Published in final edited form as:

Methods. 2017 May 15; 121-122: 138–145. doi:10.1016/j.ymeth.2017.05.006.

Generation of chromosomal translocations that lead to conditional fusion protein expression using CRISPR-Cas9 and homology-directed repair

Fabio Vanoli^a and Maria Jasin^{a,*}

^aDevelopmental Biology Program, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065 USA

Abstract

Recurrent chromosomal translocations often lead to expression of fusion proteins associated with oncogenic transformation. To study translocations and downstream events, genome editing techniques have been developed to generate chromosomal translocations through non-homologous end joining of DNA double-strand breaks introduced at the two participating endogenous loci. However, the frequencies at which these events occur is usually too low to efficiently clone cells carrying the translocation. This article provides a detailed method using CRISPR-Cas9 technology and homology-directed repair to efficiently isolate cells harboring a chromosomal translocation. For an additional level of control, the resulting fusion protein is conditionally expressed to allow early events in oncogenic transformation to be studied. We focus on the generation of the *EWSR1-WT1* fusion using human mesenchymal cells, which is associated with the translocation found in desmoplastic small round cell tumors.

Keywords

double-strand break; HDR; chromosomal translocation; CRISPR-Cas9; EWSR1-WT1

1. Introduction

Chromosomal translocations are thought to arise by the inappropriate joining of DNA ends from contemporaneous double-strand breaks (DSBs) occurring on two different chromosomes, and have been considered as the primary cause of several tumor types, including some lymphomas, leukemias, sarcomas, and carcinomas, due to either the generation of oncogenic fusion proteins or enhanced expression of proto-oncogenes [1-3].

Because DSBs induce chromosomal translocations, the development of programmable nucleases has greatly expanded interest in studying mechanisms and cellular consequences of translocation formation. Initial studies of the DSB repair mechanisms underlying

*Corresponding author address m-jasin@ski.mskcc.org (M. Jasin).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

translocation formation employed the rare-cutting homing endonuclease I-SceI to introduce two DSBs at chromosomal loci previously modified in mouse cells to bear I-SceI recognition sites [4-6]. These studies demonstrated that DNA ends from two chromosomes could join through non-homologous end joining (NHEJ) (Fig. 1A,B), although the use of canonical and alternative NHEJ pathways may differ in different cell types [7-9]. Subsequent studies used the programmable nucleases ZFNs (zinc-finger nucleases) and TALENs (transcription activator-like effector nucleases) [10], which target endogenous sites and as such do not require prior modification of chromosomal loci. Thus, programmable nucleases make the approach more flexible and have allowed the generation of cancer-relevant translocations in several human cell lines [3, 9, 11, 12].

More recently, the spectrum of modeled translocations was further expanded through the use of the RNA-guided Cas9 endonuclease (CRISPR-Cas9) [9, 13-15]. Endonuclease activity of the Cas9 protein requires a guide RNA (gRNA) with ~19 nucleotides of complementarity to a DNA target site in the genome which is followed by a short sequence important for initial binding of the Cas9 protein (NGG for SpCas9, termed the Protospacer Adjacent Motif or PAM) [16, 17]. Given the ease in programming Cas9 cleavage through a gRNA and the minimal genomic sequence requirement (e.g., GG in the SpCas9 PAM), CRISPR-Cas9 has become the preferred approach for inducing chromosomal translocations.

Clonal selection of cells harboring a translocation is important for studying steps leading to tumorigenesis, but two impediments exist with current approaches. Nuclease-induced translocations that occur by NHEJ are readily detected by breakpoint-junction PCR; however, the low frequency of chromosomal translocation formation ($\sim 10^{-3}$ - 10^{-4}) [11, 12] may require extensive sib-selection for the recovery of clones, unless immediate cellular transformation occurs. Although this procedure has proved to be feasible in a transformed cell line [11], it has limitations when using non-immortalized cells because of their limited ability to be passaged. Further, although a fusion protein arising from a translocation is potentially oncogenic, its immediate expression without “enabling” mutations may lead to cell cycle progression defects.

Homology-directed repair (HDR) is a critical DSB repair pathway whereby the damaged locus is typically repaired using homologous sequences on the sister chromatid as a template, although homologous sequences in other contexts can also be used [18-20]. In particular, as initially shown with I-SceI but later also with ZFNs, TALENs, and CRISPR-Cas9, induction of a DSB at a target locus increases HDR with a donor DNA fragment or plasmid orders of magnitude [18, 21-23]. Thus, both HDR and NHEJ can be used to modify a target locus. Advantages of HDR include the generation of a defined allele to facilitate genotyping of recovered clones and, compared with traditional gene targeting methods, the homology arms on the donor fragment can be short (<1 kb), which facilitates donor plasmid construction. Moreover, HDR events can be readily selected if the donor fragment contains a promoterless selectable marker (i.e., a gene trap) [24].

Here we describe in detail a recently published method involving HDR and CRISPR-Cas9 to select clones harboring a chromosomal translocation in which the fusion protein is

conditionally expressed Fig. 1C,D), using non-immortalized human mesenchymal progenitor cells [25].

2. Methods

2.1 Overall Strategy

The strategy involves induction of DSBs in the two loci involved in a chromosomal translocation, followed by integration of a donor fragment by HDR (Fig. 1C). The donor fragment has sequences homologous to the two loci (homology arms) flanking a selectable marker coding sequence preceded by a splice acceptor. Fusion protein expression is conditional and dependent on removal of the selectable marker cassette by expression of Cre recombinase (Fig. 1D). Non-immortalized cells are used to be able to understand the steps leading to oncogenic transformation.

2.2 Translocation Strategy

While chromosomal translocations typically arise by NHEJ to generate fusion genes, the method presented in this paper is designed to generate a translocation by HDR. DSBs are generated within introns of two genes present on different chromosomes at sites determined by the RNA-guided Cas9 endonuclease (scissors) (Fig. 1A). Sequences targeted by Cas9 are not present in the homology arms (green and red shading, Fig. 1C) of the donor fragment, but are adjacent to the homologous sequences in the chromosomes. Homologous integration of the 5' and 3' homology arms with the two different chromosomes generates the desired translocation. The reciprocal translocation, if formed, can form instead by NHEJ, as in Fig. 1B.

Because introns often contain repetitive sequences, it is particularly important to choose gRNAs which are predicted to target with high specificity and have few potential off-target sites (see Section 2.3.1 and Discussion). Whenever possible, repetitive sequences in the homology arms should be avoided. Repetitive elements in the genome can often be quite diverged, however, and even a small amount of divergence reduces HDR [26].

The homology arms are positioned 5' and 3' of a promoterless selectable marker (*Hyg*-) which is flanked by LoxP sites (blue arrowheads, Fig. 1C). An upstream splice acceptor (SA) allows in-frame expression of the marker upon homologous integration of the donor fragment at the 5' homology arm (*Hyg*⁺). The self-cleaving T2A peptide sequence just downstream of the SA minimizes the length of the peptide fused to the selectable marker. Expression of Cre recombinase and consequent excision of the selectable marker coding sequence allows expression of the fusion gene (Fig. 1D).

The method described here models the *EWSR1-WT1* translocation, which is the t(11;22)(p13;q12) chromosomal rearrangement found in desmoplastic small round cell tumor (DSRCT), a sarcoma of mesenchymal origin that affects primarily young adults (Fig. 2A) [27, 28]. This translocation is characterized by fusion of the first 7 exons of *EWSR1* on chromosome 22 with the last three exons (8 to 10) of *WT1* on chromosome 11. The resulting protein fusion is a transcription factor composed of the transactivator domain (TA) of *EWSR1* and three zinc-finger motifs from *WT1* that bind DNA (Fig. 2B). Because of the

presumed mesenchymal origin of DSRCT, experiments are performed in nonimmortalized human mesenchymal stem/progenitor cells (hMSC) derived from human embryonic stem cells [29]. Although these cells have a limited lifespan, they provide the long-term opportunity to determine the requirements for cellular transformation.

2.3 Molecular Cloning

2.3.1 Cloning of guide RNA expression plasmids—Cas9 and gRNAs are expressed from plasmid pSpCas9(BB)-2A-GFP (PX458 Addgene # 48138). gRNA target sequences of 19 bp for *EWSR1* and *WT1* are cloned into the empty vector at the BbsI restriction site. Two gRNA for each gene will be tested in this chapter.

The choice of gRNA targets is based on: a) >200 bp from *EWSR1* and *WT1* exons to reduce the chance of indels in the coding sequence of the non-targeted alleles and b) high target specificity and low numbers of off-target sites, as determined using the online CRISPR Design Tool (<http://tools.genome-engineering.org>) (Fig. S1).

Oligos for *EWSR1* and *WT1* gRNAs:

gRNA ^{<i>EWSR1-2</i>}	[25, 30]
Sense	5'-CACCGGGGCATCCAAGATGTTAGC-3'
Antisense	5'-AAACGCTAACATCTTGGATGCCCC-3'
gRNA ^{<i>EWSR1-9</i>}	
Sense	5'-CACCGATGCCCTAAAGATGTGTCC-3'
Antisense	5'-AAACGGACACATCTTTAGGGCATC-3'
gRNA ^{<i>WT1-1</i>}	[25]
Sense	5'-CACCGTGAGCACGCCTTCTATGCC-3'
Antisense	5'-AAACGGCATAGAAGGCGTGCTCAC-3'
gRNA ^{<i>WT1-2</i>}	[25]
Sense	5'-CACCGGGCTGAGCCCTTTATGTGA-3'
Antisense	5'-AAACTCACATAAAGGGCTCAGCCC-3'

Underlined sequences represent DNA sequences bound by the gRNAs. Note: it is not necessary to phosphorylate the oligos.

2.3.1.1 Digestion of PX458: Digest PX458 (1 µg) with BbsI restriction enzyme (NEB catalog # R0539S). A 20 µl reaction is incubated for 1 h at 37°C according to the manufacturer's recommendations and the digested product is gel purified. This linearized plasmid can be stored at -20°C and used for subsequent cloning.

2.3.1.2 Oligo annealing (Modified from [31])

gRNA oligos sense and antisense (100 µM)	1+1 µl
ddH ₂ O	8 µl
Total	10 µl

Incubate in a thermocycler at the following parameters: 37°C for 30 min, 95°C for 5 min. Ramp down to 25°C at 5°C/min.

2.3.1.3 Ligation (Roche catalog # 04898117001):

Digested PX458 (from 2.3.1.1) (50 ng)	x µl
Annealed oligos (1:250)	1 µl
DNA dilution buffer	2 µl
ddH ₂ O	up to 10 µl

Mix, add 10 µl T4 DNA ligation buffer and 1 µl T4 DNA ligase. Incubate 5' at room temperature and use 5 µl for bacterial transformation with 10 µl of chemically competent *E. coli*. Correct integration was verified by Sanger sequencing.

Sequencing primer: PX458SeqVector 5'-AACGCGGCCTTTTACGGTT-3'

2.3.2 Donor plasmid—The donor plasmid is constructed from pBluescript II SK+ and contains a modified promoterless hygromycin resistance gene (*Hyg*-) site with a splice acceptor-T2A (SA-T2A) module flanked by LoxP sites [32] (Fig. 1C). The *Hyg*- gene was modified to remove its ATG to further cripple the gene. Keep in mind that hygromycin resistance depends on an inframe fusion after splicing with the upstream exon (in this case *EWSR1* exon 7).

Homology arms are suggested to be ~500 to 800 bp in length (see also [24]). Chromosomal coordinates of homology arms for the targeting strategies are as follows:

gRNA^{*EWSR1-2*}: hg38 29287068-29287592

gRNA^{*WT1-1*}: hg38 32393083-32392564

gRNA^{*EWSR1-9*}: hg38 29286942-29287561

gRNA^{*WT1-2*}: hg38 32393056-32392437

Homology arms can be obtained either by PCR amplification of genomic DNA from cells of interest (hMSCs in this study) or by DNA synthesis (Life Technologies) of the region of interest using the reference genome sequence. The synthetic fragment is simpler as it can be designed to have compatible restriction sites for cloning and can be modified to alter the T2A reading frame after splicing or some other aspect of the selectable marker exon. However, since polymorphisms can exist in the genome of the cells of interest compared to the reference genome, leading to mismatches with the homology arms and reducing HDR efficiency, it is recommended that the homology region be amplified directly from genomic DNA and sequenced prior to synthesizing a homology arm fragment.

2.4 Non-immortalized human mesenchymal cell line transfection

Non-immortalized hMSCs are grown on gelatinized plates in α-MEM (Life Technologies, GIBCO 12571-063) media with 10% Hyclone FBS (Thermo Fisher Scientific SH30070.03). Cells are transfected using the Amaxa nucleofector and homemade (HM) buffer [24] with 2.5 µg each gRNA (*EWSR1* and *WT1*) and donor fragment plasmid or with empty vectors as controls. Plasmids are prepared using Invitrogen Purelink midiprep (catalog # K210014).

To assess transfection efficiency, GFP fluorescence from the Cas9-GFP fusion encoded by PX458 can be measured after 24 or 48 h by flow cytometry (FACScan Becton-Dickinson).

To assess gRNA efficiency, NHEJ-based translocation breakpoint junctions are PCR amplified 48 h after transfection of *EWSR1* and *WT1* gRNA pairs without the donor fragment plasmid.

To obtain translocations, selection is carried out with 50 µg/ml hygromycin (TOKU-E catalog # HO11), a concentration determined to be sufficient to kill parental cells.

Nucleofection and selection (Fig. 3)

1. Split low passages cells 24 h before transfection from a 60 mm plate (~95% confluency) to 2 × 60 mm plates.
2. Trypsinize and resuspend cells in prewarmed 10 ml media in 15 ml tubes.
3. Spin down 1×10^6 cells at 1000 rpm for 5 min. Wash once with 5 ml PBS.
4. Resuspend in 100 µl of HM buffer, combine with plasmid DNA and transfer to a Gene Pulser Cuvette, 0.2 cm electrode gap (Bio-Rad catalog # 165-2086).
5. Transfect cells with Amaxa nucleofector II (B-016 program).
6. Transfer cells immediately to 10 ml media, count and seed 3 × 96 wells with 1000 cells/well in media containing DNA-PKcs inhibitor (NU7441) at a final concentration of 0.5 µM.
7. 2 days after transfection remove media and replace with fresh media without the inhibitor.
8. Once the wells are confluent (~3-4 days) start selection with media containing hygromycin.

To gauge a suitable DNA-PKcs inhibitor (NU7441) concentration, cells were exposed for 7 days to increasing concentrations (2 µM), which were previously shown to increase HDR [33], with no observed cell death. In the experimental condition, cells were exposed to 0.5 µM NU7441 for 2 days to minimize any mutagenesis resulting from DNA-PKcs inhibition.

2.5 Translocation breakpoint analysis

As cells become confluent, remove media from wells and wash with PBS. Cells are trypsinized with 30 µl trypsin, and 7 µl of trypsinized cells added to an equal amount of 2X lysis buffer (10 mM Tris pH 8, 0.45% NP40, 0.45% Tween 20, 100 µg/ml Proteinase K) in 0.2 ml PCR tubes followed by incubation at 55 °C for 2 h and then 95 °C for 5 min. The remaining cells from each well are transferred to a single well in a 24 well plate in media containing hygromycin for further expansion.

A 10× stock solution of lysis buffer can be prepared and maintained at 4°C, while proteinase K needs to be added fresh each time. Genotyping is performed using 0.5 µl of cell lysate or 40 ng of purified genomic DNA in a 15 µl PCR reaction. Genotyping of the HDR-generated *EWSR1-Hyg-WT1* and *EWSR1-WT1* 5' and 3' breakpoint junctions (out-in and in-out PCR), *WT1-EWSR1* reciprocal translocation junctions, and NHEJ-generated *EWSR1-WT1* breakpoint junctions is performed using Dream Taq Green PCR Master Mix 2X (Thermo Scientific catalog # K1081), while amplification across the insert (out-out PCR) is done with

a combination of Taq DNA polymerases (Fisher Scientific catalog # FEREP0406) and Pfu turbo (Agilent Technologies catalog # 600254) with Jeffreys' PCR Buffer (450 mM Tris HCl pH 8.8, 110 mM (NH₄)₂SO₄, 45 mM MgCl₂, 67 mM β-mercaptoethanol, 44 μM EDTA, 10 mM dATP, 10 mM dCTP, 10 mM dGTP, 10 mM dTTP, 1.13 mg/ml BSA, 12.5 mM Tris).

PCR primers for EWSR1-Hyg-WT1 and EWSR1-WT1 breakpoint junctions:

*EWSR1*ext1: 5'-TCTCAGCAGAACACCTATGG-3' [25]

*WT1*ext1: 5'-AGGAGGAACATCTCCAGAGA-3' [25]

*EWSR1*ext2: 5'-GGATGTCTGTGTACATGGT-3'

*WT1*ext2: 5'-CGCTACAAATTGGATTCCGC-3'

Hygro-For1: 5'-GTATCACTGGCAAAGTGTGATGG-3' [25]

Hygro-For2: 5'-GCGACGTCTGTGCGAGAAGTT-3'

Hygro-Rev: 5'-CCACTATCGGCGAGTACTTC-3' [25]

PCR primers for WT1-EWSR1 reciprocal translocation junctions:

*WT1*intron7: 5'-GGATTCTCCTAAGAAGGTGG-3' [25]

*EWSR1*exon8: 5'-GTTATCAGGGCCACTCATGC-3' [25]

PCR program

1. 96°C, 3 min
2. 96°C, 20 s
3. 60°C, 30 s (58° for out-out PCR)
4. 65°C, 2 min (3.5 min for out-out PCR)
5. Repeat steps 2-4 30 times.
6. 65°C, 5 min
7. 12°C, holding temperature

2.6 Cre recombinase viral infection

Cells in which the translocation is verified by PCR across the breakpoint junction (out-out) are infected with or without a self-deleting lentivirus expressing Cre recombinase [34] and harvested for PCR and RT-PCR analysis.

2.6.1 Virus production

1. Day 0: seed HEK293T cells at 6 million per 10 cm dish in 10 ml DMEM 10% FBS.
2. Day 1: transfect packaging cells using Lipofectamine 2000 (Thermo Fisher Scientific catalog # 11668027).
 - a. add to 500 μl OPTIMEM:
 - i. 4.5 μg of psPAX2 (Addgene # 12260)
 - ii. 1.0 μg of pCMV-VSV-G (Addgene # 8454)

- iii. 4.5 µg of Lenti-Cre-SD [34].
 - b. Add 30 µl Lipofectamine reagent to another tube with 500 µl OPTIMEM.
 - c. Mix the two tubes and let sit for 5 min.
 - d. Mix the reagents in the two tubes by pipetting up and down then let sit for 25 min.
 - e. Add directly to cells drop by drop, followed by 10 ml media.
3. Day 3: virus is harvested 48 h and 72 h after transfection. Store at 4 °C up to a week, aliquot and store at -80 °C for long term.

2.6.2 Transduction

1. Day 0: plate cells at ~ 30000 cells per well in a 12 wells plate.
2. Day 2: infect cells. Prepare an infection solution containing virus and polybrene (final concentration 8 µg/ml) to increase infection efficiency. The amount of virus needs to be optimized by titration.
3. Day 3: 24 h after transduction, remove media and add fresh media.
4. Day 5: Harvest cells and proceed with genomic DNA and RNA extraction.

2.7 RNA extraction and RT-PCR

For RNA extraction the RNeasy mini kit was used (Qiagen) and cDNA was synthesized from 100 ng RNA using the Superscript III First Strand cDNA Syhthesis Kit (Life Technologies catalog # 18080051). RT-PCR amplification was obtained using Taq Green PCR master mix with 30 ng of cDNA as template.

PCR primers [25]:

*EWSR1*ext1: 5'-TCTCAGCAGAACACCTATGG-3'

Hygro-Rev: 5'-CCACTATCGGCGAGTACTTC-3'

WT1-exon10 5'-GACCGGGCAAACCTTTTCTG-3'

EWSR1-ex6 5'-GTAACACAGTTATCCCCAG-3'

EWSR1-ex8 5'-GTTATCAGGGCCACTCATGC-3'

*GAPDH-F*5'-GAGGGGCCATCCACAGTCTTCT-3'

*GAPDH-R*5'-GGAGCCAAAAGGGTCATCATCT-3'.

PCR Program

1. 96°C, 3 min
2. 96°C, 20 s
3. 60°C, 30 s
4. 65°C, 30 s
5. Repeat steps 2-4 30 times.

6. 65°C, 5 min
7. 12°C, holding temperature

3. Results

We verified the functionality of the gRNAs pairs by PCR amplification of NHEJ-generated *EWSR1-WT1* breakpoint junctions. hMSCs were transfected with Cas9 and gRNA expression vectors for each gene and genomic DNA was isolated 48 h later for breakpoint junction PCR. All four pairs of gRNAs gave rise to PCR products of the size expected for translocations (Fig. 4A). No such products were observed using a single gRNA.

We first focused on one pair, gRNA^{*EWSR1-2*} and gRNA^{*WT1-1*} [25], which gave particularly good scores for target specificity (Fig. S1). NHEJ-mediated translocation frequency was estimated by PCR amplification using serial dilutions of genomic DNA [13] (Fig. 4A). Translocation frequency was 1.25×10^{-4} , in agreement with previous measurements carried out in different human cell lines [9, 11, 12].

Given the low frequency of translocations, the HDR strategy was employed to enrich for cells that would have undergone a translocation (Fig. 1C). hMSCs were transfected with Cas9 and gRNA expression vectors plus the donor plasmid and seeded in small pools (Fig. 3). NHEJ was transiently inhibited using a DNA-PKcs inhibitor and when cells were confluent hygromycin was added. Cells in 25 wells from three 96-well plates grew in hygromycin; the low frequency of positive wells suggests that cells in most wells arise from a single event, i.e., are clonal.

After selection with hygromycin, resistant cells were lysed and genotyped by PCR for the correct integration of the donor fragment at the 5' and 3' junctions (out-in and in-out PCR; Fig. 4B). All 25 clones analyzed showed the correct integration at the 5' junction, demonstrating the efficacy of the gene trap strategy, and 40% were also positive for the 3' junction (10 positive clones out of 25 analyzed).

Cells were further tested for the *EWSR1-Hyg-WT1* fusion (out-out PCR), and 3 positive clones were obtained from the 10 clones positive for the 3' junction (e.g., Fig. 4B), indicating a translocation. Clones positive for the 5' and 3' integration but not the *EWSR1-Hyg-WT1* fusion could be explained by independent integration of the donor fragment separately at both the *EWSR1* and *WT1* loci; alternatively, integration of the donor fragment may have rendered PCR detection inefficient (e.g., a tandem integration of the donor fragment).

One of the clones harboring the *EWSR1-Hyg-WT1* fusion, clone 2, could be further expanded and cells were infected with lentivirus expressing Cre. Viral transduction was found to be necessary in order to deliver Cre to the highest number of cells in one round. Analysis of genomic DNA showed that cells exposed to the Cre lentivirus had lost the marker, as noted by shift in the PCR product size, generating the *EWSR1-WT1* fusion (Fig. 4C).

To demonstrate that expression of the *EWSR1-WT1* fusion product is dependent on Cre-mediated removal of the marker cassette, we performed RT-PCR. The *EWSR1-Hyg* transcript was only detected prior to Cre expression, demonstrating the high Cre lentiviral transduction efficiency, while the *EWSR1-WT1* fusion transcript was detected in cells only after Cre expression, demonstrating that conditional expression was achieved (Fig. 4D). Furthermore, conditional expression of the EWSR1-WT1 fusion protein was detected by Western blot analysis, and strong induction of a known target of the EWSR1-WT1 fusion, PDGF-A, was also observed [25].

The chromosomal translocation was confirmed by fluorescence in situ hybridization, demonstrating a reciprocal translocation and unarranged chromosomes 11 and 22 [25]. The reciprocal translocation could be generated through NHEJ giving rise to the *WT1-EWSR1* fusion (Fig. 2A). This fusion was observed by breakpoint junction PCR, which, as expected, was not affected by Cre expression (Fig. 4E). Sequencing of the PCR fragment demonstrated joining of the *WT1* and *EWSR1* ends [25].

We also examined expression of the unarranged *EWSR1* and *WT1* genes. *EWSR1* is normally expressed in hMSCs, and we observed its continued expression in cells with the translocation (Fig. 4F). *WT1* is not ubiquitously expressed in all tissues [35], including hMSCs, and its expression was not observed in cells with the translocation.

We also generated the *EWSR1-WT1* translocation using a different pair of gRNAs, gRNA^{*EWSR1-9*} and gRNA^{*WT1-2*}, and a donor fragment with somewhat different homology arms. (Fig. 4G). In this case, 19 hygromycin-resistant clones were obtained, 18 of which showed the correct 5' junction, and 4 of the 18 were also positive for the 3' junction. One of these 4 clones, clone 18, gave rise to the PCR product expected for the translocation leading to *EWSR1-Hyg-WT1* (out-out PCR).

4. Discussion

The CRISPR-Cas9 translocation strategy described here is based on integration of a promoterless selection marker by HDR at DNA ends from two different chromosomes. By using a gene trap approach, this system allows for the enrichment of clones containing a translocation. Further, conditional expression of the fusion product is achieved through Cre expression.

Compared with earlier systems involving direct fusion of two genes by NHEJ after nuclease-generated DSB formation, the system presented here allows the effective recovery of clones harboring the desired translocation. The system also presents several advantages compared to those that use patient-derived cell lines or ectopic expression of the gene fusion. For example, fusion gene expression is conditional, allowing the analysis of the early steps of tumor development. Further, once induced by Cre, fusion gene expression is regulated by the endogenous promoter (e.g., *EWSR1* in the example shown here). The protocol described here generates the *EWSR1-WT1* translocation in non-immortalized hMSCs, but the system has also proved to be effective in immortalized hMSCs and transformed HEK293 cells [25,

30]. Theoretically, the approach is applicable to unrelated translocations and genomic rearrangements as well as other cell lines.

To attempt to increase HDR events, we transiently exposed cells to sublethal doses of a DNA-PKcs inhibitor immediately after transfection. The rationale for this approach is that inhibition of NHEJ repair of Cas9-generated DSBs increases the frequency of HDR [33, 36]. Thus, the expectation was that integration of the donor by HDR would be increased by transiently impairing the NHEJ pathway. When cells were not treated with a DNA-PKcs inhibitor, we previously observed a reduction in the number of hygromycin-resistant clones, fewer of these had the correct 3'- junction, and none was positive for amplification across the translocation [25].

We expect that the efficacy of the protocol could be improved by additional modifications. For example, a puromycin-based expression plasmid to deliver the gRNA and Cas9 (PX459, Addgene #48139) would allow the transient selection of cells that had been successfully transfected. In principle, transfected cells could be also enriched by sorting cells expressing GFP from PX458, although cell sorting may place unnecessary stress on primary cells.

Finally, as for all approaches using genome-editing techniques, it is important to consider the specificity of Cas9. Specificity can be increased by using Cas9 variants eSpCas9 and SpCas9-HF1 that retain on-target activity but reduce off-target activity [37, 38]. Alternatively, instead of wild-type Cas9, paired Cas9 nickases, which efficiently stimulate gene targeting [17] and have been shown to give rise to translocations [9], can also be considered for generating translocations by HDR. Careful selection of gRNA target sites is also important. For this purpose we used one of the available bioinformatics tool (see section 2.3.1) that provides a list of potential target sites with a score indicating the faithfulness (i.e., on-target activity) and the predicted number of off-target sites. Other software, which uses different algorithms, e.g., Cas-OFFinder [39] and GuideScan [40], is also available for the design of gRNAs. (For a more complete list, see [41]).

5. Conclusions

The protocol described here represents a flexible and inexpensive method for the generation of cells harboring chromosomal translocations and other rearrangements, and may become an important tool for the scientific community interested in cancer and other diseases whose study has been hindered by the paucity of available cell lines with desired genomic rearrangements.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors wish to thank Travis White for critical reading of the manuscript, and Marc Ladanyi and Lee Spraggon for initial discussions and reagents. MSK research is supported by NIH/NCI Cancer Center support grant P30 CA008748. This work was supported by an Alex's Lemonade Stand Innovation Award, NIH R01CA185660 and R35GM118175 (M.J.).

References

1. Mani RS, Chinnaiyan AM. *Nat Rev Genet.* 2010; 11:819–829. [PubMed: 21045868]
2. Elliott B, Jasin M. *Cell Mol Life Sci.* 2002; 59:373–385. [PubMed: 11915950]
3. Lieber MR. *Nat Rev Cancer.* 2016; 16:387–398. [PubMed: 27220482]
4. Richardson C, Jasin M. *Nature.* 2000; 405:697–700. [PubMed: 10864328]
5. Elliott B, Richardson C, Jasin M. *Molecular cell.* 2005; 17:885–894. [PubMed: 15780943]
6. Weinstock DM, Elliott B, Jasin M. *Blood.* 2006; 107:777–780. [PubMed: 16195334]
7. Simsek D, Jasin M. *Nature structural & molecular biology.* 2010; 17:410–416.
8. Simsek D, Brunet E, Wong SY, Katyal S, Gao Y, McKinnon PJ, Lou J, Zhang L, Li J, Rebar EJ, Gregory PD, Holmes MC, Jasin M. *PLoS genetics.* 2011; 7:e1002080. [PubMed: 21655080]
9. Ghezraoui H, Piganeau M, Renouf B, Renaud JB, Sallmyr A, Ruis B, Oh S, Tomkinson AE, Hendrickson EA, Giovannangeli C, Jasin M, Brunet E. *Molecular cell.* 2014; 55:829–842. [PubMed: 25201414]
10. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, Miller JC, Urnov FD, Gregory PD, Holmes MC. *Nat Methods.* 2011; 8:74–79. [PubMed: 21131970]
11. Brunet E, Simsek D, Tomishima M, DeKolver R, Choi VM, Gregory P, Urnov F, Weinstock DM, Jasin M. *Proceedings of the National Academy of Sciences of the United States of America.* 2009; 106:10620–10625. [PubMed: 19549848]
12. Piganeau M, Ghezraoui H, De Cian A, Guittat L, Tomishima M, Perrouault L, Rene O, Katibah GE, Zhang L, Holmes MC, Doyon Y, Concordet JP, Giovannangeli C, Jasin M, Brunet E. *Genome Res.* 2013; 23:1182–1193. [PubMed: 23568838]
13. Renouf B, Piganeau M, Ghezraoui H, Jasin M, Brunet E. *Methods Enzymol.* 2014; 546:251–271. [PubMed: 25398344]
14. Torres R, Martin MC, Garcia A, Cigudosa JC, Ramirez JC, Rodriguez-Perales S. *Nat Commun.* 2014; 5:3964. [PubMed: 24888982]
15. Choi PS, Meyerson M. *Nat Commun.* 2014; 5:3728. [PubMed: 24759083]
16. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. *Science.* 2012; 337:816–821. [PubMed: 22745249]
17. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. *Science.* 2013; 339:819–823. [PubMed: 23287718]
18. Rouet P, Smih F, Jasin M. *Proceedings of the National Academy of Sciences of the United States of America.* 1994; 91:6064–6068. [PubMed: 8016116]
19. Liang F, Han M, Romanienko PJ, Jasin M. *Proceedings of the National Academy of Sciences of the United States of America.* 1998; 95:5172–5177. [PubMed: 9560248]
20. Moynahan ME, Jasin M. *Nat Rev Mol Cell Biol.* 2010; 11:196–207. [PubMed: 20177395]
21. Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC. *Nature.* 2005; 435:646–651. [PubMed: 15806097]
22. Joung JK, Sander JD. *Nat Rev Mol Cell Biol.* 2013; 14:49–55. [PubMed: 23169466]
23. Mali P, Esvelt KM, Church GM. *Nat Methods.* 2013; 10:957–963. [PubMed: 24076990]
24. Zhang Y, Vanoli F, LaRocque JR, Krawczyk PM, Jasin M. *Methods.* 2014; 69:171–178. [PubMed: 24929070]
25. Vanoli F, Tomishima M, Feng W, Lamribet K, Babin L, Brunet E, Jasin M. *Proceedings of the National Academy of Sciences of the United States of America.* 2017
26. Elliott B, Richardson C, Winderbaum J, Nickoloff JA, Jasin M. *Molecular and cellular biology.* 1998; 18:93–101. [PubMed: 9418857]
27. Gerald WL, Rosai J. *Pediatr Pathol.* 1989; 9:177–183. [PubMed: 2473463]
28. Gerald WL, Rosai J. *Zentralbl Pathol.* 1993; 139:141–151. [PubMed: 8396418]
29. Barberi T, Willis LM, Socci ND, Studer L. *PLoS Med.* 2005; 2:e161. [PubMed: 15971941]
30. Spraggon L, Martelotto LG, Hmeljak J, Hitchman TD, Wang J, Wang L, Slotkin EK, Fan PD, Reis-Filho JS, Ladanyi M. *The Journal of pathology.* 2017

31. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Nat Protoc. 2013; 8:2281–2308. [PubMed: 24157548]
32. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R. Nature biotechnology. 2011; 29:731–734.
33. Vriend LE, Prakash R, Chen CC, Vanoli F, Cavallo F, Zhang Y, Jasin M, Krawczyk PM. Nucleic acids research. 2016; 44:5204–5217. [PubMed: 27001513]
34. Pfeifer A, Brandon EP, Kootstra N, Gage FH, Verma IM. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:11450–11455. [PubMed: 11553794]
35. Pritchard-Jones K, Fleming S, Davidson D, Bickmore W, Porteous D, Gosden C, Bard J, Buckler A, Pelletier J, Housman D, et al. Nature. 1990; 346:194–197. [PubMed: 2164159]
36. Pierce AJ, Hu P, Han M, Ellis N, Jasin M. Genes & development. 2001; 15:3237–3242. [PubMed: 11751629]
37. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Science. 2016; 351:84–88. [PubMed: 26628643]
38. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales AP, Li Z, Peterson RT, Yeh JR, Aryee MJ, Joung JK. Nature. 2015; 523:481–485. [PubMed: 26098369]
39. Bae S, Park J, Kim JS. Bioinformatics. 2014; 30:1473–1475. [PubMed: 24463181]
40. Perez AR, Pritykin Y, Vidigal JA, Chhangawala S, Zamparo L, Leslie CS, Ventura A. Nature biotechnology. 2017
41. Tycko J, Myer VE, Hsu PD. Molecular cell. 2016; 63:355–370. [PubMed: 27494557]

Highlights

- CRISPR-Cas9-generated DSBs induce chromosomal translocations by HDR
- An HDR strategy allows selection of clones with the desired translocation
- Expression of the fusion gene is conditional and depends on Cre recombinase
- EWSR1-WT1 translocation t(11;22)(p13;q12) found in DSRCT is modeled

Fig. 1. Strategy to generate chromosomal translocations using a gene trap and CRISPR-Cas9

A. DNA DSBs (scissors) are induced by Cas9 and gRNAs at the two loci involved in the chromosomal translocation.

B. Translocations can form by NHEJ though the direct joining of DNA ends.

C. Translocation strategy using HDR. A DNA donor fragment that contains a promoterless hygromycin-resistance gene (*Hyg*⁻) with a polyadenylation (pA) signal sequence flanked by homology arms (shaded regions) for genes on chromosome A and B is inserted at the two cleaved chromosome ends by HDR. A splice acceptor-T2A (SA-T2A) sequence before the *Hyg*⁻ selectable marker leads to its expression (*Hyg*⁺) from the endogenous promoter of the gene on chromosome A.

D. In cells harboring the HDR-mediated chromosomal translocation shown in (C), expression of the fusion depends on Cre recombinase expression, which leads to removal of the selectable marker at the LoxP sites.

Fig. 2. Generation of the EWSR1-WT1 chromosomal translocation

A. DNA cleavage within the *EWSR1* and *WT1* genes on chromosomes 22 and 11 generates the $t(11;22)(p13;q12)$ translocation thereby creating the *EWSR1-WT1* gene fusion on derivative chromosome 22 and the reciprocal *WT1-EWSR1* gene fusion on derivative chromosome 11.

B. The *EWSR1-WT1* fusion product results from joining of DSBs in intron 7 of each gene. The fusion protein has the transactivation domain (TA) from EWSR1 and three DNA-binding zinc finger motifs from WT1.

Fig. 3. Workflow for generation of chromosomal translocation in non-immortalized hMSCs

Day 0: cells are transfected with the donor fragment and Cas9-gRNAs plasmids and plated in 96-well plates (~1000cells/well) in low dose of DNA-PKcs inhibitor. Day 2: media containing the DNA-PKcs inhibitor is removed and replaced with fresh media to allow cell expansion. Day 5-6: cells are exposed to hygromycin for selection. Day 11: wells with hygromycin-resistant cells are analyzed by PCR for correct integration of the donor fragment. The initial number of cells per well can be increased to accommodate a low frequency of translocation and/or plating efficiency.

E. Reciprocal translocation analysis. PCR with primers amplifying across the *WT1-EWSR1* junction (black arrows) shows a product expected from the reciprocal translocation.

F. *EWSR1* expression from the unarranged allele is not significantly altered in cells expressing the *EWSR1-WT1* fusion compared with parental cell line. The *WT1* transcript is not detected in either the parental cells or cells harboring the translocation.

D. Targeting strategy using a different gRNA pair (9 + 2) and homology arms. One of the clones analyzed (clone 18) shows the expected PCR products for the *EWSR1-Hyg-WT1* fusion.