Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press



Cancer translocations in human cells induced by zinc finger and TALE nucleases

Marion Piganeau, Hind Ghezraoui, Anne De Cian, et al.

Genome Res. published online April 8, 2013
Access the most recent version at doi:10.1101/gr.147314.112

P<P Published online April 8, 2013 in advance of the print journal.

Accepted Manuscript Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.

Open Access Freely available online through the *Genome Research* Open Access option.

Creative Commons License This manuscript is Open Access.This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported License), as described at http://creativecommons.org/licenses/by-nc/3.0/.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**



To subscribe to *Genome Research* go to:
https://genome.cshlp.org/subscriptions

© 2013, Published by Cold Spring Harbor Laboratory Press

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Cancer translocations in human cells induced by zinc finger and TALE nucleases

Marion Piganeau[#1], Hind Ghezraoui[#1], Anne De Cian[1], Lionel Guittat[3], Mark Tomishima[2], Loic Perrouault[1], Oliver René[1], George E. Katibah[4], Lei Zhang[4], Michael C. Holmes[4], Yannick Doyon[4], Jean-Paul Concordet[5], Carine Giovannangeli[1], Maria Jasin[2]* and Erika Brunet[1]*

[1]*Museum National d'Histoire Naturelle, 43 rue Cuvier, F-75005 Paris, France; CNRS, UMR7196, 43 rue Cuvier, F-75005 Paris, France; Inserm, U565, 43 rue Cuvier, F-75005 Paris, France*

[2]*Developmental Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA;*

[3] *INSERM U978, 74 rue Marcel Cachin, 93017 Bobigny, France; Université Paris 13, Sorbonne Paris Cité, Labex "Inflamex", 74 rue Marcel Cachin, 93017 Bobigny, France.*

[4]*Sangamo BioSciences, Inc., Point Richmond Tech Center, 501 Canal Boulevard, Suite A100, Richmond, CA 94804, USA*

[5] *Institut Cochin, Universite Paris Descartes; CNRS, UMR 8104; Inserm, U1016, Paris, France*

[#] Equal contribution

* Corresponding authors:

m-jasin@ski.mskcc.org, Developmental Biology Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065 USA

Phone: 212-639-7438, Fax: 212-772-8410

ebrunet@mnhn.fr, Museum National d'Histoire Naturelle, 43 rue Cuvier, F-75005 Paris, France; CNRS, UMR7196, 43 rue Cuvier, F-75005 Paris, France; Inserm, U565, 43 rue Cuvier, F-75005 Paris, France

Phone: +33-1-40-79-37-27, Fax:+33-1-40-79-37-05

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Abstract

Chromosomal translocations are signatures of numerous cancers and lead to expression of fusion genes which act as oncogenes. The wealth of genomic aberrations found in cancer, however, makes it challenging to assign a specific phenotypic change to a specific aberration. In this report, we set out to use genome editing with zinc finger (ZFN) and tale effector (TALEN) nucleases to engineer, de novo, translocation-associated oncogenes at cognate endogenous loci in human cells. Using ZFNs and TALENs designed to cut precisely at relevant translocation breakpoints, we induced cancer-relevant t(11;22)(q24;q12) and t(2;5)(p23;q35) translocations found in Ewing sarcoma and Anaplastic Large Cell lymphoma (ALCL), respectively. We recovered both translocations with high efficiency, resulting in the expression of the *EWSR1-FLI1* and *NPM1-ALK* fusions from the formation of Ewing sarcoma and ALCL translocations. Breakpoint junctions recovered after ZFN cleavage in human ES-cell derived mesenchymal precursor cells fully recapitulated the genomic characteristics found in tumor cells from Ewing sarcoma patients. This approach with tailored nucleases demonstrates that expression of fusion genes found in cancer cells can be induced from the native promoter, allowing interrogation of both the underlying mechanisms and oncogenic consequences of tumor-related translocations in human cells. With an analogous strategy, the ALCL translocation was reverted in a patient cell line to restore the integrity of the two participating chromosomes, further expanding the repertoire of genomic rearrangements that can be engineered by tailored nucleases.

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Introduction

Despite the wide range of recurrent chromosomal translocations identified in various cancers (more than 300 genes implicated) (Mitelman et al. 2007), the evidence is not always clear for a direct link from translocation formation to tumorigenesis. In mouse and human cells, studies are mostly based on either ectopically expressing or silencing the fusion gene induced by the translocation. In the case of fusion protein expression from a cDNA (often randomly integrated into the genome), the choice of the fusion transgenic promoter is crucial as the level of fusion protein expression must often be tightly controlled to recapitulate endogenous levels or risk spurious results from over expression. And in silencing strategies, even low levels of expression remaining for the fusion protein could mask to some extent the full cellular consequences of the translocation.

DNA double-strand breaks (DSBs) are considered to be causative lesions for many genomic rearrangements, including chromosomal translocations (Richardson and Jasin 2000; Mani and Chinnaiyan 2010). With the development of tailored endonucleases like Zinc Finger Nucleases (ZFNs) (Urnov et al. 2010; Carroll 2011) and more recently Transcription Activator-Like Effector Nucleases (TALENs) (Doyon et al. 2011; Mussolino and Cathomen 2012), it is now possible to create a DSB in the genome of a human cell at any locus of interest for the purposes of gene correction and gene disruption. In addition, contemporaneous expression of two ZFNs targeting loci has led to the induction of translocations at model loci in human multipotent and stem cells (Brunet et al. 2009) and intrachromosomal rearrangements (e.g., deletions) in transformed cell lines (Lee et al. 2010; Lee et al. 2012). This approach to study translocation formation obviates the need for prior genetic manipulation or cloning of cells, significantly expanding the repertoire of human cells that can be interrogated for translocation formation.

In this work, we now investigate the formation of two specific translocations, one frequently observed in Ewing sarcoma and one found in Anaplastic Large Cell Lymphoma (ALCL) using both types of nucleases (ZFNs and TALENs). Ewing sarcoma is a prototype of a solid tumor carrying a specific chromosomal translocation; it enables the transcription of the EWSR1-FLI1 chimeric protein corresponding to the in-frame fusion of the EWSR1 amino terminus with the FLI1 carboxyl-terminus. It is well accepted that the EWSR1-FLI1 fusion protein acts as a transcriptional factor, but target genes induced or repressed by the fusion protein are not fully identified yet (Chansky et al. 2004; Prieur et al. 2004; Smith et al. 2006; Riggi et al. 2010). ALCL is an aggressive T-cell non-Hodgkin lymphoma, accounting for as much as 10-15% of children with the disease. About half of tumors exhibits the specific translocation

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

t(2;5)(p23;q35) resulting in NPM1-ALK expression and constitutive ALK tyrosine kinase activity (Morris et al. 1994; Elmberger et al. 1995; Kuefer et al. 1997).

Results

Inducing Ewing sarcoma specific translocations with ZFNs

To target reported Ewing sarcoma breakpoints, two ZFN pairs were designed within the *EWSR1* and *FLI1* genes on chromosomes 22 and 11, respectively, to induce t(11;22)(q24;q12) translocations (Fig. 1A). ZFN^{EWS} targets *EWSR1* intron 7 and ZFN^{FLI} targets *FLI1* intron 5 (Fig. 1B), which contain breakpoints for the most common type of *EWSR1-FLI1* translocation (Plougastel et al. 1993). In particular, the ZFNs target sequences at breakpoint junctions reported in two tumors, T60 for *EWSR1* and T64 for *FLI1* (Supplemental Fig. S1) (Zucman-Rossi et al. 1998).

We chose to test this system in human mesenchymal precursor cells, in particular those derived from human embryonic stem cells (hES-MP (Barberi et al. 2005)), because of the presumed mesenchymal origin of this sarcoma (Tirode et al. 2007; Riggi et al. 2008). Both ZFNs efficiently generated DSBs at the respective target sites in these cells, as evidenced by the frequent formation of indels (Fig. 1B). To detect translocations, we used a nested PCR strategy on genomic DNA from cells transfected with the ZFN expression vectors (Fig. 1C). Translocation breakpoint junctions for both derivative chromosomes, der(11) and der(22), were detected upon expression of both ZFN^{EWS} and ZFN^{FLI} (Fig. 1C). Expression of a single ZFN, however, was not sufficient to give rise to translocations. We next asked whether the *EWSR1-FLI1* fusion transcript was expressed from der(22). Strikingly, we could detect the *EWSR1-FLI1* fusion transcript by RT-PCR of RNA extracted from bulk hES-MP cell population treated with ZFN^{EWS} and ZFN^{FLI}, but not from treatment with either ZFN^{EWS} or ZFN^{FLI} alone (Fig. 1D).

The detection of *EWSR1-FLI1* fusion transcripts from the transfected cell population suggested efficient t(11;22)(q24;q12) translocation. To quantify translocations, we applied a high throughput 96-well PCR screen for translocation junctions using small pools of cells (~800 cells per well) (Brunet et al. 2009). Translocation breakpoint junctions were detected in this small pool PCR at a frequency of $3.1 \pm 0.7 \times 10^{-4}$ for der(22) and $4.2 \pm 0.45 \times 10^{-4}$ for der(11). Reciprocal translocations were frequent: ~30% of wells with one derivative chromosome also contained the other derivative chromosome. The lack of a reciprocal product in the remaining wells may indicate non-reciprocal translocation formation, as observed in a portion of tumors (Zucman-Rossi et al. 1998) or, in some cases, lack of amplification (Brunet et al. 2009).

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Breakpoint junctions recovered after ZFN cleavage fully recapitulate the genomic characteristics found in Ewing tumor cells

The occurrence of translocation breakpoints within intronic sequences means that joining of DNA ends can occur with a variety of DNA end modifications without compromising the ability to express the EWSR1-FLI1 fusion protein. To gain insight into the joining mechanism, we analyzed 174 t(11;22)(q24;q12) breakpoint junctions detected by small pool PCR after ZFN expression and compared the joining characteristics (deletions, microhomology, insertions) with those from patient-derived translocations (Zucman-Rossi et al. 1998). Small deletions were observed in most ZFN-induced translocations: 74.5% of the derivative chromosomes had deletions of <30 bp total from the two DNA ends forming the junction (Fig. 2A; Supplemental Fig. S2). Exact breakpoints are not known for patient-derived events; however, for the 36 cases where junctions from both derivative chromosomes were sequenced, the total deletion from each chromosome could be determined and was found to be similarly small, with 71% of deletions also <30 bp (Supplemental Fig. S3). Except in one tumor junction (T23 with a deletion of 8.3 kb), all deletions from patient-derived translocations were <200 bp. Thus, our PCR based screening approach, which detects deletions of <250 bp, appears to be appropriate for detecting most translocations. Overall, no significant difference in deletion length was found between ZFN-induced and patient-derived translocations (medians of 6.5 and 19 bp, respectively, p=0.188, Mann-Whitney test, with the 8.3 kb deletion removed from the analysis as an outlier).

Breakpoint junctions can occur at microhomologies (short sequence identities) between DNA ends more frequently than that expected by chance. The involvement of microhomology for joining is considered to be more typical of the alternative pathway of NHEJ than of the canonical pathway (Weinstock et al. 2007; Yan et al. 2007; Simsek and Jasin 2010; Simsek et al. 2011). We determined the microhomology present at t(11;22)(q24;q12) junctions from both ZFN-induced and patient-derived events and compared the microhomology distributions with that expected from chance joining of two random sequences (Fig. 2B; Supplemental Fig. S2 and S3). Microhomologies of ≥3 bp were observed in only a fraction of junctions from both ZFN-induced events and tumors, although somewhat more frequently that than expected by chance. Only one particularly long microhomology of 8 bp was found in one tumor (T44: Supplemental Fig. S3). These results imply that substantial pre-existing microhomology is not required for translocation formation in human cells. Microhomology of 1 bp was often present in junctions involving the ZFN overhangs (e.g., asterisks, Supplemental Fig. S2), whereas for tumor

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

translocation junctions, which derive from a variety of DNA breakpoint sequences and types of ends, the major class of events showed no microhomology at the junctions, as would be expected if microhomology was not driving these events. Despite these differences, overall the difference in microhomology length distributions between the ZFN-induced and patient-derived events was not significant (p=0.076, Mann-Whitney test), supporting the ability of ZFN induction of translocations to model events seen in patients. They also point out that subtle design changes, i.e., avoiding sites where ZFN overhangs have the possibility to anneal at even 1 base, could even more faithfully recapitulate events in tumors.

As seen in some tumors from Ewing sarcoma patients (Zucman-Rossi et al. 1998), a portion of the ZFN-induced translocation breakpoints contained inserted sequences. Although many were only a few bp, several longer insertions were also detected which were derived from a variety of sources: nearby the DSB sites on chromosomes 11 and 22, other chromosomes, and exogenous DNA (i.e., the ZFN expression plasmids). For example, a patient der(11) junction was reported to have a 140 bp insertion from chromosome 1, with otherwise little modification to the chromosome 11 and 22 ends (T89, Supplemental Fig. S3). Likewise, a ZFN-induced der(11) junction was identified with a 180 bp insertion from chromosome 5, also with little other modification to the DNA ends.

Furthermore, insertions were also identified in ZFN-induced translocations where different pieces of DNA were assembled and/or duplicated in a more complex way (Fig. 2C; Supplemental Fig. S4). For duplications, der(22)-D duplicated 139 bp from chromosome 11 that is derived from 363 bp downstream of the DSB, and der(11)-A and der(11)-B duplicated chromosome 22 sequences derived from upstream and downstream of the DSB, respectively. A more complicated insertion is found in der(22)-B, which has a duplication of 153 bp from chromosome 11 derived from directly upstream of the DSB, but in inverted orientation; this type of insertion has previously been observed in patient translocations and has been termed a locally derived inverted sequence (LDIS) (Zucman-Rossi et al. 1998). Der(22)-B additionally contains an insertion from the ZFN plasmid. The longest insertion, in der(22)-C, is 994 bp and is derived from 4 DNA segments from 3 different chromosomes, suggestive of iterative processing of DNA ends prior to the ligation step (Simsek and Jasin 2010).

Inducing ALCL specific translocations with TALENs

With ZFNs, we managed to recapitulate Ewing sarcoma translocations in human stem cells but a limitation of extending this method to other translocations is that ZFN design is not trivial. The recent development of TALENs provides a potentially more advantageous strategy to

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

generate DSBs at endogenous loci in human cells, given the simpler code of base recognition (Miller et al. 2011). We designed TALENs to target DSBs to the *NPM1* and *ALK* genes on chromosomes 5 and 2, respectively, to induce t(2;5)(p23;q35) translocations found in ALCL patients (Morris et al. 1994) (Fig. 3A). TAL^{NPM} targets *NPM1* intron 4 and TAL^{ALK} targets *ALK* intron 16, both of which contain breakpoints found in patient-derived *NPM1-ALK* translocations. In this manner, TALEN-induced chromosomal translocations would result in expression of the oncogenic *NPM1-ALK* fusion gene, preserving the N-terminal dimerization domain of NPM1 and the C-terminal cytoplasmic domain of ALK.

Two different human cell lines were tested in this system. We first expressed the TALENs in a human acute T cell leukemia line, Jurkat, as ALCL is a lymphoma of T-cell origin (Gillis and Watson 1980). Jurkat cells are pseudodiploid with two normal *NPM1* and *ALK* loci (Mathas et al. 2009). Both TALENs efficiently generated DSBs at the *NPM1* and *ALK* target sites in these cells (Fig. 3B). Using a similar nested PCR strategy as for ZFN induction of *EWSR1-FLI1* translocations, we detected translocation breakpoint junctions for both derivative chromosomes when the TALENs were co-expressed, while no PCR product was detected when a single TALEN was expressed (Fig. 3C). Applying the same high throughput 96-well format used to detect *EWSR1-FLI1* translocations, we detected *NPM1-ALK1* translocations at a frequency of ~1 x 10^{-3} per cell (see Methods). Given this relatively high frequency of translocations, we attempted direct PCR detection of translocation junctions using only one round of PCR, and found that der(2) and der(5) breakpoint junctions were readily apparent (Fig. 3E). In serial dilutions of genomic DNA, both breakpoint junctions were obtained with as little as 6.25 ng DNA (representing ~1000 cells), corroborating that translocations arose at a frequency of ~10^{-3} per cell.

Confirming the generation of translocations, the *NPM1-ALK* fusion transcript was detected by PCR from the bulk Jurkat cell population following co-expression of TAL^{NPM} and TAL^{ALK} (Fig. 3D), and the NPM1-ALK fusion protein was detected by Western blot analysis using an antibody directed against ALK (Fig. 3F). The NPM1-ALK fusion protein co-migrated with that found in the ALCL tumor SUPM2 cells and was ~1/40th the level of protein seen in the SUPM2 cells (Fig. 3F). Expanding pools of cells, junctions were observed 25 days following TALEN expression (data not shown), indicating maintenance of the translocation.

We also performed similar experiments in retinal pigment epithelial cells, RPE-1, expressing hTERT to determine if the efficiency of t(2;5)(p23;q35) induced by TALENs is seen in other cells or may be related to the lymphocytic origin of the Jurkat cells. As in Jurkat cells, both TALENs efficiently generated DSBs at the target sites (Supplemental Fig. S5A). We

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

recovered translocation junctions for der(2) and der(5) from RPE-1 cells co-expressing both TAL^{NPM} and TAL^{ALK} (Supplemental Fig. S5B); using one round of PCR on dilutions of genomic DNA from the transfected cells, translocation frequency was estimated at 0.5–1 X 10^{-2} (Supplemental Fig. S5C). The *NPM1-ALK* fusion transcript and fusion protein were detected from the bulk transfected cell population (Supplemental Fig. S5D-E). Fluorescence in situ hybridization confirmed the presence and the high frequency of the translocation (2 of 70 metaphases, Fig. 3A). Junctions were observed at 8 and 26 days following TALEN expression (Fig. 3G), indicating maintenance of the translocation. These results demonstrate that TALENs can efficiently induce t(2;5)(p23;q35) in different human cell types and that the translocation can be propagated.

TALENs and ZFNs both cleave DNA using the FokI nuclease domain but the left and right DNA binding domains have a different spacing from each other in the two types of nucleases, with the potential to influence the site(s) of DNA cleavage. For ZFN^{EWS} and ZFN^{FLI}, the spacers are 6 bp; cleavage occurs at two sites, adjacent to the DNA binding site and 1 bp into the spacer, leaving 5' overhangs of 6 and 4 bp, respectively (Smith et al. 2000) (Fig. 1A), consistent with our junction analysis (Supplemental Fig. S2). By contrast, the spacers in TAL^{NPM} and TAL^{ALK} are 18 bp (Fig. 3B). To map the TALEN cleavage sites, TAL^{NPM} and TAL^{ALK} were expressed in vitro and cleavage assays were performed on double-stranded DNA fragments labeled at either 5' end (Supplemental Fig. S6). For both TAL^{NPM} and TAL^{ALK} we identified a region of cleavage spanning 6 to 7 bp which overlapped the center of the spacer (Fig. 3B). TALEN cleavage results primarily in 5' overhangs, potentially of different lengths, although a fraction of blunt ends are also possible.

To examine joining mechanisms from TALEN-induced translocations, we analyzed 161 breakpoint junctions arising from t(2;5)(p23;q35) in Jurkat cells (Supplemental Fig. S7A). As the TALEN cleavage sites may vary in cells as in vitro, we considered the longest 5'-overhang encompassing the full set of in vitro cleavage sites. Consistent with the in vitro cleavage assays, bases predicted to be in an overhang were incorporated into some of the junctions, and deletions emanating from the predicted DSB sites were frequent. As seen in the ZFN-induced translocations, small deletions were most common in the TALEN-induced translocations: 81% of junctions had deletions <30 bp, while <5% of junctions had deletions of >100 bp.

One feature of TALEN-induced breakpoint junctions was the frequent presence of short insertions (Supplemental Fig. S7A). Almost half of breakpoint sequences showed insertions of <10 bp (47.5%), compared with 7% in the ZFN-induced translocations. A few insertions could be identified as duplications of sequences close to the DSB sites but most of them were not

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

readily evident as being templated. As shown for ZFN induced translocations, some longer insertions were observed that were derived from a variety sources: duplications of regions located near the DSB sites on chromosomes 2 and 5 as well as sites more distant from the DSBs (Mb), and the TALEN expression plasmids. Complex insertions were also identified. For example, one der(5) junction had tandem duplications from both chromosomes 2 and 5 in addition to a deletion from each DSB and another der(5) junction had duplications from chromosome 5 from both upstream and downstream of the DSB (junctions 19 and 54, respectively; Supplemental Fig. S7A).

The involvement of microhomology in the joining mechanism of TALEN-induced translocations is more problematic to analyze because of the large number of short insertions. Nevertheless, when considering only sequences without insertions, microhomologies of 3 to 6 bp were observed in only a small fraction of sequences (6.2%), suggesting a minimal dependence on microhomology in formation of these translocations as found in ZFN-induced and tumor translocations.

All of the joining characteristics observed in Jurkat cells were recapitulated in translocations from hES-MP cells, including a large fraction of junctions with short insertions (24/38 total; Supplemental Fig. S7B). Frequent small insertions were previously reported for intrachromosomal repair after expression of a single TALEN in hES cells (Hockemeyer et al. 2011), although not in other another study using U2OS cells (Reyon et al. 2012).

The homodimerizing FokI domains in the TALENs leave open the possibility that newly formed translocation junctions could be recleaved by mixed TALENs (i.e., $TAL^{NPM/ALK}$ and $TAL^{ALK/NPM}$). To circumvent this possibility, TALENs were also constructed with FokI heterodimerization domains (Doyon et al. 2011). Translocations were observed in Jurkat cells with the heterodimeric TALENs (Supplemental Fig. S8A). Junctions showed similar characteristics as with the homodimeric TALENs (Supplemental Fig. S8B), including frequent small insertions, indicating that cleavage of newly formed translocation junctions are not substantially affecting characteristics.

Reversion of the *NPM1-ALK* translocation

The ability to specifically induce translocations suggested that it should also be possible to reverse translocations with a similar approach (Fig. 4A). Genetic reversion permits the analysis of phenotypic consequences of fusion protein loss, such as whether tumor cells are “addicted” to the oncogene. To revert the t(2;5)(p23;q35) translocation, TAL^{NPM} and TAL^{ALK} were expressed in SUDHL-1 patient-derived cells, with efficient cleavage of both the *NPM1* and *ALK* loci (Fig.

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

4B). Due to the position of the TALEN cleavage sites relative to the translocation breakpoint junctions, restoration of the intact chromosomes leaves a small segment of DNA from the partner chromosome that serves as a tag to differentiate the revertant chromosomes from the wild-type chromosomes 2 and 5 (Fig. 4C). By PCR, both revertant chromosomes were detected (Fig. 4C), and sequencing confirmed that the translocation had been reversed (Fig. 4D and Supplemental Fig. S9). We also tested the reversion strategy in SUPM2 cells, another ALCL patient-derived cell line, and also obtained revertant breakpoint junctions (Supplemental Fig. S9). As with translocations, many of breakpoint junctions for both revertant chromosomes from both ALCL cell lines had small insertions. Overall then, frequent small insertions were observed at 4 different junctions (two translocation and two reversion junctions), suggesting that they do not derive from a specific sequence context from the TALEN pair we used.

Discussion

In this report, we could reproduce cancer relevant translocations and revert them by targeting concurrent DSBs to intronic sequences in human cells using two types of designer nucleases, ZFNs and TALENs. With this approach, fusion genes created during translocation formation, which are potential oncogenes, are expressed from their endogenous promoters. In addition, translocation disrupts a copy of each participating gene, reproducing the heterozygosity (and potential haploinsufficiency) found in tumor cells. Therefore, this strategy provides a more accurate model for translocation-related tumor formation than ectopic expression of a fusion protein to study effects on proliferation and other cellular phenotypes of newly translocated cells and, further, provides an approach to revert the translocation.

Despite the fact that a ZFN induces cohesive 5'-overhangs that should be easily repaired by direct ligation or by polymerase fill-in followed by ligation, breakpoint junctions of ZFN-treated cells recapitulated the features found in Ewing sarcoma cells: deletions, short insertions, complex insertions from other chromosomes and/or nearby the break regions, and an absence of long microhomology. Overall ZFN-treated cells showed a comparable proportion of each type of junction modification as patient cells, with the exception of complex duplications, which are overrepresented in tumor cells (Zucman-Rossi et al. 1998). It is possible that these latter events in patients arise from multiple breaks in the same intronic regions (fragile sites). Translocations junctions obtained with TALENs frequently showed short deletions and other

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

modifications seen in patient translocation junctions. Short insertions were overrepresented, although it is not clear if this is specific to the TALEN pair.

Relevant to Ewing sarcoma, the strategy presented here has the potential to resolve inconsistencies in previous studies which have attempted to identify target genes using ectopic expression or knock-down of the EWSR1-FLI1 fusion protein and also those in studies using mouse cells as a tumor model (reviewed in (Lessnick and Ladanyi 2012)). Der(5) NPM1-ALK driven ALCL tumorigenesis appears to be more defined, since the fusion protein is known to inhibit apoptosis and promote cell proliferation through the constitutively active ALK tyrosine kinase moiety (Kuefer et al. 1997; Greenland et al. 2001; Kasprzycka et al. 2006). Nonetheless, it must be noted that NPM1 is a multifunctional protein involved in many cellular processes (Colombo et al. 2006), including centrosome duplication through a C-terminal interaction with the CRM1 complex, such that der(2) expressing the NPM1 C terminus has the potential to participate in the lymphomagenesis of ALCL in addition to der(5) (Wang et al. 2005). And while normal levels of NPM1 induce senescence and/or apoptosis under UV damage conditions by binding p53 (Kurki et al. 2004), NPM1 overexpression can lead to cellular transformation (Li et al. 2008), although paradoxically, loss of one allele of *NPM1* leads to DNA instability allowing possible cellular transformation as well (Grisendi and Pandolfi 2005). Thus, formation of both derivative chromosomes with loss of at least one allele of *NPM1* may all be critical to faithfully recapitulate the disease initiation state.

The present work provides a novel approach to understand translocation-related tumorigenesis, and more broadly demonstrates that precise genome rearrangements can be achieved in relevant cell types with tailored nucleases. This work also expands the repertoire of genome engineering accomplished with these nucleases beyond gene mutation/correction and intrachromosomal rearrangements (Lee et al. 2010; Lee et al. 2012) and translocations in model systems (Brunet et al. 2009; Simsek et al. 2011). Unlike site-specific recombinases, endonucleases are designed to cleave endogenous sites, abrogating the need for prior genome manipulation. And because they introduce DSBs, endonucleases provide more accurate models for studying the etiology of genomic rearrangements and their potential for reversion, as well as providing more accurate cellular models for further functional investigation.

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Methods

Cell lines, transfections, FISH

hES-MP cells, derived from hES strain H1, were provided by the SKI Stem Cell Research Facility (MSKCC), and cultivated in alpha-MEM supplemented with 10% FBS and 50 ng/ml of recombinant human FGF (R&D Systems). Jurkat and SUDHL-1 and SUPM2 (DSMZ, Germany) cells were cultivated in RPMI-1640 Medium supplemented with 10% h.i. FBS. The hTERT-immortalized retinal pigment epithelial cell line, RPE-1, was cultured in DMEM:F12 Medium supplemented with 10% FBS.

Transfections were done as previously described (Brunet et al. 2009). Typically, 7.5 X 10^5 (for hES-MP and RPE-1) or 2 X 10^6 (for Jurkat, SUDHL-1 and SUPM2) cells were transfected by Amaxa technology (Lonza) using the cell line nucleofector kit V (program B-16 for hES-MP cells and X-001 for Jurkat, SUDHL-1, SUPM2, and RPE-1 cells) with 7.5 µg of each ZFN^{EWS} and 5 µg of each ZFN^{FLI1}; or 5 µg of each TAL^{NPM} and 2 µg (for Jurkat, SUDHL-1 and SUPM2 cells) to 5 µg (for RPE-1 cells) of each TAL^{ALK}.

FISH was performed on metaphases of RPE-1 cells 48 h after transfection with TAL^{NPM} and TAL^{ALK} using *ALK* break apart probes (LPS019 Amplitech) and an *NPM1* BAC probe (RP11-117L6). Images were taken with the Zeiss Axio Observer.Z1 system.

PCR based translocation detection, frequency determination, and sequencing of breakpoint junctions

Primer sequences are below. For detection of translocations in pooled cells, 100 ng genomic DNA from cells transfected with ZFNs or TALENs was amplified by nested PCR 72 h post-transfection. Where noted, the transfected cell population was kept in culture for the indicated length of time to monitor maintenance of the translocation. For translocation frequency determination, a high throughput 96-well screen using small pools of cells and nested PCR to amplify translocation junctions was used 48 h post-transfection, as described (Brunet et al. 2009). The number of PCR-positive and negative wells for translocation junctions was determined to calculate the translocation frequency. Amplified products from positive wells were directly sequenced to verify translocations and determine the junction sequences. For hES-MP cells, 96-well plates contained 8 X 10^4 transfected cells per plate, resulting in 17 to 33 positive wells per plate in 4 experiments for each derivative chromosome. For Jurkat cells, 96-well plates contained 2 X 10^5 transfected cells per plate. Centrifugation of plates with the non-adherent Jurkat cells, followed by media removal, resulted in cell loss, and so cells were requantified prior

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

to lysis, demonstrating ~1 X 10^5 cells remaining. After transfection of the TALEN homodimer expression plasmids, two 96-well plates were analyzed for each derivative chromosome; 61 ± 2 positive wells were obtained per plate, which adjusted to a beta distribution is estimated to represent about 95 translocations per plate, for a frequency of ~1 x 10^{-3}. Sequencing of PCR products from positive wells showed mixed sequences in a single well from almost half of the wells corresponding to multiple junctions, consistent with the high frequency of translocations, and were removed from the junction analysis. For TALEN heterodimer experiments, translocation frequency was half that of the homodimer experiments.

Translocation frequency was also determined for Jurkat cells and for RPE-1 cells by serial dilution of genomic DNA from cells transfected with the TALEN expression vectors followed by one round of PCR (i.e., non-nested) indicated in Fig. 3C. For Jurkat cells, genomic DNA dilutions were 1:2, for 50, 25, 12.5, 6.25, 3.125 and 1.56 ng DNA. Considering that one cell contains ~6 pg DNA (Jurkat cells being a pseudodiploid human cell line), 6.25 ng represents ~10^3 cells. PCR amplification with 6.25 ng DNA was seen in half of the experiments with the homodimeric TALENs, corresponding to a frequency ~10^{-3} per cell. For the heterodimeric TALENs, junction detection was less efficient (1 of 3 experiments with 6.25 ng DNA and 2 of 3 experiments with 12.5 ng DNA), implying about half the translocation efficiency as with homodimeric TALENs. Thus, translocation frequency determination by the 96-well format and by serial dilution of genomic DNA gave similar results. For RPE-1 cells, genomic DNA dilutions were 1:2, for 5, 2.5, 1.25, 0.625, and 0.31 ng DNA. Amplification was detected at 1.25 ng genomic DNA (2 of 3 experiments) and 0.625 (1 of 3 experiments), indicating a frequency of 0.5-1 X 10^{-2} per cell. The detection of 2 translocations by FISH from 70 total metaphases is in agreement with the high translocation frequency estimated for RPE-1 cells.

PCR for revertant chromosome detection and sequencing of breakpoint junctions

Both Chr(2) and Chr(5) revertant chromosomes were detected by PCR amplification 72 h post-transfection of the TALEN expression vectors. Chr(2) and Chr(5) amplified products were cloned with the TOPO-TA cloning system (Invitrogen) and sequenced.

T7 endonuclease 1 assay

48 h after transfection, genomic DNA was extracted with QIAamp DNA Mini Kit (Qiagen). The genomic region encompassing the ZFN and TALEN target sites was amplified with the primers listed below, typically using 50 ng genomic DNA, melted and annealed (95°C for 5 min, 95°C to 25°C at -0.5°C/30 s and 4°C for 15 min) to form heteroduplex DNA. The annealed DNA was

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

treated with 1.5 units of T7 endonuclease 1 (New England BioLabs) for 10 min at 37°C and run on a 2.4% agarose gel.

RT-PCR and cDNA amplification

RNA was extracted from transfected cells with absolutely RNA microprep Kit (Agilent technologies) and RT-PCR was done following manufacturer directions (SuperScript II First-Strand Synthesis System, Invitrogen). Primers used for cDNA amplification of *EWSR1-FLI1* and *NPM1-ALK* are listed below.

Western blotting

Whole cell extracts were prepared with protein lysis buffer (50 mM Tris-HCl pH 7.4, 1% Triton X-100, 0.1% SDS, 150 mM NaCl, 1 mM EDTA, and 1 mM DTT prepared from a 1 M DTT stock), with addition of cocktail protease inhibitor tablets (Complete, Roche). Typically 40 µg protein extract from TALEN-treated cells were run on an 8% (w/v) Tris-HCl SDS page gel, blotted, and then probed with ALK (31F12) Mouse Antibody (Cell Signal). α-tubulin Mouse antibody (Sigma) was used to detect those proteins as loading controls.

In vitro mapping of TALEN cleavage sites

TNT SP6 Quick coupled Transcription/Translation System (Promega) was used to express TAL^{NPM} and TAL^{ALK} (Mussolino et al. 2011). 68 bp fragments containing target sites of TAL^{NPM} and TAL^{ALK} were 5'-radiolabelled on one strand. Typically, 0.5 µl of each TNT lysate containing one TALEN subunit was incubated overnight at 37°C with 10 nM target duplex in NEB4 buffer (New England Biolabs) supplemented with 100 mM NaCl and 0.4 µg salmon sperm (DNA competitor) in 10 µl. The reaction was analyzed on a denaturing 12% PAGE gel.

Primers:

A-Ewing system:

- translocation

	PCR(1)		Nested-PCR (2)	
Der22	DER22-f	AGGCTGGTCTCGAACTCCTG	DER22-Nf	CAGCCACGTTTGGAGTTTTT
	DER22-b	TCTGGAAAAGCCCGTTCAGT	DER22-Nb	GAATGAGGTGGAAGGCTCTG
Der11	DER11-f	ACCCCACATACCACAAGCTC	DER11-Nf	AAGGTCCCAGAGTTTGCTGA
	DER11-b	AGGGGCTGAGCTCCATAAAT	DER11-Nb	TTTTTCCCCCTTTCCTTTTC

-T7 cleavage

EWSR1 gene :

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

EWS-T7-f GCCACGTTTGGAGTTTTTGA
EWS-T7-b GCCATTTTTCCCCTTAATGG
FLI1 gene :
FLI1-T7-f ACCTTGACATCTCACCACCC
FLI1-T7-b GAATGAGGTGGAAGGCTCTG

-Fusion *EWSR1-FLI1* cDNA

EWS-FL-RNA-f CAGCAGGGCTACAGTGCTTA
EWS-FL-RNA-b GTTGGGGTTGGGGTAGATTC

B-ALCL system:

-translocation

	PCR(1)		**Nested-PCR (2)**	
Der5	DER5-f	CAGTTGCTTGGTTCCCAGTT	DER5-Nf	GGGGAGAGGAAATCTTGCTG
	DER5-b	AGGAATTGGCCTGCCTTAGT	DER5-Nb	GCAGCTTCAGTGCAATCACA
Der2	DER2-f	TCCTTCAGTGTCCATCACGA	DER2-Nf	CCCACCCCTAGACGTCACT
	DER2-b	GAACCTTGCTACCACCTCCA	DER2-Nb	TTCACATCCTCCTCCTCTTCA

-reversion

Revertant Chr(2):
SUDHL-1 cells
DHL1-chr2-f CAGAGACATGCCCAGGACAG
DHL1-chr2-b GCCTAATGTTCACTGCCAGGT

SUPM2 cells
M2-ch2-f CTGCCGGTAGAAGGGAGATG
M2-ch2-b CCACTTGAGGCCAAGAGTCC

Revertant Chr(5):
SUDHL-1 cells
DHL-ch5-f AAGGGGTTTCGTGATGTTGG
DHL-ch5-b CCCACCCTCTAGGGTTGTCA

SUPM2 cells
M2-ch5-f GTGCACCACCAAGTCCAGTT
M2-ch5-b CCCACCCTCTAGGGTTGTCA

-T7 cleavage

NPM1 gene
T7-NPM-f ACCACCAAGTCCAGTTGCGTTT
T7-NPM-b ATCTCACTTTGGGGAGGCCAAG
ALK gene
T7-ALK-f AGATGGGCAGAGGCTTGAAAAG
T7-ALK-b TGAGGATGTTCTGGAAGGCAAA

-Fusion *NPM1-ALK* cDNA
NP-AL-RNA-f TTGTGAACTAAAGGCCGACAAA
NP-AL-RNA-b CCAGGCTGGTTCATGCTATTCT

C-Control cDNA:
GAPDH-f TGCACCACCAACTGCTTAGC

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

GAPDH-b GGCATGGACTGTGGTCATGAG

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Acknowledgments

We thank Francesca Cole and other members of the Jasin laboratory, Patrizia Alberti and Loic Ponger from the laboratory of “Regulation and dynamics of genomes” at MNHN, and Alexei Morozov in Malcolm Moore’s laboratory at MSKCC for helpful discussions. We also thank Fyodor Urnov (Sangamo BioSciences, Inc.) for his help to initiate the project. M.P. is supported by a “Le Canceropole IDF” PhD grant and H.G is supported by a “La Ligue Nationale contre le Cancer” PhD grant. This work was supported by CNRS, Inserm, MNHN, IBiSA (Infrastructures for Biology), and ANR-12-JSV6-0005-01 (E.B.) and by R01 NIHGM54668 grant (M.J.).

Author Contributions

E.B. and M.J. designed research; M.P, H.G., A.D.C., L.G, M.T, L.P, O.R and E.B. performed experiments; G-E.K, L.Z, M.H and Y.D designed new ZFN reagents; A.D.C, J-P.C and C.G designed new TALEN reagents and did analytic analysis; E.B. and M.J. did statistical analysis of data; E.B. and M.J. wrote the paper.

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

References

Barberi T, Willis LM, Socci ND, Studer L. 2005. Derivation of multipotent mesenchymal precursors from human embryonic stem cells. *PLoS Med* **2**: e161.

Brunet E, Simsek D, Tomishima M, DeKelver R, Choi VM, Gregory P, Urnov F, Weinstock DM, Jasin M. 2009. Chromosomal translocations induced at specified loci in human stem cells. *Proc Natl Acad Sci* **106**: 10620-10625.

Carroll D. 2011. Genome engineering with zinc-finger nucleases. *Genetics* **188**: 773-782.

Chansky HA, Barahmand-Pour F, Mei Q, Kahn-Farooqi W, Zielinska-Kwiatkowska A, Blackburn M, Chansky K, Conrad EU, 3rd, Bruckner JD, Greenlee TK et al. 2004. Targeting of EWS/FLI-1 by RNA interference attenuates the tumor phenotype of Ewing's sarcoma cells in vitro. *J Orthop Res* **22**: 910-917.

Colombo E, Martinelli P, Zamponi R, Shing DC, Bonetti P, Luzi L, Volorio S, Bernard L, Pruneri G, Alcalay M et al. 2006. Delocalization and destabilization of the Arf tumor suppressor by the leukemia-associated NPM mutant. *Cancer Res* **66**: 3044-3050.

Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, Miller JC, Urnov FD, Gregory PD, Holmes MC. 2011. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods* **8**: 74-79.

Elmberger PG, Lozano MD, Weisenburger DD, Sanger W, Chan WC. 1995. Transcripts of the npm-alk fusion gene in anaplastic large cell lymphoma, Hodgkin's disease, and reactive lymphoid lesions. *Blood* **86**: 3517-3521.

Gillis S, Watson J. 1980. Biochemical and biological characterization of lymphocyte regulatory molecules. V. Identification of an interleukin 2-producing human leukemia T cell line. *J Exp Med* **152**: 1709-1719.

Greenland C, Touriol C, Chevillard G, Morris SW, Bai R, Duyster J, Delsol G, Allouche M. 2001. Expression of the oncogenic NPM-ALK chimeric protein in human lymphoid T-cells inhibits drug-induced, but not Fas-induced apoptosis. *Oncogene* **20**: 7386-7397.

Grisendi S, Pandolfi PP. 2005. NPM mutations in acute myelogenous leukemia. *N Engl J Med* **352**: 291-292.

Gu J, Lu H, Tippin B, Shimazaki N, Goodman MF, Lieber MR. 2007. XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. *EMBO J* **26**: 1010-1023.

Guschin DY, Waite AJ, Katibah GE, Miller JC, Holmes MC, Rebar EJ. 2010. A rapid and general assay for monitoring endogenous gene modification. *Methods Mol Biol* **649**: 247-256.

Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC et al. 2011. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* **29**: 731-734.

Kasprzycka M, Marzec M, Liu X, Zhang Q, Wasik MA. 2006. Nucleophosmin/anaplastic lymphoma kinase (NPM/ALK) oncoprotein induces the T regulatory cell phenotype by activating STAT3. *Proc Natl Acad Sci* **103**: 9964-9969.

Kuefer MU, Look AT, Pulford K, Behm FG, Pattengale PK, Mason DY, Morris SW. 1997. Retrovirus-mediated gene transfer of NPM-ALK causes lymphoid malignancy in mice. *Blood* **90**: 2901-2910.

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Kurki S, Peltonen K, Latonen L, Kiviharju TM, Ojala PM, Meek D, Laiho M. 2004. Nucleolar protein NPM interacts with HDM2 and protects tumor suppressor protein p53 from HDM2-mediated degradation. *Cancer Cell* **5**: 465-475.

Lee HJ, Kim E, Kim JS. 2010. Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res* **20**: 81-89.

Lee HJ, Kweon J, Kim E, Kim S, Kim JS. 2012. Targeted chromosomal duplications and inversions in the human genome using zinc finger nucleases. *Genome Res* **22**: 539-548.

Lessnick SL, Ladanyi M. 2012. Molecular pathogenesis of Ewing sarcoma: new therapeutic and transcriptional targets. *Annu Rev Pathol* **7**: 145-159.

Li Z, Boone D, Hann SR. 2008. Nucleophosmin interacts directly with c-Myc and controls c-Myc-induced hyperproliferation and transformation. *Proc Natl Acad Sci* **105**: 18794-18799.

Mani RS, Chinnaiyan AM. 2010. Triggers for genomic rearrangements: insights into genomic, cellular and environmental influences. *Nat Rev Genet* **11**: 819-829.

Mathas S, Kreher S, Meaburn KJ, Johrens K, Lamprecht B, Assaf C, Sterry W, Kadin ME, Daibata M, Joos S et al. 2009. Gene deregulation and spatial genome reorganization near breakpoints prior to formation of translocations in anaplastic large cell lymphoma. *Proc Natl Acad Sci* **106**: 5831-5836.

Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ et al. 2011. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* **29**: 143-148.

Mitelman F, Johansson B, Mertens F. 2007. The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer* **7**: 233-245.

Morris SW, Kirstein MN, Valentine MB, Dittmer KG, Shapiro DN, Saltman DL, Look AT. 1994. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science* **263**: 1281-1284.

Mussolino C, Cathomen T. 2012. TALE nucleases: tailored genome engineering made easy. *Curr Opin Biotechnol DOI: 10.1016/j.copbio.2012.01.013*.

Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T. 2011. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res* **39**: 9283-9293.

Nick McElhinny SA, Havener JM, Garcia-Diaz M, Juarez R, Bebenek K, Kee BL, Blanco L, Kunkel TA, Ramsden DA. 2005. A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining. *Mol Cell* **19**: 357-366.

Plougastel B, Zucman J, Peter M, Thomas G, Delattre O. 1993. Genomic structure of the EWS gene and its relationship to EWSR1, a site of tumor-associated chromosome translocation. *Genomics* **18**: 609-615.

Prieur A, Tirode F, Cohen P, Delattre O. 2004. EWS/FLI-1 silencing and gene profiling of Ewing cells reveal downstream oncogenic pathways and a crucial role for repression of insulin-like growth factor binding protein 3. *Mol Cell Biol* **24**: 7275-7283.

Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. 2012. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol* **30**: 460-465.

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Richardson C, Jasin M. 2000. Frequent chromosomal translocations induced by DNA double-strand breaks. *Nature* **405**: 697-700.

Riggi N, Suva ML, De Vito C, Provero P, Stehle JC, Baumer K, Cironi L, Janiszewska M, Petricevic T, Suva D et al. 2010. EWS-FLI-1 modulates miRNA145 and SOX2 expression to initiate mesenchymal stem cell reprogramming toward Ewing sarcoma cancer stem cells. *Genes Dev* **24**: 916-932.

Riggi N, Suva ML, Suva D, Cironi L, Provero P, Tercier S, Joseph JM, Stehle JC, Baumer K, Kindler V et al. 2008. EWS-FLI-1 expression triggers a Ewing's sarcoma initiation program in primary human mesenchymal stem cells. *Cancer Res* **68**: 2176-2185.

Roth DB, Porter TN, Wilson JH. 1985. Mechanisms of nonhomologous recombination in mammalian cells. *Mol Cell Biol* **5**: 2599-2607.

Simsek D, Brunet E, Wong SY, Katyal S, Gao Y, McKinnon PJ, Lou J, Zhang L, Li J, Rebar EJ et al. 2011. DNA Ligase III Promotes Alternative Nonhomologous End-Joining during Chromosomal Translocation Formation. *PLoS Genet* **7**: e1002080.

Simsek D, Jasin M. 2010. Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4-ligase IV during chromosomal translocation formation. *Nat Struct Mol Biol* **17**: 410-416.

Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D. 2000. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res* **28**: 3361-3369.

Smith R, Owen LA, Trem DJ, Wong JS, Whangbo JS, Golub TR, Lessnick SL. 2006. Expression profiling of EWS/FLI identifies NKX2.2 as a critical target gene in Ewing's sarcoma. *Cancer Cell* **9**: 405-416.

Tirode F, Laud-Duval K, Prieur A, Delorme B, Charbord P, Delattre O. 2007. Mesenchymal stem cell features of Ewing tumors. *Cancer Cell* **11**: 421-429.

Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. 2010. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11**: 636-646.

Wang W, Budhu A, Forgues M, Wang XW. 2005. Temporal and spatial control of nucleophosmin by the Ran-Crm1 complex in centrosome duplication. *Nat Cell Biol* **7**: 823-830.

Weinstock DM, Brunet E, Jasin M. 2007. Formation of NHEJ-derived reciprocal chromosomal translocations does not require Ku70. *Nature cell biology* **9**: 978-981.

Yan CT, Boboila C, Souza EK, Franco S, Hickernell TR, Murphy M, Gumaste S, Geyer M, Zarrin AA, Manis JP et al. 2007. IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* **449**: 478-482.

Zucman-Rossi J, Legoix P, Victor JM, Lopez B, Thomas G. 1998. Chromosome translocation based on illegitimate recombination in human tumors. *Proc Natl Acad Sci* **95**: 11786-11791.

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Figures Legends

Figure 1. Induction of t(11;22)(q24;q12) translocations in hES-MP cells with ZFNs.

A. Ewing sarcoma translocations involve breakpoints within the *EWSR1* and *FLI1* genes on chromosomes 22 and 11, respectively, creating an *EWSR1-FLI1* fusion gene on der(22). To induce t(11;22)(q24;q12), ZFNs are expressed in hES-MP cells to create DSBs (scissors) in both genes.

B. ZFN^{EWS} and ZFN^{FLI} cleavage sites are within *EWSR1* and *FLI1* introns, respectively, relevant to the *EWSR1-FLI1* translocation. Zinc fingers in the ZFNs are designed to bind to the shaded sequences. Arrows represent the presumed DSB sites after FokI nuclease domain cleavage. ZFN cleavage activity in hES-MP cells is monitored by a T7-endonuclease assay (Guschin et al. 2010). The region around the ZFN site is amplified; the amplified product is then denatured, reannealed, and then subjected to T7 endonuclease cleavage. Insertions and deletions (indels) characteristic of imprecise DSB repair by NHEJ give rise to T7 endonuclease-cleavable DNA.

C. Nested PCR to detect derivative chromosomes der(11) and der(22) in hES-MP cells. Translocation breakpoint junctions are only detected after expression of both ZFN^{EWS} and ZFN^{FLI}.

D. RT-PCR detection of the *EWSR1-FLI1* fusion transcript after ZFN^{EWS} and ZFN^{FLI} expression in hES-MP cells. The forward primer overlaps the exon 2/3 junction of *EWSR1* and the reverse primer is within exon 9 of *FLI1*, amplifying most of the *EWSR1-FLI1* coding sequence.

Figure 2. Analysis of t(11;22)(q24;q12) breakpoint junctions.

A. Deletion lengths for der(11) and der(22) junctions from ZFN-induced translocations from hES-MP cells. Each value represents the combined deletion from both ends of an individual junction. The median deletion length is indicated by a bar on the graph and the value is given below the graph.

B. Microhomology length distributions for junctions from ZFN-induced and tumor translocation (Zucman-Rossi et al. 1998). (Der(11) and der(22) are grouped). Only junctions with simple deletions (i.e., without an insertion) are included. The probability that a junction will have *X* nucleotides of microhomology by chance assumes an unbiased base composition and is calculated as previously described (Roth et al. 1985).

C. Examples of complex breakpoint junctions with insertions derived from sequences near the DSB sites. Schematics of the derivative and parental chromosomes are shown. The DSB sites on the unrearranged chromosomes are represented by scissors. Segments that are duplicated,

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

deleted or added are represented as independent blocks and in the orientation relative to that found on the parental chromosomes. (See Supplemental Fig. S4 for the sequences).

Figure 3. Induction of t(2;5)(p23;q35) translocations with TALENs.

A. ALCL translocations have breakpoints within the *NPM1* and *ALK* genes on chromosomes 2 and 5, respectively, creating an *NPM1-ALK* fusion gene on der(5). To induce t(2;5)(p23;q35), TALENs are expressed to create DSBs (scissors) in both genes. FISH demonstrates the t(2;5)(p23;q35) translocation after TALEN expression in RPE-1 cells. Red and green signals are from an *ALK* probe that "breaks apart" upon translocation. The blue signal is from an *NPM1* probe. Of 70 metaphases screened, 2 exhibited translocations and 3 showed breaks with the *ALK* break apart probe, likely due to remaining TALEN expression at this time.

B. TAL^{NPM} and TAL^{ALK} cleavage within *NPM1* and *ALK* introns, respectively, relevant to the *NPM1-ALK* translocation. DNA binding domains of TALENs are designed to bind the shaded sequences. Arrows represent DSB sites with different tail lengths representing the efficiency of cleavage in vitro, as assayed by in vitro expression of the TALENs (see Supplemental Fig. S6). TALEN cleavage activity in Jurkat cells is monitored by the T7-endonuclease assay, as shown below the sequences.

C. Nested PCR to detect derivative chromosomes der(2) and der(5) in Jurkat cells. Translocation breakpoint junctions are only detected after expression of both TAL^{NPM} and TAL^{ALK}.

D. RT-PCR detection of the *NPM1-ALK* fusion transcript after TAL^{NPM} and TAL^{ALK} expression in Jurkat cells and in ALCL cell line SUP-M2. The forward primer is within exon 2 of *NPM1* and the reverse primer is within exon 29 of *ALK*, amplifying most of the *NPM1-ALK* coding sequence.

E. Single round PCR to detect derivative chromosomes der(2) and der(5) in Jurkat cells. Translocation breakpoint junctions are detected after expression of both TAL^{NPM} and TAL^{ALK} by PCR of the fragment marked (*) in **C** on serial dilutions of genomic DNA (50, 25, 12.5, 6.25, 3.125, and 1.56 ng). The number of times the PCR was positive for each dilution from 6 total experiments is indicated. The markers for 933 and 951 bp correspond to der(2) and der(5) junctions, respectively, without end modification. The larger and smaller fragments seen in some of the lanes likely correspond to junctions with large insertions or deletions.

F. Detection of the NPM1-ALK fusion protein in Jurkat cells after TAL^{NPM} and TAL^{ALK} co-expression and in the ALCL cell line SUPM2. The signal from 1 µg cell extract from SUPM2 cells was compared to 40 µg from Jurkat cells.

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

G. Expanding pools of RPE-1 cells treated with TALENs. The PCR product corresponding to der(5) was detected when cells were split 1 to 10 every 3 days. The number of passages at each time point is indicated below the gel.

Figure 4. Reversion of the t(2;5)(p23;q35) translocation in SUDHL-1 cells.

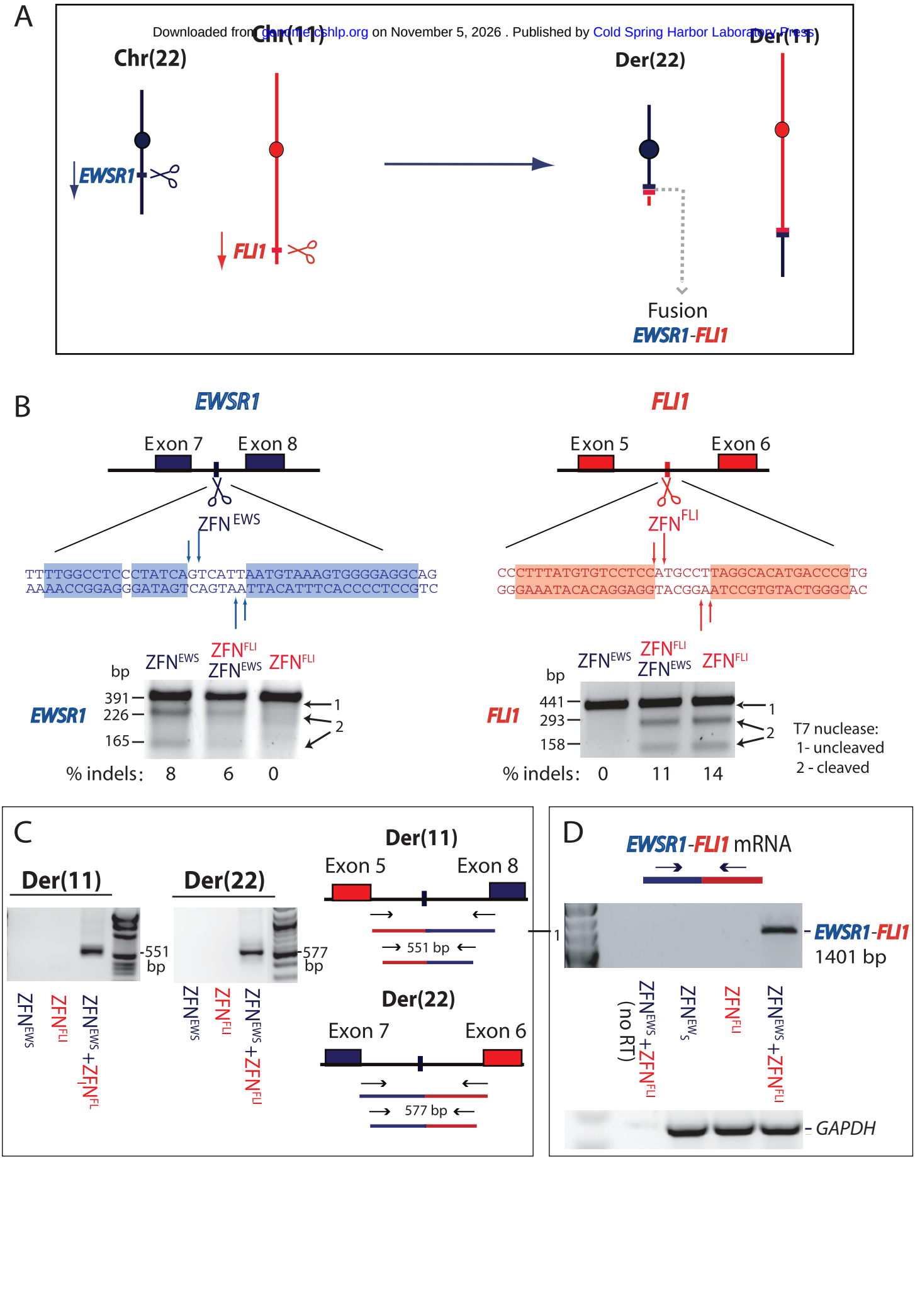
A. Patient-derived SUDHL-1 cells carry the t(2;5)(p23;q35) translocation and express the *NPM1-ALK* fusion gene from der(5). To reverse the translocation, TAL^{NPM} and TAL^{ALK} are expressed to create DSBs in both fusion genes; repair between the derivative chromosomes restores intact chromosomes 5 and 2.

B. TAL^{NPM} and TAL^{ALK} efficiently cleave target loci in SUDHL-1 cells. Cleavage is monitored by the T7-endonuclease assay directed to the *NPM1* (left) and *ALK* (right) loci.

C. PCR detection of revertant chromosomes 2 and 5. Reversion is only detected after expression of both TAL^{NPM} and TAL^{ALK}. Because the TALENs cleave to the side of the translocation breakpoint junctions in the SUDHL-1 cells, segments of the other chromosome remain to “tag” the revertant chromosomes to distinguish them from chromosomes that did not participate in the translocation.

D. Sequences of representative revertant chromosomes 5 and 2 which restore the *NPM1* and *ALK* genes, respectively.

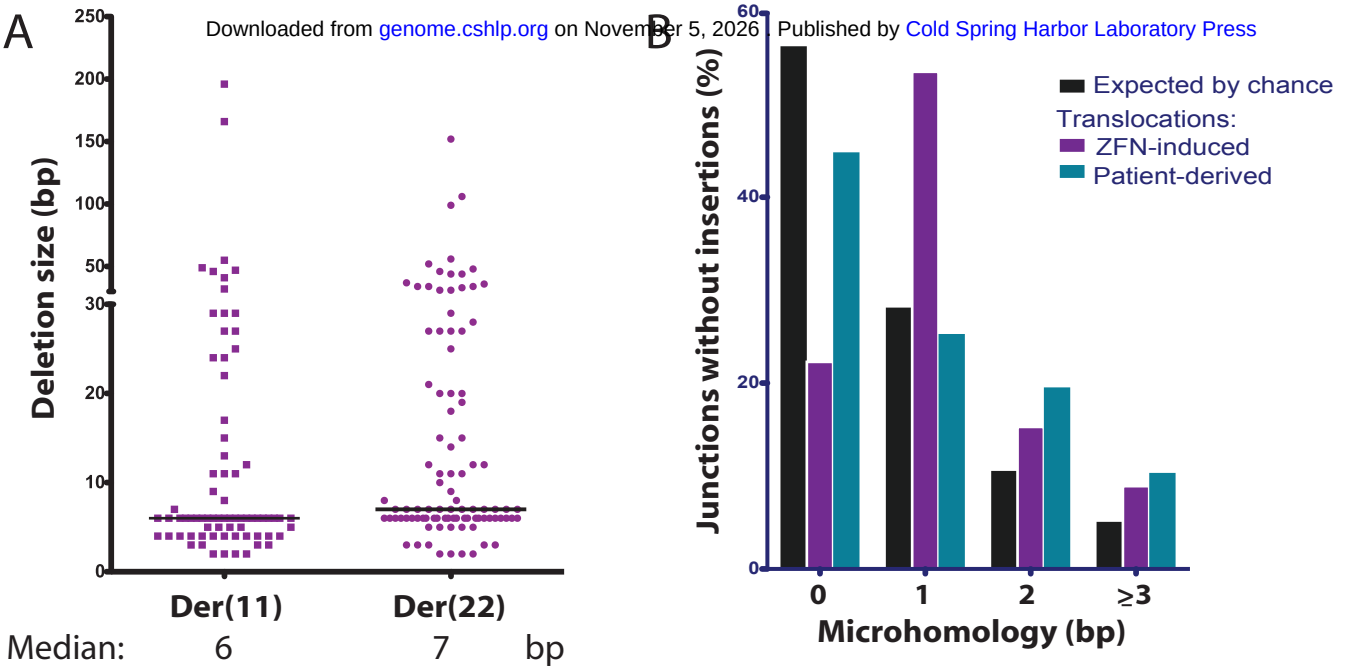
Figure 1



Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Figure 2

Downloaded from genome.cshlp.org on November 5, 2026 Published by Cold Spring Harbor Laboratory Press



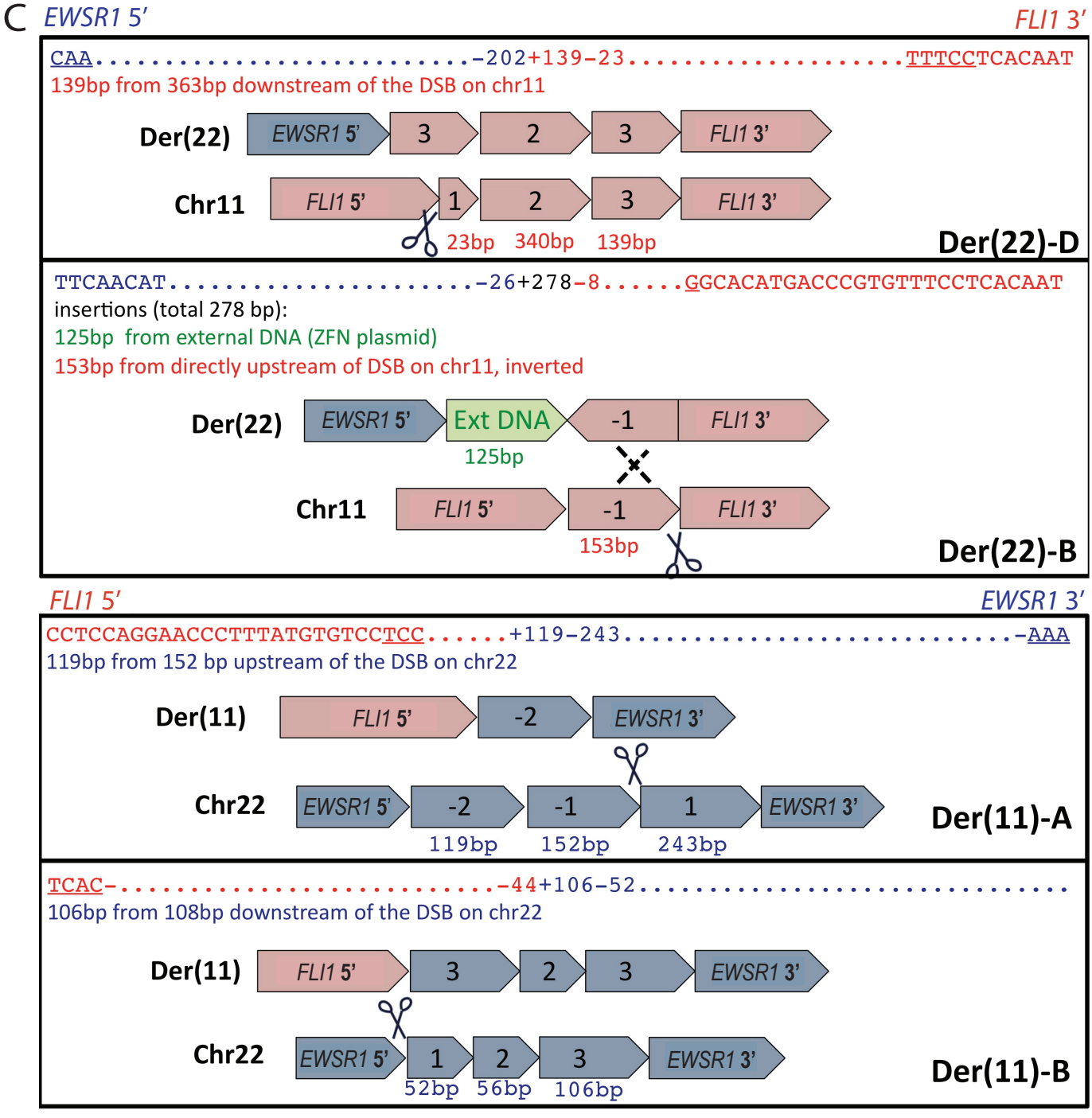
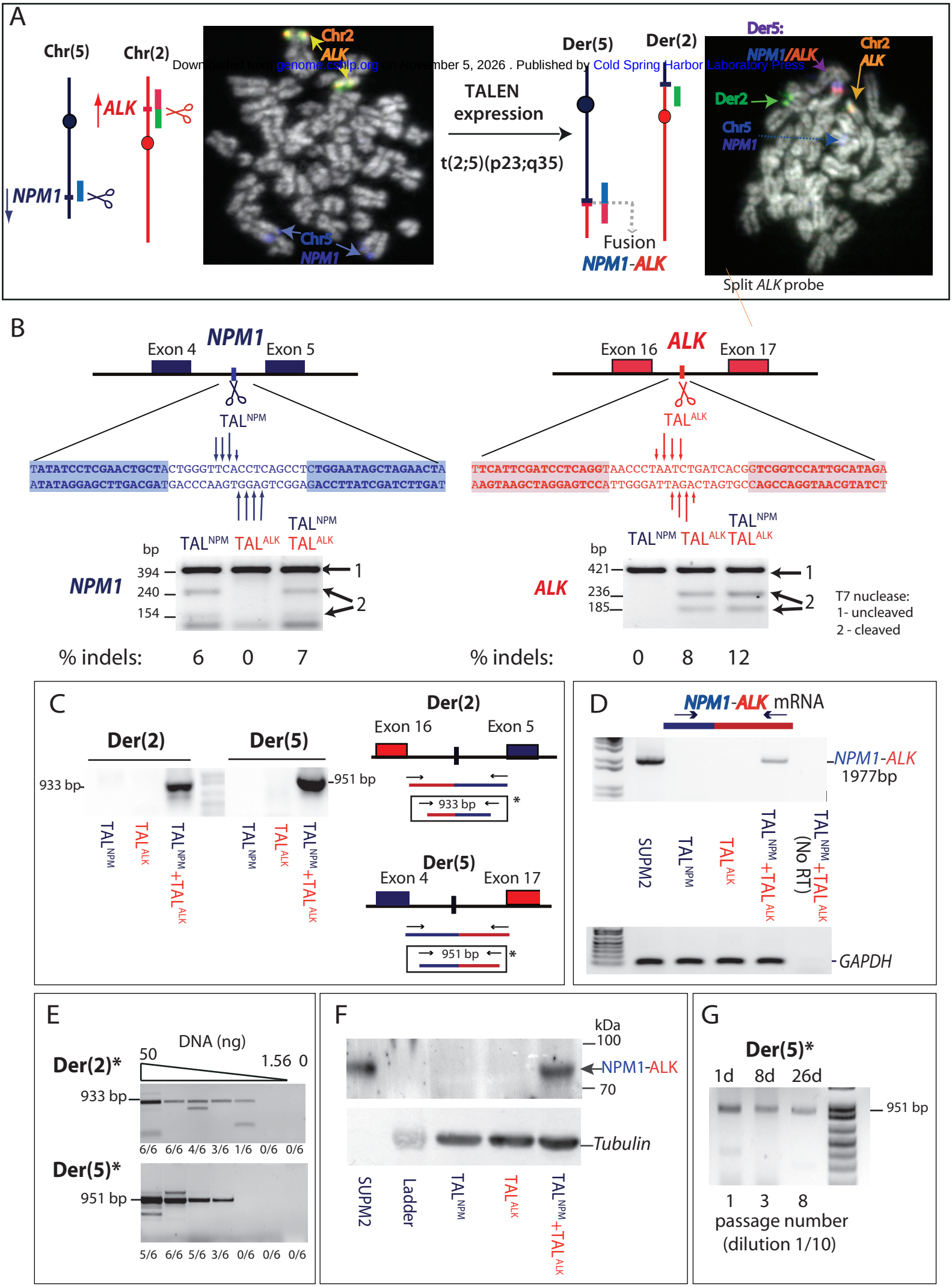


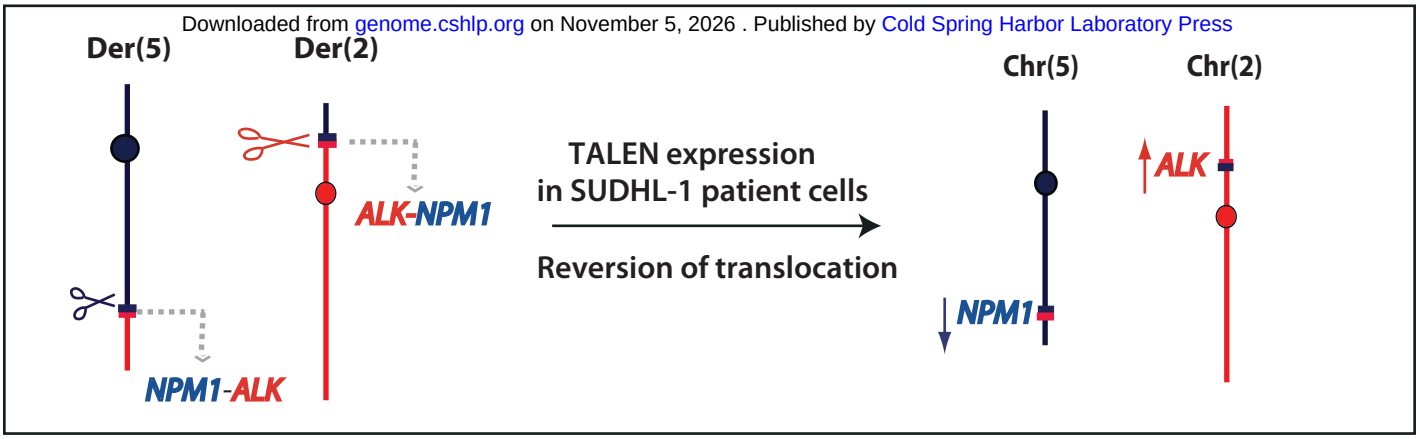
Figure 3



Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

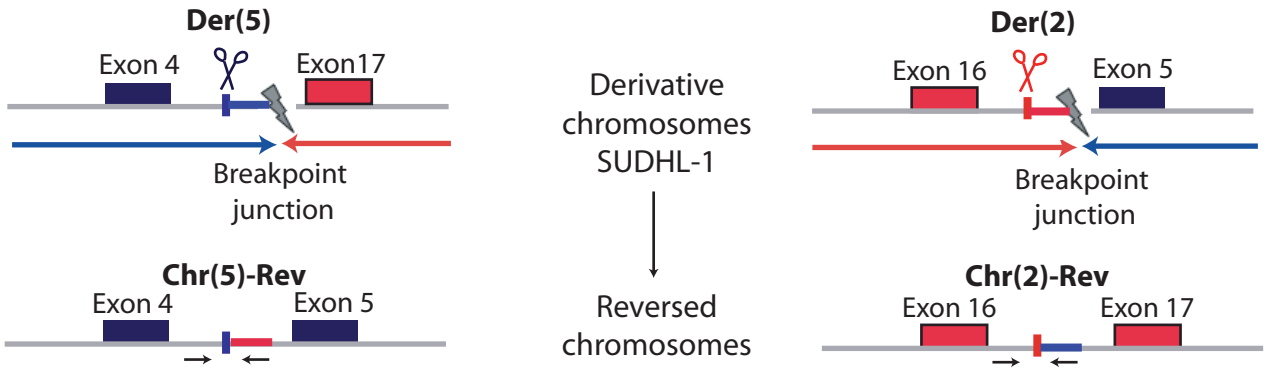
A



B

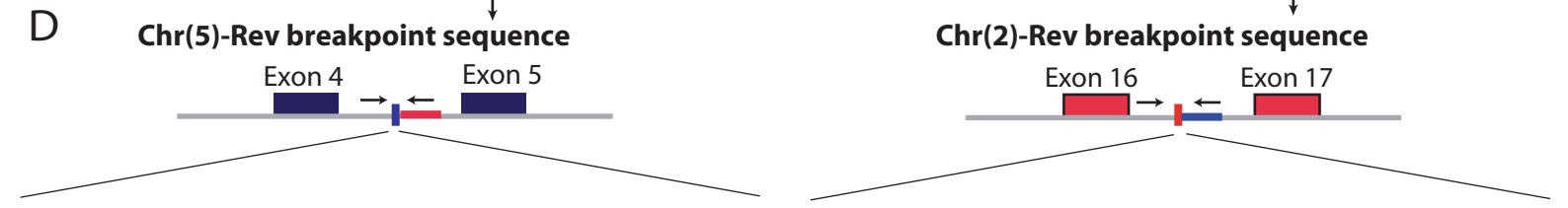


C





D




```
Template:
GCTATATCCTCGAACTGCTACTGGGTTCA(CCT    AA)TCTGATCACGGTCGGTCCATTGCATAGAG
Chr(5)-Rev:
GCTATATCCTCGAACTGCTACTGGGTTCAC..     ..TCTGATCACGGTCGGTCCATTGCATAGAG

Template:
ATTCATTCGATCCTCAGGTAACCCTAAT(Ct      TCA)CCTCAGCCTCTGGAATAGCTAGAAC
Chr(2)-Rev:
ATTCATTCGATCCTCAGGTAACCCTAAT..   G   .....TCAGCCTCTGGAATAGCTAGAAC
```

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Supplemental Information

ZFN and TALEN sequences

ZFN:

The recognition helices for ZFN^{EWS} are as follows:

ZFN	F1	F2	F3	F4	F5	F6
ZFN^{EWS-R}	QSNHRKT	DRSDLSR	RSDNLTR	RSDYLST	QRSHRNT	
ZFN^{EWS-L}	DRSHLSR	QSGHLSR	RSDHLSQ	TSANRTT	QSGALAR	TSSNRKT
ZFN^{FLI-R}	RSDNLSV	QKATRIN	DQSNLRA	QSGHLQR	QSGHLSR	
ZFN^{FLI-L}	RSDTLSE	DRSNLTR	LKQNLDA	QSGSLTR	RSDNLTT	

Each ZFN was designed to have a distinct obligate heterodimeric architecture through modification of the FokI nuclease domain (Doyon et al. 2011). These modifications decrease the potential off-target sites and promote correct heterodimer formation while reducing formation of incorrect heterodimers and homodimers.

We used obligate heterodimeric domains of FokI (Doyon et al. 2011): ELD (ZFN^{EWS-R}) /KKR (ZFN^{EWS-L}) and RDD (ZFN^{FLI-R})/DRR (ZFN^{FLI-L}).

TALEN:

TALENs were assembled by a method derived from Huang et al. (Huang et al. 2011). For each TALEN subunit, the fragment containing the 16 RVD segment was obtained from single unit vectors: A (NI), T (NG), G (NN) and C (HD), kindly provided by the laboratory of Bo Zhang. The assembled TALENs were subcloned in the pCS2 vector for *ALK* and pVax vector for *NPM1*. Sequences of encoded TALEN proteins are as follow (underline: TAL recognition domain; *italic*: wt FokI domain). Where indicated, obligate heterodimeric domains of FokI were used: ELD (TAL^{NPM-L}) /KKR (TAL^{NPM-L}) and RDD (TAL^{ALK-R})/DRR (TAL^{ALK-L}).

$TALEN^{ALK-L}$:

MAPKKKRKVYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAAHGTVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGH
GFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKI
AKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQA
LETVQRLLPVLCQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLP
VLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLT
PAQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALE
TVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVL

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

CQAHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS**NI**GGRPALESIVAQLSRPDPALAALTNDHLVALACL
GGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVA*GSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDR*
ILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHI
NPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGE
INFRS

TAL$^{ALK-R.}$:

MAPKKKRKVYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAAHGTVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGH
GFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKI
AKRGGVTAVEAVHAWRNALTGAPLNLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQA
LETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLP
VLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLT
PAQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIAS
NGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALE
TVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGKQALETVQRLLPVL
CQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPE
QVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACL
GGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVA*GSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDR*
ILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHI
NPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGE
INFRS

TALEN^{NPM-L}:

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPMVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTH
AHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPGTLDTGQLLKIAKRG
GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGGGKQALETV
QRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQ
AHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQV
VAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNNGG
KQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQR
LLPVLCQAHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVA
IASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRP
ALDAVKKGLPHAPALIKRTNRRIPERTSHRVA*GSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEM*
KVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNE
WWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFR
S

TAL^{NPM-R}:

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPMVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTH

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

AHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPGTLDTGQLLKIAKRG
GVTAVEAVHAWRNALTGAPLNLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETV
QRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQ
AHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQV
VAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGG
KQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQR
LLPVLCQAHGLTPDQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAH
GLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVA
IASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRP
ALDAVKKGLPHAPALIKRTNRRIPERTSHRVA*GSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEM*
KVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNE
WWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Supplemental Figures Legends

Supplemental Figure S1

A. Map of *EWSR1-FLI1* patient breakpoint junctions in *EWSR1* intron 7 and in a 1.5 kb region from FLI1 intron 5 from (Zucman-Rossi et al. 1998). *EWSR1* intron 7 is 1.5 kb; *FLI1* intron 5 is 23.3 kb. Locations of the ZFN binding sites are indicated by the boxes and the scissors. Translocations for which breakpoint junctions for both derivative chromosomes have been sequenced but are not adjacent to each other are specified for both der(22) and der(11).
B. Sequences of the segments shown in A. The tumor breakpoints for each junction are indicated by a slash in blue. The ZFN binding sites are highlighted in yellow. Breakpoints falling within the ZFN sites are in red.

Supplemental Figure S2

Der(22) and der(11) translocation junction sequences obtained in hES-MP cells after combined ZFN^{EWS} and ZFN^{FLI} expression. Junctions were obtained from individual wells after ZFN expression from 3 independent experiments; as such, all junctions likely arose from independent events. Junctions are arranged according to the total deletion from the DSBs. The chromosome 22 sequence at the breakpoint is in black and the chromosome 11 sequence is in red, both shown with fill-in of the 5' overhang and with the ZFN binding sequence underlined. The cleavage sites are likely to be variable as to whether they include the terminal nucleotides shown in parenthesis, i.e., T and A for der(22) or T and G for der(11). The number of deleted bases from each end (dots) is indicated. Microhomologies are underlined. Junctions seen repetitively that occur in the ZFN overhangs at 1 bp microhomology are indicated by the asterisks. Inserted sequences are indicated in green. If the deletion extends beyond the chromosome sequence indicated at the top of the figure, a few bp flanking the deletion are indicated, with microhomology underlined. Junctions with long insertions are indicated at the end, along with the inserted sequence and its derivation(s).

Supplemental Figure S3

Breakpoint junction characteristics from patients with Ewing sarcoma tumors (EWTUM) reported in Zucman-Rossi et al. (Zucman-Rossi et al. 1998) (GenBank accession numbers AJ229253 to AJ229365). Deletions and duplications (insertions) from chromosomes 11 and 22 were previously analyzed. Microhomology at the junctions of both der(11) and der(22) as well as additional insertions not from chromosomes 11 and 22 were analyzed for the current report.

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Supplemental Figure S4

Sequences of complex t(11;22)(q24;q12) breakpoint junctions with insertions obtained with ZFN expression in hES-MP cells. Junctions are derived from different sets of experiments and include some junctions from Supplemental Fig. S1 that are also described in Fig. 2. Duplications from chromosome 11 and 22 are shown in red and black; insertions from other sources are green. Microhomologies between the derivative chromosomes and the inserted sequences are underlined. Total length and DNA origins of insertions are also reported.

Supplemental Figure S5

TALEN-induced translocations in RPE-1 cells.

A. TALEN cleavage activity in RPE-1 cells, as monitored by the T7-endonuclease assay. Controls include untransfected cells (wt) and cells expressing only the left part of TAL^{ALK} ($TAL^{ALK\text{-}L}$).

B. Nested PCR to detect derivative chromosomes der(2) and der(5) in RPE-1 cells. Translocation breakpoint junctions are only detected after co-expression of TAL^{NPM} and TAL^{ALK}.

C. Single round PCR to detect derivative chromosomes der(5) in RPE-1 cells. Translocation breakpoint junctions are detected after expression of both TAL^{NPM} and TAL^{ALK} by PCR of the fragment marked (*) in Fig. 3C on serial dilutions of genomic DNA (5, 2.5, 1.25, 0.625, and 0.312 ng). The number of times the PCR was positive for each dilution from 3 total experiments is indicated. The 951 bp marker corresponds to a der(5) junctions without end modification. The smaller fragment seen in the first lane likely correspond to a junction with a large deletion.

D. RT-PCR detection of the *NPM1-ALK* fusion transcript after TAL^{NPM} and TAL^{ALK} co-expression in RPE-1 cells.

E. Western blot detection of the NPM1-ALK fusion protein in RPE-1 cells after TAL^{NPM} and TAL^{ALK} co-expression. The antibody is directed against ALK.

Supplemental Figure S6

In vitro mapping of TALEN cleavage sites. DNA oligonucleotides of 68 bp containing TAL^{NPM} (**A**) and TAL^{ALK} (**B**) binding sites were ^{32}P-labelled at the 5' end on either the top (left) or bottom (right) strand and then annealed to the complementary unlabelled strand. Cleavage reactions were carried out with in vitro translated TALENs. The positions of the spacer sequences for the labeled strands, as deduced from the G+A sequencing, are shown to the side of the gel, as are

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

the positions of the left and right binding sites of the TALENs (blue and red bars, respectively). The major cleavage sites are represented by arrows on the sequences below the gels.

Supplemental Figure S7

Translocation junction sequences for t(2;5)(p23;q35) obtained in Jurkat (**A**) and hES-MP (**B**) cells after combined TAL^{NPM} and TAL^{ALK} expression. Junctions were obtained from individual wells after TALEN expression from 3 independent experiments; as such, all junctions likely arose from independent events. Junctions are arranged according to the total deletion from the DSBs. The chromosome 5 sequence at the breakpoint is in black and the chromosome 2 sequence is in red, both shown with fill-in of the 5' overhang and with the TALEN binding sequence underlined. Cleavage sites are likely to be variable (Supplemental Fig. S6): Nucleotides indicated in italics represent the maximal 5' overhang considering both top strand and bottom strand cleavage, while the terminal nucleotides in parentheses represent the maximal 5' overhang considering just bottom strand cleavage. Deleted bases from each end are indicated by dots. Microhomologies are underlined and inserted sequences not clearly derived from chromosomes 5 or 2 are indicated in green. Junctions with long insertions are at the end, along with the inserted sequence and its derivation(s).

Supplemental Figure S8

Induction of t(2;5)(p23;q35) translocations in Jurkat cells with heterodimeric TALENs.

A. ALCL translocations induced by TALENs with heterodimeric FokI domains.

Top: Nested PCR was performed to detect derivative chromosomes der(2) and der(5) in Jurkat cells. Translocation breakpoint junctions are only detected after expression of both TAL^{NPM} and TAL^{ALK}.

Bottom: Single round PCR detects after heterodimeric TALEN expression detects der(5) in Jurkat cells, as in Fig. 3E, but at an apparent slightly lower frequency than homodimeric TALENs. In serial dilutions of genomic DNA, the der(5) breakpoint junction was obtained in 2 of 3 experiments with 12.5 ng DNA (representing ~2000 cells), implying that translocations arose at a frequency of $\sim 0.5 \times 10^{-3}$.

B. Translocation junction sequences obtained after expression of TAL^{NPM} and TAL^{ALK} containing heterdimeric FokI domains.

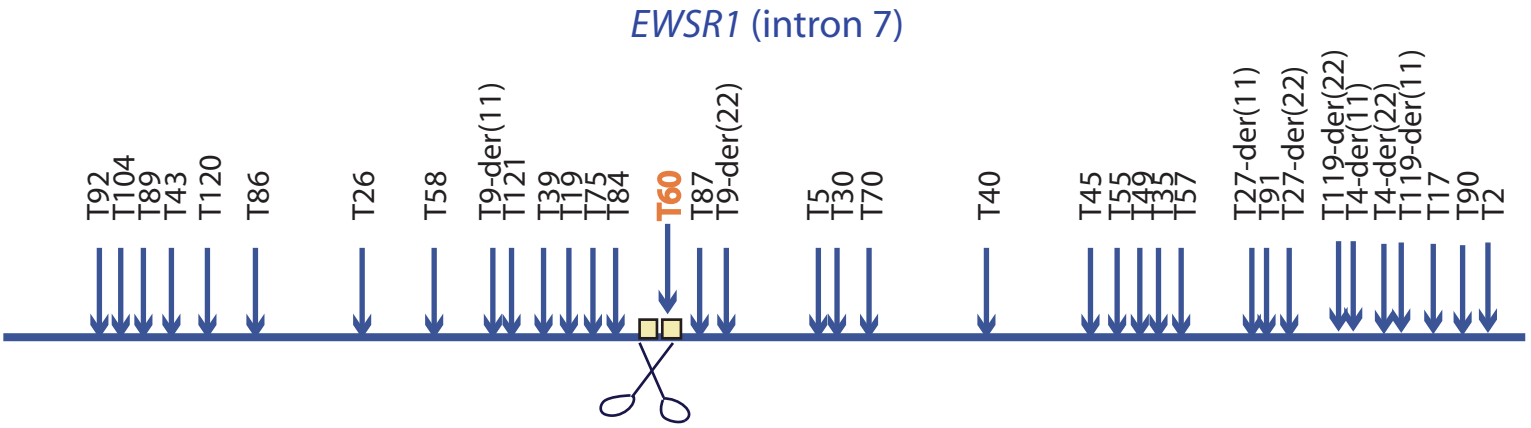
Supplemental Figure S9

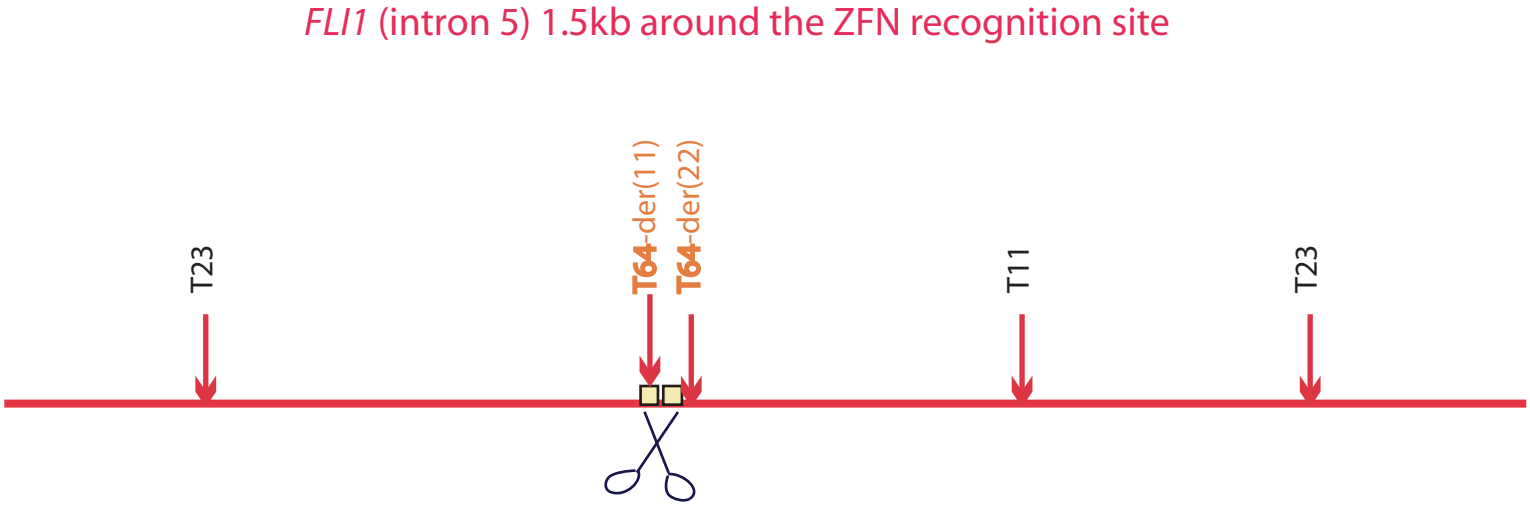
Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Chr(5) and Chr(2) reversed breakpoint sequences obtained from ALCL patient-derived cell lines SUDHL-1 and SUPM2. Junction sequences were obtained after cloning and are presented as in Supplemental Figure S7.

Supplementary Figure 1A

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press





Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Supplementary Figure 1B

EWSR1 **intron 7-8:**

T92; T104 (der(22),der(11)); T89 (der(11),der(22)); T43; T120; T86(der(11),der(22)); T26; T58; T9 (der(11)); T121; T39; T19 (der(22),der(11)); T75(der(11),der(22)); T84; **T60**; T87; T9 (der(22)); T5(der(22),der(11)); T30; T70(der(11),der(22)); T40; T45; T55; T49; T35 (der(11),der(22)); T57; T27(der(11)); T91; T27(der(22)); T119(der(22)); T4(der(11),der(22)); T119(der(11)); T17(der(22),der(11)) T90(der(22),der(11)); T2(der(22),der(11))

GTGAGTTGCTAAGAGAGAAAACCAAATAAGAATGAATGTGTTTAGAGTTTTTTTGTGGGGTTGATTTTTTTTGT
TGGTTTGTTTTTCTTTTGAGATGGAGTTTCACTCTTGGGGCT/G/AGGCTGGAGTACAGTGGTGCCAT/CTCAGC
TCA/CTGC/AGCCTCCGCCTCCTGGTTCAAGCAAT/TCTCCTGCCCCAGCCTCCCGAGTAGCTGAGATGATAGGC
/GCCTGCCACTATGCCCAGCTATTTTTTGTATTTTTCGTAGAGA/CGGGG/TTTCACCATGTTGGTCAGGCTGGT
CTCGAACTCCTGACCTCAGGTGATCCTCCCGCCTCAGCCACCCAGAGTGTTGGGATTACAGTCATGAGCCACTGC
GCCCAGCCACG/TTTGGAGTTTTTGAACAGGGGGAAATACTCTTTGCCATTGTTTGCTTTTGGAATCCAGGACACA
TCTTTAGGGCATGGCATTCCAGCT/AACATCTTGGATGCCCCAGTGAAGATTTCAGTGCTTTTTTCCAATTTGTT
TTTAGTATGCCTTGGCAAAAAAGAAAA/GAGATA/CTAATTGATCAGAAAAAATGTTTTAGACTGCT/AGC/CC/
TGCTGTC/TTTGG/GGAAGTTGT/ATGCAGTGAGTAAATTCAACATCGTTTTTGGCCTCCCTATCAGTCATTAA
/TGTAAAGTGGGGAGGCAGCTATT/GCAGGCCACTATGATTTTGA/TAGTCAAGTAAAAGCTATGTTTTTTTGT
TGCTGTTTGTTTATATCCATTAAGGGGAAAAATGGCCAGGCATGGTGGCTCACACCTGTAATCCCAGCACTTTGG
GGA/GGCCA/AGGCAGGAGGATCAC/TTGAGACCAGGAGTTTGAGACCACCCTGG/GCAACATA/GTGAGACCCT
GTCT/TTTTAAAGATAAATAAAGATTCTTTAAGAAAATAGAAAAGGAAAGGGGGAAAAATAATCCTCCTCACAGA
ATATTTGCAGTTCTTCTGTATGGAGAGAGGTAACCAAAACACATAATTATCTTC/AGAATTTAGGATCATTGCTC
TTTTTAAATTACATTCTATCCACGGGGATTTCTCAACAACTCAAAGGAGTAGTGTTGATTTATGGAGCT/CAGCC
CCTAGCAGTGTGCTAAAGCCCTAAGAA/ACCTGGCTCTGTTTTCATTTCTTGA/GTTTAGCTCAAAAGAAGC/T/
GCTTTTGGAGACATCTTAG/GATATAGAACCCTATAAACTATATGTTAGCCAGAGGGCTTTTTTTGTTGTTGTTA
TTTAGAAGAGGTTTATTAT/A/GGCTAACTATATAAGGTTATTTCAGC/TGTGATGATCCAGCAGTTTTGTTGAA
CTTAAAAGAGCCTACC/TATTA/AGGATGCTTTATCGTGATGT/A/AAGAAGATGGTGCCTTGGTTAGTGCCT/T
/GGAATTG/AGGATT/TTGTGATTCCAGGAGAAAGT/ACATCAGGCAGTGGTG/TAAATGCTGGTCCATGGCTTA
CAGATGTGACTCTTTCCTCAG

T60 A/T breakpoint junction found on der(22)

FLI1 **intron 5-6 (1.5 kb around the ZFN binding site)**

T23 (der(22)); **T64(der(11),der(22))**; T11(der(22),der(11)); T40(der(22))

AACTAGAATAAGAGAATAGGGAGAAAAGCCGATAAGAAGTCAACGATAACTATAGCCATGGCCATAATTTGGGCT
CCAAAAGAAATAATTCTGAGTGGTTACCAGATAGCTTTCAGAGAGGGTTGAAGCCTTCAGTAGACTAATGATTGA
GCAGAAAAAAATGTTTAGAAAACTTGAGACAGATTTTATCCATGCAGTAAGTCAGAAGTCAAGTCAGAAGTTCAG
ATGACAT/GTACGATTCCCTGGAATTTAGTGATTTTTTGTTTGAATGTGTTACTTCTCTTCTACTTTCTGTTCTG
CTCCCACTAGAATTATCTCCTTCCCCCAGGCCCTCAGATCCACCCCACATACCACAAGCTCATTCTTCACCACAC
CCTGTTTTCTAGCATGGCTGCAAGCGCAAAACCTCCCTTAGTTCCCACCTGCAACGCCAAGGTCCCAGAGTTTGC
TGAGCTACCATCAGCAAGACCTTGAACATGCCTCGGCCCTCTCTCCTCTCCCTCCACATTGCTCTACTTCAGTGG
TACAGCTCCTCCCATGGCTGTCTCTAGACCTCACCTTGACATCTCACCACCCACCTCTGCCTTGTTCCTACCTGA
TGCACTCTCCCCTCTCCAAAGAGTCACCTGGCCCTAGTTCCACTCACCTCTCTGTCAACCTCCAGGAACCCTTTA
TGT/GTCCTCCATGCCTTAGGCACATGACCCGTGTTTCCTC/ACAATCACAGTTACCCTCTCATACTCACATGCA
AGGACGAGGAACTTGTTTCCTATTACTTCACACCCACTCCCTGGCCATGCACACAATACAGTGCAGAAGTTGGTC
CCAAATCTATTACCAGAGGCCTAAGCAGCAGCCTCCTCCTGCTCAAAGAAAAATGTTCTGCCGCTGCCAACATTG
CACTTCCTCTGCAAGACATCTCAGAAAATGCCCATCAGAGCCTTCCACCTCATTCCAAGCTTGCAAGCAAGCTGG
TCTTTCATTTGTCTTGTTTGTTTTTAAGCAAGAAGAATCCCTTTAGAGGAGGAATTAGGAAAGAAAAAAAAGTCA
AACAGAAACAGAAGGAGTG/GAG/GTTTGTCTCTCTCTTTTGTCCCCTGCCTTGCCCTTCCTCCCTATCCCATCC
CCCTCACTTTCTGTGTTGCTTTTCATGCCAGTTGCTAAAGACTGTAATTTTAAAACTGAACGGGCTTTTCCAGAG
CCTTGAGGAATGTGCATCTGCACGTGGATTTTTCCACAAAAAAAGTGTCCCTGTCTTTTGGATCTTCTTAGCCAG
AAACAATGAGGCCAGGCAATCAGAGTCCATCCGTCTTCCCAACTC/TTATTCTAATTTGTTTGAGAGGCTCAGAA
GGGGCTGAAGGGCATGGTTGTGTTTCACCCCGTCAGACTCTGTTGTTTGATGAGGTTTGTGCTCCATTTGTT

T64: T/G breakpoint junction found on der(11) and C/A breakpoint junction found on der(22) resulting in ZFN site deletion

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Supplementary Figure 2

A. Der(22)

```
Template:
 AACATCGTTTTTGGCCTCCCTATCAGTCAT(T                A)TGCCTTAGGCACATGACCCGTGTTTCCTCA
23,61,97
 AACATCGTTTTTGGCCTCCCTATCAGTCATT                 ..GCCTTAGGCACATGACCCGTGTTTCCTCA
21,26,86,81,91
 AACATCGTTTTTGGCCTCCCTATCAGTCATT.                -2GCCTTAGGCACATGACCCGTGTTTCCTCA
79,82,50
 AACATCGTTTTTGGCCTCCCTATCAGTCAT.                 -4..CTTAGGCACATGACCCGTGTTTCCTCA
69
 AACATCGTTTTTGGCCTCCCTATCAGTC.-3                 -2GCCTTAGGCACATGACCCGTGTTTCCTCA
58
 AACATCGTTTTTGGCCTCCCTATCAGTCATT                 -5...TTAGGCACATGACCCGTGTTTCCTCA
35,49
 AACATCGTTTTTGGCCTCCCTATCAGTCATT                 -6...TAGGCACATGACCCGTGTTTCCTCA
63
 AACATCGTTTTTGGCCTCCCTATCAGTCA-2                 -4..CTTAGGCACATGACCCGTGTTTCCTCA
**2,8,13,20,23,27,28,30,34,39,45,53,60,62,71,84
 AACATCGTTTTTGGCCTCCCTATCAGTT..-4                -2GCCTTAGGCACATGACCCGTGTTTCCTCA
90
 AACATCGTTTTTGGCCTCCCTATCAGTC.-3      CGT        -3.CCTTAGGCACATGACCCGTGTTTCCTCA
77
 AACATCGTTTTTGGCCTCCCTATCAGTT..-4                -3.CCTTAGGCACATGACCCGTGTTTCCTCA
9,16,31,54,66,74,85
 AACATCGTTTTTGGCCTCCCTATCAGTCATT                 —6...TAGGCACATGACCCGTGTTTCCTCA
52
 AACATCGTTTTTGGCCTCCCTATCAGTC.-3                 -4..CTTAGGCACATGACCCGTGTTTCCTCA
5,22,27,28,44,46,55,75
 AACATCGTTTTTGGCCTCCCTATCAGTCATT                 —6....TAGGCACATGACCCGTGTTTCCTCA
36
 AACATCGTTTTTGGCCTCCCTATCAGTCAT.      G          -6....TAGGCACATGACCCGTGTTTCCTCA
48
 AACATCGTTTTTGGCCTCCCTATC.....-7                 .TGCCTTAGGCACATGACCCGTGTTTCCTCA
76
 AACATCGTTTTTGGCCTCCCTATCAGTC.-3                 -6....TAGGCACATGACCCGTGTTTCCTCA
14
 AACATCGTTTTTGGCCTCCCTATCAGT..-4                 -5...TTAGGCACATGACCCGTGTTTCCTCA
88
 AACATCGTTTTTGGCCTCCCTATCAGTCAT.                 -7.....AGGCACATGACCCGTGTTTCCTCA
18
 AACATCGTTTTTGGCCTCCCTAT......-8                 -2GCCTTAGGCACATGACCCGTGTTTCCTCA
47
 AACATCGTTTTTGGCCTCCCTATCA....-6      CG         -5...TTAGGCACATGACCCGTGTTTCCTCA
65,80
 AACATCGTTTTTGGCCTCCCTATC.....-7                 -5...TTAGGCACATGACCCGTGTTTCCTCA
67
 AACATCGTTTTTGGCCTCCCT.......-10                 .TGCCTTAGGCACATGACCCGTGTTTCCTCA
19
 AACATCGTTTTTGGCCTCCCT.......-10                 -2GCCTTAGGCACATGACCCGTGTTTCCTCA
6,17
 AACATCGTTTTTGGCCTCCCTAT......-8                 —6....TAGGCACATGACCCGTGTTTCCTCA
10
 AACATCGTTTTTGGCCTCCCTATCAG...-5                 -9.......GCACATGACCCGTGTTTCCTCA
78
 AACATCGTTTTTGGCCTCCCTATCA....-6                 -13.........ATGACCCGTGTTTCCTCA
24
 AACATCGTTTTTGGCCTCCCTA.......-9       A         -10.......CACATGACCCGTGTTTCCTCA
59
 AACATCGTTTTTGGCCTCCCTATCA....-6                 -14..........TGACCCGTGTTTCCTCA
40,56
 AACATCGTTTTTGGCCT...........-14                 -6....TAGGCACATGACCCGTGTTTCCTCA
1
 AACATCGTTTTTGGCCTCCCTATCAG...-5                 -16............ACCCGTGTTTCCTCA
26
 AACATCGTTTTTGGCCTCCCTATC.....-7                 -18..............CCGTGTTTCCTCA
```

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

```
83
 AACATCGTTTTTGGCCT...........-14                -13.........ATGACCCGTGTTTCCTCA
25,51
 AACATCGTTTTTGGC.............-16                -11........ACATGACCCGTGTTTCCTCA
29
 AACATCGTTTT.................-20                -7.....AGGCACATGACCCGTGTTTCCTCA
70
 AACATCGTTTTTGGCCTCCCTAT.....-8                 -20.................GTGTTTCCTCA
25
 AACATCGTTT..................-21                -8......GGCACATGACCCGTGTTTCCTCA
33,37
 AACATCGTTTTTGGCCTCCC........-11                -20.................GTGTTTCCTCA
96
 AACATCGTTTTTGGCC............-15                -18...............CCGTGTTTCCTCA
32,68
 AACATCGTTTTTGGCC............-15                -19................CGTGTTTCCTCA
73
 ............................-35                ATGCCTTAGGCACATGACCCGTGTTTCCTCA
24
 AACATCGTTTTTGGCCTCCCTATCA....-6                -31............................
93
 AACATC......................-25                -18...............CCGTGTTTCCTCA
84
 AACAT.......................-26                -18...............CCGTGTTTCCTCA
87
 ............................-31                -15............GACCCGTGTTTCCTCA
92
 ............................-33       GGG      -13..........ATGACCCGTGTTTCCTCA
57
 AACATCGT....................-23                -29.........................CA
94
 AACAT.......................-26                -30..........................A
7
 AACATCG.....................-24                -75............................
29
 CT..........................-61                -45............................
64
 A...........................-93                -59............................

11 Der(22)-B
 AACAT.......................-26      +278      -8......GGCACATGACCCGTGTTTCCTCA
125 bp from ZFN plasmid
TTCAACATTTTTAGTATGCCTTGGCAAAAAATGTAAAAACAAATTGGAAAAAAGCACTGTGTTTTACGGTTCCTGGCCTTTTG
CTGGCCTTTTGCTCACATGTTCTTGCTGCTTCGCGATGTACG
153 bp (insert bp 126-278) from other site of the DSB on chr11 inverted
GGCATGGAGGACACATAAAGGGTTCCTGGAGGTTGACAGAGAGGTGAGTGGAACTAGGGCCAGGTGACTCTTTGGAGAGGGGA
GAGTGCATCAGGTAGGAACAAGGCAGAGGTGGGTGGTGAGATGTCAAGGTGAGGTCTAGAGACAGCCATGG

38
 AACATCGTTTTTGGCCTCCCTATCAGTCAT.      +564       .TGCCTTAGGCACATGACCCGTGTTTCCTCA
564 bp from ZFN plasmid
GCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATGTACTGCCAAAACCGCATCACCAT
GGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGG
GCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGT
AAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCG
GGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCGGAACTCCATATATGGGCTATGAACTAATG
ACCCCGTAATTGATTACTATTAATAACTAGTCAATAATCAATGTCAACGCGTATATCTGGCCCGTACATCGCGAAGCAGCAAG
AACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC

89 Der(22)-C
 AACATCGTTTTTGGCCTCCCTATCAGT..-4      +994       -25.....................TCCTCA
136 bp from Homo sapiens genomic DNA, chromosome 8q23, clone: KB1517D11
TTGACAAGAGCAATTTAGGTTCATAAACTCTTTTGTTGCAAAAGATCCTCCTTTTTTGCTCCAGATAATGACAAAAATTATTAC
AGTAACATGAAATAGTGACTTATCTTATGGACAAAATTTCTTGGAGGGCAGGGC
213 bp (insert bp 137-350) from Homo sapiens transmembrane domain transport protein
(SLC36A1) chromosome 5q33.1
GGCAGAGGGGTTCCCACACAGCAACCCTGACTTTCCTCAGAGCCTGGACCATACCTGAGGCTCTGAGGACAGGTAACGGGGGA
CACATGGGGCAGGGGTAGGCAGCTGCAGAGACGAACCCAGATCCCTATATGAAGGCACAGGTGGAATTGATGAAGATGGGAGC
ATTGCTTGGCTGGATCAGCTCATAGAGAGCCTCATAGGTCCCCATCACA-G
228 bp (insert bp 352-580) from Homo sapiens chromosome 5 clone CTB-99A3
```

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

TTACGTAAATGGAATCATAAGGTATGCATTCCTTTTTTTGGTGACTAACTTCTTTCAACATAATGCATTTGATAGTTATCCAA
GTTGTTGCATGTAGTAgTTCTTTGTTTTTTAAATTGCTGAGAGGTATTATATTATATAGGTATacCACAGTTCATTCATTTTC
TTGCTAATTGATATTTGGATTATTTCCGTCTTTGGCTATTGTGAACAATGATACTATAGACA

412 bp (insert bp 581-994) Homo sapiens chromosome 18, clone CTD-2544E12

ACACCCACCTTGAAGGGATGTTGCAAGCATCAGGAAAAACGAGTGTTCAATGCCTGTCTTAGTAGGTCCTAAACAATGATTGA
TCCCGGAGTAGCTAGCCCAGAGCTAAGCCCAGAGTAAGTGCTCAATAAATATTTGTGAATGGGACTCTCTGAAGCCCAGTCCC
AAGGTGAGCTGGATTTCCAGGGTTTGGCAATGAAAATCCTTGGATTGTGTAACCTGAGAGCTGAATATGTGAGTGAGGGGGTT
CTTAcAGCAGAGCACCAGCACAGGCGGCCAAGGAACCCAGGCACCCAGCCTTCTAACCCCTTCCTCTGCCTCAGGGAGAGGCG
TGCATGCATCCCTGAGCCGGTCAGTGGCCACAAGGGACTcaGCCCcagGGGACAGCAGGTGTCAGATCAGCCTCCACAGACCC
TCC

95 **Der(22)-D**

.....-202 +139 -23.....TTTCCTCA

139 bp from 363 bp downstream of the DSB on chr11

CAAACAGAAACAGAAGGAGTGGAGGTTTGTCTCTCTCTTTTGTCCCCTGCCTTGCCCTTCCTCCCTATCCCATCCCCCTCACT
TTCTGTGTTGCTTTTCATGCCAGTTGCTAAAGACTGTAATTTTAAAACTGAACGGGCTTTTCC

B. Der(11)

Template:

CAGGAACCCTTTATGTGTCCTCCATGCC(T G)TCATTAATGTAAAGTGGGGAGGCAGCTA

42
CAGGAACC**CTTTATGTGTCCTCC**ATGCCT .TCATT**AATGTAAAGTGGGGAGGC**AGCTA
9,19,20,37,48
CAGGAACC**CTTTATGTGTCCTCC**ATGCCT ..CATT**AATGTAAAGTGGGGAGGC**AGCTA
4,32,76
CAGGAACC**CTTTATGTGTCCTCC**ATGCC. -2CATT**AATGTAAAGTGGGGAGGC**AGCTA
29
CAGGAACC**CTTTATGTGTCCTCC**ATGC-2 .TCATT**AATGTAAAGTGGGGAGGC**AGCTA
**3,14,16,21,39,41,46,49,52,55,66
CAGGAACC**CTTTATGTGTCCTCC**ATGCC. -3.ATT**AATGTAAAGTGGGGAGGC**AGCTA
7,54,71
CAGGAACC**CTTTATGTGTCCTCC**ATGC-2 -3.ATT**AATGTAAAGTGGGGAGGC**AGCTA
27,44
CAGGAACC**CTTTATGTGTCCTCC**ATG.-3 -2CATT**AATGTAAAGTGGGGAGGC**AGCTA
**5,8,17,24,30,33,36,38,45,50,63
CAGGAACC**CTTTATGTGTCCTCC**AT..-4 -2CATT**AATGTAAAGTGGGGAGGC**AGCTA
23,26,28,34,51,57,61,73
CAGGAACC**CTTTATGTGTCCTCC**ATGCCT -6...**.AATGTAAAGTGGGGAGGC**AGCTA
67
CAGGAACC**CTTTATGTGTCCTCC**ATGCC. AG -5...T**AATGTAAAGTGGGGAGGC**AGCTA
65
CAGGAACC**CTTTATGTGTCCTCC**ATG.-3 -3.ATT**AATGTAAAGTGGGGAGGC**AGCTA
6
CAGGAACC**CTTTATGTGTCCTCC**ATGCCT -5...**.AATGTAAAGTGGGGAGGC**AGCTA
13
CAGGAACC**CTTTATGTGTCCTCC**AT..-4 -4..TT**AATGTAAAGTGGGGAGGC**AGCTA
74,69,25
CAGGAACC**CTTTATGTGTCCTCC**AT..-4 -5...T**AATGTAAAGTGGGGAGGC**AGCTA
22
CAGGAACC**CTTTATGTGTC**.....-10 .TCATT**AATGTAAAGTGGGGAGGC**AGCTA
47
CAGGAACC**CTTTATGTGTCCTCC**....-6 CTATCACTTGTTTCC -5...T**AATGTAAAGTGGGGAGGC**AGCTA
15
CAGGAACC**CTTTATGTGTCCTCC**A...-5 -6....**AATGTAAAGTGGGGAGGC**AGCTA
72
CAGGAACC**CTTTATGTGTCCTCC**ATGCC. -12.....**AAGTGGGGAGGC**AGCTA
2
CAGGAACC**CTTTATGTGTCCTC**.....-7 G -8.....**TGTAAAGTGGGGAGGC**AGCTA
59
CAGGAACC**CTTTATGTGTCCTCC**A...-5 GGTG -12.....**AAGTGGGGAGGC**AGCTA
18
CAGGAACC**CTTTATGTGT**.....-11 -11.....**AAAGTGGGGAGGC**AGCTA
53
CAGGAACC**CTTTATGT**.....-13 -11.....**AAAGTGGGGAGGC**AGCTA

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

```
58
 CAGGAACCCTTTATGTG.........-12          GT        -12.........AAGTGGGGAGGCAGCTA
35
 CAGGAACCCTTTATGTGTCCTCCATGCC.                    -24.....................AGCTA
10
 CAGGAACCCTTTATGTG.........-12        AGGT        -15............TGGGGAGGCAGCTA
68
 CAGGAACCCTTTATGTGTCC........-9                   -18..............GGAGGCAGCTA
1,31,62
 CAGGAACCCTTTATGTG.........-12                    -17.............GGGAGGCAGCTA
60
 CAGGAACC...................-21                   -11........AAAGTGGGGAGGCAGCTA
64
 CAGGAACCC..................-20          C         -21..................GGCAGCTA
70
 CAGGAAC....................-22                   -24.....................AGCTA
11
 CAGGAACCCT.................-19                   -28........................A
56
 CAGGAAC....................-22                   -27.......................TA
12
 C..........................-28                   -27.......................TA
43
 T..........................-53                   -49.........................
40
 CAGGAACCCTTTATGTGTCCTCCA...-5                    -161........................
77
 ..........................-195                   .TCATTAATGTAAAGTGGGGAGGCAGCTA

75 Der(11)-E
 CAGGAACCCTTTATGTGTCCTCCATG.-3         +332       -8......TGTAAAGTGGGGAGGCAGCTA
332 bp from ZFN plasmid
GGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATGTACTGCCAAAACCGCATCACCATGGTAAT
AGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATT
TACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATAC
TCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAGTCAATAGGGGGT
TG
```

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Supplementary Figure 3

Tumor	Deletions(-) /Duplications (bp)		Insertions (bp)		Microhomology (bp)	
	ch22	ch11	der22	der11	der22	der11
t2	0	-25			1	1
t3	4	-3			2	0
t4	-60	-3	19	9		
t5	-5	-1			0	0
t6					0	
t7	-19	-56			0	1
t9	157	207	3			
t10					0	
t11	4	-2			0	0
t12					0	
t13					2	
t15					0	
t17	48	-21			0	0
t18					0	
t19	3	-16			5	2
t20	-2	-45			0	0
t21	49	-7			1	1
t23	467	-8262		4		
t24	-2	2			0	2
t26					2	
t27	27	24			0	1
t30					1	
t34	-51	-13	24			2
t35	1	-3			1	2
t37	-191	-43		1	4	
t39			2			
t40			1			
t41	3	-1			0	3
t43					2	
t44	255	-27			0	8
t45					2	
t47	-192	-6		2	0	
t49					3	
t50					0	
t51					0	
t53					1	
t54					0	
t55					4	
t56					0	
t57			48			
t58					0	
t59					0	
t60			1			
t61	103	-14	2			0
t62					2	
t63					1	
t64	-9	-38	34			1
t66					0	

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

t67					1	
t68					1	
t70	22	78	1			0
t71					2	
t72	-40	-170			1	1
t73	4	-5			6	1
t75	4	0			1	1
t81	-29	-25			0	2
t83	2	-4			3	0
t84					0	
t86	5	-5			1	1
t87			1			
t88					0	
t89	4	-3		140	2	
t90	-10	-4				0
t91					0	
t92			2			
t95	42	98			0	0
t96					0	
t104	-29	-24		8	1	
t106					3	
t110			1			
t112	42	-33	10	6		
t115					0	
t117	-129	1			1	2
t119	-25	13	2	5		
t120					2	

Supplementary Figure 4

A. Der(22) insertions

A

GGGAAGTTGTATGCAGTGAGTAAATTCAACATCGTT**TTTGGCCTC**CC**T**.....-10+409..CT**TAGGCACATGACCCG**TGTTTCCTCACAATCACAGTTACCCTCTCATACTC

409 bp from ZFN plasmid

TGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGA
CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGG
TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCG
TTTAAACTTAAGCTGATCCACTAGTCCAGTGTGGTGGAATTCGTCATGGTC

B

GGGAAGTTGTATGCAGTGAGTAAATTCAACAT.....-26+125/+153-8.....G**GCACATGACCCG**TGTTTCCTCACAATCACAGTTACCCTCTCATAC

125 bp from ZFN plasmid

TTCAACATTTTTAGTATGCCTTGGCAAAAAATGTAAAAACAAATTGGAAAAAAGCACTGTGTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTGCTGCTTCGCGAT
GTACG

153 bp (insert bp 126-278) from other site of the DSB on chr11 inverted

GGCATGGAGGACACATAAAGGGTTCCTGGAGGTTGACAGAGAGGTGAGTGGAACTAGGGCCAGGTGACTCTTTGGAGAGGGGAGAGTGCATCAGGTAGGAACAAGGCAGAGGTGGGTGGT
GAGATGTCAAGGTGAGGTCTAGAGACAGCCATGG

C

GGGAAGTTGTATGCAGTGAGTAAATTCAACATCGTTTTTGGCCTCCCTATCAGT..-4+994-25.....TCCTCACAATCACAGTTACCCTCTCATACTC

136 bp from Homo sapiens genomic DNA, chromosome 8q23, clone: KB1517D11

TTGACAAGAGCAATTTAGGTTCATAAACTCTTTTGTTGCAAAAGATCCTCCTTTTTGCTCCAGATAATGACAAAAATTATTACAGTAACATGAAATAGTGACTTATCTTATGGACAAAAT
TTCTTGGAGGGCAGGGC

213 bp (insert bp 137-350) from Homo sapiens transmembrane domain transport protein (SLC36A1) chromsome 5q33.1

GGCAGAGGGGTTCCCACACAGCAACCCTGACTTTCCTCAGAGCCTGGACCATACCTGAGGCTCTGAGGACAGGTAACGGGGGACACATGGGGCAGGGGTAGGCAGCTGCAGAGACGAACC
CAGATCCCTATATGAAGGCACAGGTGGAATTGATGAAGATGGGAGCATTGCTTGGCTGGATCAGCTCATAGAGAGCCTCATAGGTCCCCATCACA-G

228 bp (insert bp 352-580) from Homo sapiens chromosome 5 clone CTB-99A3

TTACGTAAATGGAATCATAAGGTATGCATTCCTTTTTTTGGTGACTAACTTCTTTCAACATAATGCATTTGATAGTTATCCAAGTTGTTGCATGTAGTAgTTCTTTGTTTTTTAAATTGC
TGAGAGGTATTATATTATATAGGTATacCACAGTTCATTCATTTTCTTGcTAATtGATAtTTGgATTATTTCCGtCTTTggCTATTGTGAACAaTGAtACTATAGACA

412 bp (insert bp 581-994) Homo sapiens chromosome 18, clone CTD-2544E12

CCCACCTTGAAGGGATGTTGCAAGCATCAGGAAAAACGAGTGTTCAATGCCTGTCTTAGTAGGTCCTAAACAATGATTGATCCCGGAGTAGCTAGCCCAGAGCTAAGCCCAGAGTAAGTG
CTCAATAAATATTTGTGAATGGGACTCTCTGAAGCCCAGTCCCAAGGTGAGCTGGATTTCCAGGGTTTGGCAATGAAAATCCTTGGATTGTGTAACCTGAGAGCTGAATATG
TGAGTGAGGGGGTTCTTAcAGCAGAGCACCAGCACAGGCGGCCAAGGAACCCAGGCACCCAGCCTTCTAACCCCTTCCTCTGCCTCAGGGAGAGGCGTGCATGCATCCCTGAGCCGGTCA
GTGGCCACAAGGGACTcaGCCCcagGGGACAGCAGGTGTCAGATCAGCCTCCACAGACCCTCC

D

.....-202+139-23.....TTTCCTCACAATCACAGTTACCCTCTCATACTC

139 bp from 363 bp downstream of the DSB on chr11

CAAACAGAAACAGAAGGAGTGGAGGTTTGTCTCTCTCTTTTGTCCCCTGCCTTGCCCTTCCTCCCTATCCCATCCCCCTCACTTTCTGTGTTGCTTTTCATGCCAGTTGCTAAAGAC
TGTAATTTTAAAACTGAACGGGCTTTTCC

B. Der(11) insertions

A

TAGTTCCACTCACCTCTCTGTCAACCTCCAGGAACC**CTTTATGTGTCCTCC**....-6+119-243.....

119 bp from 152 bp upstream of the DSB on chr22

TCCCCTTTTCCCTATCTGGTGTCTGTACTTTGATGAAGGTGAGCTATAATATTGGTGGTTAACATAATCCAGAAGGCTTAGTTCTGTGTGTGTTCTGTCCCTATATAAACTGGTAAAG
AAAAAAA

B

TAGTTCCACTCAC.....-44+106-52.....GTCAAGT

106 bp from 108 bp downstream of the DSB on chr22

TCACTTGACAGCAAGAGTTTGAGACCCCCCTGGGTGACCTACAGAGACCCTGTCTTTTTAAAGAAGAATAANAATGCTTTANAAAAATAGAAAAGGAGGGAAGGAAAAAA

C

TAGTTCCACTCACCTCTCTGTCAACCTCCAGGAACC**CTTTATGTGT**.....-11+83-36.....CACTATGATTTTGATAGTCAAGT

83 bp from 29 bp downstream of the DSB on chr11

GTGTTTCCTCACAATCACAGTTACCCTCTCATACTCACATGCAAGGACGAGGAACTTGTTTCCTATTACTTCACACCCACTCCCTGG

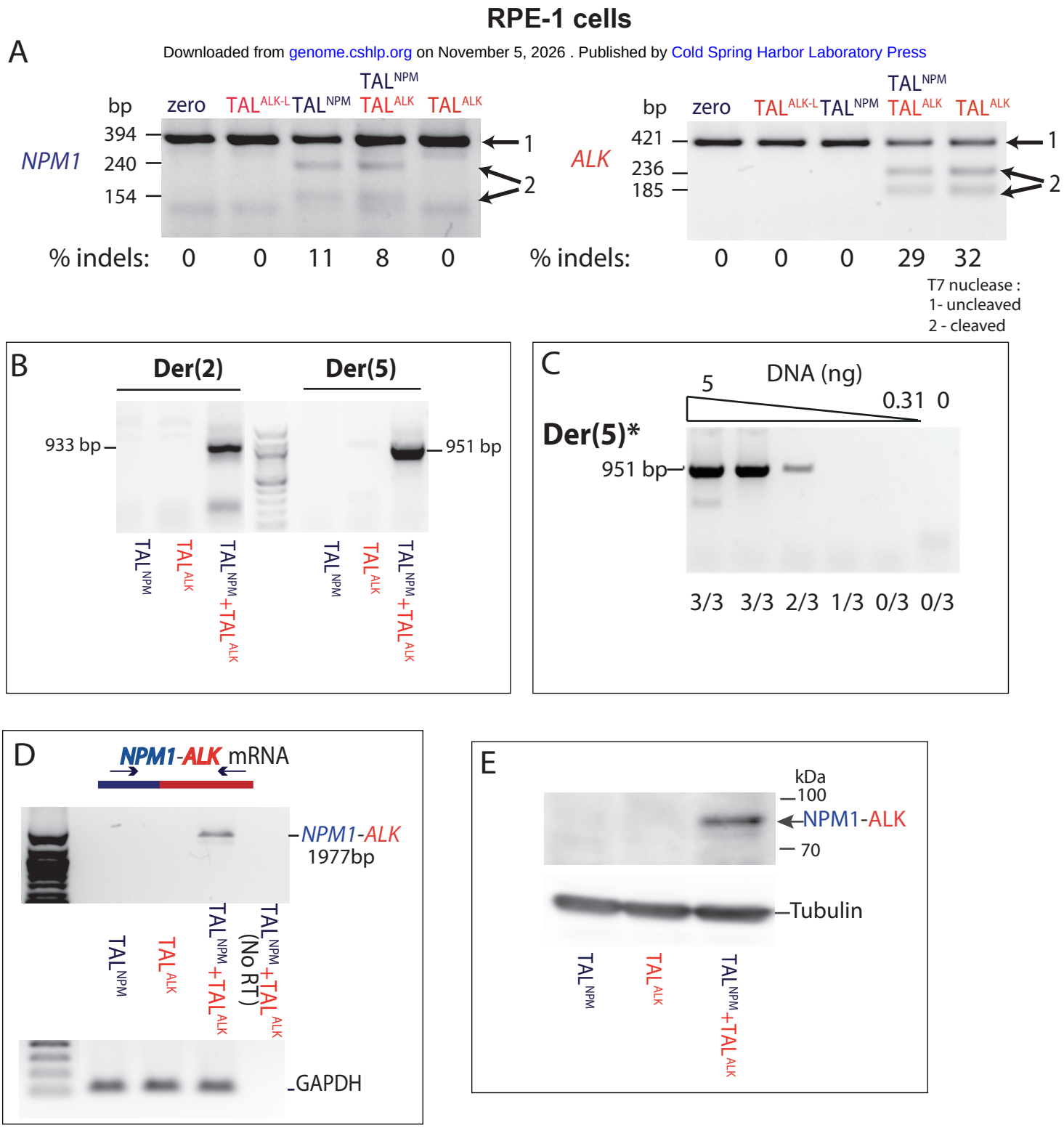
D

TAGTTCCACTCACCTCTCTGTCAACCTCCAGGAACCCTTTATGTGTCCTCCATGCC.+180....TAATGTAAAGTGGGGAGGCAGCTATTGCAGGCCACTATGATTTTGATAGTCA

180 bp from Homo sapiens nuclear receptor binding SET domain protein 1 (NSD1) on chr5

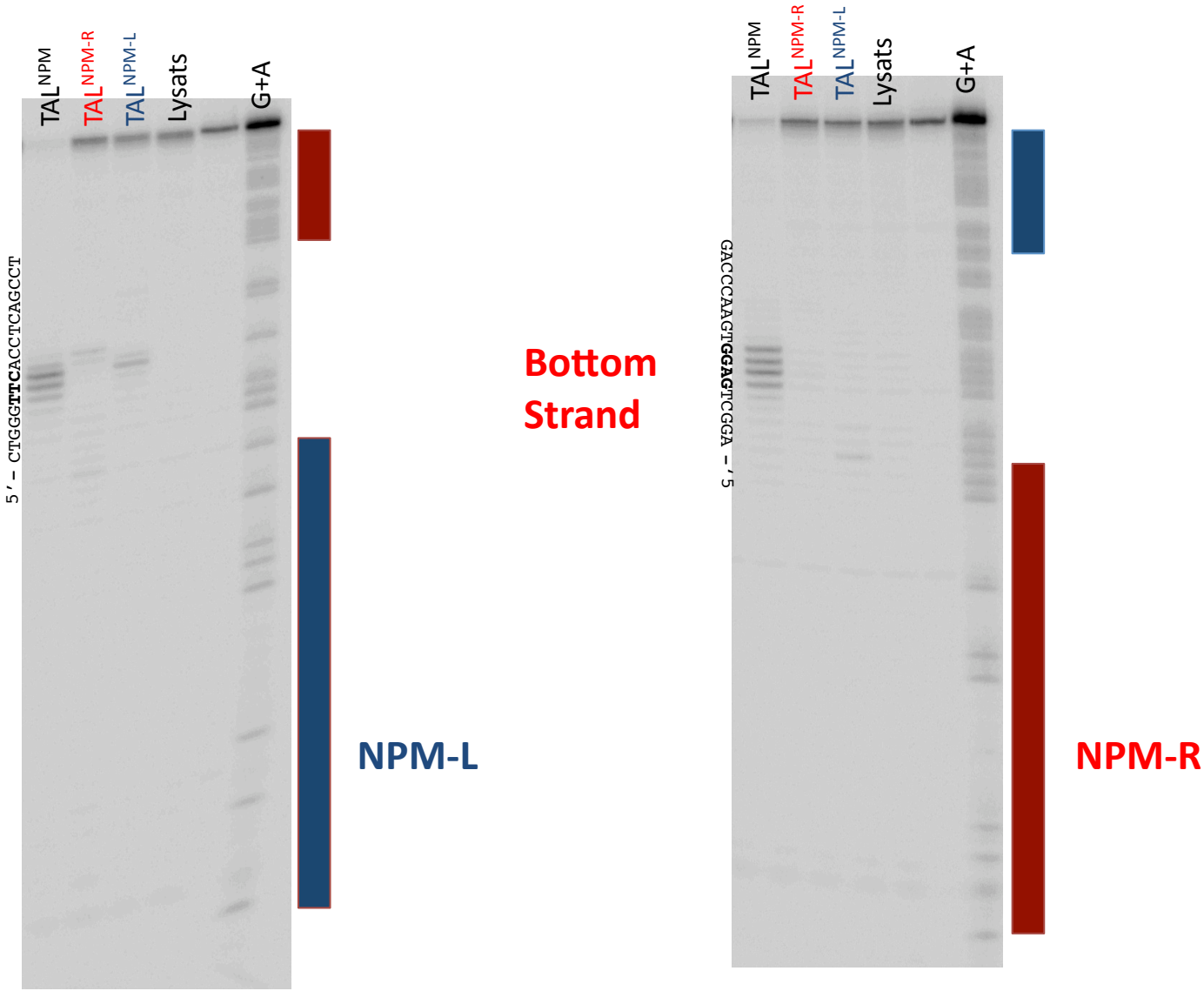
TGGTTTGAGATGGAGTCTGGCTCTGTTGCCCAGGCTGGAGTACAGTGGCACAATCTCGGTTCACTGCAGCCTCCACCTCCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGTAGC
TGGGACTACAGATGTACTCCCTCACGCCTGGCTAATTTGTTTTGTATTTTTAGTAGAGAT

Supplementary Figure 5



Supplementary Figure 6

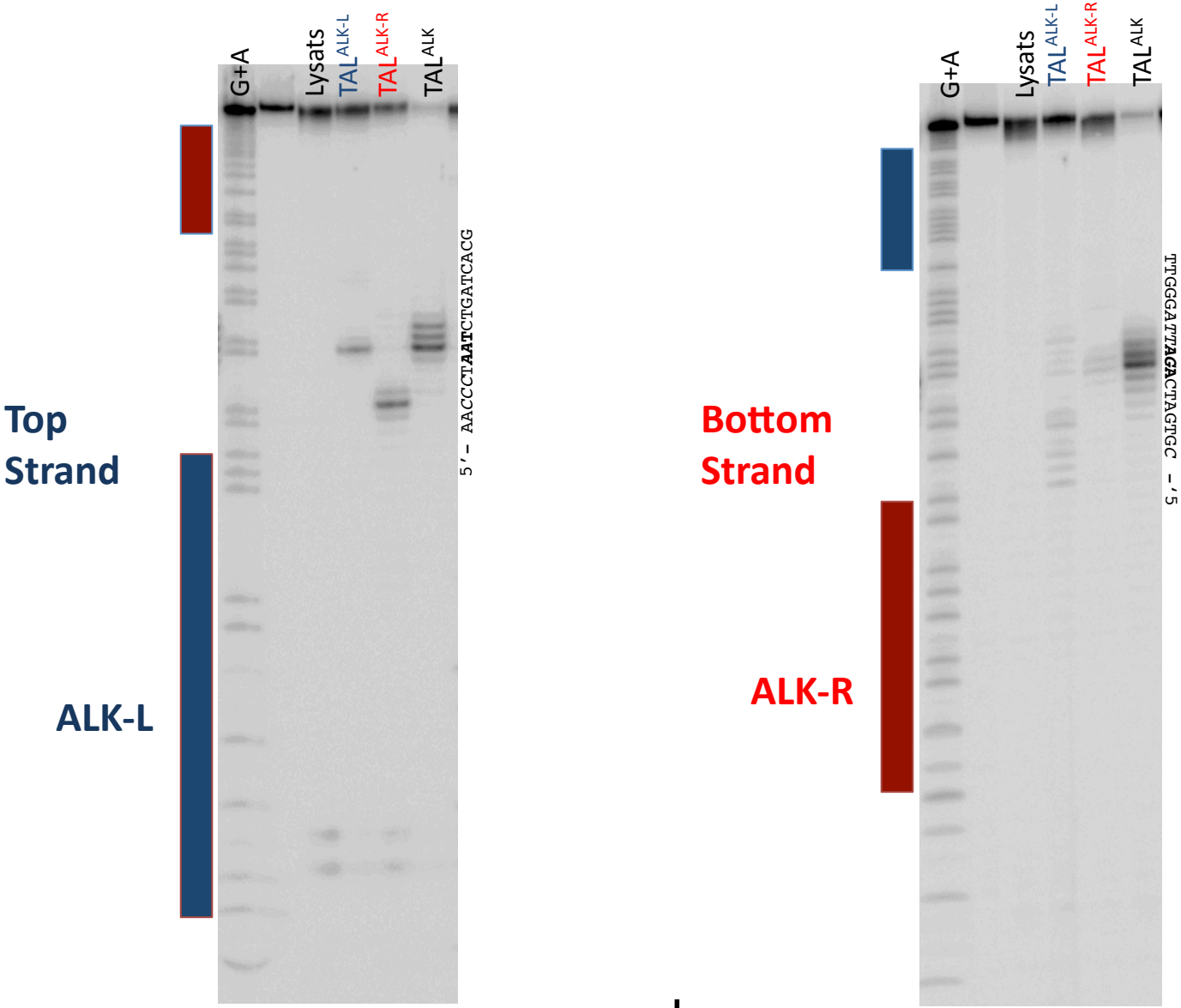
A



NPM-L AGCTTGCTATATCCTCGAACTGCTACTGGGTTCACCTCAGCCTCTGGAATAGCTAGAACTACAGGCAC

TCGAACGATATAGGAGCTTGACGATGACCCAAGTGGAGTCGGAGACCTTATCGATCTTGATGTCCGTG NPM-R

B. ALK



ALK-L ATTTCAA**TTCATTCGATCCTCAGGT**AACCCT**AAT**CTGATCACGGTCGGTCCATTGCATAGAGGAGGGA

TAAAGTTAAGTAAGCTAGGAGTCCATTGGGATTA**GA**CTAGTGC**CAGCCAGGTAACGTATCT**CCTCCCT **ALK-R**

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Supplementary Figure 7

A. Jurkat cells

Der(5)

Template: GCTATATCCTCGAACTGCTACTGGGTTCA(CCT		AA)TCTGATCACGGTCGGTCCATTGCATAGAG
93 GCTATATCCTCGAACTGCTACTGGGTTCA...		.ATCTGATCACGGTCGGTCCATTGCATAGAG
67 GCTATATCCTCGAACTGCTACTGGGTTCACC.	AGC	.ATCTGATCACGGTCGGTCCATTGCATAGAG
6 GCTATATCCTCGAACTGCTACTGGGTTC-1..		..TCTGATCACGGTCGGTCCATTGCATAGAG
32 GCTATATCCTCGAACTGCTACTGGGTTC-1..	C	AATCTGATCACGGTCGGTCCATTGCATAGAG
60 GCTATATCCTCGAACTGCTACTGGGTTCAC..		..TCTGATCACGGTCGGTCCATTGCATAGAG
74 GCTATATCCTCGAACTGCTACTGGGTTC-1..	C	..TCTGATCACGGTCGGTCCATTGCATAGAG
71 GCTATATCCTCGAACTGCTACTGGGTTC-1..	TCGGG	..TCTGATCACGGTCGGTCCATTGCATAGAG
17,29,34,38,47,81 GCTATATCCTCGAACTGCTACTGGGTTCACCT		.-3..GATCACGGTCGGTCCATTGCATAGAG
18 GCTATATCCTCGAACTGCTACTGGGTTCACC.	C	.-1CTGATCACGGTCGGTCCATTGCATAGAG
65 GCTATATCCTCGAACTGCTACTGGGTTCACC.	CCA	.-1CTGATCACGGTCGGTCCATTGCATAGAG
25 GCTATATCCTCGAACTGCTACTGGGTTCAC..	GTGGGGGC	.-1CTGATCACGGTCGGTCCATTGCATAGAG
56 GCTATATCCTCGAACTGCTACTGGGTTC-1..	GGGCC	..TCTGATCACGGTCGGTCCATTGCATAGAG
26 GCTATATCCTCGAACTGCTACTGGGTTC-1..		.-1CTGATCACGGTCGGTCCATTGCATAGAG
72 GCTATATCCTCGAACTGCTACTGGGTTC-1..	C	.-1CTGATCACGGTCGGTCCATTGCATAGAG
27 GCTATATCCTCGAACTGCTACTGGGTTC-1..	TG	.-1CTGATCACGGTCGGTCCATTGCATAGAG
45 GCTATATCCTCGAACTGCTACTGGGTTC-1..	TCCCC	.-1CTGATCACGGTCGGTCCATTGCATAGAG
55 GCTATATCCTCGAACTGCTACTGGGTTCA...	GCCT	..-2TGATCACGGTCGGTCCATTGCATAGAG
2,20,43,66 GCTATATCCTCGAACTGCTACTGGGTTC-1..		..-2TGATCACGGTCGGTCCATTGCATAGAG
21 GCTATATCCTCGAACTGCTACTGGGTTC-1..	CT	..-2TGATCACGGTCGGTCCATTGCATAGAG
35 GCTATATCCTCGAACTGCTACTGGGTTCACCT	G	..-3.GATCACGGTCGGTCCATTGCATAGAG
42 GCTATATCCTCGAACTGCTACTGGGTTCACC.		..-3.GATCACGGTCGGTCCATTGCATAGAG
24 GCTATATCCTCGAACTGCTACTGGGTT-2...	ACCGACC	..-3.GATCACGGTCGGTCCATTGCATAGAG
30 GCTATATCCTCGAACTGCTACTGGGTT-2...	AGGG	..-3.GATCACGGTCGGTCCATTGCATAGAG
31 GCTATATCCTCGAACTGCTACTGGGTTCAC..		..-7.....ACGGTCGGTCCATTGCATAGAG
1 GCTATATCCTCGAACTGCTACTGGGTTC-1..	C	..-4..ATCACGGTCGGTCCATTGCATAGAG
23 GCTATATCCTCGAACTGCTACTGGGTTCA...	AGTGAT	..-5...TCACGGTCGGTCCATTGCATAGAG
44 GCTATATCCTCGAACTGCTACTGGGTT-2...	TGCCGCC	..-4..ATCACGGTCGGTCCATTGCATAGAG
52 GCTATATCCTCGAACTGCTACTGGG..-4...	CCGGGGC	..-4..ATCACGGTCGGTCCATTGCATAGAG
39 GCTATATCCTCGAACTGCTACTGGGTTC-1..	C	..-6...CACGGTCGGTCCATTGCATAGAG
40 GCTATATCCTCGAACTGCTACTGGGTTC-1..	GGGC	..-6....CACGGTCGGTCCATTGCATAGAG

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

```
41
 GCTATATCCTCGAACTGCTACTGGGTT-2...    AGTCGACCGCCAC    ..-6....CACGGTCGGTCCATTGCATAGAG
37
 GCTATATCCTCGAACTGCTACTGGGTT-2...       GCCAGA        ..-7.....ACGGTCGGTCCATTGCATAGAG
69
 GCTATATCCTCGAACTGCTACTGGGTTCAC..                     ..-8.......GGTCGGTCCATTGCATAGAG
59
 GCTATATCCTCGAACTGCTACT.....-7...                     .ATCTGATCACGGTCGGTCCATTGCATAGAG
76
 GCTATATCCTCGAACTGCTACT.....-7...                     .-1CTGATCACGGTCGGTCCATTGCATAGAG
50
 GCTATATCCTCGAACTGCTACTGGGTTC-1..       CCCCT         ..-9.......GGTCGGTCCATTGCATAGAG
33
 GCTATATCCTCGAACTGCTA.......-9...        TA           ..-4..ATCACGGTCGGTCCATTGCATAGAG
14
 GCTATATCCTCGAACTGCTA.......-9...        TTC          ..-6....CACGGTCGGTCCATTGCATAGAG
80
 GCTATATCCTCGAACTGC.........-11...       CTT          ..-5...TCACGGTCGGTCCATTGCATAGAG
8
 GCTATATCCTCGAACTGCTACTGGGT.-3...                     ..-12.........CGGTCCATTGCATAGAG
49
 GCTATATCCTCGAACTGCTACTGGGTTC-1..                     ..-12.........CGGTCCATTGCATAGAG
72,87
 GCTATATCCTCGAACTGCTACTGGGT.-3...        GA           ..-13..........GGTCCATTGCATAGAG
88
 GCTATATCCTCGAACTGCTACTG....-6...         T           ..-11........TCGGTCCATTGCATAGAG
64
 GCTATATCCTCGAACTGCTAC.......-8...                    ..-9.......GGTCGGTCCATTGCATAGAG
89
 GCTATATCCTCGAACT..........-13...        TCT          ..-4..ATCACGGTCGGTCCATTGCATAGAG
10
 GCTATATCCTCGAACTGCTACTGGGTT-2...        CCC          ..-15............TCCATTGCATAGAG
90
 GCTATATCCTCGAAC...........-14...        CCA          .-1CTGATCACGGTCGGTCCATTGCATAGAG
74
 GCTATATCCTCGAACTGCTACTGGGTTC-1..        +23          ..-16.............CCATTGCATAGAG
23bp : GGTAAGGTCCTGAGAAGCTATTC

51
 GCTATATCCTCGAACTGCTACTGG...-5...                     ..-15............TCCATTGCATAGAG
86
 GCTATATCCTCGA.............-16...       GCTC          ..TCTGATCACGGTCGGTCCATTGCATAGAG
61
 GCTATATCCTCGA.............-16...         G           ..-2TGATCACGGTCGGTCCATTGCATAGAG
5
 GCTATATCCTCGAACTG.........-12...         T           ..-8......CGGTCGGTCCATTGCATAGAG
3
 GCTATATCCTCGAACTGCTACTGGG..-4..        CCGA          ..-17..............CATTGCATAGAG
36
 GCTATATCCTCGAACTGCTACTGGGTTCACC..                    ..-23...................ATAGAG
91
 GCTATATCC.................-20...        GG           ..-5...TCACGGTCGGTCCATTGCATAGAG
62
 GCTATATCCTCGAACTG.........-12...                     ..-14...........GTCCATTGCATAGAG
22
 GCTATATCCTCGAACTGCTACTGGGTT-2...                     ..-25.....................AGAG
11
 GCTATATCCTCGAACTGCTAC.......-8...    AGTTCTCCAACTG    ..-16.............CCATTGCATAGAG
16
 GCTATATCCTCGAACTGCTACTGGGT.-3...        CTC          ..-15............TCCATTGCATAGAG
85
 GCTATAT...................-22...        GC           .-1CTGATCACGGTCGGTCCATTGCATAGAG
53
 GCTATATCCT................-19...         G           ..-8......CGGTCGGTCCATTGCATAGAG
82
 GCTATATCCTCGAACTGCTACTGGG..-4...                     ..-30.........................
73
 GCTATATCC.................-20...         C           ..-18..............ATTGCATAGAG
4
 GC........................-27...                     ..-18..............ATTGCATAGAG
```

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

13
GCTATATCCT.....-19... +11 ..-27.....G
11 bp duplication from 42 bp upstreamstream of the DSB on chr5
CTGGAGTGCAGTG

28
G.....-28... +13 ..-21.....GCATAGAG
13 bp duplication from 23 bp downstream of the DSB on chr2
AGGAGGGAACTGAGCA

15
GCTATATCCTCGAACTG.....-12... ..-32.....
7
.....-41... TTA ..-17.....CATTGCATAGAG
9
CA.....-35... ..-43.....
68
CATAGC.....-31... ..-43.....
12
GCTATATCCTCGAACTGCTACTGGGT*TC*-1.. ..-183.....
48
GCTATATCCTCGAAC.....-14... +20 ..-143.....
20 bp : 19bp duplication inverted from 152bp downstream of DSB on chr2
CCTACATGTGAGGATGTTCC

83
.....-231... ..-19.....TTGCATAGAG
56
.....-299... CCTCC ..-11.....TCGGTCCATTGCATAGAG

19
.....-269... +11+14 ..-27.....AG
25 bp : 2 duplications from nearby the DSBs
11bp duplication from 26bp downstream of DSB on chr2
GGGAACTGAGC
14bp duplication from 266bp upstream of DSB on chr5
TCAGCCTCCCCAGT

54
CAGTGATGTG.....-43... +28 ..-9.....GGTCGGTCCATTGCATAGAG
28bp :
10bp repeated duplication from 40bp upstream of the DSB on chr5 at the breakpoint junction
CAGTGATGTG CAGTGATGTG
8bp insertion from 15bp downstream of the DSB on chr5
TAGTTCTA

84
GCTATATCCTCGAACTGCTACTGGGT*TCACCT*. +68 ..-3.*GATCACGGTCGGTCCATTGCATAGAG*
68bp : 61bp from TALEN plasmid
GGAAGT ACGCCAGGGTACATTGGCAGTACTCCCATTGACGTCAATGGCGGTAAATGGCCCGCGATGG G

46
GCTATATCCTCGAACTGCTACTGGGT*TCA*... +137 ..-17.....CATTGCATAGAG
137bp : insertion inverted from 1.27Mb upstream of the DSB on chr2
GTTCCCACTCCATATGTGAGCCCGGTCCAGTGGAGATACCCAGAATTTGTCATGGGAGACAGAGTCCTTGCTCAGACA
CCTACTATACAATATGGATACCATCAAATGAATCACCTTATCTCTTTATGCATCAATTT

68
GCTATATCCTCGAACTGC.....-11... +195 ..-3.*GATCACGGTCGGTCCATTGCATAGAG*
195 bp from TALEN plasmid
CTAGTTCAAACAGGCTCTTGAGACCGTCCAACGCCTTCTACCAGTTCTCTGTCAAGCCCACGGACTAACCCCAGCGCAAGTTG
TAGCGATTGCTAGTAATGGGGGTGGCAAACAGGCTCTTGAGACTGTTCAGCGCCTTCTACCAGTTCTCTGTCAAGCCCACGGC
CTGACGCCCGAGCAAGTTGTAGCGATTTC

57
GCTATATCCTCGAACTGCTACTGGGT*TC*-1.. +174 ..*TCTG*ATCACGGTCGGTCCATTGCATAGAG
174bp : insertion 168bp inverted from 24Mb upstream of the DSB on chr2
CCCAATGTTTACACATATATTAACACTTATATTTCCAATTTCTTCTCCTTATTATGTAATCTCTGACTCCATCGTTTCATCTC

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

CAAATTTAGAGCTCTTGGGAGCTCCATGGGTTAGAATTACTCACGTCTTTGGTGTAACTTGCTCAGGTACATTTTCCAGGATA
AAGTTGAG

70
GCTATATCCTCGAACTGCTACTGGGT*T*-2... +265 ..-7......CGGTCGGTCCATTGCATAGAG
265bp from TALEN plasmid
GGAGCTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTC
TTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC
GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAagaACAGTATTTGGTATCTGCGCTCTGCTGAAGCC
AGTTACCTTCGGACGC

Der(2)

Template:
ATTCATTCGATCCTCAGGTAACCCTAAT*(Ct* ***TCA)CCTC*AGCCTCTGGAATAGCTAGAACTA**

45
ATTCATTCGATCCTCAGGTAACCCTAA*T*. CACTCCAACC .*CACCTC*AGCCTCTGGAATAGCTAGAACTA
38
ATTCATTCGATCCTCAGGTAACCCTAA*TC* ..-1*CTC*AGCCTCTGGAATAGCTAGAACTA
43
ATTCATTCGATCCTCAGGTAACCCTA-2. GAATCCCGG .*CACCTC*AGCCTCTGGAATAGCTAGAACTA
16,31
ATTCATTCGATCCTCAGGTAACCCTA-2. ..-1*CTC*AGCCTCTGGAATAGCTAGAACTA
27
ATTCATTCGATCCTCAGGTAACCCTA-2. GAATATCAT ..-1*CTC*AGCCTCTGGAATAGCTAGAACTA
49
ATTCATTCGATCCTCAGGTAACCCTAA-*1* GG ...-2*TC*AGCCTCTGGAATAGCTAGAACTA
64
ATTCATTCGATCCTCAGGTAACCCTAA*TC* CCGG ...-4..AGCCTCTGGAATAGCTAGAACTA
50
ATTCATTCGATCCTCAGGTAACCCTA-*2*. ...-2*TC*AGCCTCTGGAATAGCTAGAACTA
51
ATTCATTCGATCCTCAGGTAACCCTAA-*1* GG ...-3.*C*AGCCTCTGGAATAGCTAGAACTA
19
ATTCATTCGATCCTCAGGTAACCCTAA-*1* AGTAAAAGCC ...-3.*C*AGCCTCTGGAATAGCTAGAACTA
17
ATTCATTCGATCCTCAGGTAACCCTA-2. ...-3.*C*AGCCTCTGGAATAGCTAGAACTA
66
ATTCATTCGATCCTCAGGTAACCCTA-2. CTTC ...-3.*C*AGCCTCTGGAATAGCTAGAACTA
20
ATTCATTCGATCCTCAGGTAACCCTA-2. TCTGGAC ...-3.*C*AGCCTCTGGAATAGCTAGAACTA
41
ATTCATTCGATCCTCAGGTAACCCTA*AT*. +8 ...-4..AGCCTCTGGAATAGCTAGAACTA
8 bp duplication 234bp upstream of DSB on chr2
ATGCCCAGGA
63
ATTCATTCGATCCTCAGGTAACCCT-2.. CG ..*ACCTC*AGCCTCTGGAATAGCTAGAACTA
39
ATTCATTCGATCCTCAGGTAACCCT-2.. ...-3.*C*AGCCTCTGGAATAGCTAGAACTA
26
ATTCATTCGATCCTCAGGTAACCC.-3.. ...*CCTC*AGCCTCTGGAATAGCTAGAACTA
28
ATTCATTCGATCCTCAGGTAACCC.-3.. CCCAAA ..-1*CTC*AGCCTCTGGAATAGCTAGAACTA
61
ATTCATTCGATCCTCAGGTAACCC.-3.. ...-4..AGCCTCTGGAATAGCTAGAACTA
42
ATTCATTCGATCCTCAGGTAACCC.-3.. GGTCCCG ..-1*CTC*AGCCTCTGGAATAGCTAGAACTA
40
ATTCATTCGATCCTCAGGTAACCC.-3.. CCGTTGCCACCTT ..-1*CTC*AGCCTCTGGAATAGCTAGAACTA
37
ATTCATTCGATCCTCAGGTAACCCT-2.. GGG ...-6....CCTCTGGAATAGCTAGAACTA
7
ATTCATTCGATCCTCAGGTAACCCT-2.. GGGAT ...-6....CCTCTGGAATAGCTAGAACTA
55
ATTCATTCGATCCTCAGGTAACCCTAA.. GT ...-7.....CTCTGGAATAGCTAGAACTA
54
ATTCATTCGATCCTCAGGTAACCCTAA.. ATCCT ...-7.....CTCTGGAATAGCTAGAACTA
60
ATTCATTCGATCCTCAGGTA.....-7.. CTCCC ...-3.*C*AGCCTCTGGAATAGCTAGAACTA

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

```
 ATTCATTCGATCCTCAGGTA.....-7..          GAC          ...-4..AGCCTCTGGAATAGCTAGAACTA
1
 ATTCATTCGATCCTCAGGTAA....-6..       TAGCTGGACCGA    ...-6....CCTCTGGAATAGCTAGAACTA
53,58
 ATTCATTCGATCCTCAGGTAACCCT-2..                       ...-11........GGAATAGCTAGAACTA
3
 ATTCATTCGATCCTCAGGT......-8..          TCCT         ...-6....CCTCTGGAATAGCTAGAACTA
25
 ATTCATTCGATCCTCAG.......-10..                       .CACCTCAGCCTCTGGAATAGCTAGAACTA
62
 ATTCATTCGATCCTCAG.......-10..         TGTGG         ...CCTCAGCCTCTGGAATAGCTAGAACTA
30
 ATTCATTCGATCCTCAGGTAACCCT-2..           C           ...-12.........GAATAGCTAGAACTA
29
 ATTCATTCGATCCTCAGGTAACCC.-3..          GC           ...-12.........GAATAGCTAGAACTA
33,47
 ATTCATTCGATCCTCAG.......-10..                       ...-6....CCTCTGGAATAGCTAGAACTA
24
 ATTCATTCGATCCTC.........-12..           T           ..ACCTCAGCCTCTGGAATAGCTAGAACTA
57
 ATTCATTCGATCCTCAGGTAACCCTAA..                       ...-15............TAGCTAGAACTA
21
 ATTCATTCGATCCTCA........-11..                       ...-6....CCTCTGGAATAGCTAGAACTA
12
 ATTCATTCGATCCTCAG.......-10..                       ...-7.....CTCTGGAATAGCTAGAACTA
65
 ATTCATTCGATCCTCAGGTAACCCTAA..                       ...-18..............CTAGAACTA
46
 ATTCATTCGATCC...........-14..                       ...-5...GCCTCTGGAATAGCTAGAACTA
5
 ATTCATTCGATCCTCAGGTAA....-6..                       ...-15............TAGCTAGAACTA
15
 ATTCATTCGATCCTCAGGTA.....-7..                       ...-17..............GCTAGAACTA
14
 ATTCATTCGATCCTCAGG.......-9..          AAG          ...-15............TAGCTAGAACTA
68
 ATTCATTCGATCCTC.........-12..                       ...-10.......TGGAATAGCTAGAACTA
2
 ATTCATTC................-19..                       ...-7.....CTCTGGAATAGCTAGAACTA
6
 ATTC....................-23..                       ...-3.CAGCCTCTGGAATAGCTAGAACTA
29
 ATTC.....  ......-23..          G           ...-2TCAGCCTCTGGAATAGCTAGAACTA
59
 ATTCATTCGATCCTCAG........-9..                       ...-18..............CTAGAACTA
4
 ATTCATTCG...............-18..         GCTGGCA       ...-12.........GAATAGCTAGAACTA
35,44
 ATTCATTCGA..............-17..                       ...-14...........ATAGCTAGAACTA
18
 ATTCATTCGATCCTCA........-11..                       ...-18..............CTAGAACTA
34
 ........................-46..                       ...-17..............GCTAGAACTA
48
 CACA....................-33..                       ...-35........................
67
 ........................-72..        TTGTTCT        ...-18..............CTAGAACTA
36
 T.......................-67..                       ...-37........................
56
 .......................-321..          GCC          ...-18..............CTAGAACTA

29
 ATTCATTCGATCCTCAGGTAACCC.-3..          +10          ...-31........................
10bp : 9bp insertion from 10 bp downstream of DSB on chr2
GGTCCATTGA

11
 ........................-42..                       ...-35........................
```

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

9
ATTCATTCGATC.....-15.. +21+7 ...-4..AGCCTCTGGAATAGCTAGAACTA
28bp : 21bp duplication from 12bp upstream of DSB on chr2
GG CATTTCAATTCATTCGATCCT ATTCG

8
ATTCATTCGATCCTCAGGTA.....-7.. +16+19 ...-2*TC*AGCCTCTGGAATAGCTAGAACTA
35 bp :
16bp TAGCTGAGGATCGATC
19bp inverted duplication directly upstream at DSB on chr5
CCCAGTAGCAGTTCGAGGAT

23
ATTCATTCGATCCTCAG.......-10.. +27 ...-4...GCCTCTGGAATAGCTAGAACTA
272bp :268bp duplication from 13bp upstream of DSB on chr2
ATCACTCAGTGTGGTCCCAACCAGCTCTTACTGCTGGCAGAGACATGCCCAGGACAGATGGGCAGAGGCTTGAAAAGGGCAGA
GGGAAAGGCTCTTGAGAGCCCTCGCAGGCCAGGCCCCTGCAGGCAAAGGGATCTGCCGGTAGAAGGGAGATGGCAGCACACAC
TGTGTCCCCATATGGTGCCATCCCTCAAAGGGACAGGATAATAGGAGCTAACACTTGTTGCATGGTTACTACGTGCTCGGCAA
TTTACACATTTCAATTCATTCGATCCG

29
ATTCATTCGATCCTCAGGTAACCCTA*AT*. +307 .*CACCTC*AGCCTCTGGAATAGCTAGAACTA
307 bp : 306 bp insertion from 6bp upstream of DSB on chr5
ATGAACCTCTCCGGGATTCAAGCGATCCTCCTGTCTCAGCCTCCCCAGTAGCTGGGGCCACAGACGTGCACCACCAAGTCCAG
TTGCGTTTCATAGTTATAGTAGAGAAGGGGTTTCGTGATGTTGGCCAGGCCGATCTTGAACTCCTGGCCTTAACGTGATCTGC
ACGCCCTGGCCTCTCAAAGTGCTAGTATTACACGTGTCAGCCACTGTGCCTGGCCTAAAAATTATTTTTTAATAAAGACAGTC
TCATTATAACGGCTGGAGTGCAGTGATGTGATCATAGCTTGCTATATCCTCGAACTGCTT

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

B. hES-MP cells

Der(5)

```
Template:
 GCTATATCCTCGAACTGCTACTGGGTTCA(CCT                    AA)TCTGATCACGGTCGGTCCATTGCATAGAG
6
 GCTATATCCTCGAACTGCTACTGGGTTCACC.        C            ..TCTGATCACGGTCGGTCCATTGCATAGAG
32
 GCTATATCCTCGAACTGCTACTGGGTTCAC..      TGCAAAG         .ATCTGATCACGGTCGGTCCATTGCATAGAG
16
 GCTATATCCTCGAACTGCTACTGGGTTC-1..        C            .ATCTGATCACGGTCGGTCCATTGCATAGAG
30
 GCTATATCCTCGAACTGCTACTGGGTTC-1..        T            .ATCTGATCACGGTCGGTCCATTGCATAGAG
5
 GCTATATCCTCGAACTGCTACTGGGTTC-1..       CC            ..TCTGATCACGGTCGGTCCATTGCATAGAG
38
 GCTATATCCTCGAACTGCTACTGGGTTC-1..        C            ..TCTGATCACGGTCGGTCCATTGCATAGAG
3,10,20,20,25,37
 GCTATATCCTCGAACTGCTACTGGGTTCAC..                     .-1CTGATCACGGTCGGTCCATTGCATAGAG
13
 GCTATATCCTCGAACTGCTACTGGGTTCAC..        A            .-1CTGATCACGGTCGGTCCATTGCATAGAG
3
 GCTATATCCTCGAACTGCTACTGGGT.-3...                     .ATCTGATCACGGTCGGTCCATTGCATAGAG
9,21
 GCTATATCCTCGAACTGCTACTGGGTTCAC..                     .-1.TGATCACGGTCGGTCCATTGCATAGAG
19,23
 GCTATATCCTCGAACTGCTACTGGGTT-2...                     ..TCTGATCACGGTCGGTCCATTGCATAGAG
17
 GCTATATCCTCGAACTGCTACTGGGTTCAC..        A            ..-2TGATCACGGTCGGTCCATTGCATAGAG
24
 GCTATATCCTCGAACTGCTACTGGGTTCAC..      GGT            ..-2TGATCACGGTCGGTCCATTGCATAGAG
22
 GCTATATCCTCGAACTGCTACTGGGTTCACC.       AA            ..-2TGATCACGGTCGGTCCATTGCATAGAG
18
 GCTATATCCTCGAACTGCTACTGGGTTCA...        G            ..-2TGATCACGGTCGGTCCATTGCATAGAG
35
 GCTATATCCTCGAACTGCTACTGGGTT-2...       TC            .-1CTGATCACGGTCGGTCCATTGCATAGAG
2
 GCTATATCCTCGAACTGCTACTGGGTTC-1..                     ..-2TGATCACGGTCGGTCCATTGCATAGAG
12
 GCTATATCCTCGAACTGCTACTGGGTTC-1..    CATGACCA         ..-2TGATCACGGTCGGTCCATTGCATAGAG
29
 GCTATATCCTCGAACTGCTACTGGGTTCACC.                     ..-3.GATCACGGTCGGTCCATTGCATAGAG
31
 GCTATATCCTCGAACTGCTACTGGGTTCACC.        A            ..-3.GATCACGGTCGGTCCATTGCATAGAG
34
 GCTATATCCTCGAACTGCTACTGGGTTCAC..    GTAGCACA          ..-3.GATCACGGTCGGTCCATTGCATAGAG
27
 GCTATATCCTCGAACTGCTACTGGGTTCA...                     ..-3.GATCACGGTCGGTCCATTGCATAGAG
1
 GCTATATCCTCGAACTGCTACTGGGTTCA...     ATGGCAA          ..-3.GATCACGGTCGGTCCATTGCATAGAG
4
 GCTATATCCTCGAACTGCTACTGGGTT-2...       ACC           ..-3.GATCACGGTCGGTCCATTGCATAGAG
8
 GCTATATCCTCGAACTGCTACTGGGTT-2...      TTAG           ..-3.GATCACGGTCGGTCCATTGCATAGAG
28
 GCTATATCCTCGAACTGCTACTGGGTTC-1..       TC            ..-3.GATCACGGTCGGTCCATTGCATAGAG
14
 GCTATATCCTCGAACTGCTACTG....-5...    TTCACAGG         ..TCTGATCACGGTCGGTCCATTGCATAGAG
7
 GCTATATCCTCGAACTG.........-11...        TG           ..-2TGATCACGGTCGGTCCATTGCATAGAG

11
 GCTATATCCTCGAACTGCTACTGGGTTCAC..        +95          ..-4..ATCACGGTCGGTCCATTGCATAGAG
95 bp from Homo sapiens chromosome 8 genomic contig, GRCh37.p5 Primary Assembly
(repeated sequences)
             ACTTATAACTTATAGCTA
TAGCTATAAGCTAACTTATAACTTATAGCTA
```

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

TAGCTATAAGCTAACTTATAACTTATAGCTA
TAGCTATAAGCTAACTTAT

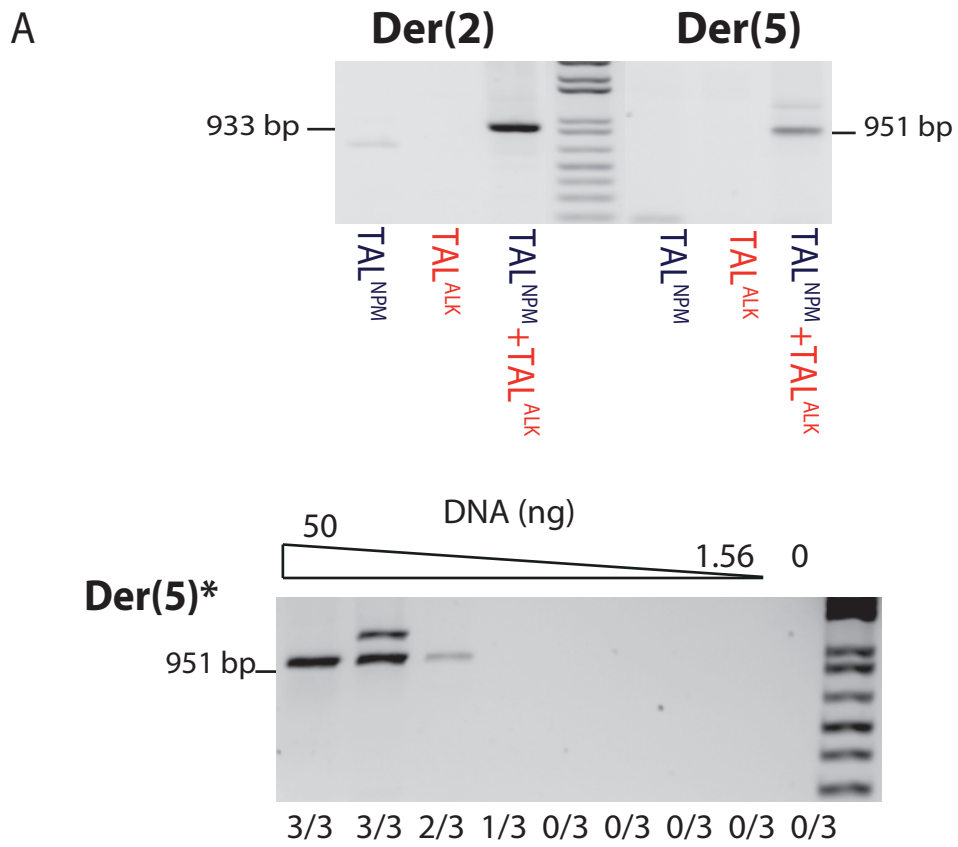
15
GCTATATCCTCGAACTGCTACTGGGT*TCA*... +47 *.ATCT*GATCACGGTCGGTCCATTGCATAGAG
47 bp from Homo sapiens chromosome 9 genomic contig, GRCh37.p5 Primary Assembly :
GCAAAGGTAGTTTCATTTTCATATGCAAAATACAATCATTACATCCC

26
GCTATATCCTCGAACTGCTACTGGGT*TC*-1.. +40 .-1*CT*GATCACGGTCGGTCCATTGCATAGAG
40bp : Insertion from 3157bp downstream the DSB on chr5
CGTTTCATAGTTTTAGTAGAGAAGGGGTTTCGTGATGTTG

36
GCTATATCCTCGAACTGCTACTGGGT*TC*AC.. +58 ..-5...TCACGGTCGGTCCATTGCATAGAG
58bp :
ATGATATATATAATATACAAATCTAATAAAGTTACAAAATTTATAATCTAGACATCGT

Supplementary Figure 8

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press



Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Supplementary Figure 8B

Jurkat cells- FokI heterodimer

Der(5)

```
Template:
 GCTATATCCTCGAACTGCTACTGGGTTCA(CCT                    AA)TCTGATCACGGTCGGTCCATTGCATAGAG
-------------------------------------------------------------------------------------
1,36
 GCTATATCCTCGAACTGCTACTGGGTTCAC..                     .ATCTGATCACGGTCGGTCCATTGCATAGAG
15
 GCTATATCCTCGAACTGCTACTGGGTTCA...                     .ATCTGATCACGGTCGGTCCATTGCATAGAG
42
 GCTATATCCTCGAACTGCTACTGGGTTCA...                     ..TCTGATCACGGTCGGTCCATTGCATAGAG
21
 GCTATATCCTCGAACTGCTACTGGGTTCACC.                     .-1CTGATCACGGTCGGTCCATTGCATAGAG
13
 GCTATATCCTCGAACTGCTACTGGGTTCA..      AAGCTGGCTCC      .-1CTGATCACGGTCGGTCCATTGCATAGAG
25,33
 GCTATATCCTCGAACTGCTACTGGGTTCACCT                     .-3..GATCACGGTCGGTCCATTGCATAGAG
45
 GCTATATCCTCGAACTGCTACTGGGTTC-1..                     ..TCTGATCACGGTCGGTCCATTGCATAGAG
17
 GCTATATCCTCGAACTGCTACTGGGTT.-2..                     ..TCTGATCACGGTCGGTCCATTGCATAGAG
31
 GCTATATCCTCGAACTGCTACTGGGTTCACC                      ..-2TGATCACGGTCGGTCCATTGCATAGAG
28
 GCTATATCCTCGAACTGCTACTGGGTTCACC                      ..-3.GATCACGGTCGGTCCATTGCATAGAG
40
 GCTATATCCTCGAACTGCTACTGGGTTCAC.          G           ..-3.GATCACGGTCGGTCCATTGCATAGAG
6,20,24,29,41
 GCTATATCCTCGAACTGCTACTGGGTTC-1..                     ..-2TGATCACGGTCGGTCCATTGCATAGAG
30
 GCTATATCCTCGAACTGCTACTGGGTT.-2..        GCC          .-1CTGATCACGGTCGGTCCATTGCATAGAG
5
 GCTATATCCTCGAACTGCTACTGGGTTC-1..                     ..-3.GATCACGGTCGGTCCATTGCATAGAG
34
 GCTATATCCTCGAACTGCTACTGGGT..-3..                     .-1CTGATCACGGTCGGTCCATTGCATAGAG
11
 GCTATATCCTCGAACTGCTACTGGGTT.-2..         G           ..-2TGATCACGGTCGGTCCATTGCATAGAG
32
 GCTATATCCTCGAACTGCTACTGGGT.-3...         C           .-1CTGATCACGGTCGGTCCATTGCATAGAG
14
 GCTATATCCTCGAACTGCTACTGGGT.-3...         TG          ..-2TGATCACGGTCGGTCCATTGCATAGAG
37
 GCTATATCCTCGAACTGCTACTG....-6...         G           ..ACTGATCACGGTCGGTCCATTGCATAGAG
4
 GCTATATCCTCGAACTGCTACTGGGT.-3...                     ..-4..ATCACGGTCGGTCCATTGCATAGAG
3
 GCTATATCCTCGAACTGCTACTGGGTT.-3..         AGG         ..-7.....ACGGTCGGTCCATTGCATAGAG
8
 GCTATATCCTCGAACTGCTACTGGGTTCACC.         G           ..-9.......GGTCGGTCCATTGCATAGAG
43
 GCTATATCCTCGAACTGCTACTGGGTTCAC..                     ..-9.......GGTCGGTCCATTGCATAGAG
22
 GCTATATCCTCGAACTGCTACTGGGTT.-3..         AGGG        ..-8......CGGTCGGTCCATTGCATAGAG
9
 GCTATATCCTCGAACTGCTACTGGG...-5..         ATAGCG      ..-9.......GGTCGGTCCATTGCATAGAG
16
 GCTATATCCTCGAACTGCTAC.......-7...                    ..-7.....ACGGTCGGTCCATTGCATAGAG
23
 GCTATATCCTCGAACTGCTACTGGGT..-4..                     ..-12.........CGGTCCATTGCATAGAG
38,44
 GCTATATCCTCGAACTGCTAC.......-7...                    ..-9.......GGTCGGTCCATTGCATAGAG
19
 GCTATATCCTCGAACTG..........-14..                     ..-4..ATCACGGTCGGTCCATTGCATAGAG
35
```

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

```
12
 GCTATATCCTCGAACT...........-14..                    ..-8......CGGTCGGTCCATTGCATAGAG
26
 GCTATATCCTCGAACTGCTA........-8...        TGA         ..-15............TCCATTGCATAGAG
27
 GCTATATCCTCGAACTGCT.........-9...         TG         ..-21..................GCATAGAG
10
 GCTATATCCTCGAAC.............-15..        +12         ..-16.............CCATTGCATAGAG
12bp : GGCTTTACTTCC
7
 GCTATATCCTCGAACTGCTACTGGGTTCACCT         +33         ..-4.GATCACGGTCGGTCCATTGCATAGAG
33bp insertion from 5bp downstream of DSB on chr5
AGCCTCTGGAATAGCTAGAACTACAGGCACACG
39
 ............................-281.       CCCT        ..-3TGATCACGGTCGGTCCATTGCATAGAG
2
 GCTATATCCTCGAACTGCTACTGGGTTCACC.         GGA         ..-350.........................
18
 GCTATATCCTCGAACTG..........-13...        +142        ..TCTGATCACGGTCGGTCCATTGCATAGAG
142bp duplication from 7.5kb upstream of NPM1 on chr5
GAGAGAGAGAAAAAGTAAGTGTATTTTTGTAGTTTTTTAAGTATTATGAGTCAAAGTTTTAAAAGTTTTACAAAGTTCAG
TAAGCTTACATTAATGAAGAAAAACATTTTTATATTTAACATAATTATACTGTGTTTATAGT
```

Der(2)

```
Template:
ATTCATTCGATCCTCAGGTAACCCTAAT(Ct                    TCA)CCTCAGCCTCTGGAATAGCTAGAACTA
26
 ATTCATTCGATCCTCAGGTAACCCTAATCT          G          ..ACCTCAGCCTCTGGAATAGCTAGAACTA
27
 ATTCATTCGATCCTCAGGTAACCCTAATC.          AC         .CACCTCAGCCTCTGGAATAGCTAGAACTA
17,28,48
 ATTCATTCGATCCTCAGGTAACCCTAATC.                     ..ACCTCAGCCTCTGGAATAGCTAGAACTA
43
 ATTCATTCGATCCTCAGGTAACCCTAA-1.                     .CACCTCAGCCTCTGGAATAGCTAGAACTA
11
 ATTCATTCGATCCTCAGGTAACCCTAA-1.                     ..ACCTCAGCCTCTGGAATAGCTAGAACTA
24
 ATTCATTCGATCCTCAGGTAACCCTAATC.                     ..-1CTCAGCCTCTGGAATAGCTAGAACTA
38
 ATTCATTCGATCCTCAGGTAACCCTAA-1.                     ...CCTCAGCCTCTGGAATAGCTAGAACTA
36
 ATTCATTCGATCCTCAGGTAACCCTAA-1.          CAC        ...CCTCAGCCTCTGGAATAGCTAGAACTA
2
 ATTCATTCGATCCTCAGGTAACCCTAATCt          A          ..-2.TCAGCCTCTGGAATAGCTAGAACTA
32
 ATTCATTCGATCCTCAGGTAACCCTA-2..                     .CACCTCAGCCTCTGGAATAGCTAGAACTA
7
 ATTCATTCGATCCTCAGGTAACCCTA-2..          CTC        .CACCTCAGCCTCTGGAATAGCTAGAACTA
3,23,30,50
 ATTCATTCGATCCTCAGGTAACCCTAATCt                     ...-3.CAGCCTCTGGAATAGCTAGAACTA
18
 ATTCATTCGATCCTCAGGTAACCCT.-3..        GGGGGGA      .CACCTCAGCCTCTGGAATAGCTAGAACTA
12
 ATTCATTCGATCCTCAGGTAACCCT.-3..                     .CACCTCAGCCTCTGGAATAGCTAGAACTA
15
 ATTCATTCGATCCTCAGGTAACCCTAA-1.          CCCGC      ...-3.CAGCCTCTGGAATAGCTAGAACTA
5
 ATTCATTCGATCCTCAGGTAACCCTAATCt                     ...-4..AGCCTCTGGAATAGCTAGAACTA
20
 ATTCATTCGATCCTCAGGTAACCC..-4..                     ..ACCTCAGCCTCTGGAATAGCTAGAACTA
25
 ATTCATTCGATCCTCAGGTAACCC..-4..                     ...CCTCAGCCTCTGGAATAGCTAGAACTA
33
 ATTCATTCGATCCTCAGGTAACCCT.-3..          CCCCT      ...-2TCAGCCTCTGGAATAGCTAGAACTA
13
 ATTCATTCGATCCTCAGGTAACCCT.-3..        GAGGATTC     ...-3.CAGCCTCTGGAATAGCTAGAACTA
```

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

ATTCATTCGATCCTCAGGTAACCCT.-3.. GGG ...-4..AGCCTCTGGAATAGCTAGAACTA
47
ATTCATTCGATCCTCAGGTAACCC..-4.. GTAAAGGAAG ...-3.CAGCCTCTGGAATAGCTAGAACTA
39
ATTCATTCGATCCTCAGGTAACCC.-3... ...-4..AGCCTCTGGAATAGCTAGAACTA
9
ATTCATTCGATCCTCAGGTAACCCTA-1. ...-6....CCTCTGGAATAGCTAGAACTA
6
ATTCATTCGATCCTCAGGTAACCCT.-3.. AGT ...-4..AGCCTCTGGAATAGCTAGAACTA
45
ATTCATTCGATCCTCAGGTAACCCT.-2.. ...-9.......CTGGAATAGCTAGAACTA
1
ATTCATTCGATCCTCAG.........-11.. ...CCTCAGCCTCTGGAATAGCTAGAACTA
37,44
ATTCATTCGATCCTCAGGTAACC...-5.. ...-8......TCTGGAATAGCTAGAACTA
49
ATTCATTCGATCC............-15.. C .CACCTCAGCCTCTGGAATAGCTAGAACTA
14
ATTCATTCGATCCTCAGGTAACCCTA-2.. GG ...-14...........ATAGCTAGAACTA
31
ATTCATTCGATCCTCAG.........-11.. ...-6....CCTCTGGAATAGCTAGAACTA
19
ATTCATTCGATCCTCAG.........-11.. ...-6....CCTCTGGAATAGCTAGAACTA
22
ATTCATTCGATCC............-15.. CC ...-3.CAGCCTCTGGAATAGCTAGAACTA
29
ATTCATTCGATCCTCAG.........-11.. ACTAA ...-11........GGAATAGCTAGAACTA
34,46
ATTCATTCGATCCTC..........-13.. ...-10.......TGGAATAGCTAGAACTA
41
ATTCATTCGAT..............-17.. ...-16.............AGCTAGAACTA
10
ATTCATT..................-21.. +21 ..-1CTCAGCCTCTGGAATAGCTAGAACTA

+21 bp insertion from 13bp upstream of DSB on chr5

TGCTATATCCTCGAACTGCTA
42
ATTCATTCGATCCTCAGGTAACCCTAATC. ...-27.....................A
8
ATT.....................-24.. ...—4..AGCCTCTGGAATAGCTAGAACTA
4
AT......................-24.. CA ...—6....CCTCTGGAATAGCTAGAACTA
40
ATT.....................-24.. ...—16.............AGCTAGAACTA
16
ATTCATTCGATCCTCAGGTAACCCTAATCt G ...-41.....................
21
ATTCATTC................-20.. +282 ..ACCTCAGCCTCTGGAATAGCTAGAACTA

+282 bp : Insertion from the TALEN expression plasmid

CCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGG
CTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG
GCAAACAAACCACCGCTGGTAGCGGTGGTGAGA

Downloaded from genome.cshlp.org on November 5, 2026 . Published by Cold Spring Harbor Laboratory Press

Supplementary Figure 9

