Manuscript for Oncogene

Mobile retroelements induced by hypomethylating agents are restricted to transpose in myeloid malignancies

Šárka Pavlová[1,2,3], Marcela Krzyžánková[1], Anastasiya Volakhava[2], Anastasia Smirnova[4,5], Tatiana Grigoreva[4,6], Zuzana Jašková[1], Hana Synáčková[1], Dennis Wahl[1*], Michaela Bohúnová[1], Libor Červinek[1], Šárka Pospíšilová[1,2,3], Ilgar Mamedov[2], and Karla Plevová[1,2,3,#]

[1]Department of Internal Medicine - Hematology and Oncology, University Hospital Brno and Faculty of Medicine, Masaryk University, Brno, Czech Republic

[2]Central European Institute of Technology (CEITEC), Masaryk University, Brno, Czech Republic

[3]Institute of Medical Genetics and Genomics, Faculty of Medicine, Masaryk University, Brno, Czech Republic

[4]Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia

[5]Skolkovo Institute of Science and Technology, Moscow, Russia

[6]Pirogov Russian National Research Medical University, Moscow, Russia

[*]Currently at Universität Rostock, Rostock, Germany

#Correspondence:

Karla Plevova, MSc, PhD

Department of Internal Medicine, Hematology and Oncology, University Hospital Brno, and Faculty of Medicine, Masaryk University.

E-mail: karla.plevova@mail.muni.cz

The authors declare no competing financial interests.

Keywords:

myelodysplastic syndrome, retrotransposition, transposable elements, LINE-1, L1, ORF1p, ORF2p, hypomethylation agent, 5'-azacytidine

Word count: 3664

Figures: 3

Supplementary figures: 3

Supplementary tables: 5

Supplementary methods are provided.

Abstract

Retroelements (RE) present in the human genome are silenced via multiple mechanisms, including DNA methylation, to prevent their potentially mutagenic effect. RE activity, demonstrated by their expression and somatic retrotransposition events, is deregulated in multiple tumor types but not in leukemia. We hypothesized that treatment with hypomethylating agents (HMA), commonly used in myelodysplastic syndromes and acute myeloid leukemia, could lead to increased RE activity and somatic retrotranspositions, and contribute to disease progression. We induced expression of ORF1p protein encoded by long interspersed nuclear element-1 (L1) after 72h treatment with HMA in DAMI and HL-60 cell lines. ORF1p was predominantly localized in the cytoplasm, as evidenced by fluorescent microscopy of the DAMI cell line. To study whether long-term HMA therapy may induce somatic retrotranspositions, we (i) treated both cell lines for four weeks, (ii) analyzed a cohort of 17 MDS patients before and on treatment with HMA. Using a previously established sensitive NGS-based method, no RE events were identified. To conclude, we show that although HMA induces the expression of L1-encoded proteins in tumor myeloid cell lines, *de novo* somatic retrotransposition events do not arise during the long-term treatment of MDS patients and myeloid cell lines with these agents.

Introduction

Long interspersed nuclear elements (LINEs), a non-LTR subfamily of retrotransposons, occupy approximately 17% of the human genome (1). Out of around 500 000 copies, the majority are inactive due to genomic rearrangements, point mutations, and 5'-truncation to prevent their harmful impact on the genome stability. The ongoing transposition is primarily caused by type 1 (LINE-1; also known as L1). Only around 100 copies of L1 elements per genome are responsible for the majority of retrotransposon activity (2, 3).

The L1 transposable element comprises two open reading frames, ORF1, encoding an RNA chaperone, and ORF2, encoding an enzyme with single-strand endonuclease and reverse transcriptase activities. After L1 transcription, the polyadenylated bicistronic L1 mRNAs are transported to the cytoplasm and translated into ORF1p and ORF2p proteins. Multiple ORF1p trimers and one or two ORF2p molecules bind the mRNA in cis to form the L1 ribonucleoprotein (L1 RNP). The strong predominance of ORF1p proteins (~30x) explains their much easier detection compared to ORF2p (4). Indeed, ORF2p is nearly undetectable in primary tumors (5). L1 RNP penetrates the nucleus during mitosis (6). Subsequently, target-primed reverse transcription occurs during the S-phase, initiated by the endonuclease nicking genomic DNA at A- and T-rich sites, followed by annealing of RNA poly(A) tail and its reverse transcription (7, 8). Using host proteins involved in DNA repair and replication, the single-stranded DNA gap at the L1 integration site is recognized, followed by the integration and ligation of a newly synthesized L1 copy into the DNA, being flanked by the target site duplications.

The process of retrotransposition is suppressed by multiple cellular mechanisms, including DNA methylation, histone acetylation, piwi RNA complexes, and p53 machinery, with a frequency of a new retrotransposition event in the human population being between 1/20

and 1/200 births (9, 10). As such, many L1 elements are present in individual genomes but absent in the haploid human genome reference. Contrary to normal cells, the expression levels of transposable elements and corresponding proteins, including ORF1p, are increased in tumor cells (11, 12). Consequently, L1 retrotranspositions have been documented in multiple cancer types (13-15). Active TEs are considered highly mutagenic and are commonly linked to the multiple steps of cancer development and progression (14, 16-18). Moreover, L1 hypomethylation has been revealed in multiple cancer types, including lung cancer (19), colorectal cancer (20), breast cancer (21), prostate cancer (22), hepatocellular carcinoma (23), ovarian cancers (24) and esophageal squamous cell carcinoma (25), and was typically associated with poor clinical outcomes, likely due to resulting genome instability and presumed TE activation.

While high levels of somatic retrotranspositions have been shown mainly in solid tumors, in hematological malignancies, the levels are considered to be generally low (15). In line, we did not identify new tumor-specific RE insertions in childhood and adult acute leukemia samples using a sensitive NGS approach allowing for the detection of Alu and L1 insertions in 1% of cells (26).

Hypomethylating agents (HMA) 5-Azacytidine (Aza) and 5-Aza-2'-deoxycytidine (decitabine, Aza-dC), either alone or in combination with other drugs, have become standard therapy for patients with high-risk myelodysplastic syndromes (MDS) (27, 28) and acute myeloid leukemia (AML) (28). Despite the unequivocal benefit for overall survival as well as improved quality of life, some patients do not respond to these drugs, and the majority of responding patients relapse. Since their use leads to hypomethylation of silenced regions in the human genome, including L1, we studied if HMA affect the activity of L1 retrotransposons in tumor myeloid

cell lines and if *de novo* L1 retrotransposition can be detected in MDS patients treated with and progressing after Aza.

Materials and methods

Cell lines and patient samples

A panel of tumor cell lines (Table S1) was used for the initial screening of ORF1 expression following cultivation instructions provided by the supplying collection (DSMZ or ATCC). For further experiments, human myeloid cell lines DAMI and HL60 were selected. The breast carcinoma MCF-7 cell line was used as a positive control for ORF1 expression. HEK293T/17 cell line was used for transient transfections.

Thirty-seven serial bone-marrow samples from 17 MDS patients treated with 5-azacytidine (Aza) were taken after written informed consent approved by the Ethical Committee of the University Hospital Brno was available following the Declaration of Helsinki. The samples were taken before treatment initiation, after several lines of Aza treatment, and/or in relapse; if a relapse sample was unavailable, the closest available sample before progression was used (Table S2). DNA was isolated from whole bone marrow leukocytes after erythrolysis.

Treatment of cell lines with 5-Aza-2'-deoxycytidine

To induce ORF1p and ORF2p expression, a panel of tumor cell lines (Table S1) was treated with 5-Aza-2'-deoxycytidine (Aza-dC; Sigma-Aldrich) for 72 hours (0.5, 1, 2 and 5 µM). Briefly, 2.5 × 10^5 cells/ml were seeded in 6-well plates (TPP) in 5 ml of cell culture medium (Table S1) + 10% fetal bovine serum (FBS). On the following day, medium was replaced with fresh medium containing Aza-dC (0.5, 1, 2, and 5 µM) and cultured for 48 hours, then the treatment was

repeated and followed by cultivation for additional 24 hours. As a control, the cell lines were processed in the same way without adding Aza-dC.

For fluorescent microscopy and flow cytometry, DAMI and HL-60 were seeded in 6-well plates (TPP) with or without sterile microscopy coverslips in density 3.0×10^6 cells/ml in 3 ml of DMEM + 10% ultra-low IgG FBS (PAN-biotech). On the following day, the cells were treated with 2 µM and 5 µM Aza-dC with media change after 48 hours, as described above.

Long-term treatment of DAMI and HL-60 was performed with 0.5 and 2 µM of Aza-dC with untreated cells as controls. Cells were harvested at days 0, 3, 7, 14 and 28.

Immunoblotting

Cells were washed with PBS and lysed in ice-cold NP-40 buffer (Thermo Fisher Scientific) with protease and phosphatase inhibitors (1:100; P8340 and P0044, Sigma Aldrich) for 30 minutes. Protein concentrations were determined by Bicinchoninic Acid Kit for Protein Determination (BCA1-1KT, Sigma Aldrich) or the Bradford Protein Assay (BioRad). The protein lysates were run on 10% sodium dodecyl sulfate-polyacrylamide gels (SDS-PAGE) and transferred to a nitrocellulose membrane (Bio-Rad) using a wet tank blotting system. The membranes were blocked with 5% non-fat milk in TBS buffer containing 0.1% Tween on a rocking shaker for 2 hours. The membranes were washed in TBS-T alone and incubated in the appropriate primary antibody overnight on a rocking shaker at 4 °C. The following day, membranes were washed and incubated in the appropriate secondary antibodies at RT for 1 hour. For the list of antibodies, see Table S3. The proteins were detected using a chemiluminescence system (Clarity Western ECL Substrate, 1705061, Bio-Rad) and imaged using a UVITEC Documentation System (Uvitec Cambridge, Mini HD9).

Intracellular detection of ORF1p and ORF2p proteins using fluorescence microscopy and flow cytometry

For fluorescent microscopy, the sterilized coverslips (13 mm/0.17 mm Menzel) coated with 0.01% poly-L-lysine (Sigma-Aldrich) were seeded with the targeted cell line at a volume of 200 µl per well.

Cultured cells were fixed with 4% formaldehyde (10 min; F8775, Sigma-Aldrich) and blocked with 3% IgG-free BSA blocking buffer (001-000-161, Jackson ImmunoResearch) with the addition of 0.25% Triton X-100 (SIALX100, Sigma-Aldrich) for 1 hour. The cells were incubated in the corresponding primary antibody for 1 hour and the secondary antibody for 30 min (Table S3). Nuclear DNA was stained with Hoechst (H1399, ThermoFisher) for 2 min. All procedures were performed at room temperature. Finally, individual stained coverslips with cells were carefully transferred with tweezers to slides with prepared mounting medium (S3023, DAKO). Fluorescence detection of ORF1p and ORF2p was performed the day after the mounting medium dried using a ZEISS 700LSM confocal microscope with plan-apochromat 40x oil objective lens. Individual images were detected using two filters with 405 and 488 nm wavelengths and 1 AU pinhole and assessed in ZEN 2009 software.

Intracellular staining for flow cytometry using FACSVerse (BD Biosciences) was performed analogically to fluorescent microscopy staining without coverslips.

Preparation of ORF1 and ORF2 Plasmids and transient transfection

E.coli DH5alpha with pBudORF1 (Addgene) and pBudORF2 (Addgene) were cultured according to the manufacturer's protocol. Plasmid DNA was isolated using EndoFree Plasmid Maxi Kit (Qiagen) according to the manufacturer's protocol.

Transient transfection for fluorescence microscopy was performed directly on coverslips coated with poly-L-lysine (see above) placed in 24-well plates using the transfection reagent Polyethylenimine (PEI; Polysciences) according to the manufacturer's protocol. Briefly, the day before transfection, HEK293T/17 cells were seeded in individual wells of a 24-well plate with

coverslips at a cell density of 0.5 × 10^6/ml with 1 ml of DMEM/F12 culture medium (PAN-biotech) and 10% ultra-low IgG FBS (PAN-biotech). On the day of transfection, the culture medium was replaced with DMEM/F12:H2O transfection medium in a 1:1 ratio with the addition of 300 µl serum-free 2mM Glutamine (PAN-biotech). The premix was added to the individual wells with prepared transfection media after being left in the box at RT for half an hour at a DNA:PEI ratio (3:9 µg) of 150 µl. The transfected cells were incubated at 37 °C and 5% CO_2. After 4 hours, the transfection medium was replaced with fresh 1ml DMEM/F12 medium + 10% ultra-low IgG FBS and plates were incubated for additional 72 hours. Fluorescence staining of the slide was carried out after PBS wash, as described above.

Transient transfection for the detection of ORF proteins by flow cytometry and western blots was performed in 6-well plates without coverslips with the identical transfection procedure, and the volumes increased accordingly, i.e., 1.5 x 10^6 in 3 ml of HEK293T/17 cells were used, and the volumes of culture medium and transfection medium were 1 ml and 500 µl, respectively. After three days of culturing in 3 ml of fresh medium, cells were washed out with PBS supplemented with 15 mM EDTA and transferred into 1.5 ml Eppendorf tubes.

Detection of DNA retrotransposition insertions in cancer cell lines and primary patient cells

A next-generation sequencing method for detecting new L1 insertions of the transpositionally-active subfamily L1HS was used according to our previous reports (26, 29). Briefly, gDNA was digested by a mixture of selected endonucleases TaqI and FspBI to generate fragments that consisted of a 3' part of L1 retroelement and its adjacent genomic region (i.e., flank). In the next step, the fragmented DNA was ligated to a stem-loop adapter containing unique molecular identifiers (UMI) that were used to quantify the number of cells bearing each insertion after sequencing. Next, a primer specific to transpositionally active L1 subfamily L1HS and a primer corresponding to the ligated adaptor were used to selectively amplify genomic

flanks adjacent to the 3' end of L1. A product of the first PCR was used in the second semi-nested PCR. Finally, an indexing PCR was carried out to introduce the sample barcodes and oligonucleotides necessary for Illumina sequencing on the NextSeq 550 machine (paired-end, 150+150). We used a custom computational pipeline (30) to map all the sequenced insertions' flanks to the human reference genome and remove various artifacts generated during library preparation and sequencing. The coordinates of the insertions in the serial samples were matched to identify insertions related to clonal propagation (i.e., compared to the corresponding pre-treatment sample). Following our previous report (26), candidate somatic L1 insertions were validated by an independent locus-specific PCR with initial gDNA as a template, and Encyclo Polymerase Mix (Evrogen) and/or Q5 High-Fidelity 2X Master Mix (New England Biolabs).

For detailed protocols, see *Supplementary Methods*.

Results

5-Aza-2'-deoxycytidine treatment induces ORF1p expression in tumor myeloid cell lines

First, we performed initial Western blot screening to reveal which cell lines express ORF1p (Table S1). While ORF1p was clearly detected in three carcinoma cell lines (MCF-7, SW-48, H1299), it was absent in all seven lymphoid and all four myeloid leukemia cell lines. Since DNA methylation is a key transposon silencing mechanism, we attempted to induce ORF1p expression using Aza-dC. Out of the tested cell lines, we were able to induce ORF1p expression in two myeloid cell lines, DAMI and HL-60 (Figure 1A), while for the remaining two cell lines

184 derived from myeloid disorders used Azd-dC dosage led to a prominent decrease in cell
185 viability (data not shown). Thus, DAMI and HL-60 cell lines were used for further experiments.
186 To validate our results with an independent method, we implemented a protocol for
187 intracellular ORF1p staining and analyzed treated and untreated DAMI and HL-60 cells with
188 flow cytometry. We confirmed the ORF1p induction in both tested cell lines (Figure 1B).
189 Similarly to the ORF1p levels detected by the Western blot analysis in bulk samples, we did
190 not observe a dose-dependent effect when increasing Aza-dC concentrations from 1 to 5 µM.
191 Next, to explore the intracellular localization of ORF1p in myeloid cell lines, we stained the
192 cells intracellularly on microscope coverslips for visualization using fluorescent microscopy. As
193 positive controls, we used MCF-7 with a high endogenous expression (Figure S1) and
194 HEK293T/17 cells transfected with pBudORF1 expression plasmid (Figure S2A). The optimized
195 procedure was used in DAMI cells that could adhere to the surface (unlike HL-60). Again, we
196 confirmed ORF1p induction in Aza-dC-treated cells (Figure 2). ORF1p signals predominated in
197 the cytoplasm, with scarce presence in the nuclei.
198 Further, we also aimed to study the expression of ORF2p, a 150 kDa protein with much lower
199 cellular abundance. For this purpose, we treated the DAMI cell line with Aza-dC and analyzed
200 it using flow cytometry. Although we observed a protein induction after Aza-dC treatment
201 (Figure 3), we were unable to confirm the induction with the fluorescence microscopy despite
202 successfully detecting ORF2p in HEK293T/17 cells transiently transfected with pBudORF2
203 expression plasmid (Figure S2B), likely due to different sensitivity of the two methods.
204 **Long-term treatment of tumor myeloid cell lines with 5-Aza-2'-deoxycytidine does not**
205 **increase the retrotransposition rate**
206 Based on our findings, showing that HMA may induce the expression of proteins encoded by
207 L1 retrotransposon, we studied whether the long-term exposition of myeloid tumor cells to

HMA also increases the retrotransposition rates. To mimic *in vitro* the therapy of hematooncology patients, we treated DAMI and HL-60 cell lines with 0.5 and 2 μM Aza-dC for a period of four weeks. On days 3, 7, 14, and 28, we collected DNA, protein and cells.

First, we assessed ORF1p induction during Aza treatment using Western blot and flow cytometry and observed the level of induction for 2 μM Aza comparable with our initial experiments (data not shown). This result prompted us to proceed with detecting novel insertions using NGS. We applied a high-sensitive NGS-based protocol enabling us to detect L1 insertions down to 1% of cells (26). Candidate *de novo* L1 insertions for each time point (3, 7, 14, and 28 days after Aza-dC addition) were determined by comparing their genomic coordinates to the list of L1 insertion sites identified in the same cell line before treatment (day 0). We additionally filtered out the insertions whose coordinates matched the insertions found in other treatment conditions. Such insertions comprise false positives as the probability of an independent insertion into the same genomic locus in different cells is infinitely low. Following our pipeline, manual filtering against the genomic sequence, and excluding regions containing known population L1 copies, we identified 358 candidate insertions. Based on read quality (R1 and R2 matching) and number of reads (>10) and UMIs (≥2), we selected 14 candidate insertions (8 in DAMI and 6 in HL-60), for which we designed locus-specific primers for PCR validation (Table S4, Figures S3A-B). None of the candidate insertions was confirmed.

Retrotransposition is uncommon in MDS patients treated with HMA

Still, we wanted to find out whether we could identify any retrotranspositions in leukocyte DNA of MDS patients treated with Aza and whether the clonal expansion of cells carrying *de novo* L1 insertions may contribute to disease relapse. The rationale behind that was that they were exposed to the hypomethylation effect for a much longer period than the cell lines in

four-week-lasting cultivation and the tumor cells eventually escape from the cell death caused by Aza. For this purpose, we again used our high-sensitive NGS-based protocol to explore 37 serial bone-marrow samples from 17 MDS patients treated with Aza. The cohort consisted of cases with high-risk disease, with five patients bearing *TP53* defects and nine progressing to AML. The samples were taken before treatment initiation, several months on treatment and/or in relapse (for details about samplings, see Table S2). The patients received a median of 11 Aza cycles (ranging from 3 to 46) between the baseline and last samples. L1 insertions of each follow-up sample were compared to the respective baseline sample of the same patient.

Following the same manual filtering as for treated tumor cell lines, we identified three candidate insertions in the genomes of three MDS patients undergoing therapy, each bearing a single candidate insertion, that were subjected to validation using locus-specific PCR (Figure 3, Table S5). Out of these candidate insertions, we were able to validate none of them, suggesting that L1 retrotransposition activity in primary patient samples is low.

Discussion

Epigenetic silencing of transposable elements is one of the key mechanisms of cellular defense against their potentially mutagenic effect. As REs become active after demethylation of their promoters (12, 23), we aimed to explore whether they become active in MDS treated with HMA and whether *de novo* insertions contribute to disease progression, at least in some patients.

MDS is a heterogeneous disease comprised of hematopoietic stem cell disorders leading to ineffective hematopoiesis, demonstrated by blood cytopenias and progression to acute myeloid leukemia in approximately one third of patients (31, 32). Besides cytogenetic changes

and gene mutations, epigenetic changes contribute to disease pathophysiology and progression. Widespread hypermethylation was correlated to impaired overall survival and progression to AML (33, 34). Although HMA are highly efficient in MDS, the mechanism of their action is not fully understood. It includes activation of tumor suppressor genes commonly inactivated in MDS, incorporation of modified cytidine derivates into RNA and DNA, and activation of tumor antigens and retroelements recognized by the immune system. (Reviewed by (35) and (36)).

Indeed, we demonstrated with several methods that ORF1p levels increase after Aza-dC treatment of tumor myeloid cell lines DAMI and HL-60, pointing to retroelement activation. Fluorescent microscopy showed that the majority of ORF1p induced by Aza-dC in DAMI is localized in the cytoplasm and scarcely in the nucleus. We were unable to localize ORF2p after treatment and could only show its expression and localization in the ectopic expression model. This is in line with other studies (5) and can be partially explained by the predominance of ORF1p proteins in L1 RNP (4). As the Aza-dC caused cell death in other myeloid cell lines tested, we could not study ORF1p and ORF2p induction and localization in other models.

Despite the apparently increased level of L1 proteins, we did not observe *de novo* retrotranspositions in cell lines and patients treated with HMA. This may be explained by the absence of a strong selective advantage that would lead to clonal expansion of affected cells. However, as we expected that novel L1 insertions could be present in minor subpopulations and the whole genome sequencing methods would have limited sensitivity to detect them, we applied a targeted-NGS-based approach with the ability to capture retrotransposition events in 1% of cells (26). This approach was previously used to demonstrate somatic retrotranspositions in brain tissue (37) and colorectal cancer (29) and the absence of RE

activity in leukemia (26). Thus, we assume that the sensitivity of the NGS method was not the reason for the negative result.

The alternative explanation for the absence of L1 retrotransposition can be that myeloid tumor cells are negatively selected for L1 expression and/or the retrotransposition process. Indeed, a comprehensive study showed that decreased L1HS expression is associated with short overall survival in AML, and it has been suggested that L1 silencing is required for oncogene-induced transformation and propagation of AML-initiating cells (38). In a model system, activation of endogenous L1s resulted in increased L1 retrotransposition and impaired leukemia cell growth *in vitro* and in xenotransplants. In line, a model proposed by (39) suggests that the impact of TEs activity evolves during myeloid transformation. While the increased activity of TEs and piRNAs (40) induces an immune response and leads to the elimination of leukemic cells in early-stage MDS, the leukemic cells finally escape the control of the immune system via suppression of TE/piRNA expression, resulting in the progression to high-risk MDS, accumulation of leukemic blasts, and AML.

The somatic retrotranspositions are absent not only in AML and high-risk MDS but also in acute lymphoblastoid leukemia, as we showed previously (26). Of note, in the report, we also analyzed a set of childhood AML that are biologically distinct from AML in adults and the elderly, and we did not detect somatic retrotranspositions either (26). Whether the intolerance of RE activity is a feature of hematopoietic stem and early progenitor cells or is inherent to the hematopoietic lineage as a whole is currently unknown. It would be of interest to explore RE activity, e.g., in mature B-cell neoplasms that are not directly related to a hematopoietic stem cell.

Significantly increased levels of ORF1p proteins and L1Hs transcripts were detected in TP53-mutated Wilms tumor and colorectal carcinoma samples, respectively (41). It has been

suggested that keeping genome integrity via preventing RE events is an ancient function of tumor suppressor p53 (42). *TP53*-mutated myeloid neoplasms have recently been recognized as an aggressive entity (43). To see if loss of p53 function enables RE events in MDS, we included five patients with *TP53* mutations in the studied cohort. The lack of RE events in these patients and the *TP53*-deficient cell lines HL-60 and DAMI suggests that intolerance of RE events in leukemic cells is p53-independent, although it is to be confirmed by future studies. Overall, we show that although HMA induce the expression of L1-encoded proteins in tumor myeloid cell lines, *de novo* somatic retrotransposition events do not arise during the treatment of MDS patients with these agents. Thus, while somatic retrotranspositions occur in carcinomas, we show they are uncommon in myeloid cells.

Funding

The present work was supported by the Czech Science Foundation grant No. 19-11299S. Further support was provided by the Ministry of Health of the Czech Republic for the conceptual development of research organization FNBr 65269705, the student project MUNI/A/1224/2022 by the Ministry of Education, Youth and Sports of the Czech Republic, and the project National Institute for Cancer Research (Programme EXCELES, ID No. LX22NPO5102), funded by the EU – Next Generation EU.

Author contributions

SPav, IM, and KP conceived the study, designed and supervised experiments, and interpreted data; MK and IM developed methods; MK, AV, TG, ZJ, HS, and DW performed experiments and analyzed data; AS carried out the bioinformatic analyses; MB collected clinical and laboratory diagnostic information; LC provided samples and clinical expertise; SPav, IM, and KP drafted the manuscript; SPosp, IM and KP obtained funding; all authors reviewed and approved the manuscript.

Competing interests

The authors declare no competing financial interests.

Data Availability Statement

Raw sequencing data are available on the Sequence Read Archive (SRA) under the accession no. SUB13719421.

References

1. Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, et al. A unified classification system for eukaryotic transposable elements. Nat Rev Genet. 2007;8(12):973-82.

2. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, et al. Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci U S A. 2003;100(9):5280-5.

3. Streva VA, Jordan VE, Linker S, Hedges DJ, Batzer MA, Deininger PL. Sequencing, identification and mapping of primed L1 elements (SIMPLE) reveals significant variation in full length L1 elements between individuals. BMC Genomics. 2015;16(1):220.

4. Taylor MS, LaCava J, Mita P, Molloy KR, Huang CR, Li D, et al. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. Cell. 2013;155(5):1034-48.

5. Ardeljan D, Wang X, Oghbaie M, Taylor MS, Husband D, Deshpande V, et al. LINE-1 ORF2p expression is nearly imperceptible in human cancers. Mob DNA. 2020;11:1.

6. Mita P, Wudzinska A, Sun X, Andrade J, Nayak S, Kahler DJ, et al. LINE-1 protein localization and functional dynamics during the cell cycle. Elife. 2018;7.

7. Cost GJ, Feng Q, Jacquier A, Boeke JD. Human L1 element target-primed reverse transcription in vitro. EMBO J. 2002;21(21):5899-910.

8. Doucet AJ, Wilusz JE, Miyoshi T, Liu Y, Moran JV. A 3' Poly(A) Tract Is Required for LINE-1 Retrotransposition. Mol Cell. 2015;60(5):728-41.

9. Ewing AD, Kazazian HH. High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. Genome Res. 2010;20(9):1262-70.

359 10. Feusier J, Watkins WS, Thomas J, Farrell A, Witherspoon DJ, Baird L, et al. Pedigree-
360 based estimation of human mobile element retrotransposition rates. Genome Res.
361 2019;29(10):1567-77.
362 11. Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S, Coffman EJ, et al. Aberrant
363 overexpression of satellite repeats in pancreatic and other epithelial cancers. Science.
364 2011;331(6017):593-6.
365 12. Rodić N, Sharma R, Zampella J, Dai L, Taylor MS, Hruban RH, et al. Long interspersed
366 element-1 protein expression is a hallmark of many human cancers. Am J Pathol.
367 2014;184(5):1280-6.
368 13. Lee E, Iskow R, Yang L, Gokcumen O, Haseley P, Luquette LJ, et al. Landscape of somatic
369 retrotransposition in human cancers. Science. 2012;337(6097):967-71.
370 14. Helman E, Lawrence MS, Stewart C, Sougnez C, Getz G, Meyerson M. Somatic
371 retrotransposition in human cancer revealed by whole-genome and exome sequencing.
372 Genome Res. 2014;24(7):1053-63.
373 15. Rodriguez-Martin B, Alvarez EG, Baez-Ortega A, Zamora J, Supek F, Demeulemeester J,
374 et al. Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by
375 LINE-1 retrotransposition. Nat Genet. 2020;52(3):306-19.
376 16. Miki Y, Nishisho I, Horii A, Miyoshi Y, Utsunomiya J, Kinzler KW, et al. Disruption of the
377 APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. Cancer Res.
378 1992;52(3):643-5.
379 17. Rodríguez-Martín C, Cidre F, Fernández-Teijeiro A, Gómez-Mariano G, de la Vega L,
380 Ramos P, et al. Familial retinoblastoma due to intronic LINE-1 insertion causes aberrant and
381 noncanonical mRNA splicing of the RB1 gene. J Hum Genet. 2016;61(5):463-6.

382 18. Wolff EM, Byun HM, Han HF, Sharma S, Nichols PW, Siegmund KD, et al.
383 Hypomethylation of a LINE-1 promoter activates an alternate transcript of the MET oncogene
384 in bladders with cancer. PLoS Genet. 2010;6(4):e1000917.
385 19. Rauch TA, Zhong X, Wu X, Wang M, Kernstine KH, Wang Z, et al. High-resolution
386 mapping of DNA hypermethylation and hypomethylation in lung cancer. Proc Natl Acad Sci U
387 S A. 2008;105(1):252-7.
388 20. Sunami E, de Maat M, Vu A, Turner RR, Hoon DS. LINE-1 hypomethylation during
389 primary colon cancer progression. PLoS One. 2011;6(4):e18884.
390 21. van Hoesel AQ, van de Velde CJ, Kuppen PJ, Liefers GJ, Putter H, Sato Y, et al.
391 Hypomethylation of LINE-1 in primary tumor has poor prognosis in young breast cancer
392 patients: a retrospective cohort study. Breast Cancer Res Treat. 2012;134(3):1103-14.
393 22. Yegnasubramanian S, Haffner MC, Zhang Y, Gurel B, Cornish TC, Wu Z, et al. DNA
394 hypomethylation arises later in prostate cancer progression than CpG island hypermethylation
395 and contributes to metastatic tumor heterogeneity. Cancer Res. 2008;68(21):8954-67.
396 23. Gao XD, Qu JH, Chang XJ, Lu YY, Bai WL, Wang H, et al. Hypomethylation of long
397 interspersed nuclear element-1 promoter is associated with poor outcomes for curative
398 resected hepatocellular carcinoma. Liver Int. 2014;34(1):136-46.
399 24. Pattamadilok J, Huapai N, Rattanatanyong P, Vasurattana A, Triratanachat S,
400 Tresukosol D, et al. LINE-1 hypomethylation level as a potential prognostic factor for epithelial
401 ovarian cancer. Int J Gynecol Cancer. 2008;18(4):711-7.
402 25. Iwagami S, Baba Y, Watanabe M, Shigaki H, Miyake K, Ishimoto T, et al. LINE-1
403 hypomethylation is associated with a poor prognosis among patients with curatively resected
404 esophageal squamous cell carcinoma. Ann Surg. 2013;257(3):449-55.

405 26. Urazbakhtin S, Smirnova A, Volakhava A, Zerkalenkova E, Salyutina M, Doubek M, et
406 al. The Absence of Retroelement Activity Is Characteristic for Childhood Acute Leukemias and
407 Adult Acute Lymphoblastic Leukemia. Int J Mol Sci. 2022;23(3).
408 27. Silverman LR, Demakos EP, Peterson BL, Kornblith AB, Holland JC, Odchimar-Reissig R,
409 et al. Randomized controlled trial of azacitidine in patients with the myelodysplastic
410 syndrome: a study of the cancer and leukemia group B. J Clin Oncol. 2002;20(10):2429-40.
411 28. DiNardo CD, Jonas BA, Pullarkat V, Thirman MJ, Garcia JS, Wei AH, et al. Azacitidine
412 and Venetoclax in Previously Untreated Acute Myeloid Leukemia. N Engl J Med.
413 2020;383(7):617-29.
414 29. Komkov AY, Urazbakhtin SZ, Saliutina MV, Komech EA, Shelygin YA, Nugmanov GA, et
415 al. SeqURE - a new copy-capture based method for sequencing of unknown Retroposition
416 events. Mob DNA. 2020;11(1):33.
417 30. Nugmanov GA, Komkov AY, Saliutina MV, Minervina AA, Lebedev YB, Mamedov IZ. [A
418 Pipeline for the Error-free Identification of Somatic Alu Insertions in High-throughput
419 Sequencing Data]. Mol Biol (Mosk). 2019;53(1):154-65.
420 31. Fenaux P, Haase D, Santini V, Sanz GF, Platzbecker U, Mey U, et al. Myelodysplastic
421 syndromes: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann
422 Oncol. 2021;32(2):142-56.
423 32. Adès L, Itzykson R, Fenaux P. Myelodysplastic syndromes. Lancet.
424 2014;383(9936):2239-52.
425 33. Shen L, Kantarjian H, Guo Y, Lin E, Shan J, Huang X, et al. DNA methylation predicts
426 survival and response to therapy in patients with myelodysplastic syndromes. J Clin Oncol.
427 2010;28(4):605-13.

428 34. Jiang Y, Dunbar A, Gondek LP, Mohan S, Rataul M, O'Keefe C, et al. Aberrant DNA
429 methylation is a dominant mechanism in MDS progression to AML. Blood. 2009;113(6):1315-
430 25.
431 35. Diesch J, Zwick A, Garz AK, Palau A, Buschbeck M, Götze KS. A clinical-molecular update
432 on azanucleoside-based therapy for the treatment of hematologic cancers. Clin Epigenetics.
433 2016;8:71.
434 36. Duchmann M, Itzykson R. Clinical update on hypomethylating agents. Int J Hematol.
435 2019;110(2):161-9.
436 37. Kurnosov AA, Ustyugova SV, Nazarov VI, Minervina AA, Komkov AY, Shugay M, et al.
437 The evidence for increased L1 activity in the site of human adult brain neurogenesis. PLoS One.
438 2015;10(2):e0117854.
439 38. Gu Z, Liu Y, Zhang Y, Cao H, Lyu J, Wang X, et al. Silencing of LINE-1 retrotransposons is
440 a selective dependency of myeloid leukemia. Nat Genet. 2021;53(5):672-82.
441 39. Merkerova MD, Krejcik Z. Transposable elements and Piwi-interacting RNAs in
442 hemato-oncology with a focus on myelodysplastic syndrome (Review). Int J Oncol. 2021;59(6).
443 40. Colombo AR, Zubair A, Thiagarajan D, Nuzhdin S, Triche TJ, Ramsingh G. Suppression
444 of Transposable Elements in Leukemic Stem Cells. Sci Rep. 2017;7(1):7029.
445 41. Wylie A, Jones AE, D'Brot A, Lu WJ, Kurtz P, Moran JV, et al. p53 genes function to
446 restrain mobile elements. Genes Dev. 2016;30(1):64-77.
447 42. Levine AJ, Ting DT, Greenbaum BD. P53 and the defenses against genome instability
448 caused by transposons and repetitive elements. Bioessays. 2016;38(6):508-13.
449 43. Arber DA, Orazi A, Hasserjian RP, Borowitz MJ, Calvo KR, Kvasnicka HM, et al.
450 International Consensus Classification of Myeloid Neoplasms and Acute Leukemias:
451 integrating morphologic, clinical, and genomic data. Blood. 2022;140(11):1200-28.

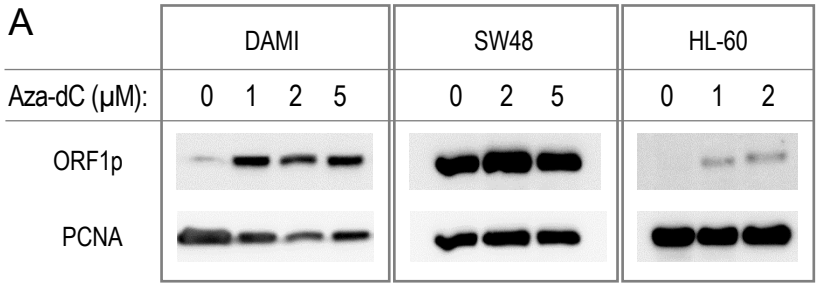
452 **Figure Legends**

453 **Figure 1.** Western blot (A) and flow cytometry (B) measurements of ORF1p after Aza-dC
454 treatment. The Western blot analysis was performed on a panel of tumor cell lines (Table S1).
455 Here only DAMI and HL-60 (myeloid cell lines) and SW48 (colorectal carcinoma cell line) are
456 shown to demonstrate the ORF1p induction in the myeloid cell lines, while in SW48, the ORF1p
457 expression was present already in the baseline sample without Aza-dC treatment. The ORF1p
458 induction in DAMI and HL-60 was further confirmed using flow cytometry. Aza-dC
459 concentrations applied for the Western blot analysis – DAMI: 0, 1, 2, 5 µM; SW48: 0, 2, 5 µM;
460 HL-60: 0, 1, 2 µM; and for flow cytometry – both DAMI and HL-60: 0, 1, 2, 5 µM.
461 Primary/secondary antibodies – Western blot: Anti-LINE-1 ORF1p (CellSignalling); flow
462 cytometry: Anti-LINE-1 ORF1p (Abcam) + Goat anti Rabbit IgG AF488 (Abcam).

463 **Figure 2.** Confocal microscopy detection of ORF1p induction after Aza-dC treatment in DAMI
464 cell line. Left panel: The immunofluorescence signal of ORF1p (green). Right panel: The
465 immunofluorescence signal of ORF1p (green) combined with the nuclei signal (DAPI, blue).
466 Primary/secondary antibodies: Anti-LINE-1 ORF1 (Abcam) + Goat anti Rabbit IgG AF488
467 (Abcam). Magnification: 40x.

468 **Figure 3.** Flow cytometric analysis of ORF2p induction after the treatment with 2 µM Aza-dC
469 in DAMI and HL-60 cell lines. The experiment was repeated in independent biological triplicate
470 or duplicate for DAMI and HL-60, respectively, with similar results. Primary/secondary
471 antibodies: Anti-LINE-1 ORF2 (Rockland) + Goat anti Chicken IgY AF488 (Invitrogen).

Figure 1



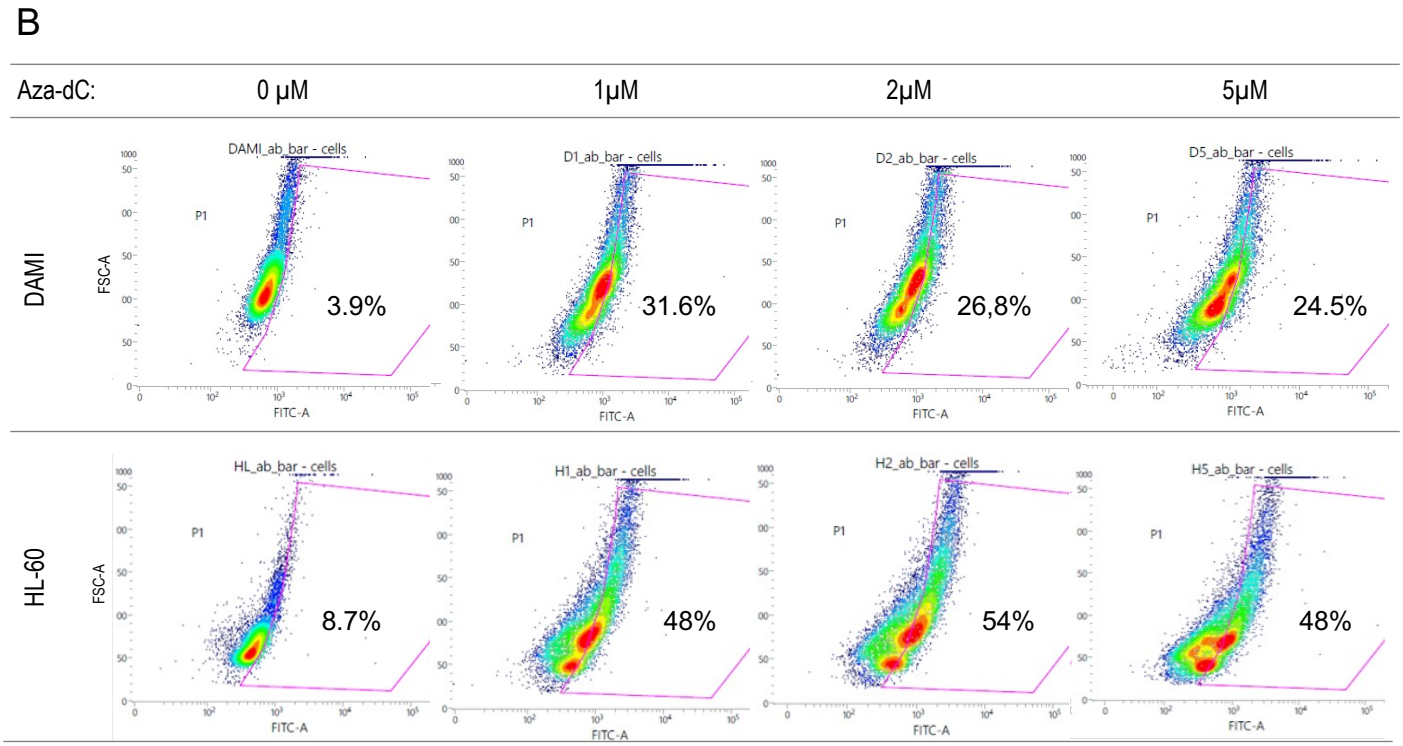


Figure 1: Western blot (A) and flow cytometry (B) measurements of ORF1p after Aza-dC treatment. The Western blot analysis was performed on a panel of tumor cell lines (Table S1). Here only DAMI and HL-60 (myeloid cell lines) and SW48 (colorectal carcinoma cell line) are shown to demonstrate the ORF1p induction in the myeloid cell lines, while in SW48, the ORF1p expression was present already in the baseline sample without Aza-dC treatment. The ORF1p induction in DAMI and HL-60 was further confirmed using flow cytometry. Aza-dC concentrations applied for the Western blot analysis – DAMI: 0, 1, 2, 5 μM; SW48: 0, 2, 5 μM; HL-60: 0, 1, 2 μM; and for flow cytometry – both DAMI and HL-60: 0, 1, 2, 5 μM. Primary/secondary antibodies – Western blot: Anti-LINE-1 ORF1p (CellSignalling); flow cytometry: Anti-LINE-1 ORF1p (Abcam) + Goat anti Rabbit IgG AF488 (Abcam).

Figure 2

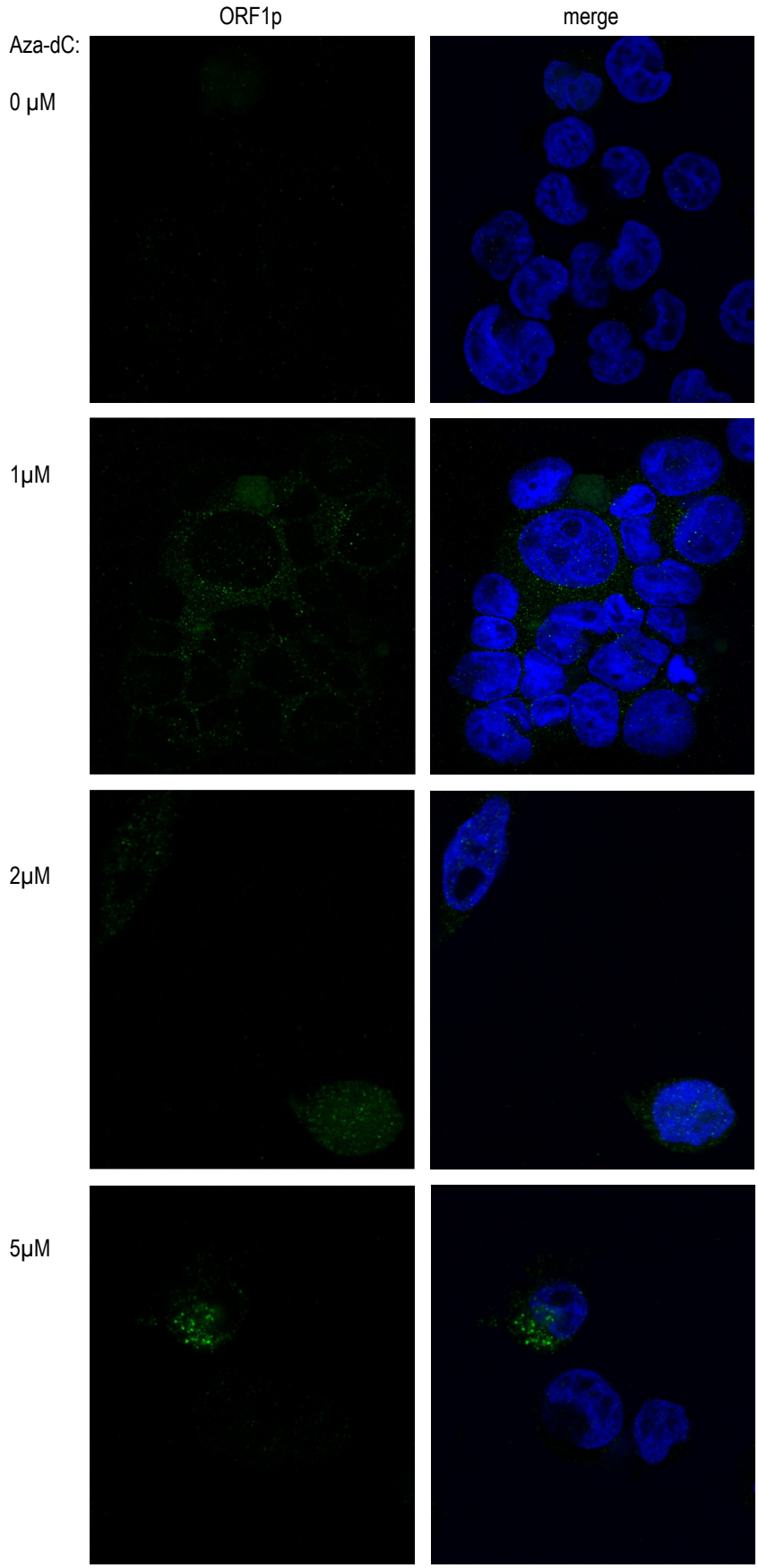


Figure 2: Confocal microscopy detection of ORF1p induction after Aza-dC treatment in DAMI cell line. Left panel: The immunofluorescence signal of ORF1p (green). Right panel: The immunofluorescence signal of ORF1p (green) combined with the nuclei signal (DAPI, blue). Primary/secondary antibodies: Anti-LINE-1 ORF1 (Abcam) + Goat anti Rabbit IgG AF488 (Abcam). Magnification: 40x.

Figure 3

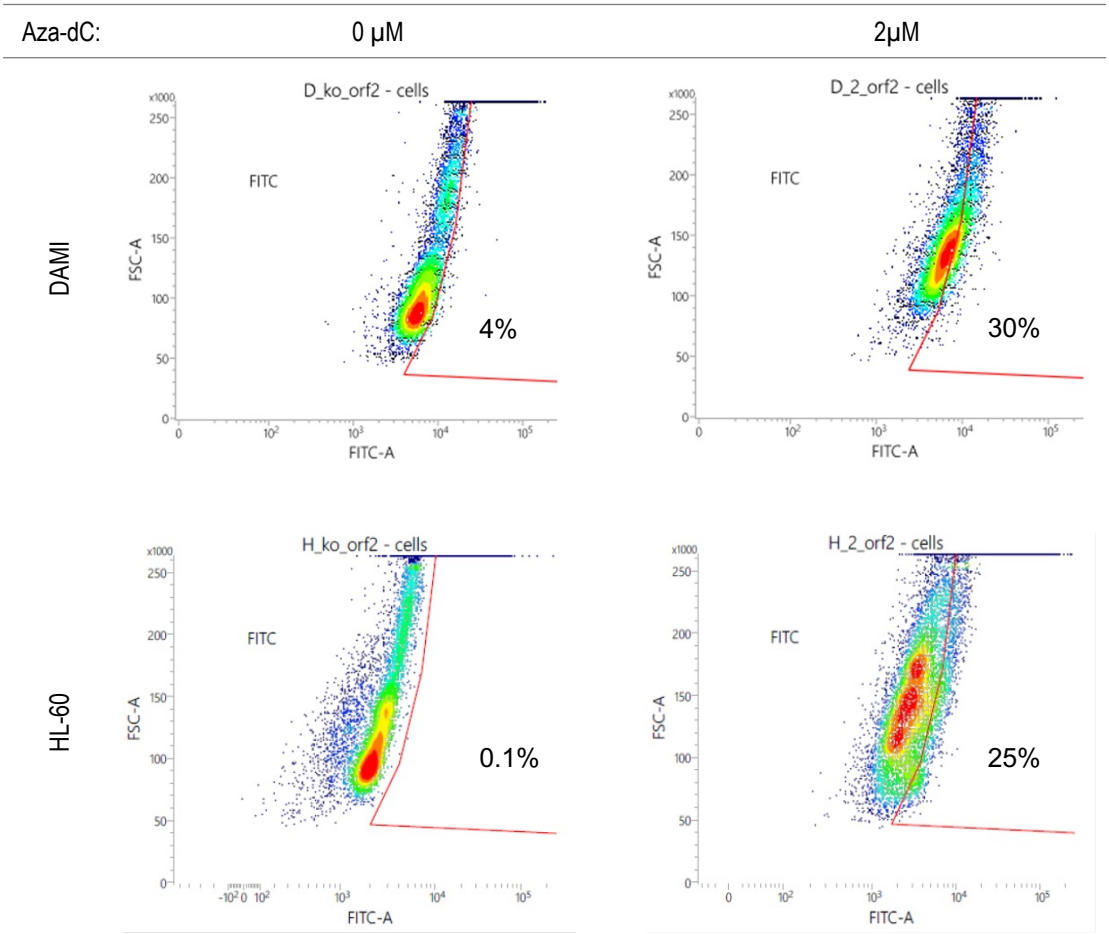


Figure 3: Flow cytometric analysis of ORF2p induction after the treatment with 2 µM Aza-dC in DAMI and HL-60 cell lines. The experiment was repeated in independent biological triplicate or duplicate for DAMI and HL-60, respectively, with similar results. Primary/secondary antibodies: Anti-LINE-1 ORF2 (Rockland) + Goat anti Chicken IgY AF488 (Invitrogen).