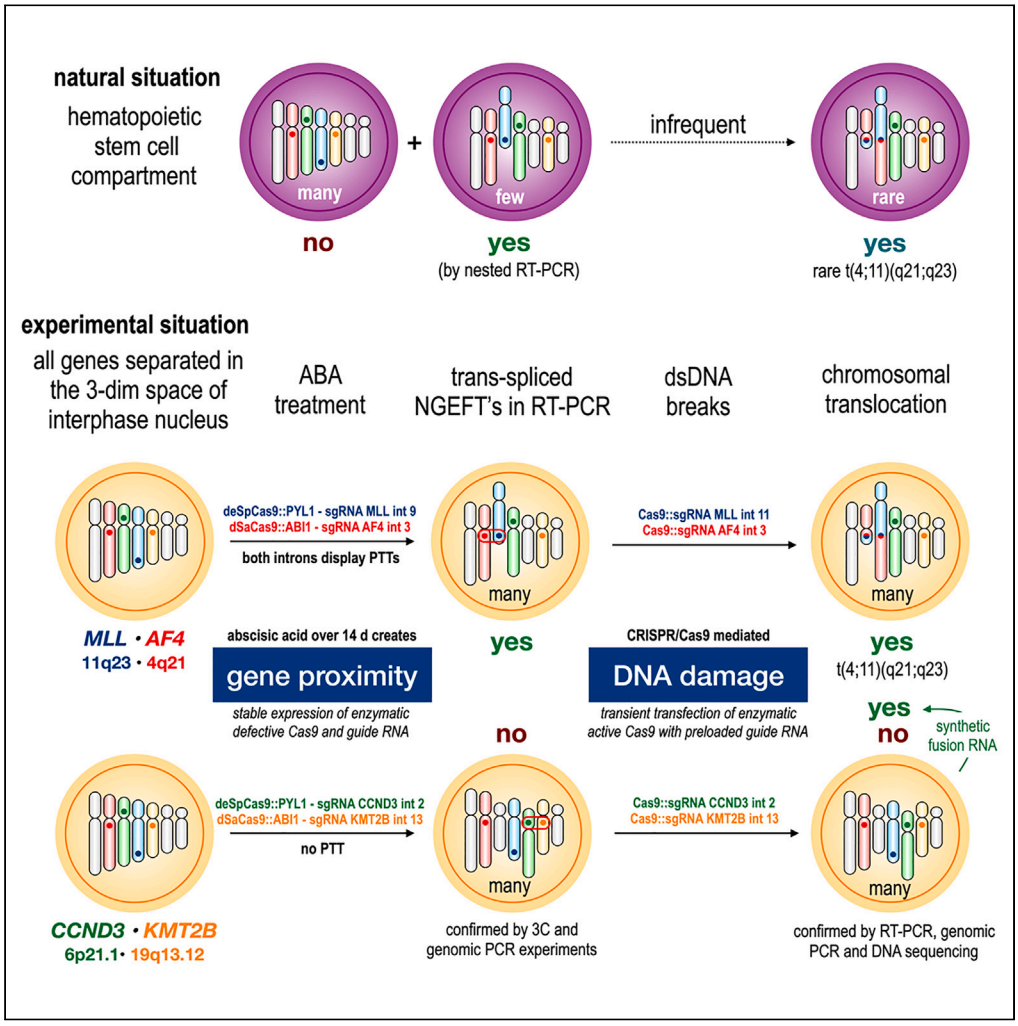
iScience

CellPress
OPEN ACCESS

Article

How chromosomal translocations arise to cause cancer: Gene proximity, *trans*-splicing, and DNA end joining



Patrick Streb, Eric Kowarz, Tamara Benz, Jennifer Reis, Rolf Marschalek

rolf.marschalek@em.uni-frankfurt.de

Highlights

Certain genes encode (PTTs

Genes encoding PTTs and in close proximity cause *trans*-spliced fusion mRNAs

Trans-spliced fusion mRNAs (NGEFTs) are essential for the onset of CTs

NGEFTs drive the DNA repair process to wrongly create specific CTs

Streb et al., iScience *26*, 106900
June 16, 2023
https://doi.org/10.1016/j.isci.2023.106900

iScience

CellPress
OPEN ACCESS

Article

How chromosomal translocations arise to cause cancer: Gene proximity, *trans*-splicing, and DNA end joining

Patrick Streb,[1] Eric Kowarz,[1] Tamara Benz,[1] Jennifer Reis,[1] and Rolf Marschalek[1,2,*]

SUMMARY

Chromosomal translocations (CTs) are a genetic hallmark of cancer. They could be identified as recurrent genetic aberrations in hemato-malignancies and solid tumors. More than 40% of all "cancer genes" were identified in recurrent CTs. Most of these CTs result in the production of oncofusion proteins of which many have been studied over the past decades. They influence signaling pathways and/or alter gene expression. However, a precise mechanism for how these CTs arise and occur in a nearly identical fashion in individuals remains to be elucidated. Here, we performed experiments that explain the onset of CTs: (1) proximity of genes able to produce prematurely terminated transcripts, which lead to the production of (2) *trans*-spliced fusion RNAs, and finally, the induction of (3) DNA double-strand breaks which are subsequently repaired via EJ repair pathways. Under these conditions, balanced chromosomal translocations could be specifically induced. The implications of these findings will be discussed.

INTRODUCTION

More than 25 years ago, it was demonstrated that chromosomal translocations (CTs) arise due to DNA double-strand breaks which are subsequently repaired via the non-homologous end-joining pathway (NHEJ).[1–3] This finding was quite important because until that time the onset of CTs was assumed to be a RAG1/RAG2-mediated event, associated with the development of leukemias or lymphomas. This initial finding has changed our view of these genetic lesions and has been validated by others.[4] Today, the end-joining DNA repair pathways have been functionally subdivided into the classical NHEJ (c-NHEJ), the alternative end-joining pathway (A-EJ), and the microhomology-mediated end-joining pathway (MMEJ).[5] The latter pathway even requires the resection of DNA sequences in order to align by microhomology the non-homologous DNA ends. All together, these pathways explain the very early findings about deletions, duplications, inversions, mini-direct repeat fusions, and the presence of filler DNA that have been identified at the junctions of the fused chromosome sequences in the derivative chromosomes of patients with leukemia.[1,2]

In principle, DNA double-strand breaks may occur anywhere in the genome, but some chromosomal regions (or certain genes) may exhibit a much higher intrinsic genetic instability.[6] Several labs have attempted to figure out the underlying mechanisms responsible for the onset of CTs, in particular those affecting the *MLL/KMT2A* gene—from now on termed only *MLL*: specific SAR/MAR chromatin loop structures in *MLL* and *AFF1*,[7–9] an *MLL* intron 11-specific DNase I hypersensitive site,[10] topoisomerase II binding sites along the breakpoint cluster region of *MLL*,[11–14] a gene internal transcription initiation site within *MLL* intron 11,[15] and a site-specific DNA cleavage within *MLL* intron 11 due to early apoptotic processes.[16–19] The latter gained much interest because classical inhibitors of apoptosis, such as Z-VAD-FMK, were able to prevent the onset of DNA double-strand breaks within the breakpoint cluster region of *MLL/KMT2A* even following treatment with chemicals such as the topoisomerase II inhibitor VP16.[16,19] All this is quite important because the *MLL* gene is the most promiscuous gene that can be found in CTs, with more than 100 known fusion partners today.[20]

The creation of CTs is an interesting genetic mechanism since it resembles the genetic mechanism underlying exon shuffling. An illegitimate recombination event between two unrelated genes results in the creation of reciprocal fusion genes. Since most translocations occur in intronic sequences, exons of both

[1]Goethe-University, Department Biochemistry, Chemistry & Pharmacy, Institute of Pharmaceutical Biology, Max-von-Laue-Street 9, 60438 Frankfurt am Main, Germany

[2]Lead contact

*Correspondence: rolf.marschalek@em.uni-frankfurt.de

https://doi.org/10.1016/j.isci.2023.106900

genes could be spliced together, thereby producing chimeric open reading frames that exhibit potent oncogenic potential.[21] According to the available data of the Sanger Cancer Census database (https://cancer.sanger.ac.uk/cosmic/census/), 735 different cancer genes have been described, of which 315 genes were identified to be involved in CTs (associated with 193 hemato-malignant and 122 solid tumors). This clearly indicates that the onset of CTs represents a general cancer-causing mechanism.

However, two questions remain: 1. What is the reason that cells fuse the open ends of two different chromosomes instead of performing intra-chromosomal DNA repair in the case of DNA damage situations? 2. Why are particular genes recurrently affected in different individuals?

From a mathematical point of view, a pure selection hypothesis was not sufficient to explain the frequency and recurrence of these genetic alterations. Therefore, other explanations are needed to solve the open questions. First of all, the nuclear architecture and specific features of genes involved in CTs are probably important. In the interphase of the cell cycle, chromosomes are relaxed and organized in "chromosome territories" within the 3-dimensional architecture of the cell's nucleus.[22,23] These chromosome territories are tightly packed with active chromatin loops protruding out, and thus, it was assumed that illegitimate recombination events could occur between two adjacent chromatin loops where active genes are brought into close proximity. Since genes in close proximity are often transcribed together in the same transcription factory,[24,25] they are probably close enough to allow the formation of a CT upon DNA damage and subsequent DNA repair. This has been elegantly demonstrated by RNA- and DNA-FISH experiments for PML and RARA,[26] IgH and cMYC,[27,28] or the genes encoding BCR and ABL proteins.[29] Thus, genes involved in CTs are located in proximity in a tissue-specific manner.

Another hint pointing in the same direction came from experiments in which peripheral blood or biopsied bone marrow cells of healthy individuals were examined. Several research groups were able to demonstrate the occurrence of known oncogenic fusion transcripts by conventional RT-PCR experiments: *MLL::AFF1* fusion transcripts or *MLL*-partial tandem duplications,[30–32] fusion transcripts for *BCR::ABL*,[33,34] *ETV6::RUNX1*,[35,36] *RUNX1::ETO*,[36] *PML::RARα*,[37] *NPM::ALK*, and *ATIC::ALK*.[38,39] All these pro-neoplastic fusion RNAs were identified in blood cells. However, these were present in the absence of any detectable genomic rearrangement of the corresponding genetic loci. Therefore, these arbitrary fusion transcripts were named "non-genomically encoded fusion transcripts" (NGEFTs). Besides these findings in blood cells, a similar phenomenon was identified in other human tissues. The *JAZF1::JJAZ1* fusion RNA—normally, created by a t(7; 17)(p15; q21) in endometrial stromal sarcomas, could be readily detected in normal endometrial stromal cells,[40,41] and moreover, this low abundance *JAZF1::JJAZ1* fusion mRNA produced a chimeric protein that exhibited anti-apoptotic effects. Therefore, these NGEFTs may encode chimeric fusion proteins that provide gain-of-survival functions for certain tissues. Similar findings were reported for the *trans*-spliced *SLC45A3::ELK4* fusion transcript identified in normal prostate and prostate cancer cells.[42] Thus, the production NGEFTs seems to be a common feature of certain genes that are all known to be involved in CTs, in particular the *MLL-AFF1* NGEFT mimicking the t(4; 11) translocation.[31,32]

In order to understand these findings in more detail, we investigated a series of genes involved in CTs for their transcriptional properties (*MLL/KMT2A*, *AFF1*, *MLLT3*, *ELL*, *and MLLT1*). It turned out that all these genes display a unique feature, namely the production of "premature terminated transcripts" (PTTs). PTTs are mostly observed from poisoned polymerase II at nucleotide +50, when awaiting the capping of the nascent RNA, but it may also occur during transcription of the gene body, mainly in intronic regions. PTTs are therefore a general feature for many genes of the human genome (~16%),[43] and many of these intronic PTTs become poly-adenylated at cryptic poly-A sites encoded by the same introns.[44] In specific qRT-PCR experiments with the above-mentioned genes, we were able to demonstrate that these intronic PTTs can be found in a relative abundance of 10%–20% when compared to the full-length poly-adenylated transcripts.[45] Moreover, the premature termination always occurred in introns that belong to the breakpoint cluster region of those genes (*MLL*-intron 8, *MLL*-intron 9, *AFF1*-intron 3, *MLLT3*-intron 5, *ELL*-intron 2, *and MLLT1*-intron 3). A second study from our laboratory revealed the same PTT features also for *ETV6*-intron 5, *EWSR1*-intron 7, *NUP98*-intron 12, *NUP98*-intron 13, and *RUNX1*-intron 6.[46] All these genes are involved in a large number of different CTs. In contrast, several other genes, such as *GAPDH*, *ACTB*, *HSPCB*, *CCND3*, *RPL3*, or *KMT2B*, did not display any PTT formation capacity. Thus, we identified two sets of genes which could be used as targets as well as control genes for further experimental studies.

Intronic PTTs are quite interesting transcripts since they contain an unsaturated splice donor site (SD) in their 3′-region, something which does not exist in full-length transcripts due to the fact that all SD sites (5′-intronic sequence) in a primary transcript could be saturated by corresponding splice acceptor sites (SA; 3′-intronic sequence) during canonical splicing events. Therefore, we established specific assays that allowed us to detect *trans*-spliced NGEFTs.[45,46] This way we revealed that a *trans*-spliced fusion transcript between *MLL/KMT2A* exon 9 and *AFF1* exon 4 can be readily identified in peripheral blood mononuclear cells isolated from healthy individuals. Other genes, such as *ENL*, displayed a full spectrum of *trans*-spliced fusions with other genes (*RPL18A, SFPQ, and ALKB7*). In summary, PTTs appear to have a specific driver function to artificially produce *trans*-spliced fusion mRNAs in low in abundance, and are synonymous with the observed NGEFTs.

Thus, we and others hypothesized that these NGEFTs are the "missing link" to explain the creation of CTs following DNA damage. We posed the idea that these NGEFTs somehow serve as templates to aberrantly drive the repair process.[45–48] However, up to that time, no experimental tools were available to set-up experiments to validate this hypothesis.

Here, we present experimental data that are based on these old observations in conjunction with state-of-the-art technologies to establish a mammalian cell culture system which allowed us to perform all the necessary steps in a sequential fashion to create specific chromosomal translocations. Resulting CTs were analyzed at the genomic DNA level by Sanger sequencing, thereby validating that they all resulted from EJ-mediated DNA repair pathways (either c-NHEJ, a-EJ, or MMEJ). The consequences of our findings will be discussed.

RESULTS

Gene proximity in combination with PTT production results in *trans*-spliced fusion mRNAs

Based on published data, we chose the HEK293T cell line where our 4 candidate genes, *MLL/KMT2A*, *AFF1*, *CCND3*, and *KMT2B*, were all not located in proximity. This was investigated by 2 criteria: (1) no NGEFT formation capacity in nested RT-PCR experiments and (2) absence of a positive PCR band in 3C experiments (data not shown).

The cell line was stably transfected with 3 different vector pairs that constitutively express two different dead Cas9 genes (dSaCas9 and deSpCas9)[49] fused in-frame with small domains that encode abscisic acid-binding pockets (dSaCas9:ABI1 and deSpCas9:PYL1; see Figures 1A–1C).[50] The two stably integrated Sleeping Beauty vector constructs also transcribe two appropriate guide RNA cassettes with corresponding spacer and scaffolds for the 2 different dCas9 that target either *MLL/KMT2A* intron 9 (11q23.3) or *AFF1* intron 3 (4q21), respectively (Figure 1A). Similarly, vectors were constructed with sgRNA cassettes targeting *CCND3* intron 2 (6p21.1) or *KMT2B* intron 13 (19q13.2) (Figure 1B), as well as a control vector pair that transcribe two non-targeting sgRNAs (Figure 1C). All 3 vector pairs were used to establish 3 stable cell lines, where the nuclear architecture could be changed over time by a treatment with abscisic acid (ABA). Importantly, the control cell line (Figure 1C) was used as a negative control for all experiments (ntRNA) following the complete workflow depicted in Figures 1D and 1E.

All 3 established cell lines were treated with ABA for 14 days (Figures 2A and 2B). Subsequently, chromosome conformation capture (3C) experiments were carried out with cells from day 0 (data not shown) and 14 (Figure 2C, left panel). These 3C experiments revealed that both gene pairs, *MLL/AFF1* and *CCND3/KMT2B*, were successfully brought into close vicinity, while the non-targeting control cells did not exhibit any PCR band, which indicated that ABA treatment alone was not causing any proximity of the 4 tested genes.

The *MLL* gene was known to produce an *MLL* intron 9 PTT, while *CCND3* and *KMT2B* do not produce any PTT. Therefore, we next investigated *trans*-splicing activity. As shown in Figure 2C (right panels), a very strong *MLL::AFF1* fusion transcript could be amplified by RT-PCR. This band of 402 bp turned out to be the typical *MLL* exon 9:*AFF1* exon 4 fusion RNA, known to derive from the *trans*-splicing activity of the *MLL* intron 9 PTT in normal blood cells. Since no *AFF1::MLL* fusion transcript was visible, we concluded that this specific NGEFT resulted from the close proximity of *MLL* and *AFF1* after the ABA treatment, and was not derived from a t(4; 11) chromosomal translocation. Importantly, no such fusion RNA was observed in the experiment with the 2 other genes, *CCND3* and *KMT2B*, respectively. This result validated our earlier findings that both *CCND3* and *KMT2B* are not able to produce any PTTs,[45] and thus, are *per se*

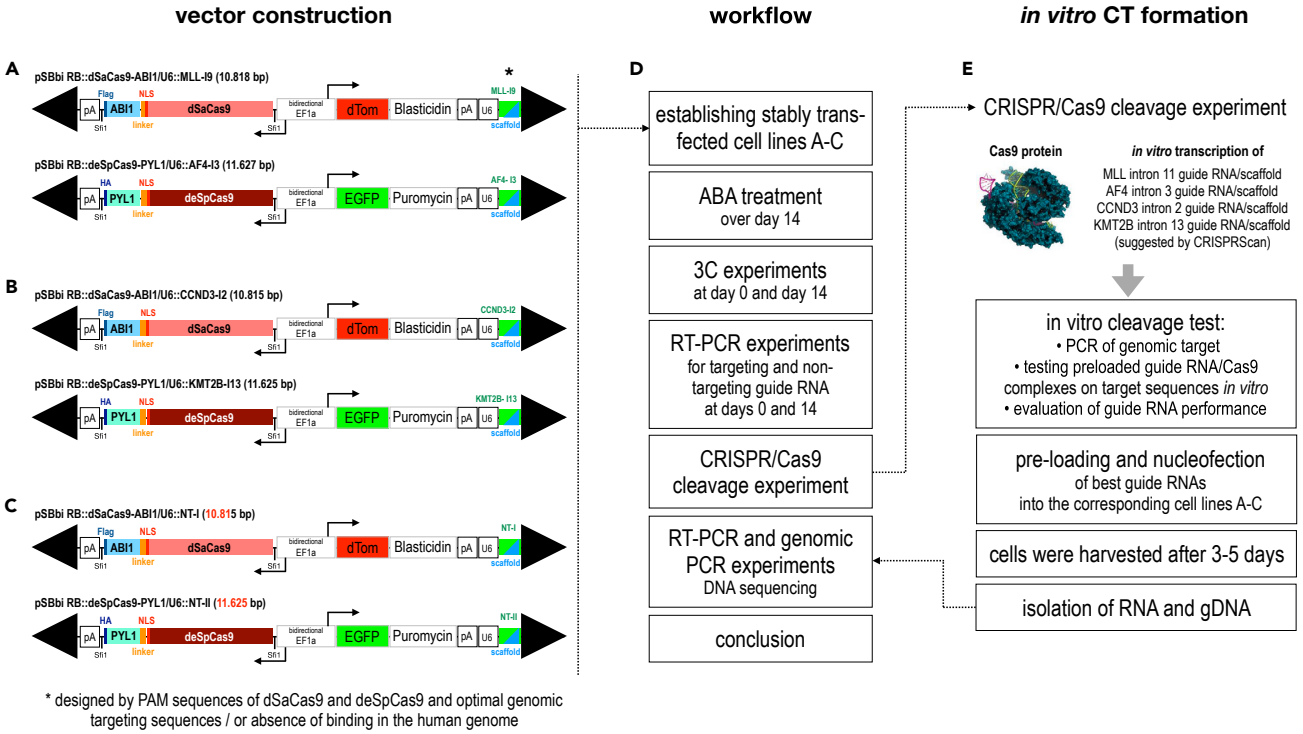


Figure 1. Vector design and workflow

The vector design of all 3 vector pairs used in our study is depicted. Sleeping Beauty vectors pSBbi-RB (AddGene #60522) and pSBbi-GP (AddGene #60511) were used as backbones. The two fusion genes dSaCas9:ABI1 and deSpCas9:PYL1 were cloned in these 2 vector backbones, respectively. Both vector backbones only differ by the constitutively expressed fluorescence marker and resistance genes.

(A) A pair of vector constructs encoding the *MLL* intron 9 and *AFF1* intron 3 single-stranded guide RNAs including the scaffold for binding either of the 2 different dead Cas9 mutant proteins.

(B) A pair of vector constructs encoding the *CCND3* intron 2 and *KMT2B* intron 13 single-stranded guide RNAs including the scaffold for binding either of the 2 different dead Cas9 mutant proteins.

(C) A pair of vector constructs encoding the non-targeting single-stranded guide RNAs including the scaffold for binding either of the 2 different dead Cas9 mutant proteins.

(D and E) short description of the experimental workflow of this project.

not able to form any *trans*-spliced NGEFT. In conclusion, gene proximity alone is not sufficient to cause the production of NGEFTs, but the combination of gene proximity and PTT formation capacity is the molecular source for the observed *MLL::AFF1* NGEFT.

Chromosomal translocations generated by CRISPR/Cas9 technology

Next, we wanted to investigate whether the proximity of the 2 corresponding genes in the cell lines can be used to successfully produce a chromosomal translocation. For this purpose, several new guide RNAs were designed using the CRISPRscan server (www.crisprscan.org) and then functionally tested in an *in vitro* DNA hydrolysis assay. The 4 genomic target sequences (parts of *MLL* intron 11, *AFF1* intron 3, *CCND3* intron 2, and *KMT2B* intron 13) were produced by genomic PCR experiments, and 1–6 potential guide RNAs were tested on each intron. The different sgRNAs were pre-incubated with Cas9 protein and subsequently incubated *in vitro* with the corresponding genomic target DNA sequence. In all cases, at least one sgRNA was identified that cleaved the cognate genomic target sequences properly (see Figure 2D). These new sgRNAs were then used in combination Cas9 protein-transfection experiments in appropriate combinations for all 3 cell lines (*MLL/AFF1*; *CCND3/KMT2B*; non-targeting cells) at day 14 of ABA treatment. After 3–5 days, aliquots of these cells were harvested, RNA and genomic DNA was isolated, and investigated by RT-PCR and genomic PCR experiments.

To our surprise, we were able to identify fusion RNAs for *AFF1* exon 3:*MLL* exon 12 (AM) and *MLL* exon 11:*AFF1* exon 4 (MA) in the targeting guide RNA experiment, while the non-targeting cell line only

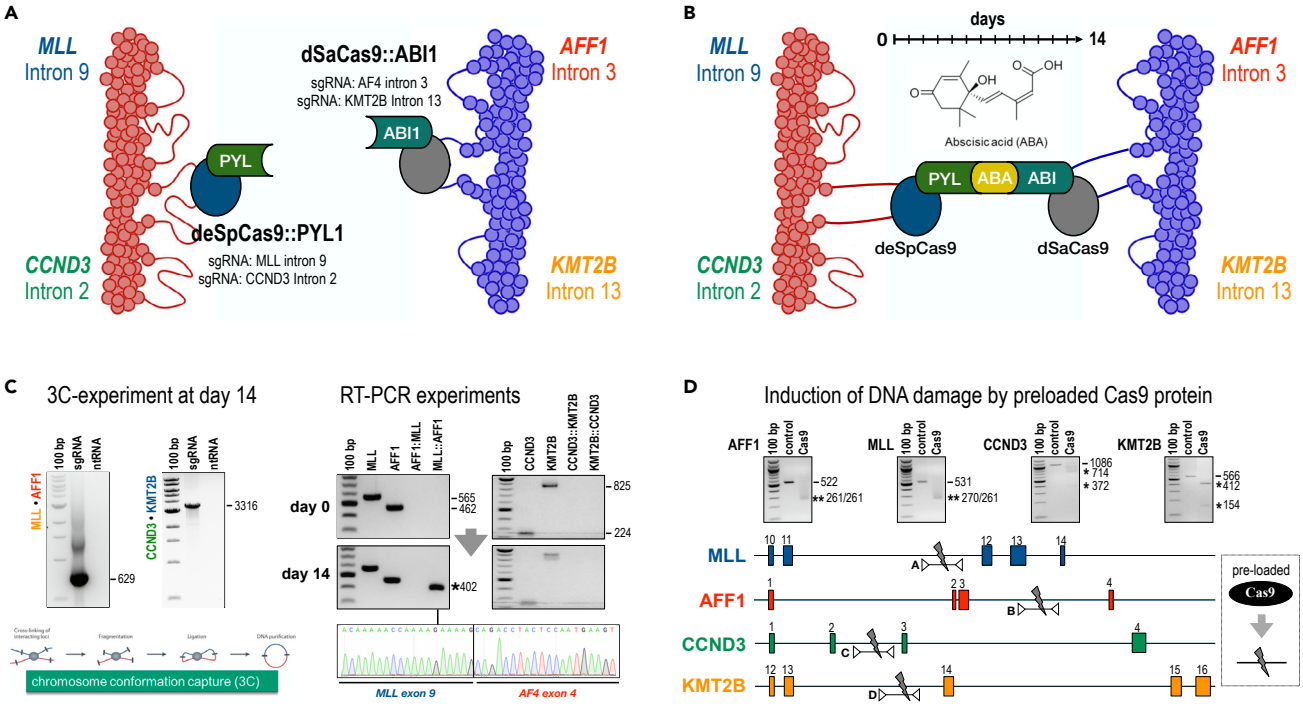


Figure 2. Proximity of genes with PTT's results in *trans*-spliced mRNAs

(A) Scheme of the initial transgenic HEK293T cells that were stably transfected with the corresponding dSaCas9 or deSpCas9 constructs, corresponding guide RNAs, or non-targeting control guide RNAs.

(B) The situation after the treatment with abscisic acid (ABA) for 14 days; both gene pairs (*MLL*/*AFF1* or *CCND3*/*KMT2B*) are indicated.

(C) Chromosome conformation capture (3C) experiments in the presence of sgRNA led to gene proximity at day 14, while the non-targeting control (ntRNA) served as negative control. The RT-PCR experiments performed at day 0 and14 demonstrated that the PTT formation capacity of *MLL* was able to produce a *trans*-spliced *MLL::AFF1* fusion mRNA (shown by the Sanger sequence below), while *CCND3* and *KMT2B* did not.

(D) Cleavage of sgRNA-preloaded Cas9 protein was tested using an *in vitro* cleavage assay using appropriate intronic DNA sequences of all 4 genes. Here, the *MLL* intron 11 was targeted with a second sgRNA to avoid any misinterpretation because of the *trans*-spliced *MLL* exon 9:*AFF1* exon 4 fusion RNA (gene proximity-induced NGEFT).

displayed the *MLL* and *AFF1* wild-type mRNA transcripts (M and A), but remained negative for the fusion mRNAs (AM and MA; Figure 3A). Of interest, the *trans*-spliced *MLL exon 9::AFF1 exon 4* fusion mRNA (NGEFT) was still visible in the day 14 cells. However, the amount of this NGEFT-derived PCR fragment seemed to be much lower when comparing to the PCR bands deriving from the rearranged chromosomes that produced the *MLL::AFF1* and *AFF1::MLL* fusion transcripts. The Sanger sequence of all 3 PCR products (numbered 1–3) is depicted in Figure 3B, validating proper splicing and identity. Unexpectedly, the targeted CRISPR/Cas9 experiment with *CCND3* (C) and *KMT2B* (K) remained negative and displayed neither the appropriate fusions mRNAs (CK or KC; Figure 3C), which indicated that no CT had occurred at the genomic DNA level. This result suggested that the combination of gene proximity and transient DNA damage was not sufficient to cause a CT.

To validate the assumption that the observed *MLL::AFF1* and *AFF1::MLL* fusion transcripts were indeed caused by a chromosomal translocation, we subsequently performed genomic PCR experiments that revealed the fusion of chromosome 4 and 11 in a reciprocal fashion (Figure 3D, 2 left panels). We also tested the genetic rearrangement between *CCND3* and KMT2B (Figure 3D, 2 right panels), but no PCR bands were visible. Several PCR bands that are visible in the non-targeting (nt) and targeting experiment (t) in the genomic *MLL/AFF1* PCR experiment turned out to be PCR artifacts of the long-range genomic PCR experiment (validated by direct sequencing). The 2 genomic PCR bands visible only in the targeting experiment (t) were cut out from the gel, subcloned, and analyzed from single bacterial clones by sequence analysis. In fact, these specific PCR bands revealed a large series of individual chromosomal translocation events with a mean size of 5143 and 4714 bp. As summarized on the left side of Figure 4, individually repaired fusion alleles derived from the translocated der(4) and der(11) were investigated. The

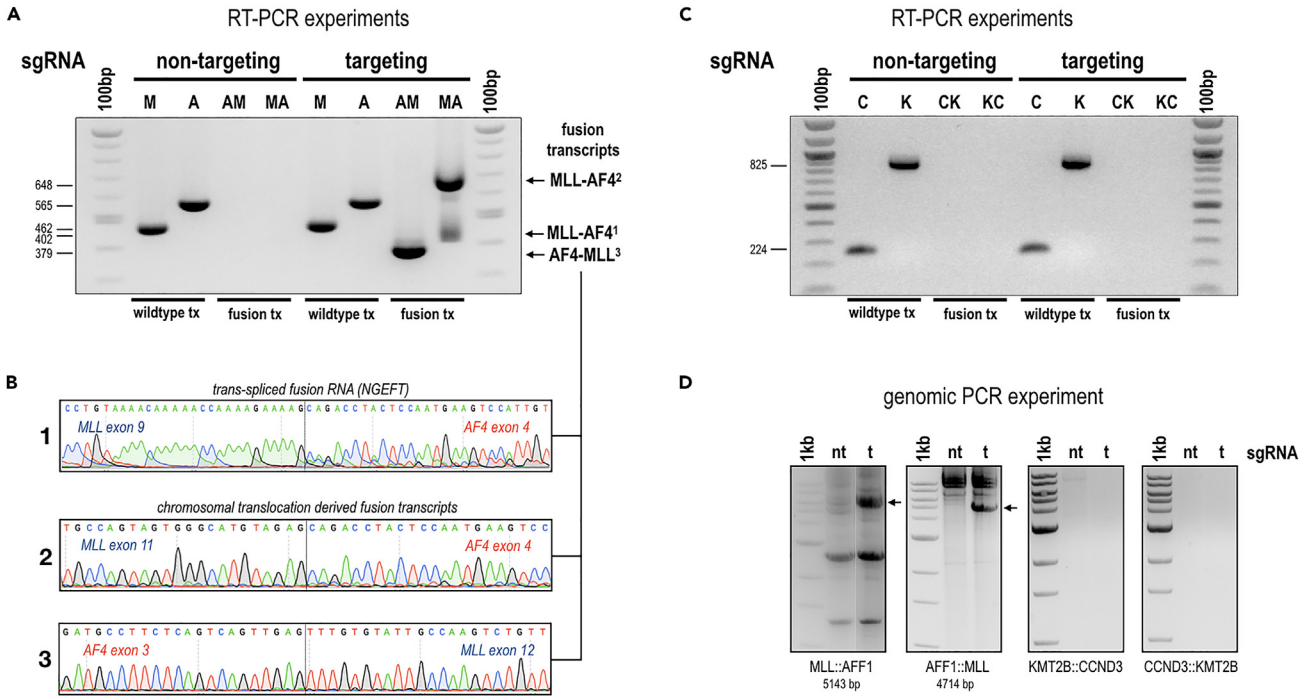


Figure 3. Proximity, NGEFT's and DNA damage cause chromosomal translocations

(A) RT-PCRT experiment after nucleofection of preloaded Cas9 protein. In cells nucleofected with sgRNAs for *MLL* intron 11 and *AFF1* intron 3, appropriate *MLL::AFF1* and *AFF1::MLL* fusion transcripts could be observed with high abundance (equal to the internal control RT-PCRs for the *MLL* or *AFF1* gene transcripts), while in the non-targeting controls, only the transcripts of both wild-type genes were observed.

(B) All 3 fusion transcript PCR bands from panel A were sequenced to validate their correct annotation.

(C) The same experiment performed for the gene pair *CCND3* and *KMT2B*. Regardless of whether non-targeting or targeting guide RNA was used, no fusion transcripts were observed.

(D) Genomic PCR experiments revealed the onset of a t(4,11) chromosomal translocation between *MLL* and *AFF1* (2 left panels). Arrows indicate the rearranged PCR fragments; mean sizes of these PCR bands are given below the panels. No rearranged alleles were observed for *KMT2B* and *CCND3* bands (2 right panels), neither in the non-targeting nor in the targeting experiments.

Cas9-mediated DNA double-strand breaks are expected to occur between nucleotides 343/344 within *MLL* intron 11, or between nucleotides 1046/1047 of *AFF1* intron 3. As shown in Figure 4, typical signs of EJ-mediated DNA repair were observed, such as small deletions (indels), mini-direct-repeats (MDR), or the presence of filler DNA (FD). These are exactly the same hallmarks that have been found in t(4; 11) patients when the first patient-specific recombination events were investigated more than 20 years ago.[1–3] On the right side of Figure 4, examples of such DNA repair processes are given in the form of DNA sequences to explain the nomenclature of the deletions, or the presence of FD or MDR-mediated DNA repair processes.

A synthetic KMT2B-CCND3 fusion RNA causes an artificial chromosomal translocation t(6:19)(p21.1; q13.12)

In a final experiment, we used the cells after 14 days of ABA treatment with *CCND3* and *KMT2B* in close proximity. These cells were protein-transfected with sgRNA/Cas9 protein to induce transient DNA damage. In addition, we co-transfected an artificial *KMT2B::CCND3* fusion RNA (*KMT2B* exons 9–13 (551bp) fused to *CCND3* exons 3–5 (465bp): 1,016 bp). Cells were harvested after 3–5 days to analyze mRNA and genomic DNA. As shown in Figure 5A, the RT-PCR experiment revealed the presence of both wild-type genes (C and K), but this time also the presence of reciprocally fused mRNAs, *CCND3::KMT2B* (CK) and *KMT2B::CCND3* (KC). This argued for the onset of an artificial t(6; 19) chromosomal translocation. In fact, the genomic PCR experiment revealed two different fusion alleles (Figure 5B), indicating the presence of two corresponding genomic fusion alleles in the Cas9-targeted introns (t), while the non-targeting control (nt) displayed neither fusion mRNAs nor the presence of rearranged fusion alleles (Figures 5A and 5B). The sequenced fusion alleles again contained small deletions in both alleles (see Figure 6), indicating an EJ-mediated DNA repair process after sgRNA/Cas9-mediated DNA cleavage. Thus, this final experiment

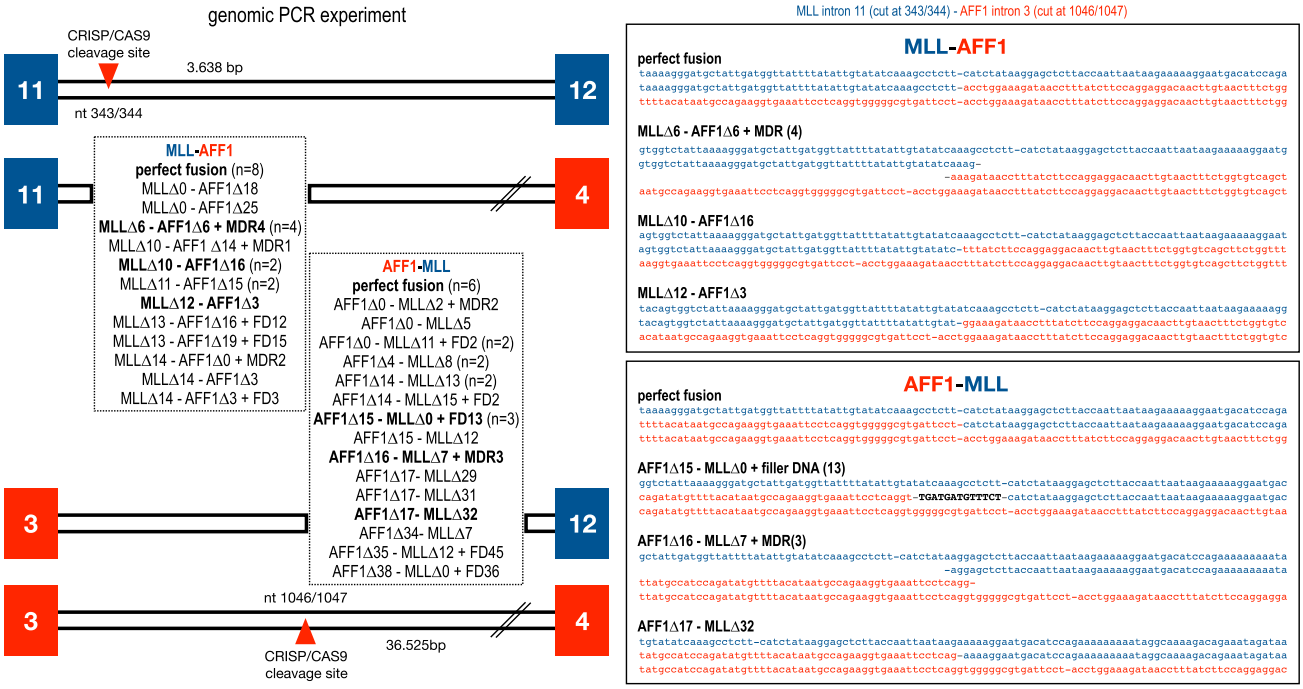


Figure 4. Induced t(4; 11) chromosomal translocations follow the rules of EJ-mediated DNA repair

The 2 PCR bands from Figure 3D were cut out from the gel, subcloned, and subjected to Sanger sequencing. As shown in the left scheme, several *MLL::-AFF1* and *AFF1::MLL* fusion alleles could be analyzed at their fusion sites. The fused chromosomes displayed small deletions, mini-direct repeats (MDR), or even the presence of filler DNA (FD). On the right side, some examples are shown to explain the nomenclature used to describe the individual fusion alleles.

demonstrated that the presence of an artificial fusion RNA, mimicking an NGEFT, was sufficient and necessary to cause a CT in human cells.

In summary, we conclude that the following molecular prerequisites are necessary to generate a chromosomal translocation: (1) gene proximity, (2) the ability to produce specific PTTs that give rise to NGEFTs, and (3) simultaneous DNA damage affecting two adjacent genes. The latter was achieved by preloaded sgRNA/Cas9 proteins that mimic timely and transient DNA damage, which was subsequently repaired by the EJ DNA repair pathway.

DISCUSSION

In the current manuscript, we used our Sleeping Beauty transposon vector technology[51] in combination with state-of-the-art CRISPR/Cas9 technologies[49,50] to establish stable cell lines that constitutively express dead Cas9 protein variants and specific targeting and non-targeting sgRNAs. This allowed us to modulate interphase chromosomes in such a way that we could recapitulate all the essential steps that are necessary to cause a specific CT. The results of our experiments point to several molecular prerequisites that are explained in detail below. Noteworthy, we validated that the EJ DNA repair pathway (either c-NHEJ, MMEJ or A-EJ) caused these CT which were mostly accompanied by small deletions (indels), microhomology-mediated repair, and the implementation of FD at repair junctions. Thus, we validated data that we originally obtained nearly 25 years ago in translocated cell lines and diagnostic material derived from patients with leukemia.[1,2]

The interphase nucleus is composed of "chromosome territories", chromatin blobs within the 3-dimensional space of the cell's nucleus. By using dead Cas9 mutant proteins linked to binding domains for ABA in combination with several cell cycles in cell culture, we were able to rearrange these chromosome territories in such a way that distinct gene pairs, such as *MLL/AFF1* or *CCND3/KMT2B*, were brought into close proximity. Noteworthy, *MLL* and *AFF1*—but not *CCND3* and *KMT2B*—have the ability to prematurely terminate their gene transcription within specific introns, a property which has been termed PTTs. These poly-adenylated PTTs represent a stable mRNA population with a single, unsaturated splice donor site.

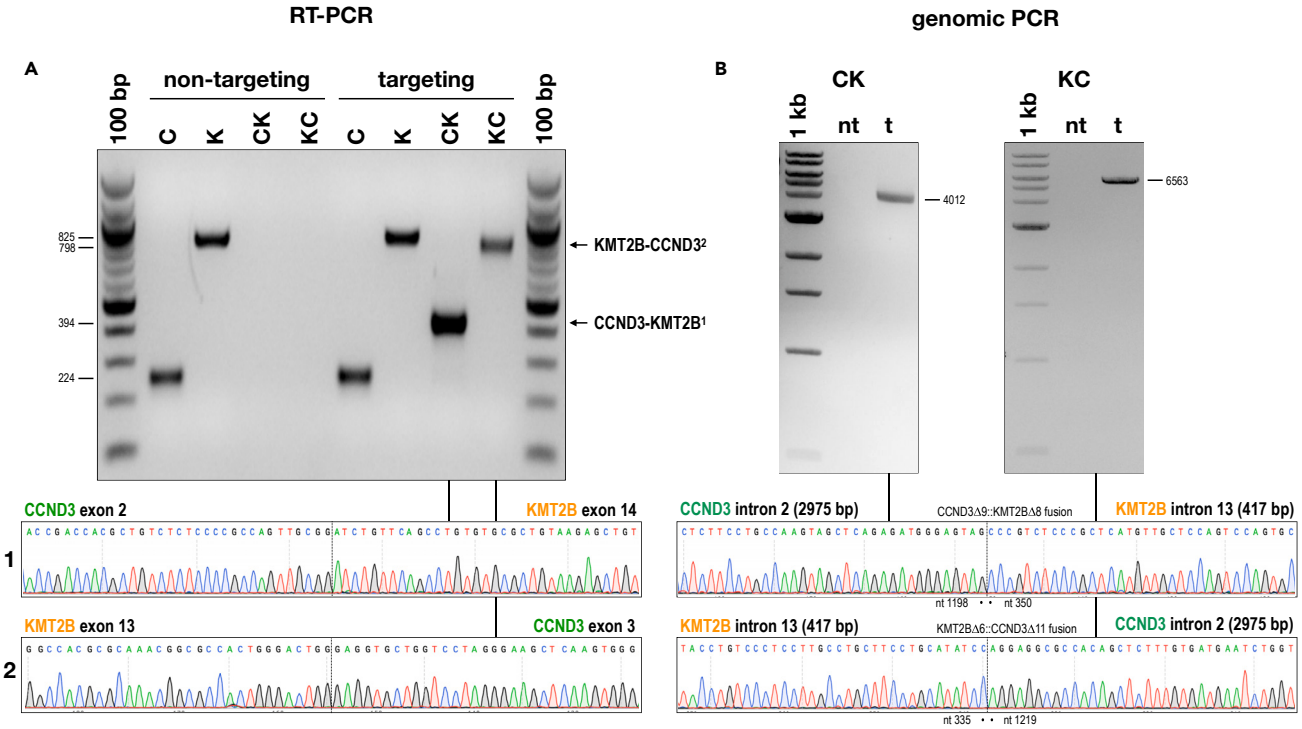


Figure 5. Induction of a specific t(6; 19) chromosomal translocation by induction of double-strand DNA breaks in the presence of a synthetic *KMT2B::CCND3* fusion RNA

Two sgRNA-preloaded Cas9 proteins for *CCND3* intron 2 and *KMT2B* intron 13 were co-transfected with an *in vitro* transcribed artificial *KMT2B::CCND3* fusion RNA (~1 kb). RNA and genomic DNA was collected from these cells and analyzed by RT-PCR and genomic PCR experiments.

(A) The analysis of isolated RNA from the treated cells revealed the presence of wild-type and *CCND3* exon 2:*KMT2B* exon14 and *KMT2B* exon 13:*CCND3* exon 3 fusion transcript RNA, which are depicted below.

(B) The analysis of genomic DNA revealed the presence of fused intronic sequences between chromosome 19 and 6 sequences.

Furthermore, only the *MLL* PTT from intron 9 has been demonstrated to be the source for *trans*-splicing activity that leads to a specific *MLL* exon 9::*AFF1* exon 4 fusion transcript.[45] Noteworthy, all PTTs so far identified in genes involved in chromosomal translocations terminate their transcription in introns that have been classified by diagnostic procedures as "breakpoint cluster region".[45,46]

Any kind of DNA damage which occurs in these breakpoint cluster regions can then lead to the onset of a specific CT. Based on data in the literature, such DNA damage is most likely to occur at very early steps of the apoptotic pathway, where high-order chromatin is cleaved at various hot spots and CpG islands into 200 kb up to megabasepair-long DNA fragments.[52] In the case of the *MLL* gene, there is an additional apoptotic hotspot and a DNase I hypersensitive site shortly upstream of *MLL* exon 12,[16–18] because of a gene-internal promoter which is located in MLL intron 11.[15] This means that the *MLL* gene is frequently broken at this site, while DNA strand breaks within the *AFF1* breakpoint cluster region are probably quite rare.

Our experimental setting has been designed in such a way that we could answer several questions at the same time. We chose 2 different gene pairs. The first set of genes were *MLL* and *AFF1* in order to investigate 2 genes known to be involved in t(4; 11) chromosomal translocations and associated with the development of high-risk acute lymphoblastic leukemia. Moreover, both genes were shown to exhibit PTT capacity (*MLL* intron 9 and *AFF1* intron 3), and thus, are both *per se* able to cause *trans*-splicing events. Of note, the *AFF1* intron 3 PTT was shown to form only exon repetitions, and never produced any NGEFT.[45] The second gene set was *CCND3* and *KMT2B*. Neither of these genes exhibit PTT capacity, and thus, are *per se* unable to participate in *trans*-splicing events. Specific sgRNA were designed for distinct introns of all 4 genes, as well as two non-targeting control sgRNAs that were both predicted *in silico* to be unable to bind anywhere to the human genome.

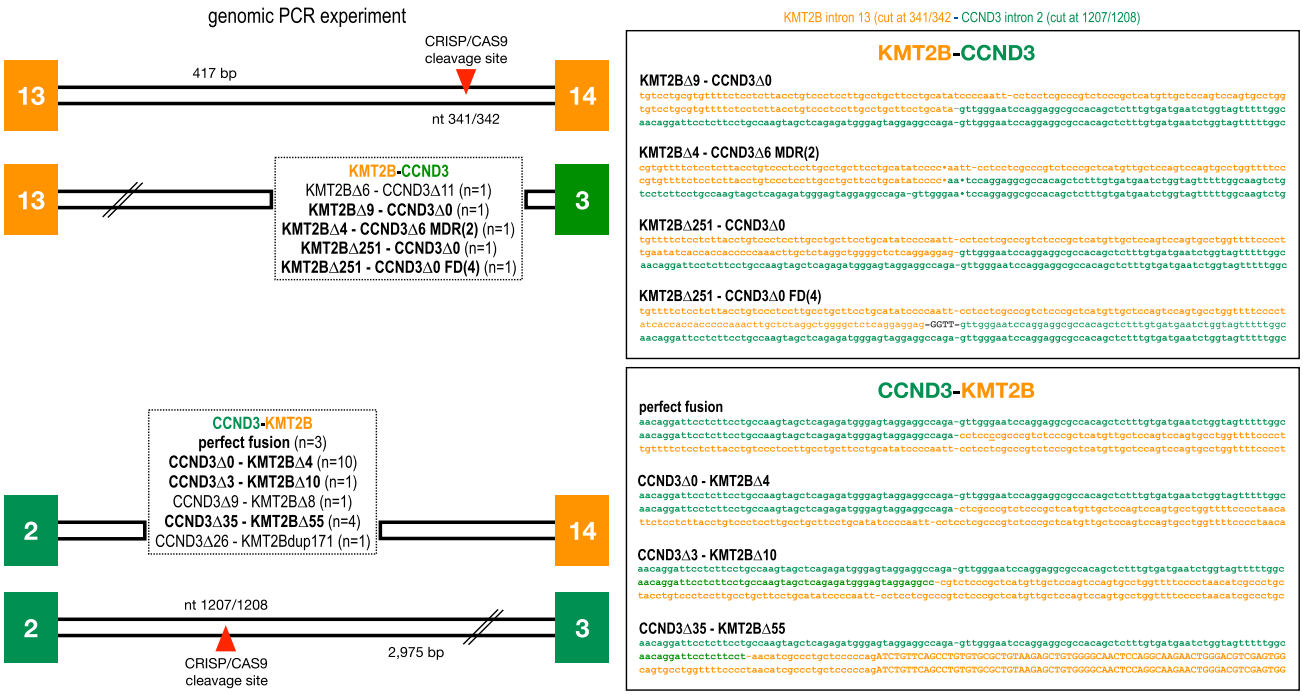


Figure 6. Induced t(6; 19) chromosomal translocations follow the rules of EJ-mediated DNA repair

The 2 PCR bands from Figure 5B were cut out from the gel, subcloned, and subjected to Sanger sequencing. As shown in the left scheme, several *CCND3-KMT2B* and *KMT2B-CCND3* fusion alleles could be analyzed at their fusion sites. The fused chromosomes displayed small deletions, mini-direct repeats (MDR), or even the presence of filler DNA (FD). On the right side, some examples are shown to explain the nomenclature used to describe the individual fusion alleles.

By using the above outlined dead Cas9 protein versions, both fused to specific ABA-binding protein domains, we were able to remodel the interphase nucleus of 2 out of the 3 cell lines. Both gene pairs, *MLL/AFF1* as well as *CCND3/KMT2B*, were brought into close proximity (see Figure 2C). This gene proximity immediately caused the production of the described *trans*-spliced *MLL* exon 9:*AFF1* exon 4 fusion mRNA, but not fusions of *CCND3* and *KMT2B* mRNA. Similarly, the control cell line which stably expressed the two different dCas9 proteins and the non-targeting control sgRNAs neither displayed any gene proximity nor the production of *trans*-spliced NGEFTs. When introducing transient DNA damage with preloaded active Cas9 protein, a specific chromosomal translocation was induced only in the cells in which the following combination was fulfilled: (1) gene proximity, (2) PTT formation capacity, and (3) induced DNA damage. As concluded from our control experiments, none of the 3 points alone was sufficient to cause a CT on its own. Even if genes were in close proximity and Cas9-mediated DNA cleavage occurred, no onset of a CT was observed when a corresponding NGEFT was missing (see absence of a chromosomal translocation between *CCND3* and *KMT2B* in Figure 3B).

This points to NGEFTs as the necessary source for the onset of CTs. This assumption was experimentally validated in the last experiment shown in Figure 5, where we used an *in vitro* transcribed, artificial *KMT2B::CCND3* fusion transcript (not capped and not polyadenylated) that was co-transfected with preloaded sgRNA/Cas9 proteins to introduce transient DNA damage in *CCND3* intron 2 and *KMT2B* intron 13. The combination of DNA damage in the presence of this artificial fusion RNA—mimicking an NGEFT—led to the formation of an artificial t(6; 19) CT, something that has never been observed in nature. This clearly pointed to the fact that the artificial *KMT2B::CCND3* fusion RNA was also able to aberrantly redirect EJ-mediated DNA repair (Figures 5A and 5B).

Consequently, the results of our experiments point to a central role for NGEFTs as the natural source for the process of CT formation. The ability to produce PTTs potentially identifies critical genes in our genome that are associated with a certain risk for genetic aberrations. Of interest, PTTs cannot be found in public databases, as they contain exonic sequences fused to intronic sequences, and thus, are probably discarded

during regular annotation processes. However, studies using specific techniques such as the 3′-Seq methods have analyzed the PTT formation capacity in a genome-wide manner, revealing this property in about 16% of all our genes.[43] PTT-producing genes together with gene proximity may result in several hundred NGEFTs in both healthy and tumor cells. Due to the nature of introns and how they disrupt open reading frames, most of these NGEFTs (2/3) should be out-of-frame without any benefit for cells, while some of them (1/3) should be in-frame and able to produce tiny amounts of fusion proteins with beneficial property for the affected tissue (see e.g. the *MLL*::*AFF1* and *JAZF1*::*JJAZ1* fusion proteins which both provide anti-apoptotic effects).[40,41,53]

Trans-splicing represents a form of exon shuffling that only occurs at the RNA level. Therefore, it would be of great interest to analyze RNA-sequencing data in more detail to identify all possible gene fusions at the RNA level in order to understand the full spectrum of *trans*-spliced RNAs in a cell type, tissue-specific, or cancer-related context. These data may help us understand and calculate the risk for the development of certain cancer types caused by recurrent CTs.

Limitations of the study

The conclusion of this study has been obtained using *in vitro* studies, where all necessary steps could be modeled by genetic manipulation of cells in culture. Since all genetic events occurred within a bulk culture, the isolated and sequenced fusion alleles represent only a small fraction of genomic events that may have taken place in the cultures. However, the analyzed genomic fusion alleles of the experimentally induced t(4; 11) or t(6; 19) chromosomal translocations displayed the same typical characteristics that have been diagnosed in the genomic DNA of patients with *MLL*-r acute leukemia. Therefore, we are quite confident that we have recapitulated in our cell culture system all the molecular steps in the correct order to cause a cancer-associated chromosomal translocation.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Material availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Animals
  - Cell culture
- METHOD DETAILS
  - Plasmid construction and oligonucleotide sequences
  - Chromosome conformation capture
  - RT-PCR experiments to monitor trans-spliced fusion RNAs
  - Construction of DNA templates for single guide and fusion RNA *in vitro* transcription
  - *In vitro* transcription of single guide and fusion RNA
  - *In vitro* cleavage assay
  - Induction of DNA double strand break
  - RT-PCR and genomic PCR experiments to validate the onset of specific chromosomal translocations
- QUANTIFICATION AND STATISTICAL ANALYSIS

ACKNOWLEDGMENTS

This study was supported by grants MA 1876/12-1 from the Deutsche Forschungsgemeinschaft (DFG), and 2020.070 from the Wilhelm Sander Foundation to R.M.

AUTHOR CONTRIBUTIONS

R.M. wrote the manuscript; P.S., E.K., and T.B. planned, performed, and interpreted experiments; P.S. generated cell lines and performed PCR experiments, E.K. and J.R. performed plasmid cloning; R.M. supervised the study and provided funding; all authors reviewed and approved the manuscript.

DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

Received: November 28, 2022
Revised: April 1, 2023
Accepted: May 12, 2023
Published: May 19, 2023

REFERENCES

1. Reichel, M., Gillert, E., Nilson, I., Siegler, G., Greil, J., Fey, G.H., and Marschalek, R. (1998). Fine structure of translocation breakpoints in leukemic blasts with chromosomal translocation t(4;11): the DNA damage-repair model of translocation. Oncogene *17*, 3035–3044. https://doi.org/10.1038/sj.onc.1202229.

2. Gillert, E., Leis, T., Repp, R., Reichel, M., Hösch, A., Breitenlohner, I., Angermüller, S., Borkhardt, A., Harbott, J., Lampert, F., et al. (1999). A DNA damage repair mechanism is involved in the origin of chromosomal translocations t(4;11) in primary leukemic cells. Oncogene *18*, 4663–4671. https://doi.org/10.1038/sj.onc.1202842.

3. Richardson, C., and Jasin, M. (2000). Frequent chromosomal translocations induced by DNA double-strand breaks. Nature *405*, 697–700. https://doi.org/10.1038/35015097.

4. Ghezraoui, H., Piganeau, M., Renouf, B., Renaud, J.B., Sallmyr, A., Ruis, B., Oh, S., Tomkinson, A.E., Hendrickson, E.A., Giovannangeli, C., et al. (2014). Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. Mol. Cell *55*, 829–842. https://doi.org/10.1016/j.molcel.2014.08.002.

5. Ramsden, D.A., and Nussenzweig, A. (2021). Mechanisms driving chromosomal translocations: lost in time and space. Oncogene *40*, 4263–4270. https://doi.org/10.1038/s41388-021-01856-9.

6. Broeker, P.L., Super, H.G., Thirman, M.J., Pomykala, H., Yonebayashi, Y., Tanabe, S., Zeleznik-Le, N., and Rowley, J.D. (1996). Distribution of 11q23 breakpoints within the MLL breakpoint cluster region in de novo acute leukemia and in treatment-related acute myeloid leukemia: correlation with scaffold attachment regions and topoisomerase II consensus binding sites. Blood *87*, 1912–1922.

7. Strissel, P.L., Strick, R., Tomek, R.J., Roe, B.A., Rowley, J.D., and Zeleznik-Le, N.J. (2000). DNA structural properties of AF9 are similar to MLL and could act as recombination hot spots resulting in MLL/AF9 translocations and leukemogenesis. Hum. Mol. Genet. *9*, 1671–1679. https://doi.org/10.1093/hmg/9.11.1671.

8. Hensel, J.P., Gillert, E., Fey, G.H., and Marschalek, R. (2001). Breakpoints of t(4;11) translocations in the human MLL and AF4 genes in ALL patients are preferentially clustered outside of high-affinity matrix attachment regions. J. Cell. Biochem. *82*, 299–309. https://doi.org/10.1002/jcb.1161.

9. Strick, R., Zhang, Y., Emmanuel, N., and Strissel, P.L. (2006). Common chromatin structures at breakpoint cluster regions may lead to chromosomal translocations found in chronic and acute leukemias. Hum. Genet. *119*, 479–495. https://doi.org/10.1007/s00439-006-0146-9.

10. Strissel, P.L., Strick, R., Rowley, J.D., and Zeleznik-Le, N.J. (1998). An in vivo topoisomerase II cleavage site and a DNase I hypersensitive site colocalize near exon 9 in the MLL breakpoint cluster region. Blood *92*, 3793–3803.

11. Aplan, P.D., Chervinsky, D.S., Stanulla, M., and Burhans, W.C. (1996). Site-specific DNA cleavage within the MLL breakpoint cluster region induced by topoisomerase II inhibitors. Blood *87*, 2649–2658.

12. Nasr, F., Macintyre, E., Venuat, A.M., Bayle, C., Carde, P., and Ribrag, V. (1997). Translocation t(4;11)(q21;q23) and MLL gene rearrangement in acute lymphoblastic leukemia secondary to anti topoisomerase II anticancer agents. Leuk. Lymphoma *25*, 399–401. https://doi.org/10.3109/10428199709114180.

13. Megonigal, M.D., Cheung, N.K., Rappaport, E.F., Nowell, P.C., Wilson, R.B., Jones, D.H., Addya, K., Leonard, D.G., Kushner, B.H., Williams, T.M., et al. (2000). Detection of leukemia-associated MLL-GAS7 translocation early during chemotherapy with DNA topoisomerase II inhibitors. Proc. Natl. Acad. Sci. USA *97*, 2814–2819. https://doi.org/10.1073/pnas.050397097.

14. Felix, C.A. (2001). Leukemias related to treatment with DNA topoisomerase II inhibitors. Med. Pediatr. Oncol. *36*, 525–535. https://doi.org/10.1002/mpo.1125.

15. Scharf, S., Zech, J., Bursen, A., Schraets, D., Oliver, P.L., Kliem, S., Pfitzner, E., Gillert, E., Dingermann, T., and Marschalek, R. (2007). Transcription linked to recombination: a gene-internal promoter coincides with the recombination hot spot II of the human MLL gene. Oncogene *26*, 1361–1371. https://doi.org/10.1038/sj.onc.1209948.

16. Stanulla, M., Wang, J., Chervinsky, D.S., Thandla, S., and Aplan, P.D. (1997). DNA cleavage within the MLL breakpoint cluster region is a specific event which occurs as part of higher-order chromatin fragmentation during the initial stages of apoptosis. Mol. Cell Biol. *17*, 4070–4079. https://doi.org/10.1128/MCB.17.7.4070.

17. Betti, C.J., Villalobos, M.J., Diaz, M.O., and Vaughan, A.T. (2001). Apoptotic triggers initiate translocations within the MLL gene involving the nonhomologous end joining repair system. Cancer Res. *61*, 4550–4555.

18. Stanulla, M., Chhalliyil, P., Wang, J., Jani-Sait, S.N., and Aplan, P.D. (2001). Mechanisms of MLL gene rearrangement: site-specific DNA cleavage within the breakpoint cluster region is independent of chromosomal context. Hum. Mol. Genet. *10*, 2481–2491. https://doi.org/10.1093/hmg/10.22.2481.

19. Sim, S.P., and Liu, L.F. (2001). Nucleolytic cleavage of the mixed lineage leukemia breakpoint cluster region during apoptosis. J. Biol. Chem. *276*, 31590–31595. https://doi.org/10.1074/jbc.M103962200.

20. Meyer, C., Larghero, P., Almeida Lopes, B., Burmeister, T., Gröger, D., Sutton, R., Venn, N.C., Cazzaniga, G., Corral Abascal, L., Tsaur, G., et al. (2023). The KMT2A recombinome of acute leukemias in 2023. Leukemia *37*, 988–1005. https://doi.org/10.1038/s41375-023-01877-1.

21. Mitelman, F., Johansson, B., and Mertens, F. (2007). The impact of translocations and gene fusions on cancer causation. Nat. Rev. Cancer *7*, 233–245. https://doi.org/10.1038/nrc2091.

22. Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat. Rev. Genet. *2*, 292–301. https://doi.org/10.1038/35066075.

23. Rosin, L.F., Crocker, O., Isenhart, R.L., Nguyen, S.C., Xu, Z., and Joyce, E.F. (2019). Chromosome territory formation attenuates the translocation potential of cells. Elife *8*, e49553. https://doi.org/10.7554/eLife.49553.

24. Jackson, D.A., Hassan, A.B., Errington, R.J., and Cook, P.R. (1993). Visualization of focal sites of transcription within human nuclei. EMBO J. *12*, 1059–1065. https://doi.org/10.1002/j.1460-2075.1993.tb05747.x.

25. Iborra, F.J., Pombo, A., Jackson, D.A., and Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. J. Cell Sci. *109*, 1427–1436. https://doi.org/10.1242/jcs.109.6.1427.

26. Neves, H., Ramos, C., da Silva, M.G., Parreira, A., and Parreira, L. (1999). The nuclear topography of ABL, BCR, PML, and RARalpha genes: evidence for gene proximity in specific phases of the cell cycle and stages of hematopoietic differentiation. Blood *93*, 1197–1207.

27. Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., and Fraser, P. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat. Genet. *36*, 1065–1071. https://doi.org/10.1038/ng1423.

28. Osborne, C.S., Chakalova, L., Mitchell, J.A., Horton, A., Wood, A.L., Bolland, D.J., Corcoran, A.E., and Fraser, P. (2007). Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS Biol. *5*, e192. https://doi.org/10.1371/journal.pbio.0050192.

29. Kozubek, S., Lukásová, E., Marecková, A., Skalníková, M., Kozubek, M., Bártová, E., Kroha, V., Krahulcová, E., and Slotová, J. (1999). The topological organization of chromosomes 9 and 22 in cell nuclei has a determinative role in the induction of t(9,22) translocations and in the pathogenesis of t(9,22) leukemias. Chromosoma *108*, 426–435. https://doi.org/10.1007/s004120050394.

30. Caldas, C., So, C.W., MacGregor, A., Ford, A.M., McDonald, B., Chan, L.C., and Wiedemann, L.M. (1998). Exon scrambling of MLL transcripts occur commonly and mimic partial genomic duplication of the gene. Gene *208*, 167–176. https://doi.org/10.1016/s0378-1119(97)00640-9.

31. Marcucci, G., Strout, M.P., Bloomfield, C.D., and Caligiuri, M.A. (1998). Detection of unique ALL1 (MLL) fusion transcripts in normal human bone marrow and blood: distinct origin of normal versus leukemic ALL1 fusion transcripts. Cancer Res. *58*, 790–793.

32. Uckun, F.M., Herman-Hatten, K., Crotty, M.L., Sensel, M.G., Sather, H.N., Tuel-Ahlgren, L., Sarquis, M.B., Bostrom, B., Nachman, J.B., Steinherz, P.G., et al. (1998). Clinical significance of MLL-AF4 fusion transcript expression in the absence of a cytogenetically detectable t(4;11)(q21;q23) chromosomal translocation. Blood *92*, 810–821.

33. Biernaux, C., Loos, M., Sels, A., Huez, G., and Stryckmans, P. (1995). Detection of major bcr-abl gene expression at a very low level in blood cells of some healthy individuals. Blood *86*, 3118–3122.

34. Bose, S., Deininger, M., Gora-Tybor, J., Goldman, J.M., and Melo, J.V. (1998). The presence of typical and atypical BCR-ABL fusion genes in leukocytes of normal individuals: biologic significance and implications for the assessment of minimal residual disease. Blood *92*, 3362–3367.

35. Eguchi-Ishimae, M., Eguchi, M., Ishii, E., Miyazaki, S., Ueda, K., Kamada, N., and Mizutani, S. (2001). Breakage and fusion of the TEL (ETV6) gene in immature B lymphocytes induced by apoptogenic signals. Blood *97*, 737–743. https://doi.org/10.1182/blood.v97.3.737.

36. Mori, H., Colman, S.M., Xiao, Z., Ford, A.M., Healy, L.E., Donaldson, C., Hows, J.M., Navarrete, C., and Greaves, M. (2002). Chromosome translocations and covert leukemic clones are generated during normal fetal development. Proc. Natl. Acad. Sci. USA *99*, 8242–8247. https://doi.org/10.1073/pnas.112218799.

37. Quina, A.S., Gameiro, P., Sá da Costa, M., Telhada, M., and Parreira, L. (2000). PML-RARA fusion transcripts in irradiated and normal hematopoietic cells. Genes Chrom. Cancer *29*, 266–275. https://doi.org/10.1002/1098-2264.

38. Beylot-Barry, M., Groppi, A., Vergier, B., Pulford, K., and Merlio, J.P. (1998). Characterization of t(2;5) reciprocal transcripts and genomic breakpoints in CD30+ cutaneous lymphoproliferations. Blood *91*, 4668–4676.

39. Maes, B., Vanhentenrijk, V., Wlodarska, I., Cools, J., Peeters, B., Marynen, P., and de Wolf-Peeters, C. (2001). The NPM-ALK and the ATIC-ALK fusion genes can be detected in non-neoplastic cells. Am. J. Pathol. *158*, 2185–2193. https://doi.org/10.1016/S0002-9440(10)64690-1.

40. Li, H., Wang, J., Mor, G., and Sklar, J. (2008). A neoplastic gene fusion mimics trans-splicing of RNAs in normal human cells. Science *321*, 1357–1361. https://doi.org/10.1126/science.1156725.

41. Li, H., Wang, J., Ma, X., and Sklar, J. (2009). Gene fusions and RNA trans-splicing in normal and neoplastic human cells. Cell Cycle *8*, 218–222. https://doi.org/10.4161/cc.8.2.7358.

42. Rickman, D.S., Pflueger, D., Moss, B., VanDoren, V.E., Chen, C.X., de la Taille, A., Kuefer, R., Tewari, A.K., Setlur, S.R., Demichelis, F., and Rubin, M.A. (2009). SLC45A3-ELK4 is a novel and frequent erythroblast transformation-specific fusion transcript in prostate cancer. Cancer Res. *69*, 2734–2738. https://doi.org/10.1158/0008-5472.CAN-08-4926.

43. Lianoglou, S., Garg, V., Yang, J.L., Leslie, C.S., and Mayr, C. (2013). Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. Genes Dev. *27*, 2380–2396. https://doi.org/10.1101/gad.229328.113.

44. Kamieniarz-Gdula, K., and Proudfoot, N.J. (2019). Transcriptional control by premature termination: a forgotten mechanism. Trends Genet. *35*, 553–564. https://doi.org/10.1016/j.tig.2019.05.005.

45. Kowarz, E., Merkens, J., Karas, M., Dingermann, T., and Marschalek, R. (2011). Premature transcript termination, trans-splicing and DNA repair: a vicious path to cancer. Am. J. Blood Res. *1*, 1–12.

46. Kowarz, E., Dingermann, T., and Marschalek, R. (2012). Do non-genomically encoded fusion transcripts cause recurrent chromosomal translocations? Cancers *4*, 1036–1049. https://doi.org/10.3390/cancers4041036.

47. Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol. *4*, e138. https://doi.org/10.1371/journal.pbio.0040138.

48. Gingeras, T.R. (2009). Implications of chimaeric non-co-linear transcripts. Nature *461*, 206–211. https://doi.org/10.1038/nature08452.

49. Hagège, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., and Forné, T. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nat. Protoc. *2*, 1722–1733. https://doi.org/10.1038/nprot.2007.243.

50. Morgan, S.L., Mariano, N.C., Bermudez, A., Arruda, N.L., Wu, F., Luo, Y., Shankar, G., Jia, L., Chen, H., Hu, J.F., et al. (2017). Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. Nat. Commun. *8*, 15993. https://doi.org/10.1038/ncomms15993.

51. Kowarz, E., Löscher, D., and Marschalek, R. (2015). Optimized Sleeping Beauty transposons rapidly generate stable transgenic cell lines. Biotechnol. J. *10*, 647–653. https://doi.org/10.1002/biot.201400821.

52. Chen, D.L., Swe, M., and Sit, K.H. (1998). G-band expression and megabase fragmentations in apoptosis. Exp. Cell Res. *240*, 293–304. https://doi.org/10.1006/excr.1998.3945.

53. Gaussmann, A., Wenger, T., Eberle, I., Bursen, A., Bracharz, S., Herr, I., Dingermann, T., and Marschalek, R. (2007). Combined effects of the two reciprocal t(4;11) fusion proteins MLL.AF4 and AF4.MLL confer resistance to apoptosis, cell cycling capacity and growth transformation. Oncogene *26*, 3352–3363. https://doi.org/10.1038/sj.onc.1210125.

54. Secker, K.A., Keppeler, H., Duerr-Stoerzer, S., Schmid, H., Schneidawind, D., Hentrich, T., Schulze-Hentrich, J.M., Mankel, B., Fend, F., and Schneidawind, C. (2019). Inhibition of DOT1L and PRMT5 promote synergistic anti-tumor activity in a human MLL leukemia model induced by CRISPR/Cas9. Oncogene *38*, 7181–7195. https://doi.org/10.1038/s41388-019-0937-9.

55. Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186–191. https://doi.org/10.1038/nature14299.

56. Kulcsár, P.I., Tálas, A., Huszár, K., Ligeti, Z., Tóth, E., Weinhardt, N., Fodor, E., and Welker, E. (2017). Crossing enhanced and high fidelity SpCas9 nucleases to optimize specificity and cleavage. Genome Biol. *18*, 190. https://doi.org/10.1186/s13059-017-1318-8.

57. Weinberg, B.H., Pham, N.T.H., Caraballo, L.D., Lozanoski, T., Engel, A., Bhatia, S., and Wong, W.W. (2017). Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. Nat. Biotechnol. *35*, 453–462. https://doi.org/10.1038/nbt.3805.

58. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W.,

Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823. https://doi.org/10.1126/science.1231143.

59. Swiech, L., Heidenreich, M., Banerjee, A., Habib, N., Li, Y., Trombetta, J., Sur, M., and Zhang, F. (2015). In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. Nat. Biotechnol. *33*, 102–106. https://doi.org/10.1038/nbt.3055.

60. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308. https://doi.org/10.1038/nprot.2013.143.

61. Moudgil, A., Wilkinson, M.N., Chen, X., He, J., Cammack, A.J., Vasek, M.J., Lagunas, T., Jr., Qi, Z., Lalli, M.A., Guo, C., et al. (2020). Self-reporting transposons enable simultaneous readout of gene expression and transcription factor binding in single cells. Cell *182*, 992–1008.e21. https://doi.org/10.1016/j.cell.2020.06.037.

62. Moreno-Mateos, M.A., Vejnar, C.E., Beaudoin, J.D., Fernandez, J.P., Mis, E.K., Khokha, M.K., and Giraldez, A.J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. Nat. Methods *12*, 982–988. https://doi.org/10.1038/nmeth.3543.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
NEB® Stable competent *E. coli*	NEB	NEB Cat# C3040H
Chemicals, peptides, and recombinant proteins		
(+)-Abscisic Acid 98%	Thermo scientific	342405000
Accutase	Capricorn Scientific	ACC-1B
AvrII	NEB	R0174L
Blasticidin	Avantor	A3784.0025
Cas9-NLS	PNA Bio	CP02
DMEM	Capricorn Scientific	DMEM-LPA
dNTPs	Boehringer Mannheim	1969064
DTT	Biomol	04010.25
FCS	Capricorn Scientific	FBS-11A
Glutamine	Capricorn Scientific	STA-B
Glycine	Carl Roth	3908.3
GoTaq® Long PCR Master Mix	Promega	M4021
HindIII-HF	NEB	R3104L
HiScribe T7 High Yield RNA Synthesis Kit	NEB	E2040S
Kapa HiFi HotStart ReadyMix	Roche	KK2602
Metafectene Pro	Biontex	T040–1.0
PBS	Capricorn Scientific	PBS-1A
PenStrep	Capricorn Scientific	PS-B
Phenol-Chloroform	Carl Roth	A156.3
Proteinase K	Carl Roth	7528.6
Puromycin	PAA Laboratories	P11-019
RNase A	Carl Roth	7164.2
RNase-Free DNase Set	Qiagen	79254
RNasin	Promega	N2515
SuperScript™ II Reverse Transcriptase	Invitrogen	18064071
T4 DNA ligase	NEB	M0202M
β-Mercaptoethanol	Carl Roth	4227.3
Critical commercial assays		
AllPrep RNA/DNA Mini Kit	Qiagen	80204
pCR™2.1-TOPO™	Invitrogen	450641
RNeasy Mini Kit	Qiagen	74104
Zymo RNA Clean & Concentrator Kit	Zymogen	R1017
Zymoclean Gel DNA Recovery Kit	Zymogen	D4008
Experimental models: Cell lines		
HEK293T	DSMZ	ACC 635
Oligonucleotides		
MLL-I9 5′-CACCGTACTCTTGAGAAGCTGAGGCAG-3′	this paper	N/A
AF4-I3 5′-CACCGTCTAATAAATGGCTTGACTTG-3′	this paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CCND3-I2 5′-CCCTTGGTAGGATATCAGGG-3′	this paper	N/A
KMT2B-I13 5′-GGGGCTCTCAGGAGGAGCAG-3′	this paper	N/A
NT-I 5′-CACCGGTAATCGCTAGCTTGTCGACT-3′	this paper	N/A
NT-II 5′-CACCGGTAATCGCTAGCTTGTCGACT-3′	this paper	N/A
MLL.I9.CCC.2.F 5′-TTGACCCCAACATCCTTTAGCAA-3′	this paper	N/A
AF4.I3.CCC.Guide2.2.R 5′-TTCCAGGGCTGAGAGGGGAAA-3′	this paper	N/A
CCND3.CCC.1.HindIII.F 5′-ATAGCCTGGGGTGGGGTCAT-3′	this paper	N/A
KMT2B.CCC.1.HindIII.R 5′-GTGTATGCAGGACAGCGAGGG-3′	this paper	N/A
AF4.3 5′-GTTGCAATGCAGCAGAAGCC-3′	this paper	N/A
AF4.5 5′-ACTGTCACTGTCCTCACTGTCA-3′	this paper	N/A
8.3 5′-CCCAAAACCACTCCTAGTGAG-3′	this paper	N/A
13.5 5′-CAGGGTGATAGCTGTTTCGG-3′	this paper	N/A
CCND3.E2.2.F 5′-TTCCCCCTGGCCATGAACTAC-3′	this paper	N/A
CCND3.E3.1.R 5′-AGGTCCCACTTGAGCTTCCC-3′	this paper	N/A
KMT2B.E8.1.F 5′-ATTCGGATGACTCGGAGCCC-3′	this paper	N/A
KMT2B.E15.1.R 5′-CATGCACCCAGTGATCGCAC-3′	this paper	N/A
CCND3.E4.2.R 5′-GCGGGTACATGGCAAAGGTAT-3′	this paper	N/A
AF4.cut.guide.F 5′-taatacgactcactataGGGGGG GCGTGATTCCTACCgttttagagctagaaatagc-3′	this paper	N/A
MLL.cut.guide.F 5′-taatacgactcactataGGAGCTCCTT ATAGATGAAGgttttagagctagaaatagc-3′	this paper	N/A
CCND3.cut.guide3.F 5′-taatacgactcactataGGGAGTA GGAGGCCAGAGTTgttttagagctagaaatagc-3′	this paper	N/A
KMT2B.cut.guide6.F 5′-taatacgactcactataGGAGACG GGCGAGGAGGAATgttttagagctagaaatagc-3′	this paper	N/A
Universal Reverse Primer 5′-AGCACCGACTCGGTGCCACT-3′	Secker et al.[54]	N/A
KMT2B.E9.T7.F 5′-TAATACGACTCACTATAAGC TGCCACTGCCAGAACCTGAGGAGC-3′	this paper	N/A
KMT2B.CCND3.fRNA.OE.R 5′-CCTAGGACCAGCACCTCCCAGT CCCAGTGGCGCCGTTTGCGC-3′	this paper	N/A
KMT2B.CCND3.fRNA.OE.F 5′-CGCAAACGGCGCCA CTGGGACTGGGAGGTGCTGGTCCTAGGG-3′	this paper	N/A
CCND3.E5.fRNA.R 5′-CTACAGGTGTATGGCTGTGACATCTGTAGGAGTGCTGG-3′	this paper	N/A
MLL.I11.BsaI.F 5′-AAACAGAGACCATTTAGCAGGTACTATTCCCTGT-3′	this paper	N/A
MLL.I11.SapI.R 5′-CCAAGAAGAGCTAACTAATCTATATTTCTTTCCACCAACA-3′	this paper	N/A
AF4.I3.BsaI.F 5′-AAACAGAGACCTAGAGCATACATTCTGTCACTGAAATT-3′	this paper	N/A
AF4.I3.SapI.R 5′-CCAAGAAGAGCGTGCATTAAATCTTCTTAAGTTAGGATCA-3′	this paper	N/A
KMT2B.E13.F 5′-GCCATGCATACCACCCGGCCTGTCT-3′	this paper	N/A
KMT2B.E14.R 5′-ATCTCCAGACCACTCGACGTCCCAGTTCTT-3′	this paper	N/A
CCND3.I2.F 5′-CAGGCCCTGATTGCTTTAAGGAAGGGCTGA-3′	this paper	N/A
CCND3.I2.3.R 5′-TCAGTGCGATGCTGAGGCAG-3′	this paper	N/A
AF4.cont.F 5′-TCCGAAGGTACTGACCGACAGCCTTAACTA-3′	this paper	N/A
MLL.con.R1 5′-AGCCAAAGAGGTATCTGCCAGGAATTTAAGAA-3′	this paper	N/A
MLL.cont.F 5′-CAGAATCAGGTGAGTGAGGAGGGCAAGAAG-3′	this paper	N/A
AF4.cont.R 5′-TTTCCCATGAGGTGAAAGGAAAGGACAGGG-3′	this paper	N/A
KMT2B.LR.1.F 5′-CGGTCCCGGCGGGGAAAGGT-3′	this paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CCND3.LR.1.R 5′-GGCTGGCCGGGCCCCTTAGTG-3′	this paper	N/A
CCND3-KMT2B.LR.1.F 5′- TGCAGCGGGAGATCAAGCCGCACA-3′	this paper	N/A
CCND3-KMT2B.LR.1.R 5′- GCGCCCGCGCCCTCACCATCTG-3′	this paper	N/A
CCND3-KMT2B.LR.3.F 5′- TGGCGGGGAAGGCGATGGGGGTG-3′	this paper	N/A
CCND3-KMT2B.LR.3.R 5′- CCAGCGTGGCCCCTGCATGCCC-3′	this paper	N/A
Recombinant DNA		
pSBbi-RB, sleeping beauty backbone with blasticidin and RFP	Kowarz et al.[51]	AddGene #60522
pSBbi-GP, sleeping beauty backbone with puromycin and GFP	Kowarz et al.[51]	AddGene #60511
pX603-AAV-CMV::NLS-dSaCas9(D10A,N580A)-NLS-3xHA-bGHpA, dead SaCas9	Ran et al.[55]	AddGene #61594
pX330-Flag-deSpCas9, dead SpCas9	Kulcsár et al.[56]	AddGene #92114
pBW2286_pCAG-PV1-FlpO-N396-L1-ABI-NLS-BGHpA, Abscisic acid binding domain ABI	Weinberg et al.[57]	AddGene #87559
pBW2287_pCAG-PV1-PYL-L1-FlpO-397C-NLS-BGHpA, Abscisic acid binding domain PYL	Weinberg et al.[57]	AddGene #87560
pX330-U6-Chimeric_BB-CBh-hSpCas9, U6 promotor and cassette for SpCas9 sgRNA	Cong et al.[58]	AddGene #42230
pX552, U6 promotor and cassette for SaCas9 sgRNA	Swiech et al.[59]	AddGene #60958
pSpCas9(BB)-2A-GFP (PX458), template plasmid for the sgRNAs	Ran et al.[60]	AddGene #48138
pSBbi RB::dSaCas9-ABI1/U6::MLL-I9, dead SaCas9 fused to ABI1 with a sgRNA for MLL intron 9	This paper	N/A
pSBbi GP::deSpCas9-PYL1/U6::AF4-I3, dead SpCas9 fused to PYL1 with a sgRNA for AF4 intron 3	This paper	N/A
pSBbi RB::dSaCas9-ABI1/U6::CCND3-I2, dead SaCas9 fused to ABI1 with a sgRNA for CCND3 intron 2	This paper	N/A
pSBbi GP::deSpCas9-PYL1/U6::KMT2B-I13, dead SpCas9 fused to PYL1 with a sgRNA for KMT2B intron 13	This paper	N/A
pSBbi RB::dSaCas9-ABI1/U6::NT-I, dead SaCas9 fused to ABI1 with a sgRNA without a target in the genome	This paper	N/A
pSBbi GP::deSpCas9-PYL1/U6::NT-II, dead SpCas9 fused to PYL1 with a sgRNA without a target in the genome	This paper	N/A
DNA template for the artificial KMT2B::CCND3 fusion RNA	This paper	N/A
Sleeping beauty transposase plasmid pcGLobin-SB100$_{XCO}$	Moudgil et al.[61]	AddGene #154887
Software and algorithms		
CRISPRscan	Moreno-Mateos et al.[62]	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Rolf Marschalek (rolf.marschalek@em.uni-frankfurt.de).

Material availability

All unique/stable reagents generated in this study are available from the lead contact with a completed Materials Transfer Agreement.

Data and code availability

- Data: All experimental data or any additional information required to reanalyze the data reported in this paper are available from the lead contact.

- Code: not applicable.
- Other items: not applicable.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Animals

Not applicable.

Cell culture

Cell culture and establishing stable cell lines

HEK293T (ACC 635, DSMZ) cells were cultivated in DMEM Low Glucose medium (DMEM-LPA, Capricorn Scientific) supplemented with 5% FCS (FBS-11A, Capricorn Scientific), 1% PenStrep (PS-B, Capricorn Scientific) and 1% glutamine (STA-B, Capricorn Scientific). The cells were split 1:10 every second day using an Accutase (ACC-1B, Capricorn Scientific) treatment (see below) and adding new fresh medium.

All stable transfections were carried out by using 1 μg of pSBbi RB::dSaCas9-ABI1/U6::MLL-I9 and 1 μg of pSBbi GP::deSpCas9-PYL1/U6::AFF1-I3 for the MLL/AFF1 targeting cell line, 1 μg of pSBbi RB::dSaCas9-ABI1/U6::CCND3-I2, 1 μg of pSBbi GP::deSpCas9-PYL1/ U6::KMT2B-I13 for the CCND3/KMT2B targeting cell line and 1 μg of pSBbi RB::dSaCas9-ABI1/U6::NT-I and 1 μg of pSBbi GP::deSpCas9-PYL1/U6::NT-II for the non-targeting control cell line. All transfections were carried out in the presence of 100 ng transposase vector $SB100_{XCO}$ to guarantee stable integration of all vectors. Selection was carried for 7 to 14 days in the presence of 20 μg/ml Blasticidin (A3784.0025, Avantor) and 2 μg/ml Puromycin (P11-019, PAA Laboratories). The selection was removed when all cells exhibited green and red fluorescence, indicating that both vector constructs were stably integrated. For the dSa/deSpCas9-dimerization experiments, a final concentration of 200 μM abscisic acid dissolved in Methanol (ABA; 342405000, Thermo Scientific) was added to the cell culture medium. Dimerization was tested at day 0 (before ABA treatment) and after 14 days of ABA treatment. Subsequently, either the 3C experiments (see below) or the induction of double strand DNA breaks by CRISPR/Cas9 was performed.

METHOD DETAILS

Plasmid construction and oligonucleotide sequences

Using our lab's Sleeping Beauty technology,[51] 6 different pSBbi vectors were designed. The backbone of pSBbi-RB and pSBbi-GP (AddGene #60522, #60511) were used to clone two fusion proteins, coding for dead SaCas9 (AddGene #61594) or dead enhanced SpCas9 proteins (AddGene #92114) in conjunction with appropriate abscisic acid binding domains, ABI1 and PYL1, respectively (AddGene #87559, #87560). In addition, a U6 promoter driven guide RNA/scaffold cassette was cloned from the AddGene plasmid #60958 for dSaCas9, or from the AddGene plasmid #42230 for the deSpCas9 protein, to create two pairs of vectors that can be used for the establishment of stable cell lines. These two plasmids were again used to clone the appropriate targeting guide RNAs for *MLL* intron 9, *AFF1* intron 3, *CCND3* intron 2, *KMT2B* intron 13 as well as two non-targeting guide RNAs that are unable to bind to human genome sequences. The guiding sequences are listed in the oligonucleotide section in Star*Methods. All 6 vector constructs used throughout the study are depicted in Figure 1A.

Chromosome conformation capture

The Chromosome Conformation Capture (3C) experiment was performed according to a published method.[49] Briefly, HEK293T (ACC 635, DMSZ) cells were harvested using 1 ml Accutase and incubation for 5 min at 37°C. Cells were then centrifuged for 5 min at 400 x g and washed once with 10 ml PBS (PBS-1A, Capricorn Scientific). For the formaldehyde crosslinking procedure, 1×10^7 cells were incubated with 9.5 ml of 2% formaldehyde/10% FCS (FBS-11A, Capricorn Scientific)/PBS (PBS-1A, Capricorn Scientific) for 10 min at room temperature. To stop the crosslinking reaction 1.425 ml of 1 M ice cold glycine (3908.3, Carl Roth) was added. Cells were centrifuged for 8 min at 225 x g at 4°C and the supernatant was removed. The pellet was lysed in 5 ml lysis buffer (10 mM Tris-HCl pH 7,5, 10 mM NaCl, 0,2% NP40 and 1x Roche Complete Mini Inhibitor (4693116001, Roche)) and incubated for 10 min on ice. Another centrifugation step for 5 min with 400 x g at 4°C was used to pellet the cell nuclei; the supernatant was discarded. Precipitated nuclei were gently dissolved in 0.5 ml of 1.2 x restriction enzyme buffer, placed at 37°C, and 7.5 μl of a 20% SDS solution was added and further incubated for 1 h at 37°C while shaking at 900 rpm. Subsequently,

50 µl of 20% Triton X-100 was added and incubated for another hour at 37°C while shaking at 900 rpm. A total of 400 U restriction enzyme (for AFF1 and MLL: AvrII (R0174L, NEB); for KMT2B and CCND3: HindIII (R3104, NEB) was added and incubated overnight at 37°C while shaking at 900 rpm. On the following morning, 40 µl of 20% SDS was added and incubated for 20 min at 65°C while shaking at 900 rpm. The digested nuclei were transferred into a 50 ml Falcon tube and 6.125 ml of 1.15x ligation buffer (10x Ligation buffer: 660 mM Tris-HCl pH 7.5, 50 mM DTT, 50 mM $MgCl_2$, 10 mM ATP) and 375 µl 20% Triton X-100 was added and incubated for 1 h at 37°C while shaking gently. 100 U T4 DNA ligase (M0202M, NEB) was added and incubated for 4h at 16°C, followed by 30 min at room temperature. Finally, 30 µl of a 10 mg/ml Proteinase K solution (7528.6, Carl Roth) was added and incubated at 65°C overnight to de-crosslink the sample. The solution was then treated with additional 30 µl of 10 mg/ml RNase A (7164.2, Carl Roth) and incubated for 45 min at 37°C. To extract the nucleic acids, 7 ml of phenol-chloroform (A156.3, Carl Roth) was added, mixed vigorously and centrifuged for 15 min at 2,200 x g at room temperature. The supernatant was transferred into a new 50 ml Falcon tube and 7 ml water, 1.5 ml of 2 M sodium acetate pH 5.6 followed by 35 ml of ethanol was added for precipitation of nucleic acids. The solution was mixed and incubated for at least 1h at -80°C (or overnight) and subsequently centrifuged for 45 min at 2,200 x g at 4°C. The supernatant was removed, the pellet washed with 10 ml of 70% ethanol and centrifuged again for 15 min at 2,200 x g at 4°C. Again, the supernatant was removed, the DNA pellet was dried at room temperature, and then dissolved in 150 µl 10 mM Tris-HCl pH 7.5. The 3C template was then used for all subsequent PCR experiments. The PCR for MLL and AFF1 was performed using the primers MLL.I9.CCC.2.F and AF4.I3.CCC.Guide2.2.R and for KMT2B and CCND3, CCND3.CCC.1.HindIII.F and KMT2B.CCC.1.HindIII.R. The sequences of the primers are listed in Star*Methods.

RT-PCR experiments to monitor trans-spliced fusion RNAs

The RNAs were extracted with the Qiagen RNeasy Mini Kit (74104, Qiagen) according to the manufacturer's protocol. Briefly, cell pellets were resuspended in 175 µl RLN-Buffer (4°C) and incubated on ice for 5 min. The samples were then centrifuged for 2 min at 300 x g at 4°C and the supernatant pipetted into a new reaction tube. 600 µl RLT-Buffer with 10% β-mercaptoethanol (4227.3, Carl Roth) was added and the tube vortexed. Then 430 µl 96% ethanol was added, mixed with a pipette and 700 µl transferred onto a RNeasy spin column. The samples were centrifuged for 15 s at 10,000 rpm, the flow through was discarded and centrifugation repeated if the sample volume was greater than 700 µl. The column was washed with 700 µl RW1-Buffer and centrifuged for 15 s at 10,000 rpm. For the on-column DNA digestion, 10 µl DNase-Stock-Solution in 70 µl RDD Buffer (Qiagen) was added to the middle of the membrane and incubated for 1 h at RT. The column was washed with 500 µl RPE-Buffer twice and centrifuged at 10,000 rpm for 15 s and 2 min. The sample was eluted in 30 µl RNase-free water by centrifugation for 1 min at 10,000 rpm.

For the cDNA synthesis 1 µg of RNA was made up to 7 µl with RNase-free water and mixed with 1 µl N6 primer (100 pmol). The sample was incubated for 10 min at 70°C and then cooled down on ice for 2 min. For the reaction, 1 µl RNasin (N2515, Promega GmbH), 4 µl 5x First Strand, 4 µl 2.5 mM dNTPs (1969064, Boehringer Mannheim), 2 µl 100 mM DTT (04010.25, Biomol) and 1 µl SuperScript™ II Reverse Transcriptase (18064071, Invitrogen) were added, incubated for 10 min at RT and then for 1 h at 42°C. To stop the reaction, 30 µl RNase-free water was added and the sample incubated for 10 min at 70°C.

To monitor trans-spliced fusion mRNAs, the cDNAs from the different cell lines transfected with the dimerization systems for *MLL* and *AFF1* or *CCND3* and *KMT2B* were analyzed at day 0 and day 14 of abscisic acid treatment via RT-PCR (n = 4). The endogenous *AFF1* gene transcript was amplified with oligonucleotides AF4.3 and AF4.5, while the endogenous *MLL* gene transcript with oligonucleotides 8.3 and 13.5, respectively. The fusion transcript *AFF1-MLL* was amplified with oligonucleotides AF4.3 and 13.5, and the *MLL-AFF1* fusion transcript was analyzed with oligonucleotides 8.3 and AF4.5. The *CCND3* gene transcript was amplified with CCND3.E2.2.F and CCND3.E3.1.R and the *KMT2B* gene transcript was analyzed with KMT2B.E8.1.F and KMT2B.E15.1.R. The *CCND3-KMT2B* fusion transcript was analyzed with CCND3.E2.2.F and KMT2B.E15.1.R, while the KMT2B-CCND3 fusion transcript was analyzed with KMT2B.E8.1.F and CCND3.E4.2.R. The PCR protocol was 94°C for 3 min for the initial denaturation, 35 cycles of 94°C for 30 s, 63°C for 30 s and 72°C for 90 s, followed by 72°C for 2 min for the final elongation. RT-PCR products were analyzed on a 2% agarose gel. The sequences of the primers are listed in Star*Methods.

Construction of DNA templates for single guide and fusion RNA *in vitro* transcription

The cloning oligos for each target gene were designed using the CRISPRscan server, and only oligos with high CRISPRscore and as few as possible off-targets were chosen. CRISPRscan can directly design primers including T7 promotor, the target complementary region and a scaffold overlap. For better primer binding a 'ATAGC' is added to the end of the primer. With these primers PCRs were performed to produce the template DNA for the sgRNA production. For this, the Kapa HiFi HotStart ReadyMix (KK2602, Roche) was used. The reaction mix consisted of 2 µl 10 µM sgRNA primer (guide RNA sequence is indicated in capital letters); for *AFF1* AF4.cut.guide.F, for *MLL* MLL.cut.guide.F, for *CCND3* CCND3.cut.guide3.F and for *KMT2B* KMT2B.cut.guide6.F in combination with 1 µl 10 µM universal reverse Primer, 1 µl 2–4 ng of the PX458 plasmid, 10 µl KAPA HiFi Mix 2x and 5 µl water. The PCR protocol used was 95°C for 3 min for the initial denaturation, 30 cycles of 98°C for 3 s, 60°C for 5 s and 72°C for 10 s, followed by 72°C for 1 min for the final elongation. The DNA was separated on a 2% agarose gel and the corresponding bands were cut out and recovered using the Zymoclean Gel DNA Recovery Kit (D4008, Zymogen) following the manufacturer's protocol.

For the synthesis of a DNA template for a synthetic and artificial fusion RNA, a PCR for *KMT2B* was performed with KMT2B.E9.T7.F and KMT2B.CCND3.fRNA.OE.R, and for *CCND3* with KMT2B.CCND3.fRNA.OE.F and CCND3.E5.fRNA.R, respectively. The two fragments were separated on a 2% agarose gel, cut out and extracted using the Zymoclean Gel DNA Recovery Kit (D4008, Zymogen) following the manufacturer's protocol. An overlap-extension PCR was then performed using the two PCR fragments in combination with the outer primers KMT2B.E9.T7.F and CCND3.E5.fRNA.R. The PCR fragment was separated on a 1% agarose gel and recovered again using the Zymoclean Gel DNA Recovery Kit (D4008, Zymogen) following the manufacturer's protocol. The sequences of the primers are listed in Star*Methods.

In vitro transcription of single guide and fusion RNA

In vitro transcription of sgRNAs was performed using the HiScribe T7 High Yield RNA Synthesis Kit (E2040S, NEB). For the transcription reaction, 400–1000 ng of template DNA were mixed with 8 µl dNTPs (ATP, GTP, CTP, UTP), 2 µl Reaction Buffer (10x), 2 µl T7 RNA Polymerase Enzyme Mix and 0–4 µl water to adjust to a final volume of 20 µl. Transcription was performed for a minimum of 5h at 37°C. Following the amplification step, the sgRNAs were purified using Zymo RNA Clean & Concentrator Kit (R1017, Zymogen) according to the manufacturer. Sample volume was made up to 50 µl with nuclease-free water and 100 µl RNA Binding Buffer added and mixed. 150 µl ethanol was added and mixed. The sample was transferred onto the Zymo-Spin™ IICR Column and centrifuged for 30 s at 10,000 x g. For the on-column DNA digestion using the RNase-Free DNase Set (79254, Qiagen), 10 µl DNase-Stock-Solution in 70 µl RDD Buffer (Qiagen) was added to the middle of the membrane and incubated for 1 h at RT. 400 µl RNA Prep Buffer was added and the tube centrifuged for 30 s at 10,000 x g, the flow through being discarded. 700 µl RNA Wash Buffer was added, the tube centrifuged and the flow through discarded. 400 µl RNA Wash Buffer was added to the column, the column centrifuged for 1 min at 10,000 x g and transferred into a new reaction tube. The sgRNAs were eluted in 25 µl nuclease-free water that were incubated for 7 min on the column before centrifugation for 1 min at 10,000 x g. Concentration was measured using a Nanodrop, and the final concentration was diluted to 1 µg/µl and stored at −80°C.

In vitro cleavage assay

Fragments for *MLL* were synthesized using MLL.I11.BsaI.F and MLL.I11.SapI.R, *AFF1* with AF4.I3.BsaI.F and AF4.I3.SapI.R, *KMT2B* using KMT2B.E13.F and KMT2B.E14.R and *CCND3* with CCND3.I2.F and CCND3.I2.3.R. For the cleavage assay, 400 ng of Cas9 protein (1 µg/µl) (CP02, PNA Bio) were mixed with 400 ng sgRNA (1 µg/µl), 2 µl Cas9 reaction buffer and water (final volume with the dsDNA was 20 µl) and incubated at 37°C for 10 min. 200 ng of the amplified DNA fragment was added to the reaction and incubated for 3 to 5 h at 37°C. The reaction mix was analysed on a 2% agarose gel to examine the cleavage. The sequences of the primers are listed in Star*Methods.

Induction of DNA double strand break

After ABA (342405000, Thermo Scientific) induction for 14 days, the double strand break was induced by Cas9 (CP02, PNA Bio) preloaded with the sgRNA. 1 µg of Cas9 protein and 1 µg of sgRNA were mixed, filled up to 10 µl with PBS (PBS-1A, Capricorn Scientific) and incubated at 37°C for 20 min. For the transfection 18 µl Metafectene Pro (T040-1.0, Biontex) and 282 µl of PBS (PBS-1A, Capricorn Scientific) were mixed. In a

different reaction tube, 10 µl of each Cas9 (CP02, PNA Bio) preincubated with the sgRNA was mixed and made up to 300 µl with PBS (PBS-1A, Capricorn Scientific). For the experiments with the artificial fusion RNA, 1 µg of RNA was added to the mixture. The solution with the Cas9 (CP02, PNA Bio) proteins (and fusion RNA) was then pipetted into the Metafectene Pro (T040-1.0, Biontex) mix and incubated for 20 min at room temperature. The mixture was then pipetted onto the cell culture dish and mixed by tilting the dish. After 4 h the medium was exchanged with fresh DMEM (DMEM-LPA, Capricorn Scientific) containing 10% FCS (FBS-11A, Capricorn Scientific), 1% glutamine (STA-B, Capricorn Scientific), 1% PenStrep (PS-B, Capricorn Scientific) and ABA (342405000, Thermo Scientific).

RT-PCR and genomic PCR experiments to validate the onset of specific chromosomal translocations

To analyse the RNA and DNA from the same culture dish the AllPrep RNA/DNA Mini Kit (80204, Qiagen) from Qiagen was used according to the manufacturer's instructions. For the cDNA synthesis, 1 µg of RNA was made up to 7 µl with RNase-free water and mixed with 1 µl N6 primer (100 pmol). The sample was incubated for 10 min at 70°C and then cooled down on ice for 2 min. For the reaction, 1 µl RNasin, 4 µl 5x First Strand, 4 µl 2.5 mM dNTPs, 2 µl 100 mM DTT and 1 µl SuperScript™ II Reverse Transcriptase (18064071, Invitrogen) were added, incubated for 10 min at room temperature and then for 1 h at 42°C. To stop the reaction, 30 µl RNase-free water was added and the sample incubated for 10 min at 70°C. The *AFF1* wild-type transcript was amplified with AF4.3 and AF4.5, the *MLL* wild-type transcript with 8.3 and 13.5, the *AFF1::MLL* fusion transcript with AF4.3 and 13.5 and the *MLL::AFF1* transcript was analyzed with 8.3 and AF4.5. The *CCND3* wild-type gene was amplified with CCND3.E2.2.F and CCND3.E3.1.R, the *KMT2B* wild-type gene with KMT2B.E8.1.F and KMT2B.E15.1.R, the *CCND3::KMT2B* fusion transcript with CCND3.E2.2.F and KMT2B.E15.1.R, and the *KMT2B::CCND3* fusion transcript was analyzed with KMT2B.E8.1.F and CCND3.E4.2.R. All experiments were performed for n = 5. The PCR protocol used was 94°C for 3 min for the initial denaturation, 35 cycles of 94°C for 30 s, 63°C for 30 s and 72°C for 90 s, followed by 72°C for 2 min for the final elongation. RT-PCR products were analyzed on a 2% agarose gel.

The chromosomal translocations were analyzed using long-range PCR in combination with a touch-down PCR protocol (n = 5). The initial denaturation was 93°C for 3 min, the touch-down was 10 cycles starting with 93°C for 15 s, 68°C for 30 s and 68°C for 10 min while reducing the annealing temperature by 1°C every cycle, followed by 32 cycles, starting with 93°C for 15 s, 63°C for 30 s and 68°C for 10 min, while increasing the elongation time by 20 s in every cycle, followed by 68°C for 10 min as final elongation. The *AFF1::MLL* translocation was analyzed with AF4.cont.F and MLL.con.R1 and the *MLL::AFF1* translocation with MLL.cont.F and AF4.cont.R. The *KMT2B::CCND3* translocation was analyzed with KMT2B.LR.1.F and CCND3.LR.1.R and the *CCND3::KMT2B* translocation with CCND3-KMT2B.LR.1.F and CCND3-KMT2B.LR.1.R. Because there were many off target bands in the PCR for *CCND3::KMT2B* a nested PCR with CCND3-KMT2B.LR.3.F and CCND3-KMT2B.LR.3.R was performed. To differentiate the sequences of one translocation from distinct cells, smaller pieces of the bands were amplified to clone them into pCR™2.1-TOPO™ (450641, Invitrogen) according to the manufacturer's manual. For the *AFF1::MLL* allele the primers AF4.I3.BsaI.F and MLL.I11.SpaI.R and for the *MLL::AFF1* allele the primers MLL.I11.BsaI.F and AF4.I3.SapI.R were used. For the *KMT2B::CCND3* allele the primers KMT2B.E13.F and CCND3.I2.3.R and for the *CCND3::KMT2B* allele the primers CCND3.I2.F and KMT2B.E14.R were used. The primers were also used for sequencing. The sequences of the primers are listed in Star*Methods.

QUANTIFICATION AND STATISTICAL ANALYSIS

Not applicable.