Derepression of retroelements in acute myeloid leukemia with 3q aberrations

Acute myeloid leukemia (AML) with chromosomal rearrangements inv(3)/t(3;3) (3q-AML) is a rare but highly fatal subtype of leukemia. It is characterized by an aberrant transcription of the proto-oncogene *EVI1* (ecotropic viral integration site 1, *MECOM*) as a result of the chromosomal 3q21q26 rearrangements that lead to the relocation of a master *GATA2* distal hematopoietic enhancer to the *EVI1* locus and deregulation of both genes.[1,2] To date, little is known about what triggers chromosomal rearrangements in 3q-AML that ultimately lead to the deregulation of *EVI1* via the repositioning of *G2DHE*. However, recent studies have shown evidence for the involvement of endogenous transposable elements of RNA family (retroelements [RE]) in the formation of complex chromosomal aberrations, including translocations, large-scale duplications and amplifications through retrotransposition across different cancer entities.[3] Furthermore, hypomethylation of RE has been linked to their pathogenic mobility in epithelial tumors.[4–6]

Here we present the results of functional genomics analysis of a cohort of 3q-AML patients. Based on our data, we hypothesized that breakpoint-associated RE (breakpoint-RE) could play an important regulatory and activating role in this AML subtype. Therefore, we performed an array of *in vitro* studies using CRISPR-Cas9 approach to dissect their role in 3q-AML.

Targeted chromosome 3q-capture sequencing of 3q-AML patient samples and cell lines with *EVI1* overexpression previously revealed a characteristic 3q21q26 pattern of patient-specific breakpoints, demarcating a leukemogenic *EVI1*-activating super-enhancer that is found uniquely in inv(3)/t(3;3) AML and contains *G2DHE*.[1] In order to identify a commonality between breakpoints relevant for super-enhancer formation, we reanalyzed 3q-capture sequencing data and found that in 38 of 41 samples, chromosomal breakpoints at 3q21.3 and 3q26.2 mapped to sequences of RE, including long interspersed elements (LINE), short interspersed elements (SINE) and long terminal repeats (LTR) (Figure 1A). Of note, RNA sequencing (RNA-Seq) of 3q-AML patients revealed a characteristic RNA readthrough spanning the large super-enhancer region at 3q21.3 (Figure 1B, top panel). A similar enhancer RNA (eRNA) signature was observed in non-3q-rearranged AML cases (Figure 1B, top panel; *Online Supplementary Figure S3*), indicating active *G2DHE* regulating *GATA2* in its native environment. In 3q-AML, however, the RNA readthrough frequently originated at 3q21.3 breakpoint sites, extending beyond the super-enhancer region. Additionally, allele-specific bisulfite amplicon sequencing performed on selected AML cases revealed focal demethylation of CpG sites around the chromosomal breakpoints exclusively on the rearranged allele, whereas the intact allele in 3q-AML and both alleles in non-rearranged leukemic cell lines did not show any hypomethylation pattern (Figure 1B, bottom panel). Focal hypomethylation around breakpoints on the rearranged allele could be the consequence of chromosomal rearrangements and super-enhancer-related epigenetic reprogramming, including the deposition of active chromatin marks and physical interaction between the *EVI1* promoter and *G2DHE*.[1]

Based on our RNA-Seq and bisulfite sequencing data, we investigated whether derepression of breakpoint-RE could possibly represent a priming event for an enhancer rearrangement by relaxation of the local chromatin compaction and may play a role in the ectopic activation of *EVI1* by the super-enhancer. In order to test this hypothesis, we performed a CRISPR-Cas9 gene editing experiment using a homology-directed repair (HDR) template to insert selected 3q21.3 breakpoint-RE sequences or *G2DHE* in the vicinity of the *EVI1* locus in the *EVI1*-positive myeloid leukemia reporter cell line K562 that does not harbor inv(3)/t(3;3) rearrangements (Figure 2A, top panel) (Ottema *et al.*, 2021, under review). The presence of a *T2A-eGFP* fusion sequence inserted downstream of *EVI1* allows for correlation of EVI1 expression with the synchronously expressed green fluorescent protein (GFP). The parental reporter cell line is tolerant of increased EVI1 levels given that its baseline expression is already increased in K562. The insertion sequences were derived from 3q21.3 breakpoints of two leukemia cases: AML 3071, a patient with inv(3) AML and MOLM-1, a near-triploid myeloid leukemia cell line harboring two chromosome 3 alleles with inv(3) (Figure 2A, bottom panel).[7] The HDR templates were inserted in the corresponding 3q26.2 breakpoint loci as found in AML 3071 and MOLM-1, that is downstream of *EVI1* and within the last *EVI1* intron, respectively.

Single-cell clones validated by polymerase chain reaction (PCR) and Sanger sequencing (*Online Supplementary Figure S1A*) harboring the ectopic *G2DHE* showed a shift in GFP fluorescence indicating successful *EVI1* activation, whereas clones with 3q21.3 breakpoint-RE sequences showed no change in the GFP signal compared with untreated cells (Figure 2B). Furthermore, single-cell clones were analyzed by quantitative PCR (qPCR) and western blot, which showed results consistent with the flow cytometry analysis (Figure 2C and D, respectively), suggesting that the ectopic activation of *EVI1* occurs via *G2DHE*, whereas breakpoint-RE themselves are insufficient to induce *EVI1* transcriptional activation in the K562 reporter cell line.

In order to further dissect a potential regulatory role of breakpoint-RE in 3q-AML, a reciprocal experimental CRISPR-Cas9 approach was applied to delete the original breakpoint-RE in MOLM-1 and UCSD-AML1, the latter being a t(3;3) AML cell line. We expressed pairs of single guide RNA (sgRNA) in stably Cas9-expressing cells to induce a segmental deletion of a fragment containing either the inverted (MOLM-1) or translocated (UCSD-AML1) breakpoint-RE at 3q26.2 on the rearranged alleles located within the last *EVI1* intron in MOLM-1, and upstream of the *EVI1* promoter in UCSD-AML1 (Figure 3A). Targeting on the non-rearranged allele was expected to result only in generation of indels at the 3q21.3 and 3q26.2 site but not segmental RE deletions. In total, two MOLM-1 and six UCSD-AML1 clones harboring the desired deletion validated by PCR and Sanger sequencing (*Online Supplementary Figure S1B*) were derived successfully from single cells. Together with the nontargeting control (NTC) clones (targeted with sgRNA against mCherry and eGFP) and the wild-type (WT) cell line, we performed phenotypic analysis of obtained deletion clones. We observed no differences in proliferation between deletion and control samples (Figure 3B). Slightly reduced EVI1 expression on mRNA and protein level was observed exclusively in the MOLM-1 deletion clones (Figure 3C, left panel).

In order to identify potential genome-wide effects of CRISPR-Cas9-induced RE deletion, we performed genomic and epigenomic analyses of the MOLM-1 and UCSD-AML1 deletion and control clones using circularized chromatin conformation capture sequencing (4C-Seq) and chromatin immunoprecipitation followed by

sequencing (ChIP-Seq). Neither 4C-Seq nor ChIP-Seq revealed an impact of RE deletion on the interaction frequency of *G2DHE* with the *EVI1* promoter or on the deposition of active chromatin marks, such as H3K27ac and H3K4me3, in any of the two edited cell lines (results for MOLM-1 shown in the *Online Supplementary Figure S2A*), making it unlikely that the expression changes in EVI1

are caused by changes in the regulatory function of *G2DHE*.

Since the reduction of EVI1 expression was specific only to the MOLM-1, but not UCSD-AML1 deletion clones, we speculated that this effect might be due to other features present in the sequence deleted in MOLM-1, rather than the consequence of the breakpoint-RE-
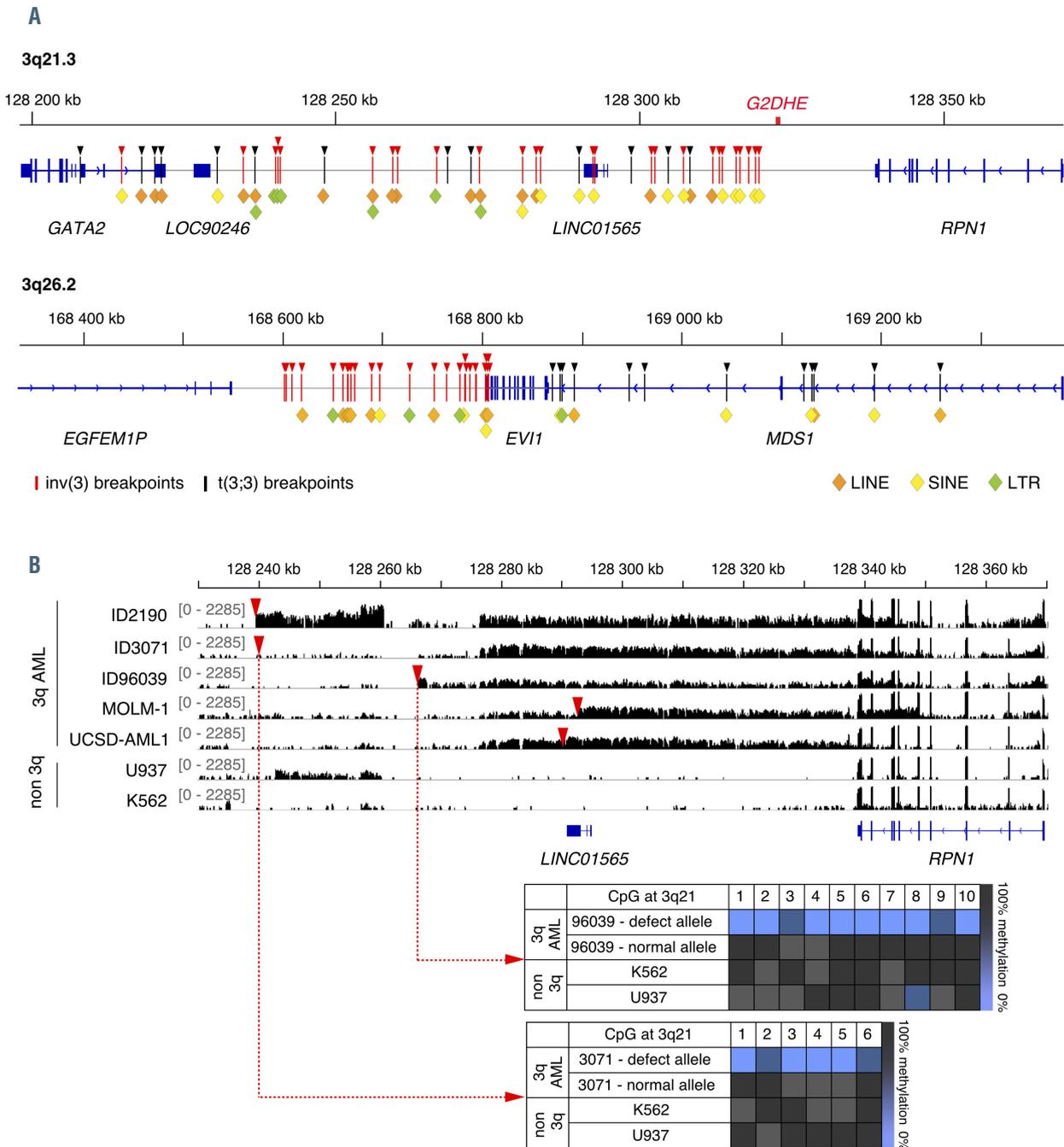
Figure 1. Chromosomal breakpoints in acute myeloid leukemia inv(3)/t(3;3) (3q-AML) are enriched in retroelements. (A) 3q-capture sequencing revealed a characteristic breakpoint pattern (red and black arrowheads) at 3q21.3 (upper tracks) and 3q26.2 (lower tracks). At 3q21.3, a breakpoint-free region downstream (left) of *RPN1* was identified as a commonly translocated segment containing *G2DHE*. At 3q26.2, breakpoints of inversion cases map exclusively downstream (left) or within the last *EVI1* intron, whereas translocation cases have breakpoints upstream of *EVI1*. Color-coded diamonds indicate the position of RE: long interspersed elements (LINE), short interspersed elements (SINE) or long terminal repeats (LTR), annotated by RepeatMasker. (B) RNA sequencing of 3q-AML samples revealed a characteristic RNA readthrough signature at 3q21.3. Bisulfite amplicon sequencing on representative 3q-AML samples showed focal hypomethylation of breakpoint regions (red arrowheads) on the rearranged allele but not on the normal allele and non-3q samples.

deletion. To this end, we reanalyzed the deleted sequences in the two cell lines. The original 3q21.3 fragment overlapping with *LINC01565* relocated to the 3q26.2 site and deleted in the MOLM-1 clones displays high degree of conservation (Figure 3A, left panel), which might indicate functional importance of this sequence. ENCODE transcription factor (TF) ChIP data generated in K562 cells revealed a plethora of TF binding to this region (*Online Supplementary Figure S2B*). Some of these TF have predicted binding sites within the *G2DHE* sequence, including IKZF1, MAX, TAL1 and MAZ (Kiehlmeier *et al.*, 2021, under review). Furthermore, EVI1 has been func-

tionally linked to some of these TF, including MTA1/2, HDAC1/2 and GATAD2B. All these proteins belong to the nucleosome remodeling and deacetylase (NuRD) complex,[8] which was shown to specifically interact with the MDS-EVI1 (PRDM3) but not the EVI1 protein,[9] while both EVI1 protein isoforms were shown to interact with HDAC1.[10] Contrary to MOLM-1, no TF binding in the region deleted in the UCSD-AML clones was found (*Online Supplementary Figure S2B*).

Taken together, the observed reduction in EVI1 expression upon breakpoint-RE deletion in MOLM-1 is more likely the consequence of TF binding loss within the
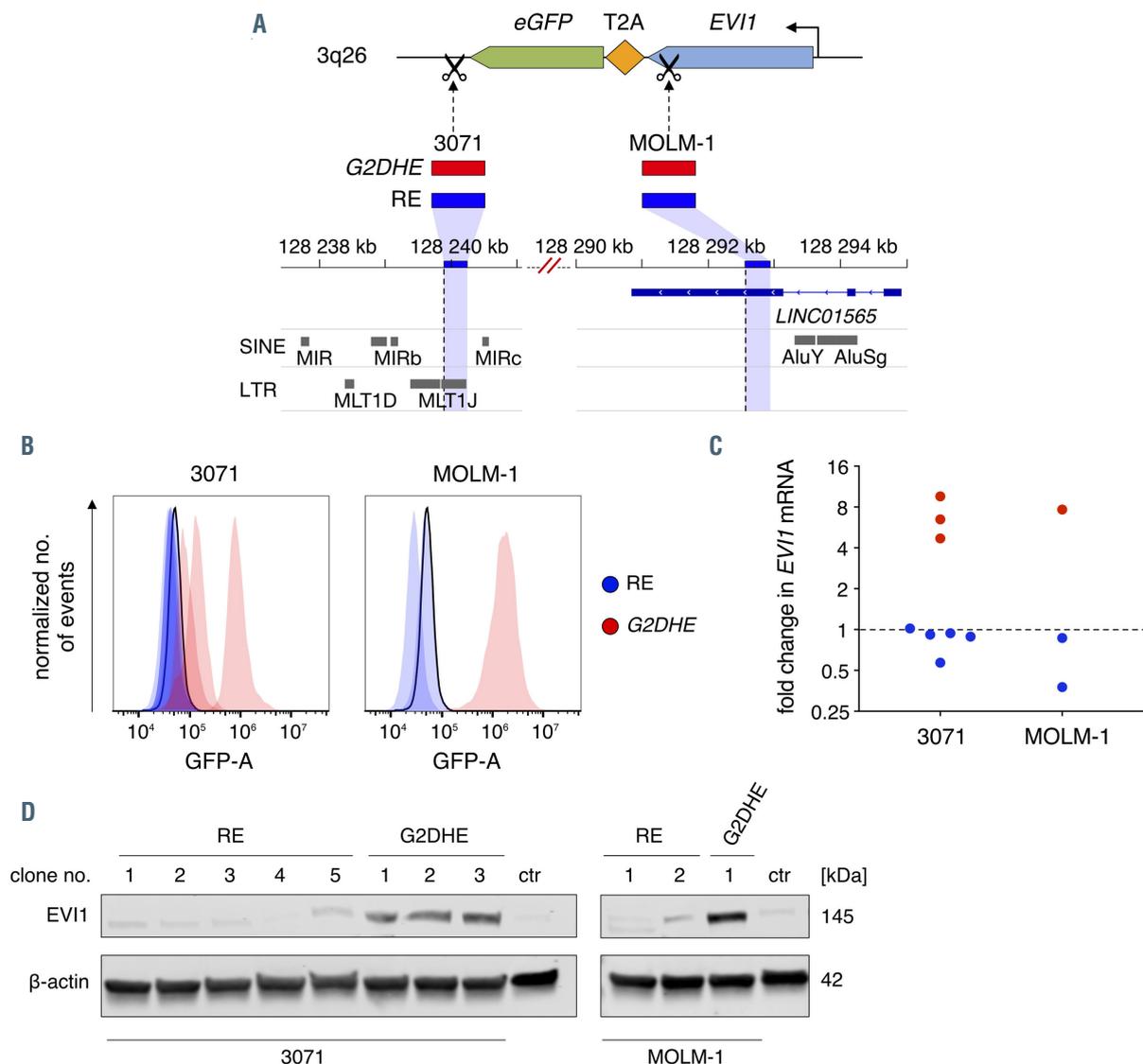
Figure 2. Ectopic activation of *EVI1* occurs via *G2DHE*, whereas breakpoint-associated retroelements do not display activating potential in the K562 cells. (A) Experimental strategy of CRISPR-Cas9-mediated genomic insertions. Donor templates containing either *G2DHE* (736 bp, red boxes) or 3q21.3 breakpoint-associated retroelements (breakpoint-RE) from selected 3q-AML cases (3071: 353 bp and MOLM-1: 280 bp, blue boxes) were inserted in the non-3q K562 reporter cell line (K562 *eGFP-T2A-EVI1*) using CRISPR-Cas9 downstream of *EVI1* and within the last *EVI1* intron. A genomic view shows the origin of 3q21.3 breakpoint sequences used in CRISPR experiments, dashed lines indicate 3q21.3 breakpoints in 3071 (left) and MOLM-1 (right). 3,071 homology-directed repair (HDR) template contains a part of MLT1J LTR element. *G2DHE* (red) and 3q21.3 breakpoint-RE samples (blue) are consistently colored throughout the figure. (B) Flow cytometry analysis on the K562 *eGFP-T2A-EVI1* single clones bearing the desired *G2DHE* or 3q21.3 breakpoint insertions. Peaks corresponding to the green fluorescent protein (GFP) signal from single clones targeted at the same region are presented together on one graph, with peaks for untreated cells shown as a black outline. (C) Quantitative polymerase chain reaction (qPCR) analysis of the *EVI1* mRNA levels in single clones shown in (B), relative to *HMBS* and normalized to the untreated *eGFP-T2A-EVI1* cells. (D) Representative western blot of the full-length EVI1 isoform from the clones and untreated cells (ctr) shown in (B) and (C). Data shown in (C) are means of three technical replicates from one independent experiment. SINE: short interspersed nuclear elements; LTR: long terminal repeats.
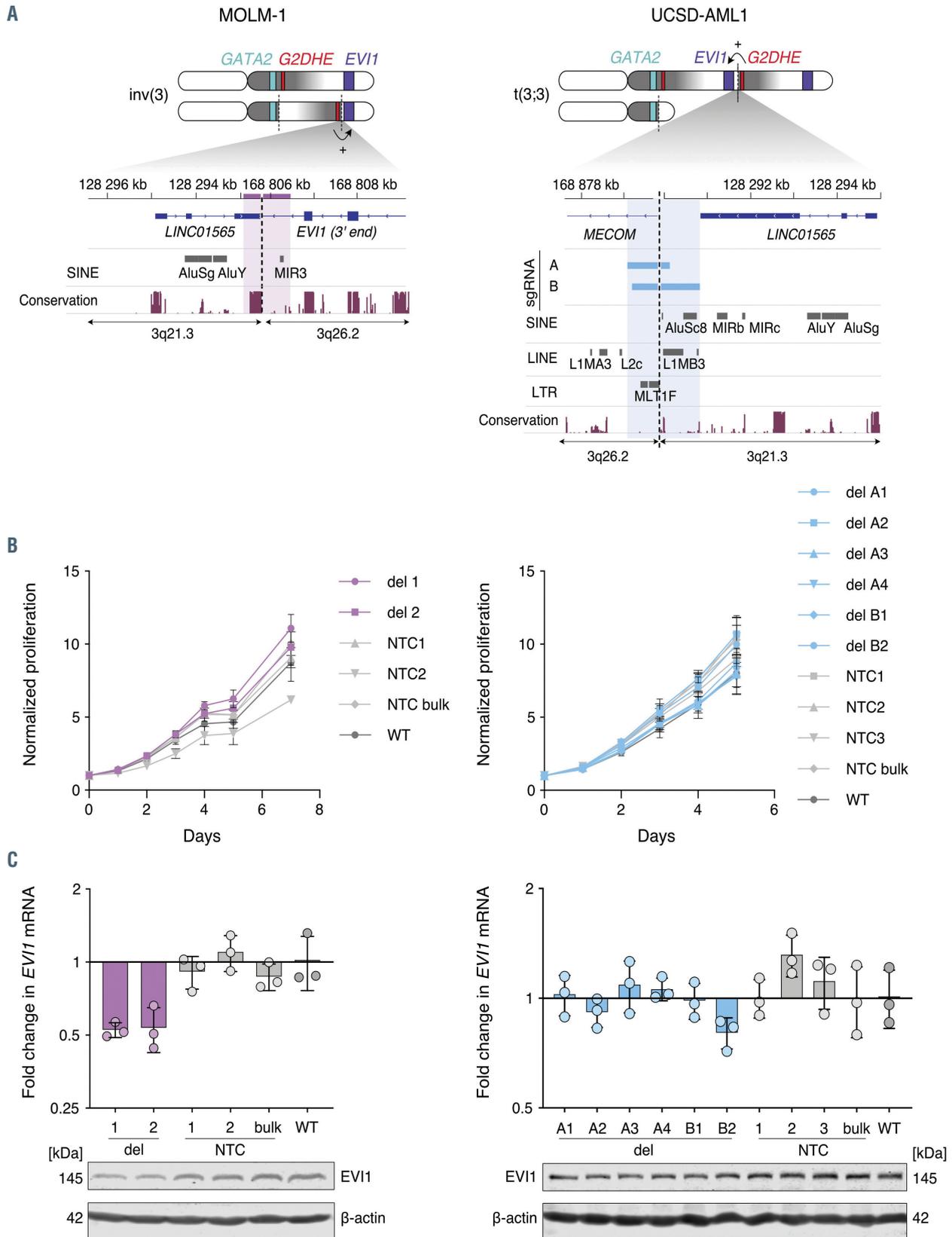
Figure 3. MOLM-1 single clones harboring the deletion of breakpoint-associated retroelements exhibit no phenotypic changes but show reduced *EVI1* expression. (A) Illustration of the CRISPR-Cas9 mediated segmental deletion surrounding the breakpoint on the rearranged allele of chromosome 3 in MOLM-1 (left) and UCSD-AML1 (right). The deleted regions (MOLM-1: purple, UCSD-AML1: blue) encompass a MIR3 SINE element in MOLM-1, and representatives of all three RE subclasses in UCSD-AML1. Normalized proliferation (B) of deletion (del), nontargeting control (NTC) clones, NTC bulk of cells and wild-type (WT) cell line. (C) Quantitative polymerase chain reaction analysis of *EVI1* mRNA levels (top panel) and representative western blot (bottom panel) of the full-length EVI1 isoform from samples shown in (B). Data shown in (B) and (C, top panel) are mean ± standard deviation from three independent experiments. Conservation across vertebrate genomes is taken from phastCons.[13] SINE: short interspersed nuclear elements; LINE: long interspersed nuclear elements; LTR: long terminal repeats.

G2DHE super-enhancer structure. However, we could not observe any specific pattern of TF binding sites commonly relocated to the *EVI1* locus in 3q-AML patient and cell line data. Meanwhile, we conclude that breakpoint-RE are not essential for the regulation and maintenance of *EVI1* expression in 3q-AML.

In summary, our data show that RE are highly enriched at inv(3)/t(3;3) breakpoint sites in AML and represent a source of genomic vulnerability without providing additional regulatory or activating signal for *EVI1* in this leukemia subtype, as evidenced by CRISPR-Cas9 editing experiments in 3q-rearranged cell lines. However, we cannot exclude the involvement of full-length source RE sequences in the formation of 3q rearrangements in earlier stages of malignant transformation. Since many retrotransposition-competent RE often become truncated or undergo internal reshuffling upon insertion in a new genomic location,[11,12] the presence of resulting chimeric breakpoint-RE sequences would bear no functional consequences, which stands in line with the lack of effects observed upon CRISPR-Cas9 editing of breakpoint-RE fusion sequences in our cell line models.

Additionally, the data from the K562 reporter cell line provide an orthogonal confirmation of the minimal *G2DHE* being sufficient for *EVI1* transcriptional activation.[1]

Jagoda Mika,[1] Sophie Ottema,[2] Sandra Kiehlmeier,[1] Sabrina Kruse,[1] Leonie Smeenk,[2] Judith Müller,[1] Sabrina Schweiggert,[1] Carl Herrmann,[3] Mathijs Sanders,[2] Ruud Delwel[2] and Stefan Gröschel[1,4,5]

[1]Molecular Leukemogenesis Group, German Cancer Research Center, Heidelberg, Germany; [2]Department of Hematology, Erasmus University Medical Center, Oncode Institute, Rotterdam, the Netherlands; [3]Health Data Science Unit, Medical Faculty Heidelberg and BioQuant, Heidelberg, Germany; [4]Internal Medicine V, Heidelberg University Hospital, Heidelberg, Germany and [5]Oncology Center Worms, Worms, Germany

Correspondence:

STEFAN GROESCHEL - stefan.groeschel@dkfz.de

References

1. Gröschel S, Sanders MA, Hoogenboezem R, et al. A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. Cell. 2014;157(2):369-381.
2. Yamazaki H, Suzuki M, Otsuki A, et al. A remote GATA2 hematopoietic enhancer drives leukemogenesis in inv(3)(q21;q26) by activating EVI1 expression. Cancer Cell. 2014;25(4):415-427.
3. Rodriguez-Martin B, Alvarez EG, Baez-Ortega A, et al. Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. Nat Genet. 2020;52(3):306-319.
4. Tubio JMC, Li Y, Ju YS, et al. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. Science. 2014;345(6196):1251343.
5. Lee E, Iskow R, Yang L, et al. Landscape of somatic retrotransposition in human cancers. Science. 2012;337(6097):967-971.
6. Cajuso T, Sulo P, Tanskanen T, et al. Retrotransposon insertions can initiate colorectal cancer and are associated with poor survival. Nat Commun. 2019;10(1):4022.
7. Matsuo Y, Adachi T, Tsubota T, Imanishi J, Minowada J. Establishment and characterization of a novel megakaryoblastic cell line, MOLM-1, from a patient with chronic myelogenous leukemia. Hum Cell. 1991;4(3):261-264.
8. Torrado M, Low JKK, Silva APG, et al. Refinement of the subunit interaction network within the nucleosome remodelling and deacetylase (NuRD) complex. FEBS J. 2017;284(24):4216-4232.
9. Ivanochko D, Halabelian L, Henderson E, et al. Direct interaction between the PRDM3 and PRDM16 tumor suppressors and the NuRD chromatin remodeling complex. Nucleic Acids Res. 2019; 47(3):1225-1238.
10. Vinatzer U, Taplick J, Seiser C, Fonatsch C, Wieser R. The leukaemia-associated transcription factors EVI-1 and MDS1/EVI1 repress transcription and interact with histone deacetylase. Br J Haematol. 2001;114(3):566-573.
11. Sassaman DM, Dombroski BA, Moran J V, et al. Many human L1 elements are capable of retrotransposition. Nat Genet. 1997; 16(1):37-43.
12. Brouha B, Badge RM, Farley AH, et al. Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci U S A. 2003;100(9):5280-5285.
13. Siepel A, Bejerano G, Pedersen JS, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 2005;15(8):1034-1050.