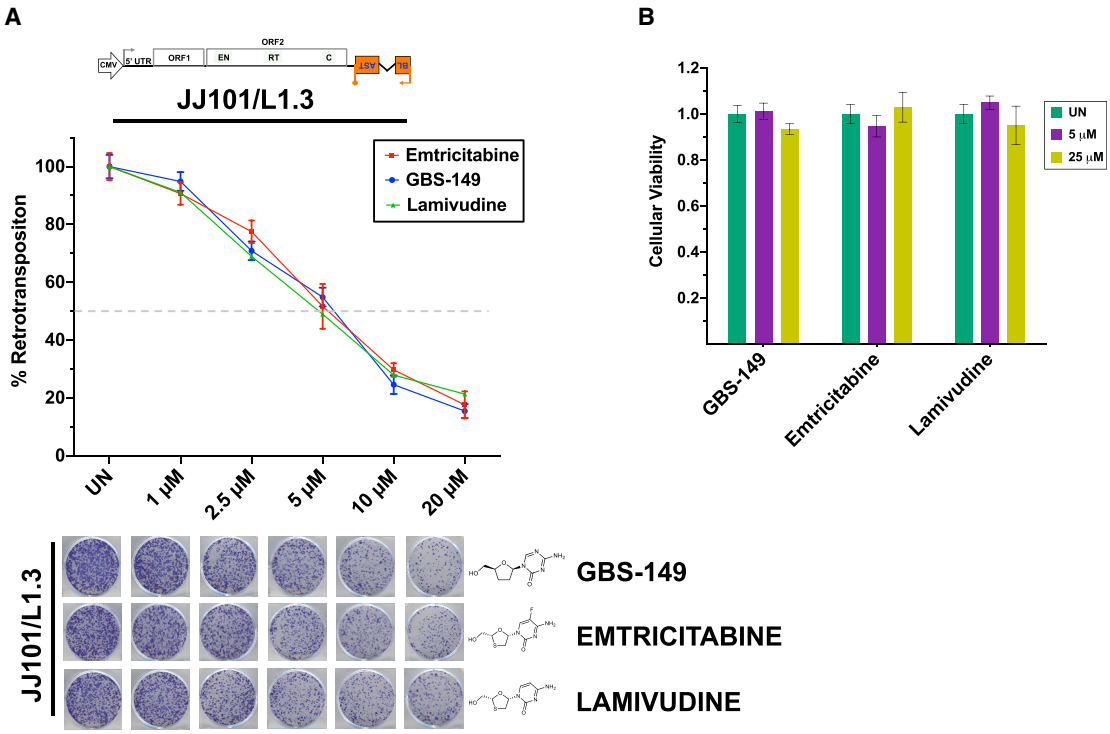


Figure 5. GBS-149, Emtricitabine, and Lamivudine Dose-Response Assays on a Human LINE-1

(A) Graph showing the effect of increasing concentrations of GBS-149 (blue), emtricitabine (red), or lamivudine (green) on human LINE-1 retrotransposition (L1.3) using HeLa cells (2×10^4 cells per well, triplicate assays). Representative retrotransposition results are shown below the graph, and a cartoon of the LINE-1 construct used is drawn on the top.

(B) MTT assays in the presence of GBS-149, emtricitabine, and lamivudine. The graph shows the results of MTT cell metabolic activity assays conducted on HeLa cells in the presence of 5 μM (purple bars) and 25 μM (yellow bars) of the indicated RTi (triplicate). Colorimetric values of untreated cells (green bars) were used to normalize and were arbitrarily designated as 1. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

See also Figures S6 and S7.

as an RTi also lead to LTR and LINE-1 mobilization inhibition. Thus, any study using AZT and tenofovir as an RTi would not allow to discern if a given biological process might be driven by LINE-1 or LTR retrotransposition. Furthermore, both AZT and tenofovir exhibit some level of toxicity to cultured cells. However, our data have demonstrated that several nucleoside analogs are specific inhibitors of mouse LINE-1 retrotransposition, including GBS-149, emtricitabine, and lamivudine. Thus, for future studies testing the role of LINE-1 retrotransposition in any biological process, and when using mouse as a model, GBS-149, emtricitabine, or lamivudine might be RTis of choice, as they specifically inhibit the retrotransposition of the three currently active mouse LINE-1 subfamilies and have no short- or long-term toxicity. Similarly, when aiming to inhibit ongoing human LINE-1 retrotransposition, GBS-149, emtricitabine, or lamivudine might be also RTis of choice as they elicit no toxicity; however, GBS-149 is the more potent inhibitor of human LINE-1 retrotransposition characterized in this study. Although it seems that ERVs no longer move in the germline of humans, there are polymorphic elements from the K class (HERV-K) partially intact that could code for all viral proteins, and active HERV-K elements have been reconstructed (Dewannieux et al., 2006; Mager and Stoye, 2015). Thus, and although no ongoing activity of HERV-K might be expected, understanding how RTis affect the RT activity of HERVs might be informative. Recently, a report tested the effect of HIV RTis on the RT activity of an HERV-K HML-2 family type element (Contreras-Galindo et al., 2017). Notably, Contreras-Galindo et al. (2017) demonstrated that several nucleoside analogs used to treat HIV can indeed inhibit the RT activity of HERV-K, including AZT, emtricitabine, lamivudine, stavudine, abacavir, and tenofovir, in a dose-response manner; this study also demonstrated that non-nucleoside HIV RTis (nevirapine, efavirenz, and etravirine) failed to inhibit HERV-K RT activity. Thus, and if the activity of HERV-K elements should be considered when using RTis to infer any retrotransposition-derived impact, we suggest that abacabir could be the RTi of choice. Indeed, abacavir efficiently inhibits the RT activity encoded by HERV-K, with an half maximal inhibitory concentration of 14.6 μM (Contreras-Galindo et al., 2017). However, and although abacavir can inhibit human LINE-1 retrotransposition, it is not a very potent inhibitor, and LINE-1 retrotransposition in HeLa cells was reduced by less than 40% using 25 μM of this RTi. On the other hand, and while we did not identify a specific RTi for mouse MusD and IAP LTR retrotransposons, tenofovir is a potent inhibitor of both ERVs, which could be used in conjunction with a specific mouse LINE-1 RTi (GBS-149, emtricitabine, or lamivudine) to discern whether a given biological phenotype is driven by LINE-1 and/or LTR

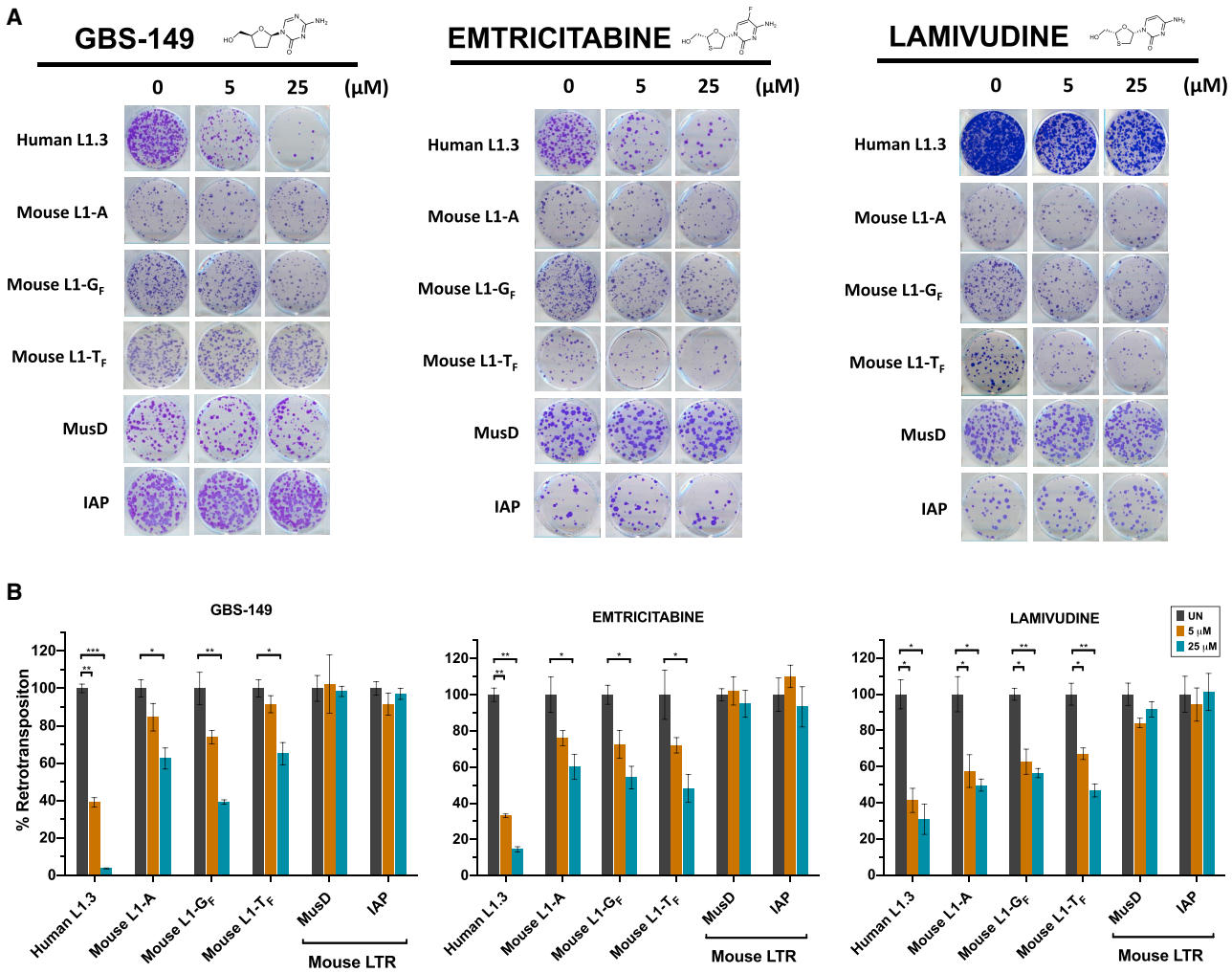


Figure 6. Specificity of GBS-149, Emtricitabine, and Lamivudine on Mammalian Retrotransposons

(A) Representative results of retrotransposition assays conducted in the presence of the indicated amount of GBS-149 (left), emtricitabine (middle), and lamivudine (right) using HeLa cells. The name of the retrotransposon tested is indicated in the left side. Note that different numbers of HeLa cells were plated depending on the construct tested (human L1.3, 2×10^4; mouse L1Md-A and L1Md-G$_F$, 2×10^4; mouse L1MdT$_F$, 1×10^4; mouse MusD, 2×10^4; mouse IAP, 1×10^4).

(B) Quantification of the above assays in triplicate (black bars, no drug; orange bars, 5 μM; green bars, 25 μM). *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001.

See also Figures S7 and S8.

retrotransposition. These data suggest that, although the catalytic domain of all characterized RTs is conserved (known as domain C, and containing bhDD, where b = big and h = hydrophobic amino acids), and presumably its 3D structure, minor changes in the structure of nucleoside analogs can have a significant effect in their inhibitory strength and specificity. Thus, it is likely that additional synthesis and testing will uncover nucleoside analog structure(s) with retrotransposition inhibition activity specific for mouse LTR retrotransposons.

From a structural view, it is worth mentioning that all active RTis to human LINE-1 retrotransposition are β-nucleoside analogs (either D- or L-series), and the most potent are pyrimidine analogs, with the exception of tenofovir. Indeed, GBS-149, emtricitabine, and lamivudine, three non-toxic and specific LINE-1 RTis, are cytidine analogs; consistently, GBS-187 and -186 are cytidine and fluorocytidine analogs, respectively, with specific inhibitory activity against human LINE-1 retrotransposons, although both elicit some level of toxicity in cultured HeLa cells. However, AZT and tenofovir, which, alongside LINE-1, are also active RTis against mouse LTR-retrotransposons, are uracil and adenine analogs, respectively. Given the mechanism of action of nucleoside analogs (Arts and Wainberg, 1996), we speculate that cytidine analogs might bind the RT domain of mammalian LINE-1s with more affinity; similarly, it is also likely that GBS-149 might bind the RT of human LINE-1s with more affinity than mouse LINE-1s. At present, there is no available crystal structure for the RT domain of any LINE-1, and no binding predictions of nucleoside analogs can be made. Similarly, the lack of a LINE-1-RT 3D structure prohibits inferring predictions of nucleoside analog structures that could potentially bind specifically and strongly to the catalytic motif of the RT domain of human L1-ORF2p. However, previous reports have shown that purifying

human L1-ORF2p is extremely challenging (Cost et al., 2002; Piskareva et al., 2003; Piskareva and Schmatchenko, 2006), which might explain why a crystal structure for mammalian L1-ORF2p is currently unavailable. Very recently, progress has been made in deciphering the 3D structure of the RT domain from two group II introns, which are retroelements evolutionarily related to LINE-1 retrotransposons (Stamos et al., 2017; Zhao et al., 2018). Thus, additional research is needed to obtain a reliable 3D structure for mammalian L1-ORF2p, but, once obtained, it could be the basis to develop potent and specific human LINE-1 RTis. It is worth mentioning that the inhibitory activity of GBS-149, a drug synthesized and characterized in this study, is on average 18-fold more active on human LINE-1s than in mouse LINE-1s (from any of the three currently active subfamilies of elements present in the mouse genome). In our study, we tested two previously characterized active human LINE-1s, L1.3 (Sassaman et al., 1997) and $L1_{RP}$ (Kimberland et al., 1999), and we showed that GBS-149 is active in both human LINE-1s, reducing retrotransposition by >30-fold at 25 μM. Currently active human LINE-1s are 99.99% identical at the nucleotide level (Beck et al., 2010; Brouha et al., 2003), and we thus speculate that GBS-149, and other nucleoside analog RTis with activity against human LINE-1s, would be effective in inhibiting the ongoing retrotransposition of any human LINE-1 present in the genome, fixed or polymorphic.

Given the potential role of LINE-1 retrotransposition in the origin and progression of human cancers (reviewed in Burns, 2017; Carreira et al., 2013; Scott and Devine, 2017), the development of non-toxic, specific, and potent inhibitors of human L1 retrotransposition might open the way for clinical trials aiming to test the role of retrotransposition on cancer progression. For most cancer cases, a fully malignant lesion is already present in the patient at the time of diagnosis, and cancer cells might carry hundreds of genomic mutations, including new retrotransposon insertions. The selection of cancer treatments depends on clinical and genomic features of tumors, but we propose that RTis could be used as adjuvant drugs to prevent the accumulation of further L1-driven genomic mutations in cancer cells. In this context, previous studies in mice have found conflicting results, as it was reported that two non-nucleoside analog RTis (efavirenz and nevirapine) with no effect on LINE-1 retrotransposition can reduce cell proliferation and reduce tumor size in animals models (Landriscina et al., 2007; Mangiacasale et al., 2003; Sciamanna et al., 2005). Although at present it is unclear why non-nucleoside analogs might aid cancer treatment, our data suggest that efavirenz is toxic to cancer cells in short- and long-term assays, which could explain their reported effect. However, future studies are needed to truly test whether human L1 RTis might help the prognosis of human cancer patients. Notably, recent data on animal and cellular models have revealed that L1 activity is activated during cellular senescence and aging, triggering a strong type I interferon response (De Cecco et al., 2019; Simon et al., 2019). Intriguingly, these studies demonstrated that nucleoside analog RTis rescued the interferon response observed in senescence cells, expanded the lifespan of aging mice models, and even reduced the sterile inflammation of tissues from aged mice (De Cecco et al., 2019; Simon et al., 2019). Thus, these data suggest that RTis could be also used to prevent aging, although future studies are required to unveil the anti-aging effect of RTis. Finally, and further considering the role of LINE-1 retrotransposons in human interferonopathies (reviewed in Crow and Manel, 2015), it might be worth to further explore non-toxic and human L1-specific RTis to treat AGS patients, as a very recent report presented data suggesting that nucleoside analog RTis with known activity to human L1 retrotransposition are effective in ameliorating part of the symptoms of these patients (Rice et al., 2018).

SIGNIFICANCE

Retrotransposons are abundant transposable elements that have successfully colonized the genome of most mammals, generating millions of insertions over evolution. As a result, almost half of the human and mouse genomic mass is made of retrotransposon-derived sequences. Although most mammalian retrotransposons are currently inactive, a small number of these retroelements continue to mobilize and can impact the genome they reside in. Retrotransposons from the long interspersed element class 1 (LINE-1 or L1) class can still mobilize in humans and mice, whereas long terminal repeat (LTR)-containing retrotransposons are only active in the mouse genome. The mobilization of LINE-1s and LTR retrotransposons in the germline can sporadically generate heritable genetic disorders, while LINE-1 activity in cancer and brain cells has the potential to influence cancer progression and brain biology. However, we know very little about the implications that retrotransposition have in these biological processes. Retrotransposition requires reverse transcription of an intermediate RNA, and previous studies have demonstrated that several reverse transcriptase inhibitors (RTis) used to treat AIDS can indeed inhibit ongoing LINE-1 retrotransposition. Here, we have conducted the most extensive analysis of the effect of RTis on mammalian retrotransposons. After testing 33 nucleoside analogs for their effect on LINE-1 and LTR retrotransposition, we have identified three RTis specific for mammalian LINE-1s, with no effect on LTR-retrotransposons. Two of these drugs were previously known and used to treat AIDS; however, in this study we have synthesized and characterized an RTi, GBS-149, that can inhibit very efficiently the mobilization of human LINE-1 retrotransposons. Although GBS-149 can inhibit the retrotransposition of mouse and humans LINE-1s, its activity on human LINE-1s is ~18-fold more effective than on mouse LINE-1s. In summary, we have identified RTis that will allow testing the functional impact of LINE-1 retrotransposition *in vivo*.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell Line Authentication
  - Cell Line Conditions
- METHOD DETAILS
  - Chemicals Used in this Study
  - Chemical Synthesis of 2',3'-dideoxynucleosides

- Expression Plasmids
- Retrotransposition Assays
- MTT Assays
- qPCR and RT-qPCR Control Assays
- Western Blot Analyses

● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.chembiol.2019.04.010.

ACKNOWLEDGMENTS

Current members of the J.L.G.-Perez lab are acknowledged for helpful discussions during the course of the study. We also acknowledge Drs Thierry Heidmann (Institut Pasteur, France) for providing the MusD- and IAP-neo^{TNF}-tagged constructs, Jef Boeke (NYU, USA) for providing plasmid pCEPL1SM, and John Goodier (John Hopkins, USA) for providing mouse LINE-1 expression vectors (from the A and G_F classes) tagged with the *mneoI* retrotransposition indicator cassette. J.A.T.'s lab has been funded by University of Granada (project CS2016-6). G.B.-S. is funded by a "Garantía Juvenil" contract from the European Social Fund in collaboration with the Andalusian Regional Government. M.B.-G. is funded by a "Formacion Profesorado Universitario" (FPU) PhD fellowship from the Government of Spain (MINECO, Ref FPU15/03294). In J.L.G.-Perez's lab, this study has been funded by an International Early Career Scientist grant from the Howard Hughes Medical Institute (IECS-55007420) and by a private donation from Ms Francisca Serrano (*Trading y Bolsa para Torpes*, *Granada*, *Spain*). J.L.G.Perez's lab is also supported by the European Research Council (ERC-Consolidator ERC-STG-2012-309433), and by MINECO-FEDER (SAF2017-89745-R).

AUTHOR CONTRIBUTIONS

Conceptualization, G.B.-S., J.A.T., and J.L.G.-Perez; Investigation, G.B.-S., L.S., V.S.-C., C.S.-P., P.T.-R., M.B.-B., M.B.-G., and S.M.; Writing – Original Draft, J.L.G.-Perez; Writing – Review & Editing, all authors; Funding Acquisition, J.A.T., S.R.H., and J.L.G.-Perez; Resources, F.F.-M., and J.L.G.-Puche; Supervision, J.A.T., S.R.H., and J.L.G.-Perez.

DECLARATION OF INTERESTS

None of the authors declare any conflict of interest.

Received: January 22, 2019
Revised: March 29, 2019
Accepted: April 19, 2019
Published: May 30, 2019

REFERENCES

Achleitner, M., Kleefisch, M., Hennig, A., Peschke, K., Polikarpova, A., Oertel, R., Gabriel, B., Schulze, L., Lindeman, D., Gerbaulet, A., et al. (2017). Lack of Trex1 causes systemic autoimmunity despite the presence of antiretroviral drugs. J. Immunol. *199*, 2261–2269.

Agyei-Aye, K., and Baker, D.C. (1988). Synthesis and evaluation of a series of 1-(3-alkyl-2,3-dideoxy-alpha,beta-D-erythro-pentofuranosyl)thymines. Carbohydr. Res. *183*, 261–275.

Arts, E.J., and Wainberg, M.A. (1996). Mechanisms of nucleoside analog antiviral activity and resistance during human immunodeficiency virus reverse transcription. Antimicrob. Agents Chemother. *40*, 527–540.

Audat, S.A., Love, C.T., Al-Oudat, B.A., and Bryant-Friedrich, A.C. (2012). Synthesis of C3' modified nucleosides for selective generation of the C3'-deoxy-3'-thymidinyl radical: a proposed intermediate in LEE induced DNA damage. J. Org. Chem. *77*, 3829–3837.

Bachiller, S., Del-Pozo-Martin, Y., and Carrion, A.M. (2017). L1 retrotransposition alters the hippocampal genomic landscape enabling memory formation. Brain Behav. Immun. *64*, 65–70.

Beck, C.R., Collier, P., Macfarlane, C., Malig, M., Kidd, J.M., Eichler, E.E., Badge, R.M., and Moran, J.V. (2010). LINE-1 retrotransposition activity in human genomes. Cell *141*, 1159–1170.

Beck-Engeser, G.B., Eilat, D., and Wabl, M. (2011). An autoimmune disease prevented by anti-retroviral drugs. Retrovirology *8*, 91.

Belshaw, R., Pereira, V., Katzourakis, A., Talbot, G., Paces, J., Burt, A., and Tristem, M. (2004). Long-term reinfection of the human genome by endogenous retroviruses. Proc. Natl. Acad. Sci. U S A *101*, 4894–4899.

Benitez-Guijarro, M., Lopez-Ruiz, C., Tarnauskaite, Z., Murina, O., Mian Mohammad, M., Williams, T.C., Fluteau, A., Sanchez, L., Vilar-Astasio, R., Garcia-Canadas, M., et al. (2018). RNase H2, mutated in Aicardi-Goutieres syndrome, promotes LINE-1 retrotransposition. EMBO J. *37*, https://doi.org/10.15252/embj.201798506.

Boeke, J.D., and Chapman, K.B. (1991). Retrotransposition mechanisms. Curr. Opin. Cell Biol. *3*, 502–507.

Boeke, J.D., Garfinkel, D.J., Styles, C.A., and Fink, G.R. (1985). Ty elements transpose through an RNA intermediate. Cell *40*, 491–500.

Brouha, B., Schustak, J., Badge, R.M., Lutz-Prigge, S., Farley, A.H., Moran, J.V., and Kazazian, H.H., Jr. (2003). Hot L1s account for the bulk of retrotransposition in the human population. Proc. Natl. Acad. Sci. U S A *100*, 5280–5285.

Burns, K.H. (2017). Transposable elements in cancer. Nat. Rev. Cancer *17*, 415–424.

Camarasa, M.J., De las Heras, F., and Perez-Perez, M.J. (1990). Aldol reaction of nucleoside 5′-carboxaldehydes with acetone. synthesis of 5′-C-chain extended thymidine derivatives. Nucleosides Nucleotides *9*, 533–546.

Carreira, P.E., Richardson, S.R., and Faulkner, G.J. (2013). L1 retrotransposons, cancer stem cells and oncogenesis. FEBS J. *281*, 63–73.

Contreras-Galindo, R., Dube, D., Fujinaga, K., Kaplan, M.H., and Markovitz, D.M. (2017). Susceptibility of human endogenous retrovirus type K to reverse transcriptase inhibitors. J. Virol. *91*, https://doi.org/10.1128/JVI.01309-17.

Cost, G.J., Feng, Q., Jacquier, A., and Boeke, J.D. (2002). Human L1 element target-primed reverse transcription in vitro. EMBO J. *21*, 5899–5910.

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Yeo, G.W., Mu, Y., Lovci, M.T., Morell, M., O'Shea, K.S., Moran, J.V., and Gage, F.H. (2009). L1 retrotransposition in human neural progenitor cells. Nature *460*, 1127–1131.

Crow, Y.J., and Manel, N. (2015). Aicardi-Goutieres syndrome and the type I interferonopathies. Nat. Rev. Immunol. *15*, 429–440.

Dai, L., Huang, Q., and Boeke, J.D. (2011). Effect of reverse transcriptase inhibitors on LINE-1 and Ty1 reverse transcriptase activities and on LINE-1 retrotransposition. BMC Biochem. *12*, 18.

De Cecco, M., Ito, T., Petrashen, A.P., Elias, A.E., Skvir, N.J., Criscione, S.W., Caligiana, A., Brocculi, G., Adney, E.M., Boeke, J.D., et al. (2019). L1 drives IFN in senescent cells and promotes age-associated inflammation. Nature *566*, 73–78.

DeBerardinis, R.J., Goodier, J.L., Ostertag, E.M., and Kazazian, H.H., Jr. (1998). Rapid amplification of a retrotransposon subfamily is evolving the mouse genome. Nat. Genet. *20*, 288–290.

Dewannieux, M., Dupressoir, A., Harper, F., Pierron, G., and Heidmann, T. (2004). Identification of autonomous IAP LTR retrotransposons mobile in mammalian cells. Nat. Genet. *36*, 534–539.

Dewannieux, M., Esnault, C., and Heidmann, T. (2003). LINE-mediated retrotransposition of marked Alu sequences. Nat. Genet. *35*, 41–48.

Dewannieux, M., Harper, F., Richaud, A., Letzelter, C., Ribet, D., Pierron, G., and Heidmann, T. (2006). Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. Genome Res. *16*, 1548–1556.

Dewannieux, M., and Heidmann, T. (2005). L1-mediated retrotransposition of murine B1 and B2 SINEs recapitulated in cultured cells. J. Mol. Biol. *349*, 241–247.

Erwin, J.A., Paquola, A.C., Singer, T., Gallina, I., Novotny, M., Quayle, C., Bedrosian, T.A., Alves, F.I., Butcher, C.R., Herdy, J.R., et al. (2016). L1-associated genomic regions are deleted in somatic cells of the healthy human brain. Nat. Neurosci. *19*, 1583–1591.

Esnault, C., Casella, J.F., and Heidmann, T. (2002). A Tetrahymena thermophila ribozyme-based indicator gene to detect transposition of marked retroelements in mammalian cells. Nucleic Acids Res. *30*, e49.

Evrony, G.D., Cai, X., Lee, E., Hills, L.B., Elhosary, P.C., Lehmann, H.S., Parker, J.J., Atabay, K.D., Gilmore, E.C., Poduri, A., et al. (2012). Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. Cell *151*, 483–496.

Fowler, B.J., Gelfand, B.D., Kim, Y., Kerur, N., Tarallo, V., Hirano, Y., Amarnath, S., Fowler, D.H., Radwan, M., Young, M.T., et al. (2014). Nucleoside reverse transcriptase inhibitors possess intrinsic anti-inflammatory activity. Science *346*, 1000–1003.

Freeman, J.D., Goodchild, N.L., and Mager, D.L. (1994). A modified indicator gene for selection of retrotransposition events in mammalian cells. Biotechniques *17*, 46, 48–49, 52.

Furukawa, Y., Yoshioka, Y., Imai, K., and Honjo, M. (1970). Studies on the synthesis of pyrimidine deoxynucleosides. I. synthesis of 2', 3'-dideoxyuridine and 1-(3-ethylthio-3-deoxy-β-D-xylofuranosyl) uracil. Chem. Pharm. Bull. (Tokyo) *18*, 554–560.

Garcia-Perez, J.L., Morell, M., Scheys, J.O., Kulpa, D.A., Morell, S., Carter, C.C., Hammer, G.D., Collins, K.L., O'Shea, K.S., Menendez, P., et al. (2010). Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. Nature *466*, 769–773.

Garcia-Perez, J.L., Widmann, T.J., and Adams, I.R. (2016). The impact of transposable elements on mammalian development. Development *143*, 4101–4114.

Goodier, J.L., Ostertag, E.M., Du, K., and Kazazian, H.H., Jr. (2001). A novel active L1 retrotransposon subfamily in the mouse. Genome Res. *11*, 1677–1685.

Goodier, J.L., Zhang, L., Vetter, M.R., and Kazazian, H.H., Jr. (2007). LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. Mol. Cell. Biol. *27*, 6469–6483.

Grimaldi, G., Skowronski, J., and Singer, M.F. (1984). Defining the beginning and end of KpnI family segments. EMBO J. *3*, 1753–1759.

Han, J.S., and Boeke, J.D. (2004). A highly active synthetic mammalian retrotransposon. Nature *429*, 314–318.

Heidmann, O., and Heidmann, T. (1991). Retrotransposition of a mouse IAP sequence tagged with an indicator gene. Cell *64*, 159–170.

Heras, S.R., Macias, S., Plass, M., Fernandez, N., Cano, D., Eyras, E., Garcia-Perez, J.L., and Caceres, J.F. (2013). The microprocessor controls the activity of mammalian retrotransposons. Nat. Struct. Mol. Biol. *20*, 1173–1181.

Horwitz, J.P., Chua, J., Noel, M., and Donatti, J.T. (1967). Nucleosides. XI. 2',3'-dideoxycytidine. J. Org. Chem. *32*, 817–818.

Hughes, J.F., and Coffin, J.M. (2004). Human endogenous retrovirus K solo-LTR formation and insertional polymorphisms: implications for human and viral evolution. Proc. Natl. Acad. Sci. U S A *101*, 1668–1672.

Ivancevic, A.M., Kortschak, R.D., Bertozzi, T., and Adelson, D.L. (2018). Horizontal transfer of BovB and L1 retrotransposons in eukaryotes. Genome Biol. *19*, 85.

Izumi, M., Miyazawa, H., Kamakura, T., Yamaguchi, I., Endo, T., and Hanaoka, F. (1991). Blasticidin S-resistance gene (bsr): a novel selectable marker for mammalian cells. Exp. Cell Res. *197*, 229–233.

Jones, R.B., Garrison, K.E., Wong, J.C., Duan, E.H., Nixon, D.F., and Ostrowski, M.A. (2008). Nucleoside analogue reverse transcriptase inhibitors differentially inhibit human LINE-1 retrotransposition. PLoS One *3*, e1547.

Kawakami, H., Ebata, T., Koseki, K., Matsumoto, K., Matsushita, H., Naoi, Y., and Itoh, K. (1990). Stereoselectivities in the coupling reaction between silylated pyrimidine bases and 1-Halo-2,3-dideoxyribose. Heterocycles *31*, 2041–2054.

Kazazian, H.H., Jr., Wong, C., Youssoufian, H., Scott, A.F., Phillips, D.G., and Antonarakis, S.E. (1988). Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature *332*, 164–166.

Khwaja, T.A., and Heidelberger, C. (1967). Fluorinated pyrimidines. XXIX. Syntheses of 2',3'-dehydro-5-fluoro-2'-deoxyuridine and 2',3'-dideoxy-5-fluorouridine. J. Med. Chem. *10*, 1066–1070.

Kim, A., Terzian, C., Santamaria, P., Pelisson, A., Purd'homme, N., and Bucheton, A. (1994). Retroviruses in invertebrates: the gypsy retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U S A *91*, 1285–1289.

Kim, C.H., Marquez, V.E., Broder, S., Mitsuya, H., and Driscoll, J.S. (1987). Potential anti-AIDS drugs. 2',3'-dideoxycytidine analogues. J. Med. Chem. *30*, 862–866.

Kimberland, M.L., Divoky, V., Prchal, J., Schwahn, U., Berger, W., and Kazazian, H.H., Jr. (1999). Full-length human L1 insertions retain the capacity for high frequency retrotransposition in cultured cells. Hum. Mol. Genet. *8*, 1557–1560.

Kopera, H.C., Moldovan, J.B., Morrish, T.A., Garcia-Perez, J.L., and Moran, J.V. (2011). Similarities between LINE-1 reverse transcriptase and telomerase. Proc. Natl. Acad. Sci. U S A *108*, 20345–20350.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Landriscina, M., Spadafora, C., Cignarelli, M., and Barone, C. (2007). Antitumor activity of non-nucleosidic reverse transcriptase inhibitors. Curr. Pharm. Des. *13*, 737–747.

Li, W., Prazak, L., Chatterjee, N., Gruninger, S., Krug, L., Theodorou, D., and Dubnau, J. (2013). Activation of transposable elements during aging and neuronal decline in *Drosophila*. Nat. Neurosci. *16*, 529–531.

Lin, T.-S., Luo, M.-Z., and Liu, M.C. (1995). Synthesis of several pyrimidine L-nucleoside analogues as potential antiviral agents. Tetrahedron *51*, 1055–1068.

Lin, T.S., Luo, M.Z., Liu, M.C., Pai, S.B., Dutschman, G.E., and Cheng, Y.C. (1994). Synthesis and biological evaluation of 2',3'-dideoxy-L-pyrimidine nucleosides as potential antiviral agents against human immunodeficiency virus (HIV) and hepatitis B virus (HBV). J. Med. Chem. *37*, 798–803.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods *25*, 402–408.

Macia, A., Widmann, T.J., Heras, S.R., Ayllon, V., Sanchez, L., Benkaddour-Boumzaouad, M., Munoz-Lopez, M., Rubio, A., Amador-Cubero, S., Blanco-Jimenez, E., et al. (2017). Engineered LINE-1 retrotransposition in nondividing human neurons. Genome Res. *27*, 335–348.

MacLennan, M., Garcia-Canadas, M., Reichmann, J., Khazina, E., Wagner, G., Playfoot, C.J., Salvador-Palomeque, C., Mann, A.R., Peressini, P., Sanchez, L., et al. (2017). Mobilization of LINE-1 retrotransposons is restricted by Tex19.1 in mouse embryonic stem cells. Elife *6*, https://doi.org/10.7554/eLife.26152.

Mager, D.L., and Stoye, J.P. (2015). Mammalian endogenous retroviruses. Microbiol. Spectr. *3*, https://doi.org/10.1128/microbiolspec.MDNA3-0009-2014.

Malki, S., van der Heijden, G.W., O'Donnell, K.A., Martin, S.L., and Bortvin, A. (2014). A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. Dev. Cell *29*, 521–533.

Mangiacasale, R., Pittoggi, C., Sciamanna, I., Careddu, A., Mattei, E., Lorenzini, R., Travaglini, L., Landriscina, M., Barone, C., Nervi, C., et al. (2003). Exposure of normal and transformed cells to nevirapine, a reverse transcriptase inhibitor, reduces cell growth and promotes differentiation. Oncogene *22*, 2750–2761.

Michelson, A.M., and Todd, A.R. (1955). Deoxyribonucleosides and related compounds. Part V. cycloThymidines and other thymidine derivatives. The configuration at the glycosidic centre in thymidine. J. Chem. Soc. 815–823.

Mitsudo, K., Kawaguchi, T., Miyahara, S., Matsuda, W., Kuroboshi, M., and Tanaka, H. (2005). Electrooxidative glycosylation through C-S bond cleavage of 1-arylthio-2,3-dideoxyglycosides. Synthesis of 2',3'-dideoxynucleosides. Org. Lett. *7*, 4649–4652.

Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D., and Kazazian, H.H., Jr. (1996). High frequency retrotransposition in cultured mammalian cells. Cell *87*, 917–927.

Morrish, T.A., Garcia-Perez, J.L., Stamato, T.D., Taccioli, G.E., Sekiguchi, J., and Moran, J.V. (2007). Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres. Nature *446*, 208–212.

Morrish, T.A., Gilbert, N., Myers, J.S., Vincent, B.J., Stamato, T.D., Taccioli, G.E., Batzer, M.A., and Moran, J.V. (2002). DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nat. Genet. *31*, 159–165.

Muotri, A.R., Chu, V.T., Marchetto, M.C., Deng, W., Moran, J.V., and Gage, F.H. (2005). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature *435*, 903–910.

Naas, T.P., DeBerardinis, R.J., Moran, J.V., Ostertag, E.M., Kingsmore, S.F., Seldin, M.F., Hayashizaki, Y., Martin, S.L., and Kazazian, H.H. (1998). An actively retrotransposing, novel subfamily of mouse L1 elements. EMBO J. *17*, 590–597.

Okabe, M., Sun, R.C., Tam, S.Y.K., Todaro, L.J., and Coffen, D.L. (1988). Synthesis of the dideoxynucleosides "ddC" and "CNT" from glutamic acid, ribonolactone, and pyrimidine bases. J. Org. Chem. *53*, 4780–4786.

Ostertag, E.M., Prak, E.T., DeBerardinis, R.J., Moran, J.V., and Kazazian, H.H., Jr. (2000). Determination of L1 retrotransposition kinetics in cultured cells. Nucleic Acids Res. *28*, 1418–1423.

Pfitzner, K.E., and Moffatt, J.G. (1964). The synthesis and hydrolysis of 2',3'-dideoxyuridine. J. Org. Chem. *29*, 1508–1511.

Piskareva, O., Denmukhametova, S., and Schmatchenko, V. (2003). Functional reverse transcriptase encoded by the human LINE-1 from baculovirus-infected insect cells. Protein Expr. Purif. *28*, 125–130.

Piskareva, O., and Schmatchenko, V. (2006). DNA polymerization by the reverse transcriptase of the human L1 retrotransposon on its own template in vitro. FEBS Lett. *580*, 661–668.

Rassu, G., Zanardi, F., Battistini, L., Gaetani, E., and Casiraghi, G. (1997). Expeditious syntheses of sugar-modified nucleosides and collections thereof exploiting furan-, pyrrole-, and thiophene-based siloxy dienes. J. Med. Chem. *40*, 168–180.

Ribet, D., Dewannieux, M., and Heidmann, T. (2004). An active murine transposon family pair: retrotransposition of "master" MusD copies and ETn trans-mobilization. Genome Res. *14*, 2261–2267.

Rice, G.I., Meyzer, C., Bouazza, N., Hully, M., Boddaert, N., Semeraro, M., Zeef, L.A.H., Rozenberg, F., Bondet, V., Duffy, D., et al. (2018). Reverse-transcriptase inhibitors in the Aicardi-Goutieres syndrome. N. Engl. J. Med. *379*, 2275–2277.

Richardson, S.R., Doucet, A.J., Kopera, H.C., Moldovan, J.B., Garcia-Perez, J.L., and Moran, J.V. (2015). The influence of LINE-1 and SINE retrotransposons on mammalian genomes. Microbiol. Spectr. *3*, https://doi.org/10.1128/microbiolspec.MDNA3-0061-2014.

Richardson, S.R., Narvaiza, I., Planegger, R.A., Weitzman, M.D., and Moran, J.V. (2014). APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. Elife *3*.

Rosowsky, A., and Pai, N.N. (1991). Synthesis of the 2′-,3′-didehydro-2′-,3′-dideoxy and 2′-,3′-dideoxy derivatives of 6-azauridine and a new route to 2′-,3′-didehydro-2′-,3′-dideoxy-5-chlorouridine. Nucleosides Nucleotides *10*, 837–851.

Sassaman, D.M., Dombroski, B.A., Moran, J.V., Kimberland, M.L., Naas, T.P., DeBerardinis, R.J., Gabriel, A., Swergold, G.D., and Kazazian, H.H., Jr. (1997). Many human L1 elements are capable of retrotransposition. Nat. Genet. *16*, 37–43.

Schauer, S.N., Carreira, P.E., Shukla, R., Gerhardt, D.J., Gerdes, P., Sanchez-Luque, F.J., Nicoli, P., Kindlova, M., Ghisletti, S., Santos, A.D., et al. (2018). L1 retrotransposition is a common feature of mammalian hepatocarcinogenesis. Genome Res. *28*, 639–653.

Sciamanna, I., Landriscina, M., Pittoggi, C., Quirino, M., Mearelli, C., Beraldi, R., Mattei, E., Serafino, A., Cassano, A., Sinibaldi-Vallebona, P., et al. (2005). Inhibition of endogenous reverse transcriptase antagonizes human tumor growth. Oncogene *24*, 3923–3931.

Scott, E.C., and Devine, S.E. (2017). The role of somatic L1 retrotransposition in human cancers. Viruses *9*, https://doi.org/10.3390/v9060131.

Simon, M., Van Meter, M., Ablaeva, J., Ke, Z., Gonzalez, R.S., Taguchi, T., De Cecco, M., Leonova, K.I., Kogan, V., Helfand, S.L., et al. (2019). LINE1 derepression in aged wild-type and sirt6-deficient mice drives inflammation. Cell Metab. *29*, 871–885.e5.

Singer, T., McConnell, M.J., Marchetto, M.C., Coufal, N.G., and Gage, F.H. (2010). LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? Trends Neurosci. *33*, 345–354.

Sivets, G.G., Klennitskaya, T.V., Zhernosek, E.V., and Mikhailopulo, I.A. (2002). Synthesis of peracylated derivatives of L-ribofuranose from D-ribose and their use for the preparation of β-L-ribonucleosides. Synthesis *2*, https://doi.org/10.1055/s-2002-19805.

Stamos, J.L., Lentzsch, A.M., and Lambowitz, A.M. (2017). Structure of a thermostable group II intron reverse transcriptase with template-primer and its functional and evolutionary implications. Mol. Cell *68*, 926–939.e4.

Stetson, D.B., Ko, J.S., Heidmann, T., and Medzhitov, R. (2008). Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell *134*, 587–598.

Takabatake, T., Ishihara, H., Ohmachi, Y., Tanaka, I., Nakamura, M.M., Fujikawa, K., Hirouchi, T., Kakinuma, S., Shimada, Y., Oghiso, Y., et al. (2008). Microarray-based global mapping of integration sites for the retrotransposon, intracisternal A-particle, in the mouse genome. Nucleic Acids Res. *36*, e59.

Takahara, T., Ohsumi, T., Kuromitsu, J., Shibata, K., Sasaki, N., Okazaki, Y., Shibata, H., Sato, S., Yoshiki, A., Kusakabe, M., et al. (1996). Dysfunction of the *Orleans* reeler gene arising from exon skipping due to transposition of a full-length copy of an active L1 sequence into the skipped exon. Hum. Mol. Genet. *5*, 989–993.

Takahashi, Y., Honda, Y., and Tsuchiya, T. (1995). Synthesis of (3'R)-3'-deoxy-3'-C-nitromethylthymidine. Carbohydr. Res. *270*, 77–83.

Thomas, C.A., Paquola, A.C., and Muotri, A.R. (2012). LINE-1 retrotransposition in the nervous system. Annu. Rev. Cell Dev. Biol. *28*, 555–573.

Thomas, C.A., Tejwani, L., Trujillo, C.A., Negraes, P.D., Herai, R.H., Mesci, P., Macia, A., Crow, Y.J., and Muotri, A.R. (2017). Modeling of TREX1-dependent autoimmune disease using human stem cells highlights L1 accumulation as a source of neuroinflammation. Cell Stem Cell *21*, 319–331.e8.

Upton, K.R., Gerhardt, D.J., Jesuadian, J.S., Richardson, S.R., Sanchez-Luque, F.J., Bodea, G.O., Ewing, A.D., Salvador-Palomeque, C., van der Knaap, M.S., Brennan, P.M., et al. (2015). Ubiquitous L1 mosaicism in hippocampal neurons. Cell *161*, 228–239.

Van Aerschot, A., Everaert, D., Balzarini, J., Augustyns, K., Jie, L., Janssen, G., Peeters, O., Blaton, N., De Ranter, C., De Clercq, E., et al. (1990). Synthesis and anti-HIV evaluation of 2',3'-dideoxyribo-5-chloropyrimidine analogues: reduced toxicity of 5-chlorinated 2',3'-dideoxynucleosides. J. Med. Chem. *33*, 1833–1839.

Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., et al. (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520–562.

Wei, W., Morrish, T.A., Alisch, R.S., and Moran, J.V. (2000). A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events. Anal. Biochem. *284*, 435–438.

Xie, Y., Rosser, J.M., Thompson, T.L., Boeke, J.D., and An, W. (2011). Characterization of L1 retrotransposition with high-throughput dual-luciferase assays. Nucleic Acids Res. *39*, e16.

Zeuthen, J., Norgaard, J.O., Avner, P., Fellous, M., Wartiovaara, J., Vaheri, A., Rosen, A., and Giovanella, B.C. (1980). Characterization of a human ovarian teratocarcinoma-derived cell line. Int. J. Cancer *25*, 19–32.

Zhao, C., Liu, F., and Pyle, A.M. (2018). An ultraprocessive, accurate reverse transcriptase encoded by a metazoan group II intron. RNA *24*, 183–195.

Zhuk, R.A., Berzinya, A.E., Silinya, V.N., Liepin'sh, E.E., and Giller, S.A. (1979). Analogs of pyrimidine mono- and polynucleotides. Chem. Heterocycl. Comp. *10*, 1468–1470.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-L1Hs-ORF1p	Merck	Cat# MABC1152
Mouse monoclonal anti-p53	Santa Cruz Biotechnology	Cat# sc-126; RRID: AB_628082
Mouse monoclonal anti-β-actin	Sigma-Aldrich	Cat# A2228; RRID: AB_476697
Horse anti-mouse IgG HRP-linked	Cell Signaling Technology	Cat# 7076; RRID: AB_330924
Bacterial and Virus Strains		
Escherichia coli DH5α [F^- ϕ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ$^-$]	Thermo Fisher Scientific	Cat# 18265017
Chemicals, Peptides, and Recombinant Proteins		
Abacavir	Hospital Clínico San Cecilio, Granada, Spain	CAS: 136470-78-5
AZT	Hospital Clínico San Cecilio, Granada, Spain	CAS: 30516-87-1
Efavirenz	Hospital Clínico San Cecilio, Granada, Spain	CAS: 154598-52-4
Emtricitabine	Hospital Clínico San Cecilio, Granada, Spain	CAS: 143491-57-0
Etravirine	Hospital Clínico San Cecilio, Granada, Spain	CAS: 269055-15-4
Lamivudine	Hospital Clínico San Cecilio, Granada, Spain	CAS: 134678-17-4
Nevirapine	Hospital Clínico San Cecilio, Granada, Spain	CAS: 129618-40-2
Stavudine	Hospital Clínico San Cecilio, Granada, Spain	CAS: 3056-17-5
Tenofovir disoproxil	Hospital Clínico San Cecilio, Granada, Spain	CAS: 147127-20-6
2′,3′-Dideoxy-3′-fluorouridine (C1)	Sigma-Aldrich	Cat# 393711
2',3'-Isopropylideneuridine (C2)	Sigma-Aldrich	Cat# I5127
2',3′-Dideoxy-3′-fluorothymidine (C4)	Sigma-Aldrich	Cat# 361275
2′,3′-Dideoxy-5-iodouridine (C5)	Sigma-Aldrich	Cat# 327859
2',3'-*O*-isopropylideneadenosine (C6)	Sigma-Aldrich	Cat# I22404
2'-Azido-2'-deoxyuridine (C7)	Sigma-Aldrich	Cat# 11544
N4-Acetyl-2'-O-methylcytidine (C8)	Carbosynth	Cat# NA08375
N4-Acetyl-2'-deoxy-2'-fluorocytidine (C9)	Carbosynth	Cat# NA02623
5'-O-Acetyl-2',3'-O-isopropylideneadenosine (C10)	Carbosynth	Cat# NA04029
Ethylaluminium dichloride, 1.8M solution in toluene	Acros Organics	Cat# 428041000
Tetrabutylammonium fluoride, 1M solution in THF	Acros Organics	Cat# 433511000
Lipofectamine 2000	Thermo Fisher Scientific	Cat# 11668019
FuGENE 6	Promega	Cat# E2691
DNase I, Amplification Grade	Invitrogen	Cat# 18068015
RQ1 RNase-Free DNase	Promega	Cat# M610A
RNasin Plus RNase Inhibitor	Promega	Cat# N261A
TRIzol Reagent	Invitrogen	Cat# 15596026
Critical Commercial Assays		
Dual-Luciferase Reporter Assay System	Promega	Cat# E1910
Micro BCA Protein Assay Kit	Thermo Fisher Scientific	Cat# 23235

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Clarity Western ECL Substrate	Bio-Rad	Cat# 1705061
High-Capacity cDNA Reverse Transcription Kit	Applied Biosystems	Cat# 4368814
GoTaq qPCR Master Mix	Promega	Cat# A6002
Experimental Models: Cell Lines		
Human: HeLa cells (HeLa S3)	ATCC	CCL-2.2
Human: PA-1 cells	ATCC	CRL-1572
Human: HEK293T cells (HEK 293T/17)	ATCC	CRL-11268
Oligonucleotides		
qPCR primer GADPH-s (5'CGTTCCCAAAGTCCTCCTGT)	Sigma-Aldrich	N/A
qPCR primer GADPH-as (5'AGGTGATCGGTGCTGGTTC)	Sigma-Aldrich	N/A
qPCR and RT-qPCR primer EBNA-1-s (5'CGTCATCTCCGTCATCACC)	Sigma-Aldrich	N/A
qPCR and RT-qPCR primer EBNA-1-as (5'AGATTTGCCTCCCTGGTTTC)	Sigma-Aldrich	N/A
RT-qPCR primer GAPDH-s (5'TGCACCACCAACTGCTTAGC)	Sigma-Aldrich	N/A
RT-qPCR primer GAPDH-as (5'GGCATGGACTGTGGTCATGAG)	Sigma-Aldrich	N/A
RT-qPCR primer NEOjunct2-s (5'TGCCTCGTCCTGAAGCTC)	Sigma-Aldrich	N/A
RT-qPCR primer NEOjunct2-as (5'CAATCGGCTGCTCTGATG)	Sigma-Aldrich	N/A
Recombinant DNA		
Plasmid JM101/L1.3	Sassaman et al.1997. From John V. Moran Lab.	N/A
Plasmid pCEPL1SM	Han and Boeke, 2004. From Jef D. Boeke Lab.	N/A
Plasmid pCEP-TG_F21	Goodier et al., 2001. From Haig H. Kazazian Lab.	N/A
Plasmid pCEP-A101	Goodier et al., 2001. From Haig H. Kazazian Lab.	N/A
Plasmid pCMV-MusD-6neo^{TNF}	Ribet et al., 2004. From Thierry Heidmann Lab.	N/A
Plasmid pIAP-92L23neo^{TNF}	Dewannieux et al., 2004. From Thierry Heidmann Lab.	N/A
Plasmid pU6ineo	Richardson et al., 2014. From John V. Moran Lab.	N/A
Plasmid JJ101/L1.3	Kopera et al., 2011. From John V. Moran Lab.	N/A
Plasmid JJL1SM	MacLennan et al., 2017. From Jose L. Garcia-Perez Lab.	N/A
Plasmid pXY014	Xie et al., 2011. From Wenfeng An Lab.	N/A
Plasmid pXY015	Xie et al., 2011. From Wenfeng An Lab.	N/A
Software and Algorithms		
GraphPad Prism	GraphPad	https://www.graphpad.com/scientific-software/prism/
Mnova	Mestrelab	https://mestrelab.com/software/mnova/
Other		
PLC Silicagel 60 F_{254}, 0,5 mm	Merck	Cat# 105744
Discovery DSC-8 SPE Tube, 500 mg, 3 mL	Supelco	Cat# 52713-U

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jose L. Garcia-Perez (jose.garcia-perez@igmm.ed.ac.uk).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Line Authentication

The cell lines used in this study were originally obtained from ATCC. Their identity was confirmed by STR analyses at least once a year at Lorgen, Granada, Spain. Furthermore, the absence of *Mycoplasma spp*. was confirmed every month using a PCR-based assay (Minerva).

Cell Line Conditions

All cell lines used in this study were grown at 37°C, 5% CO_2 and atmospheric O_2, and passaged using Trypsin 0.05% (from Gibco, used for HeLa and PA-1 cells) or by pipetting up/down (for HEK293T cells).

HeLa Cells

HeLa cells (Moran et al., 1996) were grown on Dulbecco's Modified Eagle Medium-High Glucose (DMEM, Invitrogen) supplemented with 1x penicillin/streptomycin/glutamine (Invitrogen) and 10% Fetal Bovine Serum (FBS, Hyclone).

HEK293T Cells

HEK293T cells were grown on DMEM-High Glucose (Invitrogen) supplemented with 1x penicillin/streptomycin/glutamine (Invitrogen) and 10% FBS (Hyclone).

PA-1 Cells

PA-1 cells (Zeuthen et al., 1980) were cultured as described (Garcia-Perez et al., 2010), using Minimum Essential Media (MEM, Invitrogen) supplemented with 1x penicillin/streptomycin/glutamine (Invitrogen), 0.1 mM non-essential amino acids (Invitrogen), and 10% heat-inactivated FBS (Hyclone).

METHOD DETAILS

Chemicals Used in this Study

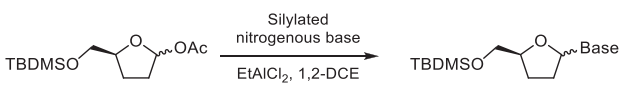
RTis used to treat HIV/AIDS (Abacavir, AZT, efavirenz, emtricitabine, etravirine, lamivudine, nevirapine, stavudine, and tenofovir) were obtained from Hospital Clinico San Cecilio, Granada, Spain. The name, provider and CAS number of commercially available nucleoside analogues used in this study are indicated in Table 1. The procedure for the synthesis of nucleoside analogues used in this study is included in Method details, and all compounds (GBS-147, 146, 197, 128, 187, 188, 186, 185, 149, 148, 127, 177, 196, 130, 179, 133, 192, and 191) were dissolved in water. Aliquots of all chemicals were prepared to avoid repeated freeze-thaw cycles.

Chemical Synthesis of 2',3'-dideoxynucleosides

Although some nucleoside analogues have been previously synthesized in other studies, below we describe the general procedures used to obtain the 18 nucleoside analogues tested in this study.

All solvents and chemicals were used as purchased without further purification. Thin layer chromatography was performed on pre-coated silica gel 60 F_{254} plates (Merck), and silica gel 60 (230-400 mesh, Aldrich) was used for flash column chromatography (FCC). Preparative TLC was developed on pre-coated silica gel 60 F_{254}, using 0.5 mm glass plates (Merck). Optical rotatory power was determined employing a JASCO DIP-370 polarimeter. NMR spectra are provided in Data S2, and were recorded using the following spectrometers: Varian Inova Unity 300 MHz, Varian Direct Drive 400 MHz, Varian Direct Drive 500 MHz and Varian Direct Drive 600 MHz. Chemical shifts (δ) are reported in parts per million relative to the residual peak of the deuterated solvent. High-resolution mass spectra were recorded on a Micromass LCT time-of-flight instrument using electrospray ionization (ESI). Low-resolution mass spectra (LRMS) were obtained operating in an electrospray ionization mode (ESI) coupled to high resolution liquid chromatography in a simple Quadrupole Agilent 6110 instrument, provided with a Zorbax Eclipse XDB-C18 4.6 x 150 mm column.

General Procedure for Sugar-Base Coupling

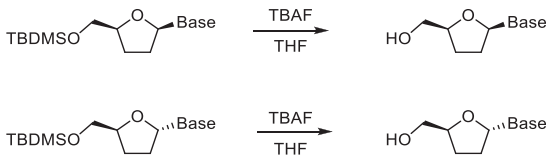


1 Protected nucleosides

Coupling reactions between sugar moiety and nitrogenous base were carried out as previously described (Okabe et al., 1988). A mixture of the nitrogenous base (0.73 mmol), HMDS (1.5 mL, 7.30 mmol), and $(NH_4)_2SO_4$ (cat.) was refluxed during 2 hours until a clear solution was obtained. The reaction mixture was cooled and the solvent evaporated in vacuum. The residue was dissolved

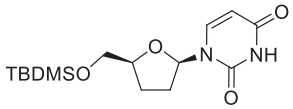
in dry 1,2-dichloroethane (3 mL), under argon atmosphere, and a solution of acetate 1 (100 mg, 0.36 mmol) in dry 1,2-dichloroethane (2 mL) was added. The reaction mixture was cooled in an ice-bath, treated with $EtAlCl_2$ in toluene (1.8 M in toluene, 222 μL, 0.4 mmol) and allowed to stir for 16 hours at room temperature. The reaction was then quenched with saturated aqueous $NaHCO_3$ solution and partitioned between CH_2Cl_2 and water. The organic layer was dried over Mg_2SO_4, filtered and concentrated. The residue was initially purified by FCC to obtain a mixture of isomeric nucleosides that was next separated by preparative thin layer chromatography on a glass plate. The silica was scraped off and percolated through a pad of silica gel using DCM/MeOH (90:10).

General Procedure for the Desilylation of Nucleosides



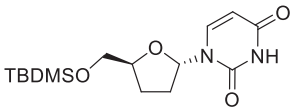
Each protected nucleoside (1.0 equiv) was dissolved in THF and treated with TBAF (1 M in THF, 1.5 equiv). After stirring during 15-30 minutes, the solvent evaporated under reduced pressure. The residue was dissolved in water and passed through a C8 SPE (Sulpelco, bed wt. 500 mg, volume 3 mL) eluting with water. Each tube was analyzed with an LC-MS instrument in order to collect those tubes containing the unprotected nucleoside and lacking tetrabutylammonium salts.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-uracil (GBS-109-P1). 1-[5'-O-(*tert*-butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-uracil (GBS-109-P1) has been previously synthesized (Mitsudo et al., 2005).



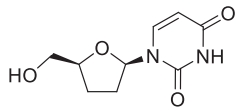
TLC diethyl ether/hexane (9:1); preparative TLC diethyl ether/hexane (1:1). Colorless syrup (28 mg, 23%). $[\alpha]_D^{28}$ +23.9° (*c* 1.2, CH_3Cl); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.21 (s, 1H), 8.08 (d, *J* = 8.1 Hz, 1H), 6.07 (dd, *J* = 6.5, *J* = 3.1 Hz, 1H), 5.65 (d, 1H, *J* = 8.1 Hz), 4.18 – 4.13 (m, 1H), 4.05 (dd, *J* = 11.5, *J* = 2.4 Hz, 1H), 3.70 (dd, *J* = 11.6, *J* = 2.4 Hz, 1H), 2.43 – 2.34 (m, 1H), 2.10 – 2.04 (m, 1H), 2.03 – 1.96 (m, 1H), 1.92 – 1.86 (m, 1H), 0.90 (s, 9H), 0.09 (2s, 6H) (Mitsudo et al., 2005); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 163.8, 150.5, 140.7, 101.5, 86.4, 82.2, 63.7, 33.6, 26.0, 24.4, 18.6, -5.33, -5.43; **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_{15}H_{27}N_2O_4Si$, 327.1740; found, 327.1746.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-uracil (GBS-109-P2). 1-[5'-O-(*tert*-butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-uracil (GBS-109-P2) has been previously synthesized (Mitsudo et al., 2005).



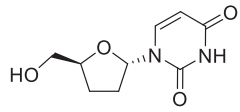
TLC diethyl ether/hexane (9:1); preparative TLC diethyl ether/hexane (1:1). White solid (22 mg, 18%). **mp:** 126-128°C; $[\alpha]_D^{29}$ -42.6° (*c* 1, CH_3Cl); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.15 (s, 1H), 7.35 (d, *J* = 8.1 Hz, 1H), 6.05 (dd, *J* = 6.2, *J* = 3.1 Hz, 1H), 5.72 (d, *J* = 8.2 Hz, 1H), 4.40 (m, 1H), 3.68 (dd, *J* = 10.9, *J* = 4.3 Hz, 1H), 3.63 (dd, *J* = 10.9, *J* = 4.1 Hz, 1H), 2.51 (m, 1H), 2.03 – 1.95 (m, 3H), 0.90 (s, 9H), 0.07 (s, 6H) (Mitsudo et al., 2005); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 163.6, 150.3, 139.4, 101.9, 88.1, 82.2, 65.4, 32.9, 26.0, 25.8, 18.4, -5.21, -5.27; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_{15}H_{26}N_2O_4NaSi$, 349.1560; found 349.1563.

2',3'-dideoxy-β-D-uridine (GBS-127). 2',3'-dideoxy-β-D-uridine (GBS-127) has been previously synthesized (Pfitzner and Moffatt, 1964).



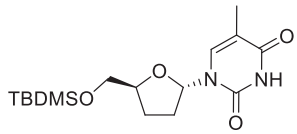
Obtained as a white solid (6 mg, quant.) from **GBS-109-P1** (9 mg, 0.027 mmol). TLC dichloromethane/methanol (9:1). **mp:** 122-124°C (Rassu et al., 1997); $[\alpha]_D^{28}$ +52.3 (*c* 0.5, MeOH) (Rassu et al., 1997); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.55 (s, 1H), 7.79 (d, *J* = 8.1 Hz, 1H), 6.09 (dd, *J* = 6.9, *J* = 3.8 Hz, 1H), 5.71 (d, 1H), 4.23 – 4.16 (m, 1H), 4.01 (dd, *J* = 11.8, *J* = 2.7 Hz, 1H), 3.74 (dd, *J* = 11.9, *J* = 3.8 Hz, 1H), 2.48 – 2.38 (m, 1H), 2.20 – 2.07 (m, 2H), 2.06 – 1.95 (m, 2H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 163.3, 150.3, 140.6, 102.1, 86.7, 81.7, 63.5, 32.7, 25.1; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_9H_{12}N_2O_4Na$, 235.0695; found 235.0683.

2',3'-dideoxy-α-D-uridine (GBS-177). 2',3'-dideoxy-α-D-uridine (GBS-177) has been previously synthesized (Kawakami et al., 1990)



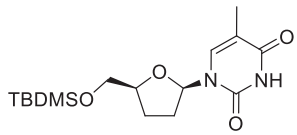
Obtained as a waxy solid (8 mg, quant.) from **GBS-109-P2** (12 mg, 0.037 mmol). TLC dichloromethane/methanol (9:1). **mp:** 96-98°C (Rassu et al., 1997); $[\alpha]_D^{28}$ -28.9 (*c* 0.5, MeOH) (Rassu et al., 1997); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.99 (s, 1H), 7.37 (d, *J* = 8.1 Hz, 1H), 6.10 (dd, *J* = 6.2, *J* = 4.5 Hz, 1H), 5.74 (d, *J* = 8.1 Hz, 1H), 4.50 – 4.42 (m, 1H), 3.76 (dd, *J* = 11.9, *J* = 3.2 Hz, 1H), 3.58 (dd, *J* = 11.9, *J* = 5.7 Hz, 1H), 2.60 – 2.47 (m, 1H), 2.12 – 2.01 (m, 3H), 1.99 – 1.88 (m, 1H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 163.4, 150.4, 139.5, 102.3, 87.7), 82.1, 64.8, 32.9, 25.9; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_9H_{12}N_2O_4Na$, 235.0695; found 235.0694.

*1-[5'-O-(tert-*butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-thymine (GBS-111-P1).



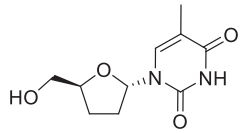
TLC diethyl ether/hexane (9:1); preparative TLC diethyl ether/hexane (1:1). Colorless syrup (16 mg, 13%). $[\alpha]_D^{28}$ -26.8 (*c* 0.7, CH_3Cl); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.70 (s, 1H), 7.15 (s, 1H), 6.07 (dd, *J* = 6.3, *J* = 3.8 Hz, 1H), 4.41 (m, 1H), 3.68 (dd, *J* = 10.9, *J* = 4.3 Hz, 1H), 3.63 (dd, *J* = 10.9, *J* = 4.1 Hz, 1H), 2.53 – 2.45 (m, 1H), 2.06 – 1.95 (m, 3H), 1.93 (s, 3H), 0.90 (s, 9H), 0.07 (s, 6H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 163.9, 150.2, 135.3, 110.4, 87.5, 81.9, 65.4, 32.8, 26.0, 26.0, 18.4, 12.8, -5.20, -5.26. **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_{16}H_{28}N_2O_4NaSi$, 363.1716; found 363.1717.

*1-[5'-O-(tert-*butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-thymine (GBS-111-P2). 1-[5'-O-(*tert-*butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-thymine (GBS-111-P2) has been previously synthesized (Camarasa et al., 1990).



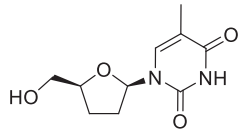
TLC diethyl ether/hexane (9:1); preparative TLC diethyl ether/hexane (1:1). White solid (28 mg, 22%). **mp:** 117-119°C (Camarasa et al., 1990); $[\alpha]_D^{29}$ +11.4 (*c* 1, CH_3Cl) (Takahashi et al., 1995); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.31 (s, 1H), 7.58 (d, *J* = 1.1 Hz, 1H), 6.07 (dd, *J* = 6.4, *J* = 4.5 Hz, 1H), 4.19 – 4.11 (m, 1H), 3.99 (dd, *J* = 11.4, *J* = 2.6 Hz, 1H), 3.71 (dd, *J* = 11.4, *J* = 2.9 Hz, 1H), 2.43 – 2.29 (m, 1H), 2.02 – 1.96 (m, 3H), 1.92 (d, *J* = 1.1 Hz, 3H), 0.93 (s, 9H), 0.11 (s, 6H) (Takahashi et al., 1995); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 164.1, 150.5, 135.8, 110.3, 85.9, 81.1, 64.6, 32.7, 26.0, 25.5, 18.6, 12.7, -5.17, -5.20 (Audat et al., 2012); **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_{16}H_{28}N_2O_4NaSi$, 363.1716; found 363.1724.

2',3'-dideoxy-α-D-thymidine (GBS-130). 2',3'-dideoxy-α-D-thymidine (GBS-130) has been previously synthesized (Agyei-Aye and Baker, 1988).



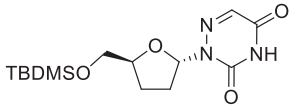
Obtained as a white solid (6 mg, quant.) from **GBS-111-P1** (9 mg, 0.026 mmol). TLC dichloromethane/methanol (9:1). **mp:** 102-104°C (Rassu et al., 1997); $[\alpha]_D^{29}$ -17.3 (*c* 0.2, MeOH) (Rassu et al., 1997); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.48 (s, 1H), 7.15 (s, 1H), 6.12 (t, *J* = 5.7 Hz, 1H), 4.48 (m, 1H), 3.76 (dd, *J* = 11.9, *J* = 3.2 Hz, 1H), 3.58 (dd, *J* = 11.9, *J* = 5.7 Hz, 1H), 2.55 – 2.47 (m, 1H), 2.14 – 1.99 (m, 2H), 1.98 – 1.88 (m, 1H), 1.94 (s, 3H) 1.63 (s, 1H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 163.7, 150.3, 135.2, 110.9, 87.2, 81.8, 64.9, 32.7, 26.1, 12.8. **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_{10}H_{14}N_2O_4Na$, 249.0851; found 249.0837.

2',3'-dideoxy-β-D-thymidine (GBS-196). 2',3'-dideoxy-β-D-thymidine (GBS-196) has been previously synthesized (Michelson and Todd, 1955).



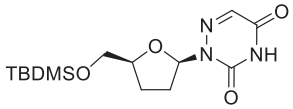
Obtained as a white solid (10 mg, quant.) from **GBS-111-P2** (15 mg, 0.026 mmol). TLC dichloromethane/methanol (9:1). **mp:** 150-152°C (Rassu et al., 1997); $[\alpha]_D^{28}$ +35.9 (*c* 0.6, MeOH) (Rassu et al., 1997); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.88 (s, 1H), 7.52 (d, *J* = 1.1 Hz, 1H), 6.10 (dd, *J* = 7.0, *J* = 4.0 Hz, 1H), 4.20 – 4.15 (m, 1H), 3.99 (dd, *J* = 12.0, *J* = 2.7 Hz, 1H), 3.73 (dd, *J* = 12.0, *J* = 4.0 Hz, 1H), 2.44 – 2.35 (m, 1H), 2.26 (s, 1H), 2.12 – 1.98 (m, 3H), 1.90 (d, *J* = 1.1 Hz, 3H) (Audat et al., 2012); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 164.0, 150.5, 136.4, 110.7, 86.3, 81.3, 63.6, 32.2, 25.3, 12.7 (Audat et al., 2012); **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_{10}H_{14}N_2O_4Na$, 249.0851; found 249.0866.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-6-azauracil (GBS-121-P1).



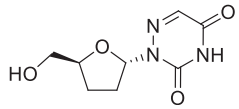
TLC diethyl ether/hexane (5:1); preparative TLC diethyl ether/hexane (1:2). Colorless syrup (28 mg, 23%). $[\alpha]_D^{28}$ +10.6 (*c* 1.1, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.35 (s, 1H), 7.43 (s, 1H), 6.44 (dd, *J* = 7.1, *J* = 2.7 Hz, 1H), 4.45 – 4.33 (m, 1H), 3.65 (dd, *J* = 10.8, *J* = 4.3 Hz, 1H), 3.61 (dd, *J* = 10.8, *J* = 4.0 Hz, 1H), 2.40 – 2.31 (m, 1H), 2.30 – 2.18 (m, 2H), 1.97 – 1.86 (m, 1H), 0.90 (s, 9H), 0.06 (s, 6H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 155.9, 147.8, 135.4, 87.7, 81.9, 65.5, 30.3, 26.8, 26.0, 18.5, -5.16, -5.25; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_{14}H_{25}N_3O_4NaSi$, 350.1512; found 350.1506.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-6-azauracil (GBS-121-P2).



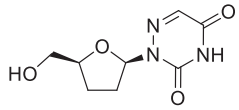
TLC diethyl ether/hexane (5:1); preparative TLC diethyl ether/hexane (1:2). Colorless syrup (36 mg, 30%). $[\alpha]_D^{30}$ -69.6 (*c* 0.5, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.03 (s, 1H), 7.42 (s, 1H), 6.38 (dd, *J* = 6.6, *J* = 3.5 Hz, 1H), 4.17 – 4.11 (m, 1H), 3.72 (dd, *J* = 10.6, *J* = 5.3 Hz, 1H), 3.65 (dd, *J* = 10.6, *J* = 5.5 Hz, 1H), 2.32 – 2.23 (m, 2H), 2.07 – 1.99 (m, 2H), 0.88 (s, 9H), 0.04 (2s, 6H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 155.7, 147.8, 135.3, 86.8, 82.3, 65.6, 30.1, 27.4, 26.0, 18.5, -5.13, -5.17; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_{14}H_{25}N_3O_4NaSi$, 350.1512; found 350.1515.

2',3'-dideoxy-α-D-6-azauridine (GBS-133).



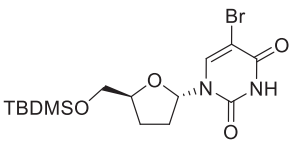
Obtained as a colorless syrup (14 mg, quant.) from **GBS-121-P1** (21 mg, 0.064 mmol). TLC dichloromethane/methanol (9:1). $[\alpha]_D^{28}$ +128.3 (*c* 1, MeOH); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.95 (s, 1H), 7.43 (s, 1H), 6.46 (t, *J* = 5.4 Hz, 1H), 4.46 – 4.39 (m, 1H), 3.74 (dd, *J* = 12.0, *J* = 2.9 Hz, 1H), 3.53 (dd, *J* = 12.0, *J* = 5.4 Hz, 1H), 2.58 (s, 1H), 2.39 – 2.32 (m, 2H), 2.29 – 2.20 (m, 1H), 1.91 – 1.82 (m, 1H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 156.2, 148.3, 135.7, 87.3, 81.9, 64.7, 30.0, 26.5; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_8H_{11}N_3O_4Na$, 236.0647; found 236.0658.

2',3'-dideoxy-β-D-6-azauridine (GBS-179). 2',3'-dideoxy-β-D-6-azauridine (GBS-179) has been previously synthesized (Rosowsky and Pai, 1991).



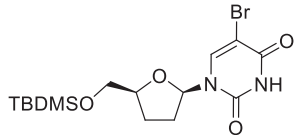
Obtained as a colorless syrup (5 mg, quant.) from **GBS-121-P2** (8 mg, 0.024 mmol). TLC dichloromethane/methanol (9:1). $[\alpha]_D^{28}$ -117.3 (*c* 0.2, MeOH); **^{1}H NMR** (600 MHz, $CDCl_3$) δ 8.80 (s, 1H), 7.46 (s, 1H), 6.40 (dd, *J* = 6.9, *J* = 3.8 Hz, 1H), 4.29 – 4.23 (m, 1H), 3.85 (dd, *J* = 12.0, *J* = 2.9 Hz, 1H), 3.61 (dd, *J* = 12.0, *J* = 4.9 Hz, 1H), 2.38 – 2.27 (m, 2H), 2.20 – 2.11 (m, 1H), 2.06 – 2.00 (m, 1H), 1.60 (s, 1H) (Rosowsky and Pai, 1991); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 155.4, 147.6, 135.8, 87.2, 82.4, 64.8, 31.0, 26.1; **HRMS** (*m/z*): $[M - H]^-$ calcd. for $C_8H_{10}N_3O_4$, 212.0671; found 212.0666.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-5-bromouracil (GBS-124-P1).



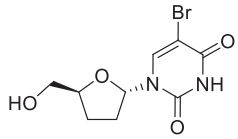
TLC diethyl ether/hexane (5:1); preparative TLC diethyl ether/hexane (1:1). White solid (24 mg, 16%). **mp:** 147-149°C; $[\alpha]_D^{28}$ -20.9 (*c* 0.7, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.84 (s, 1H), 7.67 (s, 1H), 6.02 (dd, *J* = 6.1, *J* = 3.0 Hz, 1H), 4.50 – 4.41 (m, 1H), 3.71 (dd, *J* = 11.0, *J* = 4.1 Hz, 1H), 3.64 (dd, *J* = 11.0, *J* = 4.0 Hz, 1H), 2.61 – 2.49 (m, 1H), 2.06 – 1.96 (m, 3H), 0.91 (s, 9H), 0.08 (s, 6H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 159.1, 149.5, 139.1, 96.3, 88.7, 82.4, 65.3, 33.2, 26.0, 25.6, 18.4, -5.21, -5.26; **HRMS** (*m/z*): $[M - H]^-$ calcd. for $C_{15}H_{24}N_2O_4SiBr$, 403.0689; found 403.0679.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-5-bromouracil (GBS-124-P2).



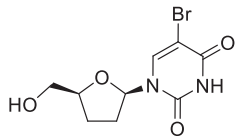
TLC diethyl ether/hexane (5:1); preparative TLC diethyl ether/hexane (1:1). White solid (32 mg, 21%). **mp:** 153-155°C; $[\alpha]_D^{29}$ -3.5 (*c* 1.8, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.44 (s, 1H), 8.17 (s, 1H), 6.02 (dd, *J* = 6.4, *J* = 4.3 Hz, 1H), 4.21 – 4.15 (m, 1H), 4.03 (dd, *J* = 11.6, *J* = 2.3 Hz, 1H), 3.70 (dd, *J* = 11.6, *J* = 2.7 Hz, 1H), 2.44 – 2.36 (m, 1H), 2.08 – 1.91 (m, 3H), 0.93 (s, 9H), 0.13 (s, 6H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 159.4, 149.9, 139.8, 96.4, 87.0, 81.9, 64.4, 33.3, 26.2, 25.0, 18.7, -5.07, -5.09; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_{15}H_{25}N_2O_4NaSiBr$, 427.0665; found 427.0671.

2',3'-dideoxy-α-D-5-bromouridine (GBS-128).



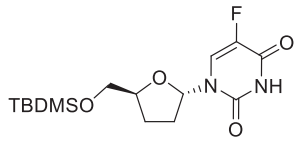
Obtained as a colorless syrup (13 mg, quant.) from **GBS-124-P1** (18 mg, 0.044 mmol). TLC dichloromethane/methanol (9:1). $[\alpha]_D^{28}$ -15.8 (*c* 1, MeOH); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.07 (s, 1H), 7.67 (s, 1H), 6.07 (dd, *J* = 6.1, *J* = 4.4 Hz, 1H), 4.54 – 4.48 (m, 1H), 3.79 (dd, *J* = 12.0, *J* = 3.2 Hz, 1H), 3.60 (dd, *J* = 12.0, *J* = 5.6 Hz, 1H), 2.62 – 2.51 (m, 1H), 2.14 – 2.01 (m, 2H), 2.00 – 1.90 (m, 1H), 1.70 (s, 1H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 159.2, 149.6, 139.1, 96.7, 88.2, 82.3, 64.7, 33.1, 25.7; **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_9H_{11}N_2O_4Br$, 312.9800; found 312.9785.

2',3'-dideoxy-β-D-5-bromouridine (GBS-197). 2',3'-dideoxy-β-D-5-bromouridine (GBS-197) has been previously synthesized (Furukawa et al., 1970).



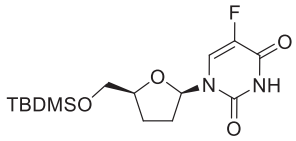
Obtained as a colorless syrup (12 mg, quant.) from **GBS-124-P2** (17 mg, 0.044 mmol). TLC dichloromethane/methanol (9:1). **mp:** 182-184°C (Furukawa et al., 1970); $[\alpha]_D^{27}$ +25.6 (*c* 0.7, MeOH); **^{1}H NMR** (600 MHz, CD_3OD) δ 8.64 (s, 1H), 6.00 (dd, *J* = 6.7, *J* = 2.9 Hz, 1H), 4.19 – 4.14 (m, 1H), 3.93 (dd, *J* = 12.3, *J* = 2.8 Hz, 1H), 3.68 (dd, *J* = 12.2, *J* = 3.2 Hz, 1H), 2.44 – 2.36 (m, 1H), 2.15 – 2.09 (m, 1H), 2.02 – 1.92 (m, 2H); **^{13}C NMR** (150 MHz, CD_3OD) δ 161.8, 151.6, 142.4, 96.2, 88.3, 84.0, 63.0, 34.0, 25.2; **HRMS** (*m/z*): $[M - H]^-$ calcd. for $C_9H_{10}N_2O_4Br$, 288.9824; found 288.9823.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-5-fluorouracil (GBS-135-P1).



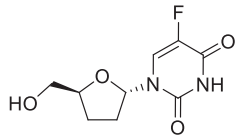
TLC diethyl ether/hexane (5:1); preparative TLC diethyl ether/hexane (1:1). White solid (19 mg, 15%). **mp:** 106-108°C (Lin et al., 1994); $[\alpha]_D^{29}$ -43.3 (*c* 1, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.53 (s, 1H), 7.42 (d, *J* = 6.1 Hz, 1H), 6.04 – 6.01 (m, 1H), 4.48 – 4.36 (m, 1H), 3.69 (dd, *J* = 11.0, *J* = 4.2 Hz, 1H), 3.63 (dd, *J* = 11.0, *J* = 4.1 Hz, 1H), 2.59 – 2.48 (m, 1H), 2.04 – 1.93 (m, 3H), 0.90 (s, 9H), 0.07 (s, 6H) (Lin et al., 1994); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 157.2 (d, *J* = 26.6 Hz), 148.9, 140.5 (d, *J* = 236.9 Hz), 123.8 (d, *J* = 33.7 Hz), 88.2, 82.3, 65.3, 32.9, 26.0, 25.7, 18.4, -5.23, -5.29; **HRMS** (*m/z*)=: $[M + Na]^+$ calcd. for $C_{15}H_{25}N_2O_4NaFSi$, 367.1465; found 367.1450.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-5-fluorouracil (GBS-135-P2).



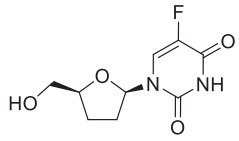
TLC diethyl ether/hexane (5:1); preparative TLC diethyl ether/hexane (1:1). White solid (35 mg, 28%). **mp:** 144-146°C (Lin et al., 1994); $[\alpha]_D^{29}$ +21.2 (*c* 1.25, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.00 (s, 1H), 8.26 (d, *J* = 6.4 Hz, 1H), 6.07 – 6.02 (m, 1H), 4.20 – 4.14 (m, 1H), 4.10 (dd, *J* = 11.7, *J* = 2.2 Hz, 1H), 3.71 (dd, *J* = 11.6, *J* = 2.1 Hz, 1H), 2.45 – 2.36 (m, 1H), 2.11 – 2.01 (m, 2H), 1.96 – 1.89 (m, 1H), 0.93 (s, 9H), 0.12 (2s, 6H) (Lin et al., 1994); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 157.1 (d, *J* = 27.1 Hz), 148.9, 140.3 (d, *J* = 235.5 Hz), 125.0 (d, *J* = 34.3 Hz), 86.7, 82.3, 63.9, 33.6, 26.1, 24.4, 18.7, -5.4; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_{15}H_{25}N_2O_4NaFSi$, 367.1465; found 367.1450.

2',3'-dideoxy-α-D-5-fluorouridine (GBS-191). 2',3'-dideoxy-α-D-5-fluorouridine (GBS-191) has been previously synthesized (Zhuk et al., 1979).



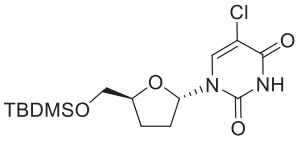
Obtained as a white solid (20 mg, quant.) from **GBS-135-P1** (30 mg, 0.087 mmol). TLC dichloromethane/methanol (9:1). **mp:** 132-134°C (Zhuk et al., 1979); $[\alpha]_D^{28}$ -41.7 (*c* 1.25, MeOH); **^{1}H NMR** (500 MHz, CD_3OD) δ 7.80 (d, *J* = 6.6 Hz, 1H), 6.06 – 6.02 (m, 1H), 4.51 – 4.45 (m, 1H), 3.64 (dd, *J* = 11.9, *J* = 3.8 Hz, 1H), 3.52 (dd, *J* = 11.9, *J* = 5.1 Hz, 1H), 2.51 – 2.42 (m, 1H, H), 2.13 – 2.03 (m, 2H), 1.93 – 1.84 (m, 1H); **^{13}C NMR** (125 MHz, CD_3OD) δ 159.6 (d, *J* = 26.0 Hz), 150.7, 141.8 (d, *J* = 232.7 Hz), 126.0 (d, *J* = 34.3 Hz), 89.1, 83.4, 65.1, 33.2, 26.7; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_9H_{11}N_2O_4FNa$, 253.0601; found 253.0604.

2',3'-dideoxy-β-D-5-fluorouridine (GBS-192). 2',3'-dideoxy-β-D-5-fluorouridine (GBS-192) has been previously synthesized (Khwaja and Heidelberger, 1967).



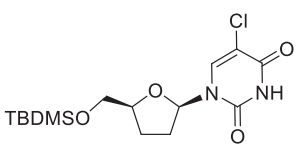
Obtained as a white solid (10 mg, quant.) from **GBS-135-P2** (15 mg, 0.043 mmol). TLC dichloromethane/methanol (9:1). **mp:** 123-125°C (Lin et al., 1994); $[\alpha]_D^{28}$ +59.4 (*c* 0.62, MeOH) (Lin et al., 1994); **^{1}H NMR** (500 MHz, CD_3OD) δ 8.39 (d, *J* = 7.0 Hz, 1H), 6.03 – 5.99 (m, 1H), 4.17 – 4.11 (m, 1H), 3.91 (dd, *J* = 12.3, *J* = 2.8 Hz, 1H), 3.68 (dd, *J* = 12.2, *J* = 3.3 Hz, 1H), 2.44 – 2.34 (m, 1H), 2.14 – 2.06 (m, 1H), 2.03 – 1.92 (m, 2H); **^{13}C NMR** (125 MHz, CD_3OD) δ 159.7 (d, *J* = 26.1 Hz), 150.8, 141.5 (d, *J* = 231.3 Hz), 126.6 (d, *J* = 35.0 Hz), 87.9, 83.7, 63.3, 33.7, 25.4; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_9H_{11}N_2O_4FNa$, 253.0601; found 253.0625.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-5-chlorouracil (GBS-136-P1).



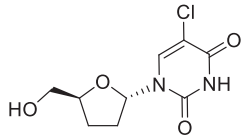
TLC diethyl ether/hexane (5:1); preparative TLC diethyl ether/hexane (1:1). White solid (24 mg, 18%). **mp** = 157-159°C; $[\alpha]_D^{28}$ -39.0 (*c* 0.1, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.01 (s, 1H), 7.56 (s, 1H), 6.02 (dd, *J* = 6.1, *J* = 3.1 Hz, 1H), 4.49 – 4.42 (m, 1H), 3.71 (dd, *J* = 10.9, *J* = 4.1 Hz, 1H), 3.64 (dd, *J* = 11, *J* = 4.0 Hz, 1H), 2.60 – 2.50 (m, 1H), 2.06 – 1.96 (m, 3H), 0.90 (s, 9H), 0.08 (s, 6H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 159.1, 149.3, 136.5, 108.7, 88.6, 82.4, 65.3, 33.1, 26.0, 25.6, 18.4, -5.21, -5.27; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_{15}H_{25}N_2O_4NaClSi$, 383.1170; found 383.1154.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-5-chlorouracil (GBS-136-P2).



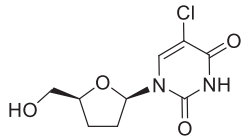
TLC diethyl ether/hexane (5:1); preparative TLC diethyl ether/hexane (1:1). White solid (44 mg, 33%). **mp** = 183-185°C; $[\alpha]_D^{28}$ +2.45 (*c* 1.1, CH_3Cl); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.93 (s, 1H), 8.15 (s, 1H), 6.03 (dd, *J* = 6.3, *J* = 4.0 Hz, 1H), 4.22 – 4.16 (m, 1H), 4.06 (dd, *J* = 11.6, *J* = 2.3 Hz, 1H), 3.71 (dd, *J* = 11.6, *J* = 2.5 Hz, 1H), 2.46 – 2.36 (m, 1H), 2.09 – 1.93 (m, 3H), 0.94 (s, 9H), 0.14 (2s, 6H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 159.1, 149.5, 137.4, 108.6, 87.0, 82.0, 64.3, 33.4, 26.1, 24.9, 18.7, -5.15, -5.16; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_{15}H_{25}N_2O_4NaClSi$, 383.1170; found 383.1159.

2',3'-dideoxy-α-D-5-chlorouridine (GBS-147).



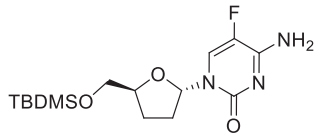
Obtained as a white solid (13 mg, quantitative) from **GBS-136-P1** (19 mg, 0.053 mmol). TLC (dichloromethane/methanol 9:1). **mp:** 165-167°C; $[\alpha]_D^{28}$ -26.50 (*c* 0.7, MeOH); **^{1}H NMR** (500 MHz, CD_3OD) δ 7.87 (s, 1H), 6.03 (dd, *J* = 6.3, *J* = 4.1 Hz, 1H), 4.53 – 4.47 (m, 1H), 3.66 (dd, *J* = 11.9, *J* = 3.7 Hz, 1H), 3.54 (dd, *J* = 11.9, *J* = 5.2 Hz, 1H), 2.54 – 2.43 (m, 1H), 2.14 – 2.05 (m, 2H), 1.95 – 1.86 (m, 1H); **^{13}C NMR** (126 MHz, CD_3OD) δ 161.7, 151.3, 138.9, 109.1, 89.6, 83.6, 65.1, 33.4, 26.7; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_9H_{11}N_2O_4NaCl$, 269.0305; found 269.0328.

2',3'-dideoxy-β-D-5-chlorouridine (GBS-146). 2',3'-dideoxy-β-D-5-chlorouridine (GBS-146) has been previously synthesized (Van Aerschot et al., 1990).



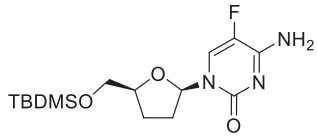
Obtained as a white solid (13 mg, quantitative) from **GBS-136-P2** (19 mg, 0.053 mmol). TLC (dichloromethane/methanol 9:1). **mp:** 159-161°C (Van Aerschot et al., 1990); $[\alpha]_D^{28}$ +40.87 (*c* 0.55, MeOH); **^{1}H NMR** (400 MHz, CD_3OD) δ 8.56 (s, 1H), 6.03 (dd, *J* = 6.7, *J* = 2.9 Hz, 1H), 4.23 – 4.14 (m, 1H), 3.95 (dd, *J* = 12.3, *J* = 2.8 Hz, 1H), 3.71 (dd, *J* = 12.3, *J* = 3.2 Hz, 1H), 2.49 – 2.36 (m, 1H), 2.20 – 2.10 (m, 1H), 2.05 – 1.94 (m, 2H); **^{13}C NMR** (100 MHz, CD_3OD) δ 161.8, 151.3, 139.8, 108.6, 88.2, 83.9, 63.0, 34.0, 25.2; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_9H_{11}N_2O_4NaCl$, 269.0305; found 269.0331.

1-[5'-O-(*tert*-butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-5-fluorocytosine (GBS-138-P1).



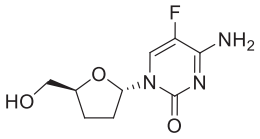
TLC dichloromethane/methanol (19:1); preparative TLC acetone/dichloromethane (1:1). Colorless syrup (37 mg, 29%). $[\alpha]_D^{29}$ -59.0 (*c* 1, MeOH); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.05 (s, 1H), 7.45 (d, *J* = 6.2 Hz, 1H), 6.02 – 5.94 (m, 1H), 5.54 (s, 1H), 4.45 – 4.35 (m, 1H), 3.69 – 3.60 (m, 2H), 2.59 – 2.48 (m, 1H), 2.07 – 1.97 (m, 1H), 1.97 – 1.87 (m, 2H), 0.89 (s, 9H), 0.06 (s, 6H) (Lin et al., 1994); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 158.1 (d, *J* = 13.5 Hz), 154.0, 136.5 (d, *J* = 240.9 Hz), 124.7 (d, *J* = 31.8 Hz), 88.8, 82.2, 65.4, 33.0, 26.0, 25.5, 18.4, -5.21, -5.25; **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_{15}H_{27}N_3O_3SiF$, 344.1806; found. 344.1823.

1-[5'-O-(*tert*-butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-5-fluorocytosine (GBS-138-P2).



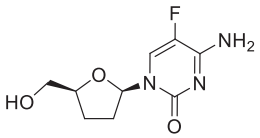
TLC dichloromethane/methanol (19:1); preparative TLC acetone/dichloromethane (1:1). White solid (40 mg, 32%). **mp:** 180-182°C (Lin et al., 1994); $[\alpha]_D^{29}$ +51.6 (*c* 0.75, MeOH); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.28 (d, *J* = 6.6 Hz, 1H), 7.51 (s, 1H), 6.05 – 5.99 (m, 1H), 5.41 (s, 1H), 4.17 – 4.13 (m, 1H), 4.11 (dd, *J* = 11.5, *J* = 2.3 Hz, 1H), 3.72 (dd, *J* = 11.6, *J* = 2.2 Hz, 1H), 2.48 – 2.38 (m, 1H), 2.15 – 2.06 (m, 1H), 2.01 – 1.89 (m, 1H), 1.88 – 1.78 (m, 1H), 0.93 (s, 9H), 0.12 (2s, 6H) (Lin et al., 1994); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 157.9 (d, *J* = 13.7 Hz), 154.1, 136.3 (d, *J* = 239.0 Hz), 126.3 (d, *J* = 32.2 Hz), 87.2, 82.5, 63.6, 33.8, 26.1, 24.0, 18.7, -5.39, -5.42; **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_{15}H_{27}N_3O_3SiF$, 344.1806; found 344.1824.

2',3'-dideoxy-α-D-5-fluorocytidine (GBS-185).



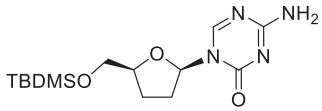
Obtained as a colorless syrup (12 mg, quantitative) from **GBS-138-P1** (18 mg, 0.052 mmol). TLC dichloromethane/methanol (9:1). $[\alpha]_D^{28}$ -86.1 (*c* 0.62, MeOH); **^{1}H NMR** (500 MHz, CD_3OD) δ 7.79 (d, *J* = 6.6 Hz, 1H), 6.02 – 5.97 (m, 1H), 4.52 – 4.45 (m, 1H), 3.64 (dd, *J* = 11.8, *J* = 3.9 Hz, 1H), 3.53 (dd, *J* = 11.8, *J* = 5.3 Hz, 1H), 2.55 – 2.46 (m, 1H), 2.08 – 1.96 (m, 2H), 1.92 – 1.83 (m, 1H); **^{13}C NMR** (125 MHz, CD_3OD) δ 159.6 (d, *J* = 13.9 Hz), 156.5, 138.4 (d, *J* = 242.0 Hz), 126.2 (d, *J* = 32.4 Hz), 89.8, 83.4, 65.2, 33.8, 26.5. **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_9H_{13}N_3O_3F$, 230.0941; found 230.0925.

2',3'-dideoxy-β-D-5-fluorocytidine (GBS-186). 2',3'-dideoxy-β-D-5-fluorocytidine (GBS-186) has been previously synthesized (Kim et al., 1987).



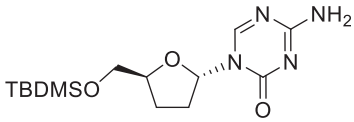
Obtained as a white foam (8 mg, quantitative) from **GBS-138-P2** (12 mg, 0.035 mmol). TLC dichloromethane/methanol (9:1). **mp:** 157-159°C (Lin et al., 1994); $[\alpha]_D^{28}$ +96.4 (*c* 0.5, MeOH) (Lin et al., 1994); **^{1}H NMR** (500 MHz, CD_3OD) δ 8.41 (d, *J* = 7.0 Hz, 1H), 5.99 – 5.94 (m, 1H), 4.18 – 4.12 (m, 1H), 3.93 (dd, *J* = 12.2, *J* = 2.8 Hz, 1H), 3.70 (dd, *J* = 12.3, *J* = 3.3 Hz, 1H), 2.47 – 2.36 (m, 1H), 2.09 – 2.01 (m, 1H), 1.96 – 1.88 (m, 2H); **^{13}C NMR** (125 MHz, CD_3OD) δ 159.5 (d, *J* = 14.1 Hz), 156.5, 138.2 (d, *J* = 240.7 Hz), 127.2 (d, *J* = 33.3 Hz), 88.6, 83.9, 63.2, 34.3, 25.1; **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_9H_{13}N_3O_3F$, 230.0941; found 230.0923.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-5-azacytosine (GBS-139-P1). 1-[5'-O-(*tert*-butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-5-azacytosine (GBS-139-P1) has been previously synthesized (Lin et al., 1995).



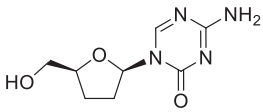
TLC dichloromethane/methanol (19:1); preparative TLC dichloromethane/methanol (97:3). White solid (26 mg, 22%). **mp** = 183-185°C (Lin et al., 1995); $[\alpha]_D^{28}$ +53.9 (*c* 0.25, CH_3Cl); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.72 (s, 1H), 6.64 (s, 1H), 6.02 (dd, *J* = 6.6, *J* = 2.4 Hz, 1H), 5.65 (s, 1H), 4.23 – 4.16 (m, 1H), 4.08 (dd, *J* = 11.6, *J* = 2.5 Hz, 1H), 3.72 (dd, *J* = 11.6, *J* = 2.4 Hz, 1H), 2.54 – 2.42 (m, 1H), 2.20 – 2.13 (m, 1H), 2.01 – 1.92 (m, 1H), 1.91 – 1.84 (m, 1H), 0.92 (s, 9H), 0.11 (2s, 6H) (Lin et al., 1995); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 166.6, 156.2, 154.3, 87.5, 82.9, 63.4, 33.9, 26.1, 23.9, 18.6, -5.28, -5.39; **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_{14}H_{27}N_4O_3Si$, 327.1852; found 327.1862.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-5-azacytosine (GBS-139-P2). 1-[5'-O-(*tert*-butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-5-azacytosine (GBS-139-P2) has been previously synthesized (Lin et al., 1995).



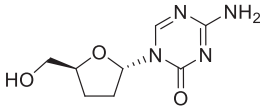
TLC dichloromethane/methanol (19:1); preparative TLC dichloromethane/methanol (97:3). White solid (26 mg, 22%). **mp** = 197-199°C (Lin et al., 1995); $[\alpha]_D^{28}$ -50.7 (*c* 0.6, CH_3Cl); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.10 (s, 1H), 6.85 (s, 1H), 6.01 (dd, *J* = 6.2, *J* = 2.9 Hz, 1H), 5.74 (s, 1H), 4.48 – 4.39 (m, 1H), 3.69 (dd, *J* = 10.9, *J* = 4.3 Hz, 1H), 3.64 (dd, *J* = 10.9, *J* = 4.2 Hz, 1H), 2.63 – 2.52 (m, 1H), 2.14 – 2.04 (m, 1H), 2.00 – 1.92 (m, 2H), 0.90 (s, 9H), 0.07 (s, 6H) (Lin et al., 1995); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 166.6, 154.8, 154.0, 89.1, 82.5, 65.3, 33.1, 26.0, 25.3, 18.4, -5.20, -5.26; **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_{14}H_{27}N_4O_3Si$, 327.1852; found 327.1841.

2',3'-dideoxy-β-D-5-azacytidine (GBS-149). 2',3'-dideoxy-β-D-5-azacytidine (GBS-149) has been previously synthesized (Lin et al., 1995).



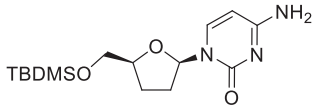
Obtained as a white solid (8 mg, quantitative) from **GBS-139-P1** (12 mg, 0.037 mmol). TLC dichloromethane/methanol (9:1). **mp:** 240°C (Lin et al., 1995); $[\alpha]_D^{29}$ +46.3 (*c* 0.1, MeOH) (Lin et al., 1995); **^{1}H NMR** (300 MHz, D_2O) δ 8.63 (s, 1H), 6.02 (dd, *J* = 6.9, *J* = 2.3 Hz, 1H), 4.36 – 4.25 (m, 1H), 3.93 (dd, *J* = 12.7, *J* = 3.0 Hz, 1H), 3.76 (dd, *J* = 12.7, *J* = 4.9, 1H), 2.59 – 2.43 (m, 1H), 2.28 – 2.17 (m, 1H), 2.11 – 1.99 (m, 1H), 1.93 – 1.75 (m, 1H); **^{13}C NMR** (150 MHz, CD_3OD) δ 168.0, 157.4, 156.6, 88.8, 84.4, 63.1, 34.3, 25.0; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_8H_{12}N_4O_3Na$, 235.0807; found 235.0809.

2',3'-dideoxy-α-D-5-azacytidine (GBS-148). 2',3'-dideoxy-α-D-5-azacytidine (GBS-148) has been previously synthesized (Lin et al., 1995).



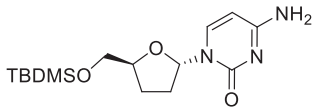
Obtained as a white solid (8 mg, quantitative) from **GBS-139-P2** (12 mg, 0.037 mmol). TLC dichloromethane/methanol (9:1). **mp:** 145-147°C (Lin et al., 1995); $[\alpha]_D^{29}$ -52.8 (*c* 0.2, MeOH) (Lin et al., 1995); **^{1}H NMR** (500 MHz, CD_3OD) δ 8.26 (s, 1H), 5.99 (dd, *J* = 6.3, *J* = 3.5 Hz, 1H), 4.53 – 4.48 (m, 1H), 3.65 (dd, *J* = 11.9, *J* = 3.9 Hz, 1H), 3.54 (dd, *J* = 11.9, *J* = 5.2 Hz, 1H), 2.57 – 2.49 (m, 1H), 2.16 – 2.03 (m, 2H), 1.95 – 1.88 (m, 1H); **^{13}C NMR** (125 MHz, CD_3OD) δ 168.2, 156.62, 156.59, 90.1, 83.7, 65.1, 33.7, 26.4; **HRMS** (*m/z*): $[M + Na]^+$ calcd. for $C_8H_{12}N_4O_3Na$, 235.0807; found 235.0811.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-cytosine (GBS-145-P1). 1-[5'-O-(*tert*-butyldimethylsilyl)-2',3'-dideoxy-β-D-ribofuranosyl]-cytosine (GBS-145-P1) has been previously synthesized (Okabe et al., 1988).



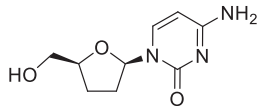
TLC dichloromethane/methanol (19:1); preparative TLC chloroform/isopropanol (90:10). White solid (27 mg, 23%). **mp:** 198-200°C (Okabe et al., 1988); $[\alpha]_D^{28}$ +41.1 (*c* 1, MeOH) (Okabe et al., 1988); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.10 (d, *J* = 7.4 Hz, 1H), 6.07 (dd, *J* = 6.6, *J* = 2.8 Hz, 1H), 5.63 (d, *J* = 7.3 Hz, 1H), 4.15 – 4.09 (m, 1H), 4.04 (dd, *J* = 11.5, *J* = 2.6 Hz, 1H), 3.71 (dd, *J* = 11.5, *J* = 2.6 Hz, 1H), 2.46 – 2.35 (m, 1H), 2.10 – 2.02 (m, 1H), 1.95 – 1.88 (m, 1H), 1.87 – 1.79 (m, 1H), 0.91 (s, 9H), 0.09 (2s, 6H) (Okabe et al., 1988); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 165.9, 156.1, 141.8, 93.3, 87.1, 82.2, 63.7, 33.9, 26.0, 24.3, 18.5, -5.29, -5.39; **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_{15}H_{28}N_3O_3Si$, 326.1900; found 326.1891.

1-[5'-O-(tert-butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-cytosine (GBS-145-P2). 1-[5'-O-(*tert*-butyldimethylsilyl)-2',3'-dideoxy-α-D-ribofuranosyl]-cytosine (GBS-145-P2) has been previously synthesized (Okabe et al., 1988).



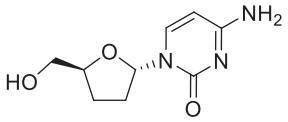
TLC dichloromethane/methanol (19:1); preparative TLC chloroform/isopropanol (90:10). White solid (25 mg, 21%). **mp:** 186-188°C (Okabe et al., 1988);$[\alpha]_D^{28}$ -50.0 (*c* 0.75, MeOH) (Okabe et al., 1988); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.43 (d, *J* = 7.4 Hz, 1H), 6.04 (dd, *J* = 6.1, *J* = 3.3 Hz, 1H), 5.74 (d, *J* = 7.4 Hz, 1H), 4.42 – 4.34 (m, 1H), 3.64 (d, *J* = 4.6 Hz, 2H), 2.56 – 2.47 (m, 1H), 2.06 – 1.95 (m, 1H), 1.94 – 1.87 (m, 2H), 0.89 (s, 9H), 0.06 (s, 6H) (Okabe et al., 1988); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 166.0, 156.1, 140.3, 93.8, 88.7, 82.0, 65.5, 33.0, 26.0, 25.5, 18.5, -5.19, -5.24; **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_{15}H_{28}N_3O_3Si$, 326.1900; found 326.1890.

2',3'-dideoxy-β-D-cytidine (GBS-187). 2',3'-dideoxy-β-D-cytidine (GBS-187) has been previously synthesized (Horwitz et al., 1967).



Obtained as a white solid (9 mg, quantitative) from **GBS-145-P1** (14 mg, 0.043 mmol). TLC dichloromethane/methanol (9:1). **mp:** 223-225°C (Rassu et al., 1997);$[\alpha]_D^{27}$ +82.1 (*c* 0.77, MeOH) (Rassu et al., 1997); **^{1}H NMR** (500 MHz, CD_3OD) δ 8.12 (d, *J* = 7.5 Hz, 1H), 6.02 (dd, *J* = 6.7, *J* = 3.2 Hz, 1H), 5.87 (d, *J* = 7.5 Hz, 1H), 4.18 – 4.11 (m, 1H), 3.87 (dd, *J* = 12.2, *J* = 3.1 Hz, 1H), 3.69 (dd, *J* = 12.2, *J* = 4.0 Hz, 1H), 2.46 – 2.37 (m, 1H), 2.05 – 1.83 (m, 3H) (Sivets et al., 2002); **^{13}C NMR** (125 MHz, CD_3OD) δ 167.6, 158.3, 142.8, 95.2, 88.4, 83.6, 63.7, 34.1, 25.7; **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_9H_{14}N_3O_3$, 212.1035; found 212.1029.

2',3'-dideoxy-α-D-cytidine (GBS-188). 2',3'-dideoxy-α-D-cytidine (GBS-188) has been previously synthesized (Okabe et al., 1988).



Obtained as a white solid (9 mg, quantitative) from **GBS-145-P2** (14 mg, 0.043 mmol). TLC dichloromethane/methanol (9:1). **mp:** 168-170°C (Rassu et al., 1997); $[\alpha]_D^{27}$ -74.5 (*c* 0.77, MeOH) (Okabe et al., 1988); **^{1}H NMR** (500 MHz, CD_3OD) δ 7.68 (d, *J* = 7.5 Hz, 1H), 6.05 (dd, *J* = 6.2, *J* = 3.7 Hz, 1H), 5.91 (d, *J* = 7.5 Hz, 1H), 4.49 – 4.44 (m, 1H), 3.63 (dd, *J* = 11.8, *J* = 4.0 Hz, 1H), 3.54 (dd, *J* = 11.8, *J* = 5.2 Hz, 1H), 2.54 – 2.45 (m, 1H), 2.07 – 1.96 (m, 2H), 1.93 – 1.84 (m, 1H); **^{13}C NMR** (125 MHz, CD_3OD) δ 167.56, 158.05, 142.10, 95.57, 89.69, 83.38, 65.22, 33.83, 26.51; **HRMS** (*m/z*): $[M + H]^+$ calcd. for $C_9H_{14}N_3O_3$, 212.1035; found 212.1028.

Expression Plasmids

Plasmids used in this study were propagated in *Escherichia coli* strain DH5α (ThermoFisher) and were purified using Plasmid Midi kits from Qiagen following manufacturer instructions. Plasmid DNAs were analysed by electrophoresis (0.7% agarose-ethidium bromide gels) and we only used highly supercoiled preparations of plasmid DNA for transfections.

***JM101/L1.3**

(Human L1.3), has been described previously (Sassaman et al., 1997). It contains a full-length copy of the human L1.3 element (L1.3, accession number #L19088) tagged with the *mneoI* indicator cassette (Freeman et al., 1994; Moran et al., 1996); it is cloned in vector pCEP4 (Life Technologies).

***pCEPL1SM**

(Mouse L1-T_F) has been described previously (Han and Boeke, 2004). It contains a full-length mouse T_F LINE-1 element (L1Md-$L1_{Orl}$, (Takahara et al., 1996)) where the coding sequence of the LINE-1 ORFs (L1-ORF1 and L1-ORF2) has been codon optimized, and is tagged with the *mneoI* indicator cassette (Freeman et al., 1994). It is cloned in vector pCEP4 (Life Technologies).

***pCEP-TG$_F$21**

(Mouse L1-G_F) has been described previously (Goodier et al., 2001). It contains a full-length mouse G_F LINE-1 element (L1Md-G_F21, accession number #AC021631.6, positions 62229-68991) tagged with the *mneoI* indicator cassette (Freeman et al., 1994). It is cloned in vector pCEP4 (Life Technologies).

***pCEP-A101**

(Mouse L1-A) has been described previously (Goodier et al., 2001). It contains a full-length mouse A LINE-1 element (L1Md-A101, accession number #AY053455) tagged with the *mneoI* indicator cassette (Freeman et al., 1994). It is cloned in vector pCEP4 (Life Technologies).

***pCMV-MusD-6neo^{TNF}**

(Mouse MusD) has been described previously (Ribet et al., 2004). It contains a nearly full-length copy of a mouse MusD element (lacks the U3 sequence from the 5′LTR, accession number #AC124426, positions 9078–16,569 (+)) tagged with the neo^{TNF} indicator cassette (Esnault et al., 2002) and is cloned in vector pCMVbeta (Clontech).

***pIAP-92L23neo^{TNF}**

(Mouse IAP) has been described previously (Dewannieux et al., 2004). It contains a full-length copy of a mouse IAP element (accession number # AC012382, positions 161,601–168,684, (+)) tagged with the neo^{TNF} indicator cassette (Esnault et al., 2002) at IAP nucleotide position 5744, and is cloned in vector pGL3basic (Promega).

***pU6ineo**

**pU6ineo has been described previously (Richardson et al., 2014). It contains the neomycin phosphotransferase (NEO) expression cassette from pEGFP-N1 (Clontech) cloned into a modified pBSKS-II(+) (Stratagene) that contains a U6 promoter in the multi-cloning site.

***JJ101/L1.3**

(Human L1.3) has been described previously (Kopera et al., 2011). It contains a full-length copy of the human L1.3 element (L1.3, accession number #L19088) tagged with the *mblastI* indicator cassette (Goodier et al., 2007; Morrish et al., 2002) and is cloned in pCEP4 (Life Technologies).

***JJL1SM**

(Mouse L1-T_F) has been described previously (MacLennan et al., 2017). It contains a full-length mouse T_F LINE-1 element (L1Md-$L1_{Orl}$, (Takahara et al., 1996)) where the coding sequence of the LINE-1 ORFs (L1-ORF1 and L1-ORF2) has been codon optimized, and is tagged with the *mblastI* indicator cassette (Goodier et al., 2007; Morrish et al., 2002). It is cloned in vector pCEP4 (Life Technologies).

***pXY014**

(Human $L1_{RP}$) has been described previously (Xie et al., 2011). It contains a full-length copy of the human $L1_{RP}$ element (accession number #AF148856.1, (Kimberland et al., 1999)) tagged with the *mflucI* indicator cassette (Xie et al., 2011) and is cloned in a modified pCEP4 (Life Technologies) that contains a Renilla firefly expression cassette.

**pXY015

**pXY015 has been described previously (Xie et al., 2011). It is derived from plasmid pXY014 but the cloned $L1_{RP}$ element contains two missense mutations in the RNA binding domain of L1-ORF1p (RR261/62AA). This plasmid was used as a negative control of the luciferase-based retrotransposition assays.

Retrotransposition Assays

All retrotransposition and clonability assays conducted in this study were conducted at least in duplicate and several independent times (>3). Clonability and retrotransposition assays using *mneoI* or neo^{TNF} tagged retrotransposons on HeLa cells were carried out as previously described (Benitez-Guijarro et al., 2018; Heras et al., 2013; Wei et al., 2000). HeLa cells were plated in 6-well plates (Corning) at the indicated number ($2x10^4$ cells for plasmids JM101/L1.3, pCEP-TG_F21, pCEP-A101, and pCMV-MusD-$6neo^{TNF}$; $1x10^4$ cells for plasmids pCEPL1SM, pIAP-92L23neo^{TNF}, and pU6ineo). Eighteen hours after plating, DNA transfections were carried out using FuGene 6 transfection reagent (Promega) and Opti-MEM (Life Technologies) following the protocol provided by the manufacturer (for a well of a 6-well plate: 3 μL of FuGene, 97 μL of Opti-MEM and 1 μg of plasmid DNA). The day after transfection, media was replaced with fresh media containing the indicated amount of each RTi (0, 5, and 25 μM unless otherwise indicated). Neomycin selection was started 72 h post-transfection using 400 μg/mL G418 (Life Technologies) and the indicated amount of each RTi. Selection media containing the indicated RTi was replaced every other day and selection was continued for 11 additional days. HeLa cells were then washed with 1x PBS (Gibco), fixed with 2% paraformaldehyde/0.4% glutaraldehyde, and stained with 0.1% (w/v) crystal violet solution as described (Moran et al., 1996; Wei et al., 2000), to visualize and count foci representing successful retrotransposition events.

Retrotransposition assays using *mblastI* tagged LINE-1 vectors (JJ101/L1.3 and JJL1SM) on HeLa cells were carried out as previously described (Benitez-Guijarro et al., 2018; Heras et al., 2013; Morrish et al., 2007). $2x10^4$ (JJ101/L1.3) or $1x10^4$ (JJL1SM) HeLa cells were plated per well of a 6-well tissue culture plate and transfected 18h later using 1 μg of each plasmid and 3 μL of FuGene 6 as described above. The day after transfection, media was replaced with fresh media containing the indicated amount of each RTi (0, 5, and 25 μM unless otherwise indicated), and cells cultured for 5 additional days changing the media every other day. Blasticidin S selection (5 μg/mL, Life Technologies) in the presence of the indicated RTi was started 120 h post-transfection, and selection was continued for 9 additional days. HeLa cells were then washed with 1x PBS (Gibco), fixed with 2% paraformaldehyde/0.4% glutaraldehyde, and stained with 0.1% (w/v) crystal violet solution as described (Moran et al., 1996; Wei et al., 2000), to visualize and count foci representing successful retrotransposition events.

Retrotransposition assays using *mflucI* tagged LINE-1 constructs (pX014 and pX015 as a negative control) were carried out in HEK293T cells as previously described (Benitez-Guijarro et al., 2018). $1x10^5$ cells per well were plated on 24-well tissue culture plates (Corning), and cells transfected 16-18 hours later using Lipofectamine 2000 (Invitrogen) and Opti-MEM (Gibco), and following the protocol provided by the manufacturer. Prior to transfection, the culture media was replaced with antibiotic-free culture media, and 200 ng of each LINE-1 construct and 1 μL of Lipofectamine 2000 were used per well (of a 24-well tissue culture plate). 6 h after transfection, media was replaced with fresh complete media containing the indicated amount of each RTi (0, 2,5, 5, 10, 15, and 20 μM unless otherwise indicated). Puromycin selection (1 μg/mL, Sigma) was started 24 h post-transfection, using culture media containing the indicated amount of each RTi. Selection was continued for 3 additional days, and then Firefly and Renilla luciferase activities were measured using the Dual-Luciferase Reporter Assay System (Promega) following manufacturer's instructions and a GloMax-Multi Detection System (Promega). Luciferase values for RTi untreated wells were designated as 1.

MTT Assays

Assays were conducted in triplicate using $1x10^3$ cells/well (HeLa or PA-1 cells) plated on 96-well tissue culture plates (Corning). 16-18 h after plating, culture media was replaced with fresh media containing the indicated amount of each RTi (0, 5, and 25 μM unless otherwise indicated), and cells cultured for 72 h. MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, Life Technologies) was dissolved in 1x PBS (5 mg/mL), and 10 μL/well were added to 96-well tissue culture plates. Plates were incubated during 3 hours at 37°C, media removed, and 100 μL DMSO were then added to each well, followed by an incubation at 37°C during 15 minutes. After the incubation, the absorbance was measured using a GloMax-Multi Detection System (Promega) with a test wavelength of 560 nm and a reference wavelength of 750 nm to obtain sample signal (OD_{570}-OD_{630}). Blank wells for background containing only media were included as controls, as well as blank wells containing each RTi and culture media (to discard the possible contribution of RTis to absorbance values). Data is represented as the mean values with standard deviation, where untreated cells were arbitrarily designated as 1 for comparisons.

qPCR and RT-qPCR Control Assays

To determine whether RTi treatments affect the stability of LINE-1 constructs in cultured cells, we transfected a human L1 overexpression plasmid (JM101/L1.3) in HeLa cells, and we then quantified the amount of plasmid DNA after RTi treatment (see Figures S5K–S5M). $8x10^4$ HeLa cells were plated per well of a 6-well tissue culture plate in triplicate; after 16-18 hours, cells were transfected with 1 μg of plasmid JM101/L1.3 using 3 μL of FuGene 6 (see above). After 24 h, cells were feed with fresh media containing 25 μM of the indicated RTi (GBS-149, Emtricitabine, Lamivudine, Tenofovir, C5, and AZT) or vehicle, and cells cultured for 48h. After 48h, genomic DNA from transfected cells was extracted using phenol:chloroform:isoamyl alcohol (25:24:1) as described (Benitez-Guijarro et al., 2018). Next, 50 ng of each extracted DNA were used in qPCR reactions, using a StepOnePlus Real-Time PCR System

(Applied Biosystems), GoTaq qPCR Master Mix Kit (Promega) and the following program: 1x (95°C, 10 min); 40x (95°C, 15 s; 60°C, 60 s). To quantify the relative amount of L1 plasmids after RTi treatments, we amplified the unique EBNA-1 region included in plasmid JM101/L1.3 (Figure S5K) and a portion of the human GAPDH gene to normalize for copy number differences, and using pair of primers whose amplification efficiencies do not differ more than 5%. To do that, we first generated a standard curve by serially diluting genomic DNA or plasmid JM101/L1.3, starting with 1 μg and making five five-fold dilutions. We then quantified copy numbers from three technical replicates of each reaction using the StepOnePlus Software v2.3. We used the EBNA-1/GAPDH ratio to calculate differences in copy number. Untransfected controls were used to discard plasmid DNA contaminations.

Similarly, and to analyse if RTi treatments affect the expression of L1 RNAs from transfected plasmids, we transfected a human L1 overexpression plasmid (JM101/L1.3) in HeLa cells, and we then quantified the amount of expressed L1 RNAs after RTi treatment (see Figures S5K–S5N). To quantify expressed L1 RNAs, we used: i) a NEO primer pair designed to only amplify spliced *mneoI*-tagged L1 transcripts; ii) a EBNA-1 primer pair designed to detect expressed EBNA-1 RNAs from transfected plasmids; and iii) GAPDH to normalize for expression differences (Figure S5L). Briefly, HeLa cells were plated, transfected and treated with RTis as described above, and total RNA was isolated using TRIzol (Invitrogen). Next, 1 μg of total RNA from each treatment condition (GBS-149, Emtricitabine, Lamivudine, Tenofovir, C5, AZT, or vehicle) was treated twice with 10 units of RNase-free DNaseI (Invitrogen), to completely remove the transfected plasmid DNA from extracted RNAs. cDNAs were then synthesized using a High-Capacity cDNA Reverse Transcription kit (Applied Biosystems) following the instructions provided by the manufacturer. Diluted cDNAs (1/5 and 1/10, and in triplicate) were then analyzed using a StepOne Real-Time PCR System (Applied Biosystems), GoTaq qPCR MasterMix (Promega) and 0.15 μM of each primer (Sigma). We included an internal control (no RT added) in all subsequent qPCR reactions. The qPCR cycling conditions were: 1x (95°C, 10 min); 40x (95°C, 15 s; 60°C, 60 s); a melting curve was recorded to confirm the identity of amplified products. To calculate differences in L1 RNA expression from transfected plasmids after RTi treatments, we used the NEO/EBNA-1 ratio and the comparative C_T ($\Delta\Delta C_T$) method (Livak and Schmittgen, 2001).

Western Blot Analyses

PA-1 cells were plated ($2x10^5$ cells/well of a 6-well tissue culture plate) and treated with 25 μM of the indicated RTi during 96 h. After the treatment, cells were harvested using a cell scrapper and duplicate Whole Cell Extracts (WCE) prepared using RIPA buffer (Sigma) supplemented with 1x Complete Mini EDTA-free Protease Inhibitor cocktail (Roche), 0.1% Phosphatase Inhibitor 1&2 (Sigma), 1 mM (PMSF) (Sigma) and 0.25% β-mercaptoethanol (Sigma), by incubating cells during 10 min on ice. Cellular debris was removed by centrifugation (1,000 g for 5 min at 4°C) and total protein concentration was determined using the Micro BCA Protein Assay Kit (Thermo) following standard procedures. Equal amounts of protein lysates were run on 10% SDS-PAGE gels and transferred to nitrocellulose membranes (Bio-Rad). Membranes were blocked in 5% milk/1x TBST (TBS + 0.2% Tween-20 (v/v)) and incubated with primary antibodies diluted in 5% milk/1xTBST overnight at 4°C. Membranes were then washed 3 times with 1xTBST, incubated with secondary antibodies for 1 h at RT, washed again and developed. A chemiluminescent detection system (Clarity Western ECL Substrate, Bio-Rad; ImageQuant LAS 4000, GE Healthcare) was used, following manufacturer's recommendations. The following antibodies were used for immunoblotting (at indicated dilutions): mouse anti-L1Hs-ORF1p (1:1000, Merck Millipore); mouse anti-p53 (1:300; Santa Cruz Biotechnology); mouse anti–β-actin (1:20,000; Sigma); horse anti-mouse HRP-linked secondary antibody (1:20,000, Cell Signaling Technology).

QUANTIFICATION AND STATISTICAL ANALYSIS

We used GraphPad Prism (version 6; GraphPad Software, Inc.) for statistical analyses. Data from multiple independent experiments is reported as retrotransposition mean ± SD, and we arbitrarily designed untreated cells as 100% Retrotransposition. Student t-test was used to calculate statistical significance and * indicates $p< 0.05$; **, $p< 0.01$, ***, $p< 0.001$, and ****, $p< 0.0001$.