



Article

Cruciform DNA Structures Act as Legible Templates for Accelerating Homologous Recombination in Transgenic Animals

Huan Ou-Yang ^{1,2} , Shiao-Hsuan Yang ^{1,3}, Wei Chen ⁴, Shang-Hsun Yang ^{5,6} , Abdulkadir Cidem ^{1,7} ,
Li-Ying Sung ^{2,*} and Chuan-Mu Chen ^{1,8,9,*}

- ¹ Program in Translational Medicine, Department of Life Sciences, National Chung Hsing University, Taichung 402, Taiwan; huan4096052151@alumni.nchu.edu.tw (H.O.-Y.); 178241@cch.org.tw (S.-H.Y.); cidema.kadir@gmail.com (A.C.)
 - ² Institute of Biotechnology, College of Bioresources and Agriculture, National Taiwan University, Taipei 106, Taiwan
 - ³ Reproductive Medicine Center, Department of Gynecology, Changhua Christian Hospital, Changhua 515, Taiwan
 - ⁴ Division of Pulmonary and Critical Care Medicine, Chia-Yi Christian Hospital, Chiayi 600, Taiwan; peteralfa2004@gmail.com
 - ⁵ Department of Physiology, National Cheng Kung University, Tainan 701, Taiwan; syang@mail.ncku.edu.tw
 - ⁶ Institute of Basic Medical Sciences, College of Medicine, National Cheng Kung University, Tainan 701, Taiwan
 - ⁷ Department of Molecular Biology and Genetics, Erzurum Technical University, Erzurum 25250, Turkey
 - ⁸ The iEGG and Animal Biotechnology Center, National Chung Hsing University, Taichung 402, Taiwan
 - ⁹ Rong-Hsing Translational Medicine Research Center, Taichung Veterans General Hospital, Taichung 407, Taiwan
- * Correspondence: liyingsung@ntu.edu.tw (L.-Y.S.); chchen1@dragon.nchu.edu.tw (C.-M.C.); Tel.: +886-2-33666015 (L.-Y.S.); +886-4-22856309 (C.-M.C.)



Citation: Ou-Yang, H.; Yang, S.-H.; Chen, W.; Yang, S.-H.; Cidem, A.; Sung, L.-Y.; Chen, C.-M. Cruciform DNA Structures Act as Legible Templates for Accelerating Homologous Recombination in Transgenic Animals. *Int. J. Mol. Sci.* **2022**, *23*, 3973. <https://doi.org/10.3390/ijms23073973>

Academic Editors: Mariarita De Felice and Mariarosaria De Falco

Received: 26 February 2022

Accepted: 29 March 2022

Published: 2 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Inverted repeat (IR) DNA sequences compose cruciform structures. Some genetic disorders are the result of genome inversion or translocation by cruciform DNA structures. The present study examined whether exogenous DNA integration into the chromosomes of transgenic animals was related to cruciform DNA structures. Large imperfect cruciform structures were frequently predicted around predestinated transgene integration sites in host genomes of microinjection-based transgenic (Tg) animals (α LA-LPH Tg goat, $Akr1A1^{eGFP/eGFP}$ Tg mouse, and $NF\kappa B$ -Luc Tg mouse) or CRISPR/Cas9 gene-editing (GE) animals (α LA-API GE mouse). Transgene cassettes were imperfectly matched with their predestinated sequences. According to the analyzed data, we proposed a putative model in which the flexible cruciform DNA structures acted as a legible template for DNA integration into linear DNAs or double-strand break (DSB) alleles. To demonstrate this model, artificial inverted repeat knock-in (KI) reporter plasmids were created to analyze the KI rate using the CRISPR/Cas9 system in NIH3T3 cells. Notably, the KI rate of the 5' homologous arm inverted repeat donor plasmid (5'IR) with the ROSA gRNA group (31.5%) was significantly higher than the knock-in reporter donor plasmid (KIR) with the ROSA gRNA group (21.3%, $p < 0.05$). However, the KI rate of the 3' inverted terminal repeat/inverted repeat donor plasmid (3'ITRIR) group was not different from the KIR group (23.0% vs. 22.0%). These results demonstrated that the legibility of the sequence with the cruciform DNA existing in the transgene promoted homologous recombination (HR) with a higher KI rate. Our findings suggest that flexible cruciform DNAs folded by IR sequences improve the legibility and accelerate DNA 3'-overhang integration into the host genome via homologous recombination machinery.

Keywords: inverted repeat (IR); cruciform; transgene; double-strand break (DSB); DNA repair; homologous recombination (HR); gene edit; CRISPR/Cas9; knock-in; genome stability

1. Introduction

Directed repeats (DRs) and inverted repeats (IRs) create genome instability and have many functions via the formation of specific secondary structures in the genome [1]. The genome instability of IRs causes cruciform extrusion [2,3]. IRs are partially mutated or deleted during replication failure or enzyme digestion *in vivo* [4]. Some environmental factors affect DNA conformation, such as salt conditions, epigenetic modification, or DNA-binding protein interactions. For example, 14-3-3 cruciform-binding protein binds to cruciform DNAs and regulates eukaryotic DNA replication [5]. The nucleotide composition of the IR structure is approximately 150 to 200 bp. However, the existence of mismatches and spacers in the IRs' structures improve the stability of IRs [6].

The instability of the DNA structure due to directed or inverted repeat sequences is the primary cause of genetic disorders, such as Huntington's disease or type A hemophilia. Type A hemophilia occurs because of clotting FVIII deficiency. Approximately 50% of FVIII-deficient patients have large fragment inversions at intron 1 (int1) and intron 22 (int22) in their *FVIII* gene (*F8*). Notably, a large IR region was found at int22 homolog 2 (int22 h2) and int22 h3 [7]. Although int1 also contains a hotspot for *F8* inversion, only one int1 h region was found. The int1 inversion rate was much lower than the int22 inversion rate [8,9]. There are many inversion hotspots, including the inverted repeat region in autosomal chromosomes, such as hVIPR2 [10] and human chromosome 15q11-q13 [11]. These reports suggest that non-allelic homologous recombination (HR) generally occurs at large inverted repeats. Therefore, we investigated whether cruciform DNA was easier to identify as a legible template for accelerating HR processing.

Current studies revealed that most virus genomes integrate into host chromosomes using inverted repeat sequence machinery, such as inverted tandem repeats (ITRs) from adeno-associated virus (AAV) families [12] and long terminal repeats (LTR) from retrovirus and lentivirus families [13]. Notably, the piggyback transposon system, which was isolated from *Cabbage looper*, contains two terminal repeats [14]. All of the integration mechanisms require specific DNA-binding proteins to import viral DNA into nuclei and recombinant viral DNA into the host genome. However, the AAV vector integrates at double-strand break (DSB) sites without DNA-binding proteins [12,15]. The AAV2 ITR mimics the DSB site to gather DNA repair enzymes [16]. These reports showed that IR sequences acted as motifs of DNA-binding proteins from viruses and interacted with the repair mechanisms of host cells. Taken together, DNA structures or specific recombinases participate in exogenous DNA integration.

Except for viral systems, other transgenic methods in animal fields showed a low efficiency of exogenous DNA integration or required a long time and a high cost to select successful recombinant cells or individuals [17]. Therefore, the discovery of new methods to improve the transgenic efficiency in animal genetic engineering is an important issue. However, the exogenous DNA integration mechanisms were not clearly elucidated using pronuclear microinjection methods with linear DNA or gene editing with CRISPR/Cas9 tools. Newly developed gene-editing tools provide more powerful methods to create site-specific gene knockout (KO) or knock-in (KI) animals. However, the KI efficiency for the large exogenous fragments of gene insertions remains low in all platforms. Single-strand DNA (ssDNA) is the most popular material to increase the KI rate for the generation of gene-edited animals [18–21]. Previous studies demonstrated that the legibility of ssDNA is better than double-stranded DNA (dsDNA) in DNA repair mechanisms.

The present study hypothesized that the stability of dsDNA with cruciform structures was a factor in template alleles for DNA repair mechanisms. Four transgenic (Tg) or gene-edited (GE) animals were used to study the structural characteristics of the surrounding sequences of integration sites before and after transgene insertion or KI using next-generation sequencing (NGS) and the bioinformatics tool of the UNAFold Web Server "DNA folding form". We constructed different artificial IR structures in the 5'- or 3'-end of transgenes to elucidate the mechanism of template allele determination that accelerates the KI rate in HR processing.

2. Results

2.1. Cruciform DNA Structures Are a Hotspot of HR in Transgenic Animal Genomes

We were interested in whether some structures surround predestinated transgene (Tg) integration sites of the three Tg animals. The Tg cassettes randomly inserted into the host genomes because those Tg animals were generated by pronuclear microinjection with linear DNA. The host genome sequences were analyzed using NGS and secondary structures were evaluated by the UNAFold Web Server to identify the repair mechanisms used for exogenous DNA integration. Notably, imperfect, large, stem-loop structures were found at approximately 2000 nucleotides (nts) around predestinated Tg integration sites from host genome sequences (Figure 1A–D). The results showed that these regions were hotspots of HR because cruciform structures were present upstream and/or downstream of the target site to create DNA instability. According to the results, we hypothesized that the Tg cassettes integrated into the host genomes via HR. Moreover, the linear Tg cassettes should be 5'- to 3'-resected by the Mre11-Rad50-Nbs1 complex (MRN complex). The linear Tg cassettes should match the predestinated template sequences in the host genomes. Therefore, we fused the sequences of the transgene (Tg) cassette 3'-overhang and the predestinated matched template sequence to predict the secondary structures for finding the match regions. Imperfect stem-loop structures were formed between the Tg cassette 3'-overhang and the predestinated matched template sequence. The data suggested that Tg cassettes imperfectly matched the template DNAs in three different analyzed transgenic animals, including α LA-LPH Tg goat (Figure 1a,b), *NFκB-Luc* Tg mouse (Figure 1c,d), and *Akr1A1^{eGFP/eGFP}* Tg mouse (Figure 1e,f). The following three observation rules were shown: First, the integration sites occurred in the loop region of cruciform structures. Second, the intervals of predestinated matched templates were replaced. Third, the Tg cassettes remained complete after integration. DNA integration may be highly tolerant of large cis-stem-loop structures using the HR mechanism.

2.2. Stability of the Secondary Structure May Relate with Sequence Modification after the HR Mechanism in the Gene-Edited Mouse Genome and α LA-LPH Tg Goat

The cruciform DNA structure frequently presented near the predestinated Tg integration sites of transgenic animals. Therefore, we further determined whether the gene-edited animals generated using knock-in methods also showed the same pattern as the Tg animals that were generated via pronuclear microinjection (Figure 2). We analyzed the secondary structure of gene knock-in surrounding sequences in the mouse ROSA locus. A strong stem-loop secondary structure was identified in the ROSA locus (Figure 2a). The ICR-*Gt(ROSA)26Sor^{em(αLa-AP1x6)BM2/M}* GE mouse was prepared using the clustered, regularly interspaced, short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9) system to induce a DSB and knock-in the α LA-AP1x6 Tg cassette in the ROSA26 locus following the HR mechanism. The repair result should be the most precise in our donor Tg cassette. However, a large unpredictable insertion (U.P. ins.) DNA sequence was observed between the Tg cassette 3'-terminus and the ROSA 3'-homologous arm (HA) (Figure 2d,e). Notably, the Gibbs free energy (dG) value of the stem-loop structure of the integrated genome sequence was higher than the vector sequence. The same situation was observed between the α LA-LPH Tg goats with or without the unpredictable insertion (Figure 3). These data suggested that dsDNA is more stable with U.P. ins. than before the insertion. This finding suggests that other repair mechanisms stabilize junction sites by modifying sequences after HR processing. The α LA-LPH Tg goat full-genome sequencing data showed that repair mechanisms recombined more than one chromosome fragment in the insertion site. These results suggest that the HR mechanism generated the U.P. ins. (Figure S1).

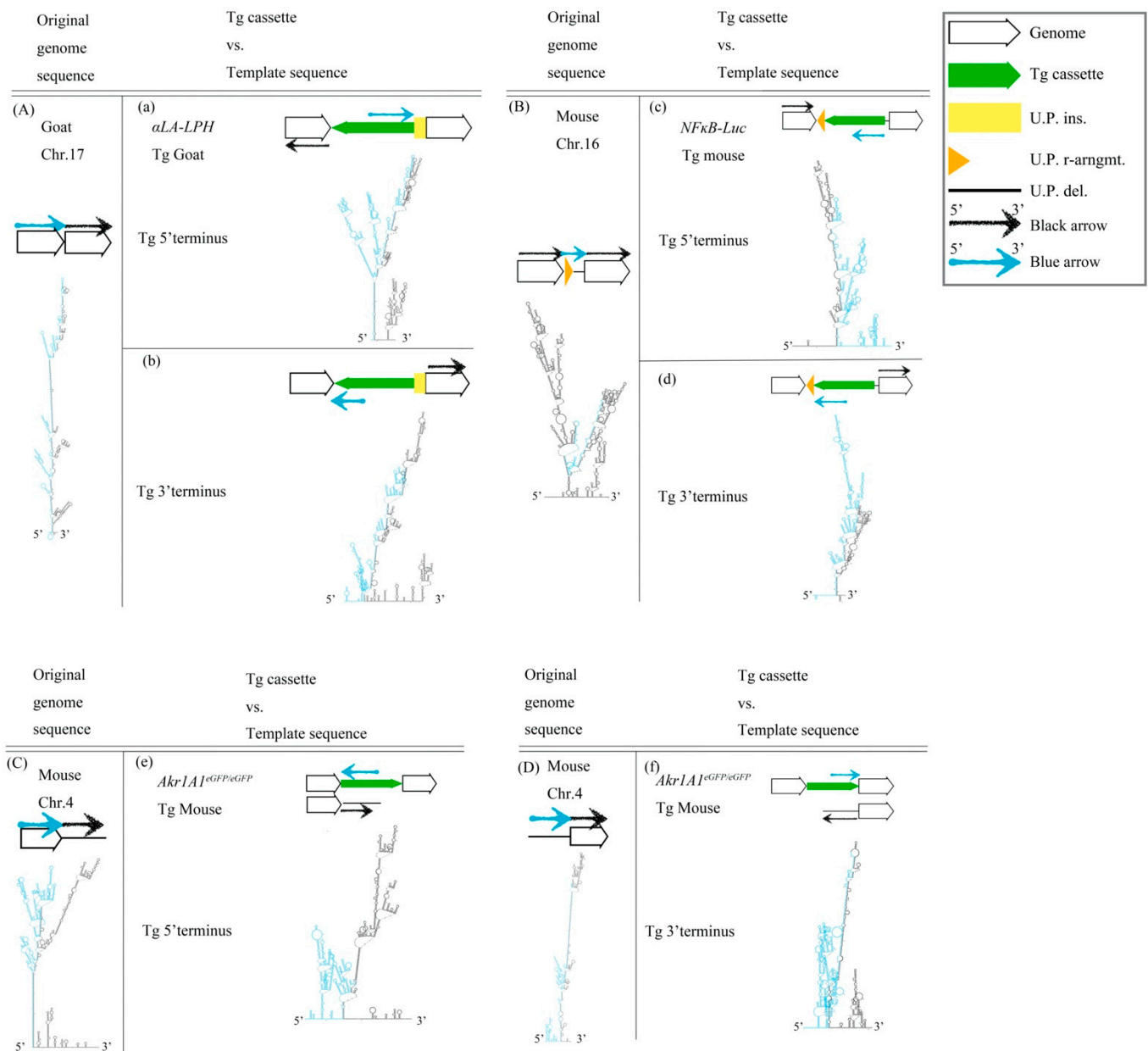


Figure 1. DNA structure of the Tg animal genome before and after Tg integration. (A) Goat chr.17 locus, (B) Mouse chr.16 locus, (C) 5' terminus, and (D) 3' terminus of mouse chr.4. (a–f) Descriptions of predicted secondary structures of fusion sequences from the Tg cassette and the predetermined template allele. These secondary structures show the situation of the two sequence assemblies. All sequences are provided in Table S1. Tg cassette, transgene cassette; U.P. ins., unpredictable insertion; U.P. del., unpredictable deletion; U.P. r-arrgmt., unpredictable rearrangement. Blue arrow, blue sequences in secondary structures. Black arrow, black sequences in secondary structures.

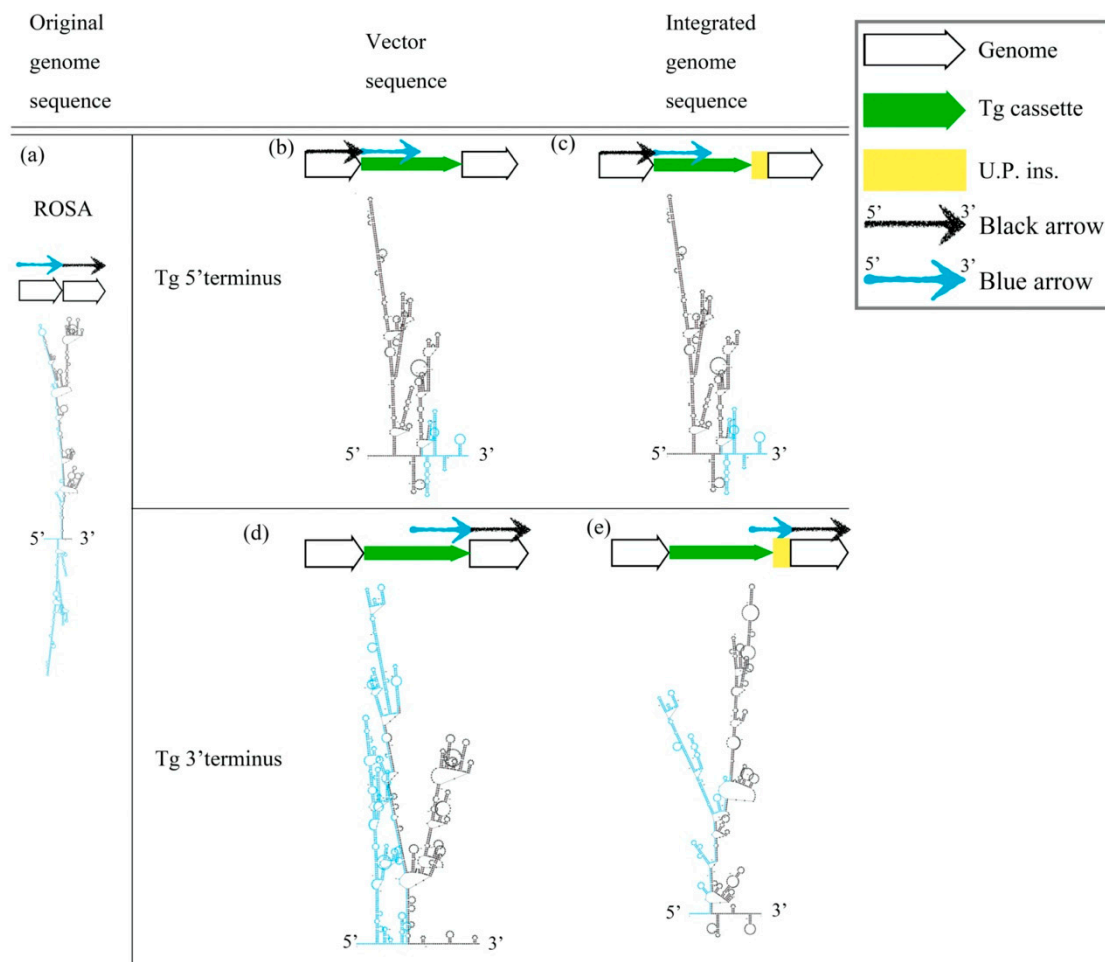


Figure 2. Prediction of the secondary structure of knock-in Tg in the ROSA locus. (a) The secondary structure of the mouse ROSA26 locus. (b) The 5' terminal secondary structures of the pDonor- α LA-Ap1x6-mROSA plasmid and (d) the 3' terminal secondary structures. Secondary structures of the genome sequence without U.P. ins. (c) and with U.P. ins. (e) from ICR-*Gt(ROSA)26Sor^{em}(α LA-Ap1x6)BM2/M*. All sequences are provided in Table S1. Tg cassette, transgene cassette; U.P. ins., unpredictable insertion. Blue arrow, blue sequences in secondary structures. Black arrow, black sequences in secondary structures.

2.3. The Cruciform DNA Structure in HA Improves KI Efficiency

We wanted to explain whether the cruciform DNA structure is a legible template for HR. Therefore, three artificial inverted repeat (IR) HA donor plasmids, including a standard knock-in reporter (KIR), a 5'-inverted repeat knock-in reporter (5'IR), and a 3'-extra inverted terminal repeat (ITR)/inverted repeat knock-in reporter (3'ITRIR), were designed to elucidate whether the cruciform DNA actually improved the legibility of HR efficiency in NIH/3T3 cells (Figure 4a). Our hypothesis was that the artificial inverted repeat HA in the donor plasmid would accelerate HR efficiency when a double-strand break (DSB) was created by the CRISPR/Cas9 machinery. The data showed that the KI efficiency of the 5'IR donor with ROSA gRNA (31.5%) was significantly higher than that with the KIR donor (21.3%) ($p < 0.05$). However, the KI efficiency of the 5'IR donor without the ROSA gRNA group was not different from the KIR without the ROSA gRNA group (17.7% vs. 13.7%) (Figure 4a). These results suggest that the 5' inverted repeat HA improves the HR efficiency in cells. The improvement of KI efficiency only occurred after DSB induction using gene-editing tools.

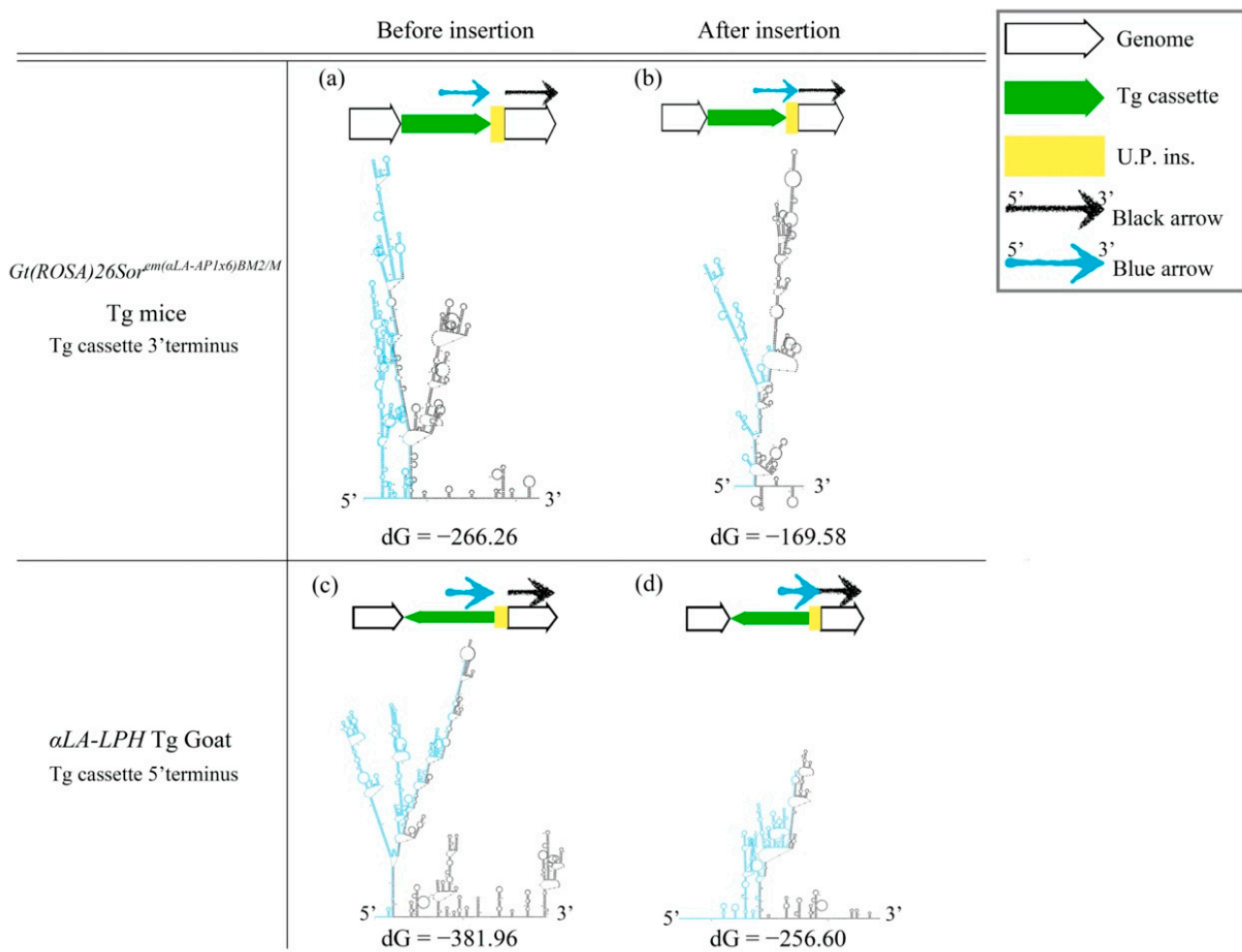


Figure 3. DsDNA stability improved after insertion. (a,c) Predictions of secondary structures and dG values of the mouse ROSA26 locus and α LA-LPH goat Tg locus without the insertion. (b,d) Predictions of secondary structures and dG values of the mouse ROSA26 locus and α LA-LPH goat Tg locus with the insertion. All sequences are provided in Table S1. Tg cassette, transgene cassette; U.P. Ins., unpredictable insertion. Blue arrow, blue sequences in secondary structures. Black arrow, black sequences in secondary structures. dG, Gibbs free energy.

2.4. Only the Presence of the Cruciform Structure in HA Improves the KI Efficiency

We further examined whether the location of the cruciform structure was an important issue in KI efficiency. An inverted repeat AAV2-ITR sequence (ITRIR) was constructed outside of the ROSA 3'HA. After transfection, the KI efficiency of ROSA 3'HA (3'ITRIR) did not show improvement compared to the KIR with ROSA gRNA group (23.0% vs. 22.0%) ($p = 0.318$) (Figure 4b). Our data suggested that the KI efficiency of the HR mechanism was only improved by the cruciform structure located in the HA sequence.

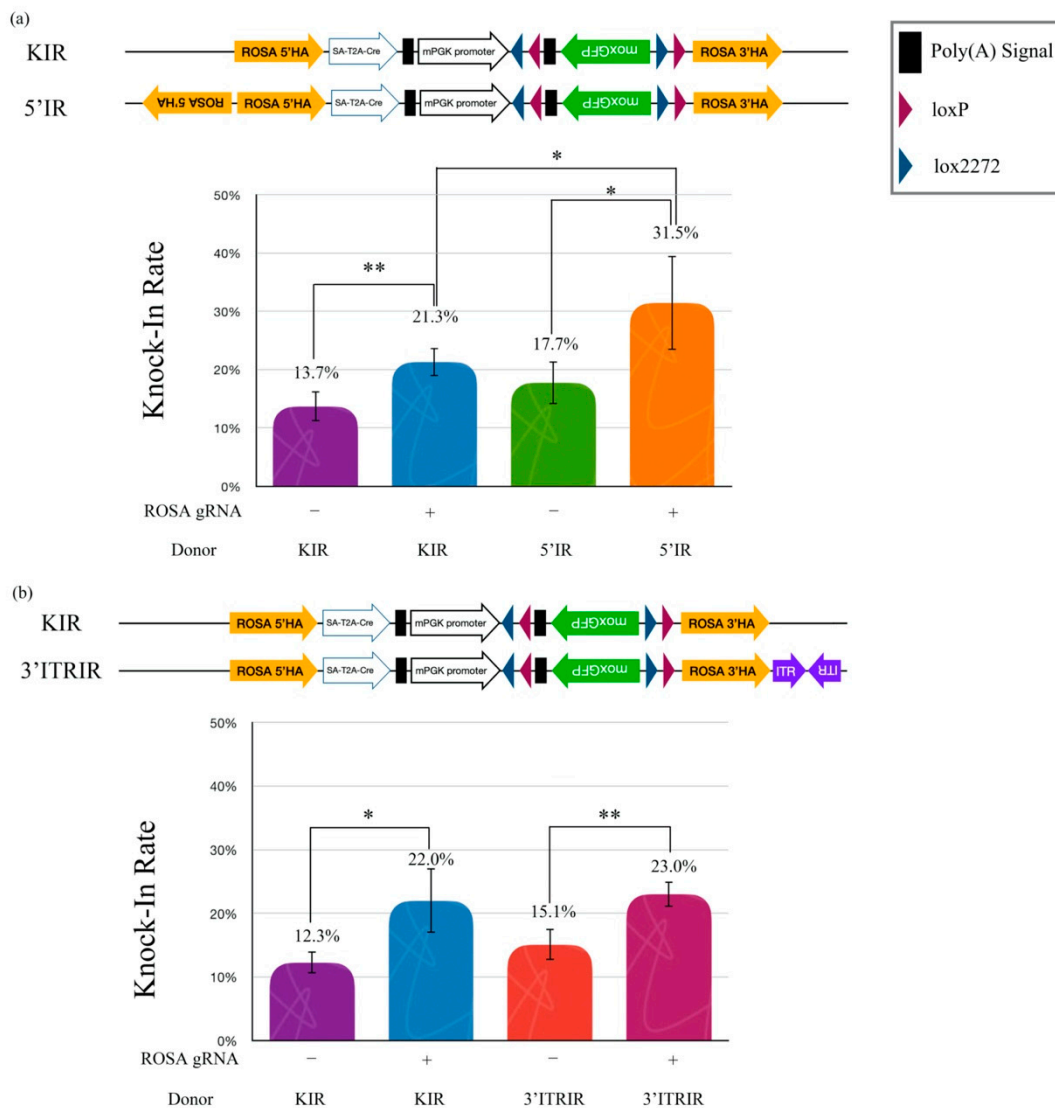


Figure 4. Inverted repeat HA improves the knock-in rate. (a) Effects of the knock-in rate by different DNA structures on the HA. (b) Effects of the knock-in rate by different DNA structures beyond the HA. Quantitative data of the knock-in rate tests obtained using the flow cytometry. The graph shows the means \pm SD of at least two replicates and four individual samples ($n = 4$) for each group. The asterisks (* and **) indicate a significant difference as determined by one-way ANOVA with Fisher's least significant difference (LSD) test.

3. Discussion

This study provides three main findings: (a) We observed an inverted repeat structure that was frequently present at Tg integration sites and genome inversion hotspots. (b) We are the first group to hypothesize a putative model of the transgene integration mechanism via HR processing and use the gene-editing tool to analyze the effects of cruciform DNA structures in the HR mechanism. (c) We successfully demonstrated that the presence of an extra inverted repeat structure in the HA region of the 5'IR plasmid construct significantly increased the KI efficiency under gRNA guidance, but not in the 3'ITRIR plasmid construct that contained an inverted ITR repeat outside of the HA region.

According to our findings, we proposed a putative model for exogenous DNA insertion cross-reacting via the stem-loop structure of host genomes. The instability of the template DNA with several inverted repeat cruciform DNA structures improves recombination enzyme recognition. The template DNA was easily integrated via the 3'-overhang

of linear exogenous DNAs (Figure 5a–c). The linear DNAs were linked with gDNA via asymmetric Holliday junction resolution (Figure 5d–g). The HR or other repair mechanisms may insert unpredicted sequences into junctions of Tg cassettes and host genomes [22]. The addition of dG values after unpredicted insertion may provide the evidence that the gDNA stability was related to the unpredicted modification.

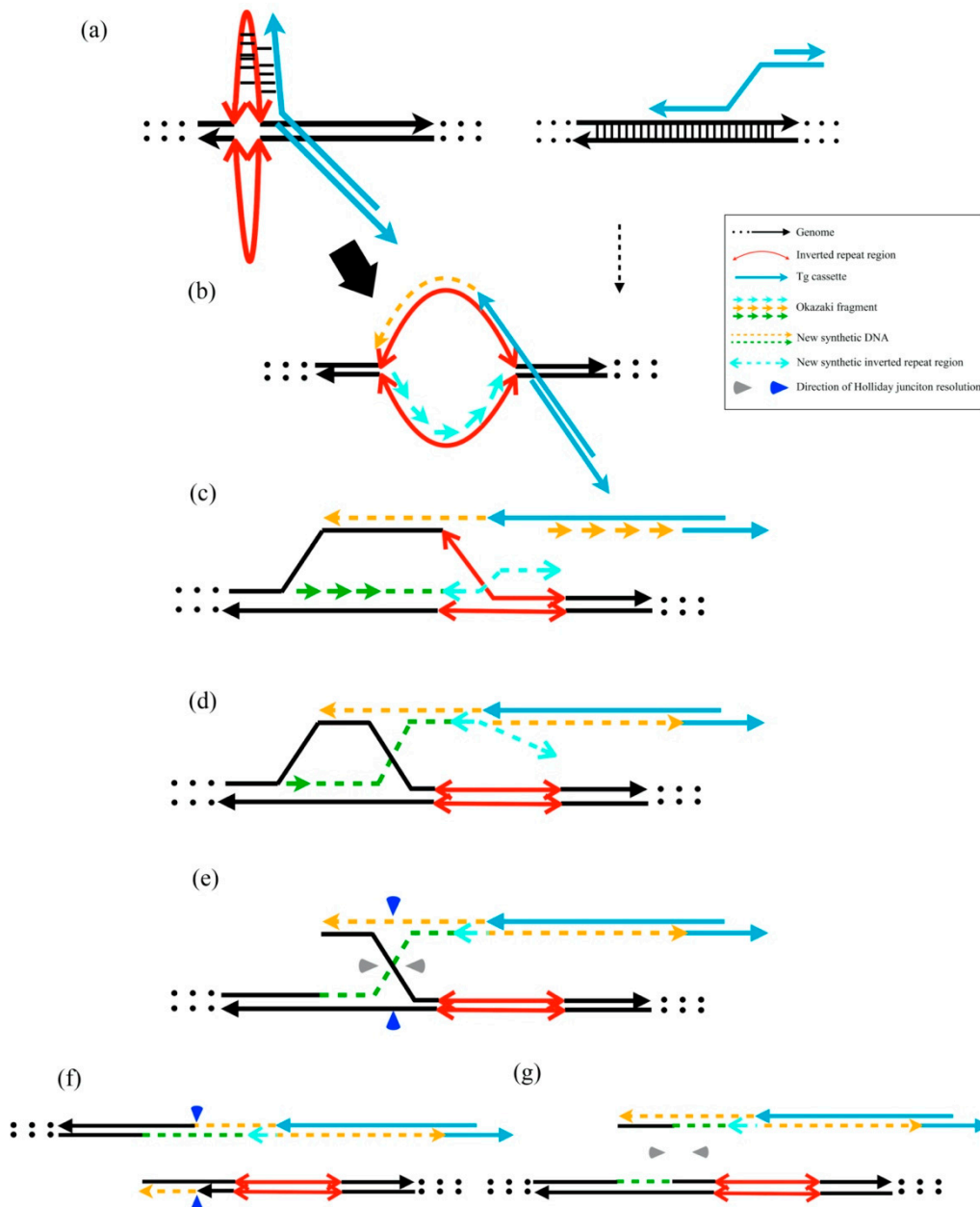


Figure 5. Putative model of transgene cassette integration via the HR mechanism. (a) Representative cruciform extrusions from inverted repeat regions and normal HR template without cruciform structures. The 3'-overhang of linear DNA integration into the extruded cruciform DNA is easier than the dsDNA after 5' to 3' resection by the MRN complex. (b,c) Initiation of new DNA synthesis and D-loop migration. (d) The lagging strand 3' terminus assembles the leading strand. (e) Holliday junction formation after DNA synthesis. (f) When the Holliday junction is resolved through the vertical axis, the Tg cassette joins the gDNA. (g) When the Holliday junction is resolved through the horizontal axis, the Tg cassette does not join the gDNA.

Several mechanisms may explain the improvement in template DNA legibility by inverted repeat sequences. First, inverted repeat sequences are relatively unstable *in vivo* [23,24]. Therefore, HR enzymes, especially Rad51 and Rad54, may easily identify inverted repeat sequences. Second, inverted repeat sequences are likely the origin of DNA replication [5,25]. The DNA replication mechanism may help the template DNA amplify itself and maintain the half-life in cells. HR enzymes are more activated during DNA replication in S phase [26]. Therefore, HR enzymes are more likely to recognize the template DNA. Third, many types of cruciform DNA-binding proteins (CBPs), such as Rad51AP1 [27,28], Rad54 [29], 14-3-3 [30], PARP-1 [31], p53 [32], and BRCA1 [33], play roles in DNA replication and repair in mammalian cells [34]. These reports suggested that the structural conformation of inverted repeat DNA strongly interacted with mechanisms of DNA replication and repair. Our data provided two new pieces of evidence. First, gDNA sequences around the predestinated Tg integration site generally formed stem-loop structures. Second, the artificial cruciform DNA structure played a role in improving HR for the template DNA *in vivo*. Therefore, the cruciform DNA structure increased the DNA legibility *in vivo* via the HR mechanism.

AAV viruses are useful tools for long-term maintenance in mammalian cells without integration into host genomes. The recombinant AAV (rAAV) system is generally episomal concatemerization (>99.99%) *in vivo* [35], which includes cruciform DNA structures at ITR-IR regions of episomes. The cruciform structure was formed via HR after second-strand DNA synthesis. The episomal concatemer is highly persistent in nuclei in a chromatin-like structure [36]. AAV-ITR-IR may play a key role in increasing plasmid persistence *in vivo* [37]. The AAV ITR sequence includes transcriptional activity *in vivo* [38]. Notably, the transcriptional silencing of rAAV in mammalian cells is related to the interaction between the ATM/MRN complex and T-shaped ITRs [39–41]. These reports indicate that the ITR structure is the key that mediates rAAV interaction with host cells. Notably, some inverted repeat sequences, such as AAV2 ITR, AAVS1, and Alu long inverted repeats [42], slightly induce the HR repair mechanism [16]. This previous study suggested that cruciform DNA structures induced HR and directly replaced the structure by inducing a DSB. Our study also showed that all four transgene cassettes integrated into cruciform DNA structures of insertion sites in Tg animals. However, we found that predestinated matched templates, which imperfectly matched with Tg cassettes, were near integration sites in host chromosomes without DSB. All predestinated matched templates were replaced by Tg cassettes. Therefore, assembly regions were not found after DSB. Our data suggested that Tg cassettes integrated and imperfectly matched with assembly regions before DSB. Therefore, cruciform DNA structures act as templates for Tg integration in the HR mechanism, but cruciform DNA structures did not induce DSB before Tg cassette integration.

Platforms for large IR plasmid construction and amplification are not stable. Therefore, it is a critical to improve methods of IR plasmid production to deeply analyze the relationship between DNA structures and the HR mechanism. Some commercially competent cells are compatible with plasmids, including unstable structures, such as Stbl strains (Invitrogen), SURE strains (Agilent), Endura strains (Lucigen), and NEB stable strains (New England Biolabs). Although commercially competent cells may be used, the stability and yield of plasmid production is dependent on plasmids. The plasmid yield of 5'IR in the SURE2 strain was lower than in the DH5 α strain (ECOS101TM) in our study (data not shown). The closed linear plasmid from Lucigen may overcome issues of unstable DNA construction and amplification. Although these companies use their products to overcome unstable plasmid production, individual customers should test the precise protocol.

Gene-editing methods were developed over the past decade. KO and KI induce DNA double-strand breaks using protein- or RNA-guided site-specific endonucleases, such as zinc-finger nucleases (ZFNs) [43], transcription activator-like effector nucleases (TALENs) [44], and CRISPR/Cas9 [45]. After DSB, cells repair their DNA via non-homologous end-joining (NHEJ) or HR [46]. Most NHEJ repairs result in insertion–deletion mutations. Genes generally lose functions when NHEJ repair sites occur at the coding DNA sequence (CDS), which is called the gene KO. The HR mechanism uses multiple steps to repair

DNA. The donor DNA integrates into the DSB locus when the donor DNA is used as the HA template. However, the HR efficiency is much lower than NHEJ in cells [47,48], which explains the lower KI rate compared to the KO rate [49]. Therefore, understanding the NHEJ/HR switch mechanism is the foundation for improving HR efficiency. Table 1 summarizes the factors involved in the NHEJ/HR switch mechanism, including repair enzymes and DNA conformation. For repair enzyme factors, the activities of DSB repair enzymes were manipulated to improve the KI efficiency in recent reports, such as HR enzyme overexpression or RS-1 addition [50–52] for approximately 0–20%, NHEJ inhibitor addition [53–55] for approximately 0–40%, and small molecule addition to arrest the cell cycle [56,57] for approximately 20–40% of KI efficiency improvements. However, these reports showed a wide range of results and were not consistent in KI improvement.

For the DNA conformation factors, the distance between the template and the host DNA was manipulated to improve KI efficiency using a modified biotin–streptavidin approach to localize repair templates to target sites. This study hypothesized that the distance of two alleles would significantly affect HR efficiency [58]. However, their results showed a high variation in KI efficiency. Current experimental data suggest that the variation in KI efficiency is largely sequence-dependent [59]. However, these data provide no clear answers of how DNA sequences affect KI efficiency. Moreover, we were interested in which is sequence-dependent in the HR mechanism, either DSB alleles or template alleles. Notably, the KI efficiency of the 5'IR plasmid was better than the KIR plasmid in our report. These data suggested that the template DNA containing the unstable cruciform DNA structure improved the HR efficiency. Therefore, the structure/stability-dependency may be more important in the template allele than in the DSB allele.

Table 1. Factors of the NHEJ/HR switch mechanism.

Factors	Summary	References
Repair Enzymes		
Activation of HR pathway	Weak improvement of KI efficiency by HR enzyme overexpression or RS-1 addition (0~20%)	[50–52]
Inhibition of NHEJ pathway	Moderate improvement of KI efficiency by NHEJ inhibitors (0~40%)	[53–55]
Arrest of cell cycle	Strong improvement of KI efficiency by using small molecules to arrest cell cycle (20~40%)	[56,57]
DNA conformation		
The distance between the template DNA and the host DNA	Strong improvement of KI efficiency by modifying donor DNA (20~40%)	[58]
The structure of the template DNA	Indirect evidence showed that the cruciform structure from ITR affected KI efficiency	[60,61]
	Our data suggested that the cruciform structure improved KI efficiency (10~20%)	This study
The transcription activity of the DSB allele	High level of transcription activity around DSB sites induced HR via Rad52 activation and 53BP1 inhibition	[62–64]
	DNA:RNA hybrid forms are related to DNA repair mechanisms at DSB loci	[65,66]

4. Materials and Methods

4.1. Transgenic Animal Production

Four transgenic animals were created using pronuclear microinjection (PN inj.) in our laboratory. Bovine alpha-lactalbumin promoter-driven lactase-phlorizin hydrolase (α LA-LPH) transgenic (Tg) goats, *Akr1A1* gene knockout with *CMV-eGFP* knock-in (*Akr1A1*^{eGFP/eGFP}) Tg

mice [67], and NF κ B promoter-driven luciferase (*NF κ B-Luc*) Tg mice [68] were created via the traditional PN inj. using linearized DNAs. The ICR-*Gt(ROSA)26Sor^{em(α LA-APIx6)BM2/M}* Tg mice were created via PN inj. using the CRISPR/Cas9 gene-editing tool and circular donor DNA. All animal experimental procedure was approved by the Institutional Animal Care and Use Committee of National Chung Hsing University (code: IACUC No. 108-029; date: 1 July 2019).

4.2. Analysis of Transgenic Cassettes in the Integration Sites of Chromosomes

4.2.1. α LA-LPH Tg Goats

The whole-genome sequencing (WGS) method was used to analyze the Tg integration site in the α LA-LPH Tg goat genome. Genomic DNA was isolated from the blood of α LA-LPH Tg goats using a QIAamp[®] DNA Blood Midi kit (51185; QIAGEN Sciences Inc., Germantown, MD, USA). The sequencing library was generated using a Truseq Nano DNA Library Prep Kit (#20010965; Illumina Inc., San Diego, CA, USA) following the manufacturer's recommendations, and index codes were added to each sample. Briefly, the genomic DNA sample was fragmented by sonication to approximately 350 bp. The DNA fragments were end-polished, A-tailed, and ligated with the full-length adapter for Illumina sequencing, followed by further PCR amplification. PCR products were purified in an AMPure XP system (A63881; Beckman Coulter Life Science, Indianapolis, IN, USA), and libraries were analyzed for size distribution using an Agilent bioanalyzer with an Agilent High-Sensitivity DNA kit (#5067-4626; Agilent Technologies Inc., Santa Clara, CA, USA) and quantified using a Qubit fluorometer with dsDNA HS assay kit reagent (Q32854; Thermo Fisher Scientific Inc., Waltham, MA, USA). The DNA libraries were sequenced on an Illumina NovaSeq 6000 platform (Illumina Inc., San Diego, CA, USA), and 150 bp paired-end reads were generated by Genomics BioSci & Tech Co. (Taipei, Taiwan).

4.2.2. *Akr1a1*^{eGFP/eGFP} Tg Mice

Our previous report used the compatible ends ligation inverse-PCR (CELI-PCR) method to analyze the *Akr1a1*^{eGFP/eGFP} Tg integration site. Briefly, genomic DNA was isolated from the tail snip of *Akr1a1*^{eGFP/eGFP} Tg mice. The genomic DNA was digested with *Bgl*III and *Bam*HI (NEB, Ipswich, MA, USA). Digested DNA fragments were ligated into a circular form. The targeted DNA was amplified using inverse PCR from circular DNA. PCR products were cloned into the pGEM-T Easy vector (A1360; Promega, Madison, WI, USA). Plasmid DNAs from a single colony were isolated and used for Sanger sequencing.

4.2.3. *NF κ B-Luc* Tg Mice

The targeted locus amplification–next-generation sequencing (TLA-NGS) method was used to analyze the genome of *NF κ B-Luc* Tg mice in this study. We isolated primary tail-snip fibroblasts from *NF κ B-Luc* Tg mice. The cells were fixed with 2% formaldehyde, and the genomic DNA (gDNA) was subjected to in-cell digestion by *Hin*1III-FD (FD1834; Thermo Fisher Scientific Inc., Waltham, MA, USA). The digested gDNAs were further ligated using T4 ligase (15224041; Thermo Fisher Scientific Inc.). After re-crosslinking formaldehyde using the phenol/chloroform and ethanol precipitation method, DNAs were digested by *Xce*I-FD (FD1474; Thermo Fisher Scientific Inc.). The TLA template was prepared via a second ligation using T4 ligase (15224041; Thermo Fisher Scientific Inc.). We used the primer sets TLA-L and TLA-R to amplify the target region using PfuUltra II Fusion HS DNA polymerase (600672; Agilent Technologies Inc., Santa Clara, CA, USA). The Celero PCR Workflow with enzymatic fragmentation library preparation Kit (9363; TECAN, Seestrasse, Männedorf, Switzerland) was used to construct the DNA library. The NGS library was analyzed via paired-end sequencing with a MiSeq system (SY-410-1003; Illumina Inc., San Diego, CA, USA) by Tri-I Biotech Inc. (Taipei, Taiwan).

4.2.4. *Gt(ROSA)26Sor^{em(aLA-AP1x6)BM2/M}* Tg Mice

The transgene junction of ICR-*Gt(ROSA)26Sor^{em(aLA-AP1x6)BM2/M}* was amplified using PfuUltra II Fusion HS DNA polymerase (600672; Agilent Technologies Inc.). PCR products were analyzed via Sanger sequencing.

4.3. Secondary Structure Prediction

The UNAFold Web Server “DNA folding form” (<http://www.unafold.org/mfold/applications/dna-folding-form.php>, accessed on 19 August 2021) was used to analyze the secondary structure of approximately 2000 nt of the surrounding sequences in Tg integration sites [69]. We chose “polymer” in correction type, “Untangle with loop fix” in structure draw mode, and “Flat” or “Flat-alt” in exterior loop type for all analyses. Other options were used in the default setting. All sequences were deposited into the National Center for Biotechnology Information (NCBI) database (<https://www.ncbi.nlm.nih.gov> accessed on 22 October 2021). The analytic intervals are provided in Table S1.

4.4. Construction of Knock-In Reporter Plasmids

We established a KI efficiency reporter system using the FLEX system. The SA-T2A-Cre cassette was constructed downstream of ROSA 5'HA in the pDonor MCS Rosa26 (a gift from Charles Gersbach; Addgene plasmid #37200; <http://n2t.net/addgene:37200> (accessed on 30 November 2015); RRID: Addgene_37200). The FLEX system was constructed downstream of the SA-T2A-Cre cassette. The KIR donor plasmid included ROSA 5'HA, SA-T2A-Cre, the FLEX system, and ROSA 3'HA. The 5'IR donor plasmid was constructed via the addition of the inverse ROSA 5'HA to the KIR donor plasmid. The 3'ITRIR donor plasmid was constructed by adding an inverted AAV2 ITR sequence to the KIR donor plasmid.

4.5. Cell Culture and Transfection

NIH/3T3 cells (CRL-1658; ATCC, Manassas, VA, USA) were cultured in Dulbecco's modified Eagle's medium (11965; Thermo Fisher Scientific Inc.) supplemented with 10% fetal bovine serum (26140079; Thermo Fisher Scientific Inc.) at 37 °C in air containing 5% CO₂. For KI rate measurement, NIH/3T3 cells were seeded in 24-well plates and co-transfected with CRISPR/Cas9 plasmid (pU6-(BbsI) CBh-Cas9-T2A-mCherry or pU6-(sgROSA)_CBh-Cas9-T2A-mCherry) (a gift from Ralf Kuehn; Addgene plasmid #64324; <http://n2t.net/addgene:64324> (accessed on 30 November 2015); RRID: Addgene_64324) and donor plasmid using Lipofectamine™ 2000 (11668030; Thermo Fisher Scientific Inc.). The sgROSA gRNA target sequence is “GACTCCAGTCTTTCTAGAAGA”.

4.6. Quantitation of Knock-In Rates

KI rates were measured 24 h post-transfection using flow cytometry (Accuri™ C6 Plus; BD Bioscience, East Rutherford, NJ, USA) with 20,000 transfected cells per sample. Red fluorescence-positive cells were used as the control for transfection efficiency. Green (G) and red (R) fluorescence-positive cells (G + R) indicated successful transfection and KI. The KI rate was determined by (G + R)/R. The background signals of the knock-in reporter system and transfection efficiencies were tested for finding the optimal condition to ignore the leak expression (Figures S2–S4) [70]. We used an optimal transfection condition of co-transfection of the 50 ng donor plasmid and 1.8 µg CRISPR/Cas9-mCherry all-in-one plasmid per well in 24 wells (Figure S3).

4.7. Statistical Analysis

The results are expressed as the means ± SD (standard deviation). Statistical analyses were performed using one-way ANOVA with Fisher's least significant difference (LSD) test in SPSS software (Version 20, IBM Corp., Armonk, NY, USA). $p < 0.05$ was considered statistically significant (* $p < 0.05$ and ** $p < 0.01$).

5. Conclusions

We first reported the presence of cruciform DNA structures of genome sequences in the transgene integration sites analyzed in four different strains of Tg animals. We found that the ROSA26 locus exhibited a similar cruciform DNA structure in the knock-in site. We analyzed the potential function of cruciform DNA structures during HR processing using the CRISPR/Cas9 knock-in system. Notably, our data suggested that the cruciform DNA structure in the HA region improved the knock-in rate. In conclusion, the inverted repeat sequence may be more legible and recognized as a template for the HR mechanism based on the structural instability of the cruciform DNA.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23073973/s1>.

Author Contributions: Conceptualization, H.O.-Y., W.C., L.-Y.S. and C.-M.C.; methodology, H.O.-Y. and S.-H.Y. (Shiao-Hsuan Yang); software, S.-H.Y. (Shang-Hsun Yang); validation, S.-H.Y. (Shang-Hsun Yang), L.-Y.S. and C.-M.C.; formal analysis, H.O.-Y.; investigation, H.O.-Y. and A.C.; resources, C.-M.C.; data curation, L.-Y.S.; writing—original draft preparation, H.O.-Y.; writing—review and editing, C.-M.C.; visualization, W.C.; supervision, C.-M.C.; project administration, L.-Y.S.; funding acquisition, C.-M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science and Technology of Taiwan (MOST-108-2313-B-005-039-MY3) and was partially supported by the iEGG and Animal Biotechnology Center from the Feature Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE-111-S-0023-A) in Taiwan.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Acknowledgments: We thank our colleagues, Gary R. L. Chang and Tung-Chou Tsai, in the Molecular Embryology and DNA Methylation Laboratory, Department of Life Sciences, National Chung Hsing University, for their help with discussions and technical issues.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; the collection, analyses, or interpretation of data; the writing of the manuscript, or in the decision to publish the results.

References

1. Bikard, D.; Loot, C.; Baharoglu, Z.; Mazel, D. Folded DNA in action: Hairpin formation and biological functions in prokaryotes. *Microbiol. Mol. Biol. Rev.* **2010**, *74*, 570–588. [[CrossRef](#)] [[PubMed](#)]
2. Lilley, D.M. The kinetic properties of cruciform extrusion are determined by DNA base-sequence. *Nucleic Acids Res.* **1985**, *13*, 1443–1465. [[CrossRef](#)] [[PubMed](#)]
3. Ramreddy, T.; Sachidanandam, R.; Strick, T.R. Real-time detection of cruciform extrusion by single-molecule DNA nanomanipulation. *Nucleic Acids Res.* **2011**, *39*, 4275–4283. [[CrossRef](#)] [[PubMed](#)]
4. Pearson, C.E.; Zorbas, H.; Price, G.B.; Zannis-Hadjopoulos, M. Inverted repeats, stem-loops, and cruciforms: Significance for initiation of DNA replication. *J. Cell. Biochem.* **1996**, *63*, 1–22. [[CrossRef](#)]
5. Zannis-Hadjopoulos, M.; Yahyaoui, W.; Callejo, M. 14-3-3 cruciform-binding proteins as regulators of eukaryotic DNA replication. *Trends Biochem. Sci.* **2008**, *33*, 44–50. [[CrossRef](#)]
6. Dutra, B.E.; Lovett, S.T. Cis and trans-acting effects on a mutational hotspot involving a replication template switch. *J. Mol. Biol.* **2006**, *356*, 300–311. [[CrossRef](#)]
7. Jamil, M.A.; Sharma, A.; Nuesgen, N.; Pezeshkpoor, B.; Heimbach, A.; Pavlova, A.; Oldenburg, J.; El-Maarri, O. F8 inversions at Xq28 causing hemophilia A are associated with specific methylation changes: Implication for molecular epigenetic diagnosis. *Front. Genet.* **2019**, *10*, 508. [[CrossRef](#)]
8. Margaglione, M.; Castaman, G.; Morfini, M.; Rocino, A.; Santagostino, E.; Tagariello, G.; Tagliaferri, A.R.; Zanon, E.; Bionchi, M.P.; Castaldo, G.; et al. The Italian AICE-Genetics hemophilia A database: Results and correlation with clinical phenotype. *Haematologica* **2008**, *93*, 722–728. [[CrossRef](#)]
9. Lyu, C.; Xue, F.; Liu, X.; Liu, W.; Fu, R.; Sun, T.; Wu, R.; Zhang, L.; Li, H.; Zhang, D.; et al. Identification of mutations in the F8 and F9 gene in families with haemophilia using targeted high-throughput sequencing. *Haemophilia* **2016**, *22*, e427–e434. [[CrossRef](#)]

10. Beri, S.; Bonaglia, M.C.; Giorda, R. Low-copy repeats at the human VIPR2 gene predispose to recurrent and nonrecurrent rearrangements. *Eur. J. Hum. Genet.* **2013**, *21*, 757–761. [[CrossRef](#)]
11. Gimelli, G.; Pujana, M.A.; Patricelli, M.G.; Russo, S.; Giardino, D.; Larizza, L.; Cheung, J.; Armengol, L.; Schinzel, A.; Estivill, X.; et al. Genomic inversions of human chromosome 15q11–q13 in mothers of Angelman syndrome patients with class II (BP2/3) deletions. *Hum. Mol. Genet.* **2003**, *12*, 849–858. [[CrossRef](#)] [[PubMed](#)]
12. Young, S.M., Jr.; Samulski, R.J. Adeno-associated virus (AAV) site-specific recombination does not require a Rep-dependent origin of replication within the AAV terminal repeat. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13525–13530. [[CrossRef](#)] [[PubMed](#)]
13. Suzuki, Y.; Chew, M.L.; Suzuki, Y. Role of host-encoded proteins in restriction of retroviral integration. *Front. Microbiol.* **2012**, *3*, 227. [[CrossRef](#)] [[PubMed](#)]
14. Wilson, M.H.; Coates, C.J.; George, A.L., Jr. PiggyBac transposon-mediated gene transfer in human cells. *Mol. Ther.* **2007**, *15*, 139–145. [[CrossRef](#)] [[PubMed](#)]
15. Miller, D.G.; Petek, L.M.; Russell, D.W. Adeno-associated virus vectors integrate at chromosome breakage sites. *Nat. Genet.* **2004**, *36*, 767–773. [[CrossRef](#)] [[PubMed](#)]
16. Hirsch, M.L. Adeno-associated virus inverted terminal repeats stimulate gene editing. *Gene Ther.* **2015**, *22*, 190–195. [[CrossRef](#)] [[PubMed](#)]
17. Laible, G. Production of transgenic livestock: Overview of transgenic technologies. In *Animal Biotechnology 2: Emerging Breeding Technologies*; Niemann, H., Wrenzycki, C., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 95–121.
18. Yoshimi, K.; Kunihiko, Y.; Kaneko, T.; Nagahora, H.; Voigt, B.; Mashimo, T. ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. *Nat. Commun.* **2016**, *7*, 10431. [[CrossRef](#)]
19. Chen, F.; Pruett-Miller, S.M.; Huang, Y.; Gjoka, M.; Duda, K.; Taunton, J.; Collingwood, T.N.; Frodin, M.; Davis, G.D. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat. Methods* **2011**, *8*, 753–755. [[CrossRef](#)]
20. Bedell, V.M.; Wang, Y.; Campbell, J.M.; Poshusta, T.L.; Starker, C.G.; Krug, R.G., 2nd; Tan, W.; Penheiter, S.G.; Ma, A.C.; Leung, A.Y.; et al. In vivo genome editing using a high-efficiency TALEN system. *Nature* **2012**, *491*, 114–118. [[CrossRef](#)]
21. Meyer, M.; Ortiz, O.; Hrabe de Angelis, M.; Wurst, W.; Kuhn, R. Modeling disease mutations by gene targeting in one-cell mouse embryos. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9354–9359. [[CrossRef](#)]
22. Piazza, A.; Wright, W.D.; Heyer, W.D. Multi-invasions are recombination byproducts that induce chromosomal rearrangements. *Cell* **2017**, *170*, 760–773. [[CrossRef](#)] [[PubMed](#)]
23. Leach, D.R. Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. *Bioessays* **1994**, *16*, 893–900. [[CrossRef](#)] [[PubMed](#)]
24. Baker, E.S.; Dupuis, N.F.; Bowers, M.T. DNA hairpin, pseudoknot, and cruciform stability in a solvent-free environment. *J. Phys. Chem. B* **2009**, *113*, 1722–1727. [[CrossRef](#)] [[PubMed](#)]
25. Alvarez, D.; Callejo, M.; Shoucri, R.; Boyer, L.; Price, G.B.; Zannis-Hadjopoulos, M. Analysis of the cruciform binding activity of recombinant 14-3-3zeta-MBP fusion protein, its heterodimerization profile with endogenous 14-3-3 isoforms, and effect on mammalian DNA replication in vitro. *Biochemistry* **2003**, *42*, 7205–7215. [[CrossRef](#)]
26. Carballar, R.; Martinez-Lainez, J.M.; Samper, B.; Bru, S.; Ballega, E.; Mirallas, O.; Ricco, N.; Clotet, J.; Jimenez, J. CDK-mediated Yku80 phosphorylation regulates the balance between non-homologous end joining (NHEJ) and homologous directed recombination (HDR). *J. Mol. Biol.* **2020**, *432*, 166715. [[CrossRef](#)]
27. Modesti, M.; Budzowska, M.; Baldeyron, C.; Demmers, J.A.; Ghirlando, R.; Kanaar, R. RAD51AP1 is a structure-specific DNA binding protein that stimulates joint molecule formation during RAD51-mediated homologous recombination. *Mol. Cell* **2007**, *28*, 468–481. [[CrossRef](#)]
28. Pires, E.; Sharma, N.; Selemenakis, P.; Wu, B.; Huang, Y.; Alimbetov, D.S.; Zhao, W.; Wiese, C. RAD51AP1 mediates RAD51 activity through nucleosome interaction. *J. Biol. Chem.* **2021**, *297*, 100844. [[CrossRef](#)]
29. Mazina, O.M.; Rossi, M.J.; Thoma, N.H.; Mazin, A.V. Interactions of human rad54 protein with branched DNA molecules. *J. Biol. Chem.* **2007**, *282*, 21068–21080. [[CrossRef](#)]
30. Alvarez, D.; Novac, O.; Callejo, M.; Ruiz, M.T.; Price, G.B.; Zannis-Hadjopoulos, M. 14-3-3sigma is a cruciform DNA binding protein and associates in vivo with origins of DNA replication. *J. Cell. Biochem.* **2002**, *87*, 194–207. [[CrossRef](#)]
31. Chasovskikh, S.; Dimtchev, A.; Smulson, M.; Dritschilo, A. DNA transitions induced by binding of PARP-1 to cruciform structures in supercoiled plasmids. *Cytom. Part A* **2005**, *68*, 21–27. [[CrossRef](#)]
32. Coufal, J.; Jagelska, E.B.; Liao, J.C.; Brazda, V. Preferential binding of p53 tumor suppressor to p21 promoter sites that contain inverted repeats capable of forming cruciform structure. *Biochem. Biophys. Res. Commun.* **2013**, *441*, 83–88. [[CrossRef](#)] [[PubMed](#)]
33. Naseem, R.; Sturdy, A.; Finch, D.; Jowitt, T.; Webb, M. Mapping and conformational characterization of the DNA-binding region of the breast cancer susceptibility protein BRCA1. *Biochem. J.* **2006**, *395*, 529–535. [[CrossRef](#)] [[PubMed](#)]
34. Holkers, M.; de Vries, A.A.; Goncalves, M.A. Nonspaced inverted DNA repeats are preferential targets for homology-directed gene repair in mammalian cells. *Nucleic Acids Res.* **2012**, *40*, 1984–1999. [[CrossRef](#)] [[PubMed](#)]
35. Buck, T.M.; Wijnholds, J. Recombinant adeno-associated viral vectors (rAAV)-vector elements in ocular gene therapy clinical trials and transgene expression and bioactivity assays. *Int. J. Mol. Sci.* **2020**, *21*, 4197. [[CrossRef](#)]
36. Penaud-Budloo, M.; Le Guiner, C.; Nowrouzi, A.; Toromanoff, A.; Cherel, Y.; Chenuaud, P.; Schmidt, M.; von Kalle, C.; Rolling, F.; Moullier, P.; et al. Adeno-associated virus vector genomes persist as episomal chromatin in primate muscle. *J. Virol.* **2008**, *82*, 7875–7885. [[CrossRef](#)]

37. Duan, D.; Sharma, P.; Yang, J.; Yue, Y.; Dudus, L.; Zhang, Y.; Fisher, K.J.; Engelhardt, J.F. Circular intermediates of recombinant adeno-associated virus have defined structural characteristics responsible for long-term episomal persistence in muscle tissue. *J. Virol.* **1998**, *72*, 8568–8577. [[CrossRef](#)]
38. Earley, L.F.; Conatser, L.M.; Lue, V.M.; Dobbins, A.L.; Li, C.; Hirsch, M.L.; Samulski, R.J. Adeno-associated virus serotype-specific inverted terminal repeat sequence role in vector transgene expression. *Hum. Gene Ther.* **2020**, *31*, 151–162. [[CrossRef](#)]
39. Zhou, Q.; Tian, W.; Liu, C.; Lian, Z.; Dong, X.; Wu, X. Deletion of the B-B' and C-C' regions of inverted terminal repeats reduces rAAV productivity but increases transgene expression. *Sci. Rep.* **2017**, *7*, 5432. [[CrossRef](#)]
40. Cataldi, M.P.; McCarty, D.M. Hairpin-end conformation of adeno-associated virus genome determines interactions with DNA-repair pathways. *Gene Ther.* **2013**, *20*, 686–693. [[CrossRef](#)]
41. Lentz, T.B.; Samulski, R.J. Insight into the mechanism of inhibition of adeno-associated virus by the Mre11/Rad50/Nbs1 complex. *J. Virol.* **2015**, *89*, 181–194. [[CrossRef](#)]
42. Lobachev, K.S.; Gordenin, D.A.; Resnick, M.A. The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* **2002**, *108*, 183–193. [[CrossRef](#)]
43. Urnov, F.D.; Rebar, E.J.; Holmes, M.C.; Zhang, H.S.; Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **2010**, *11*, 636–646. [[CrossRef](#)] [[PubMed](#)]
44. Christian, M.; Cermak, T.; Doyle, E.L.; Schmidt, C.; Zhang, F.; Hummel, A.; Bogdanove, A.J.; Voytas, D.F. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **2010**, *186*, 757–761. [[CrossRef](#)] [[PubMed](#)]
45. Jinek, M.; East, A.; Cheng, A.; Lin, S.; Ma, E.; Doudna, J. RNA-programmed genome editing in human cells. *eLife* **2013**, *2*, e00471. [[CrossRef](#)]
46. Jachimowicz, R.D.; Goergens, J.; Reinhardt, H.C. DNA double-strand break repair pathway choice—From basic biology to clinical exploitation. *Cell Cycle* **2019**, *18*, 1423–1434. [[CrossRef](#)]
47. Ingram, S.P.; Warmenhoven, J.W.; Henthorn, N.T.; Smith, E.A.K.; Chadwick, A.L.; Burnet, N.G.; Mackay, R.I.; Kirkby, N.F.; Kirkby, K.J.; Merchant, M.J. Mechanistic modelling supports entwined rather than exclusively competitive DNA double-strand break repair pathway. *Sci. Rep.* **2019**, *9*, 6359. [[CrossRef](#)]
48. Shen, W.; Ma, Y.; Qi, H.; Wang, W.; He, J.; Xiao, F.; Zhu, H.; He, S. Kinetics model of DNA double-strand break repair in eukaryotes. *DNA Repair* **2021**, *100*, 103035. [[CrossRef](#)]
49. Cristea, S.; Freyvert, Y.; Santiago, Y.; Holmes, M.C.; Urnov, F.D.; Gregory, P.D.; Cost, G.J. In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. *Biotechnol. Bioeng.* **2013**, *110*, 871–880. [[CrossRef](#)]
50. Tran, N.-T.; Bashir, S.; Li, X.; Rossius, J.; Chu, V.T.; Rajewsky, K.; Kühn, R. Enhancement of precise gene editing by the association of Cas9 with homologous recombination factors. *Front. Genet.* **2019**, *10*, 365. [[CrossRef](#)]
51. Song, J.; Yang, D.; Xu, J.; Zhu, T.; Chen, Y.E.; Zhang, J. RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat. Commun.* **2016**, *7*, 10548. [[CrossRef](#)]
52. Kurihara, T.; Kouyama-Suzuki, E.; Satoga, M.; Li, X.; Badawi, M.; Baig, D.N.; Yanagawa, T.; Uemura, T.; Mori, T.; Tabuchi, K. DNA repair protein RAD51 enhances the CRISPR/Cas9-mediated knock-in efficiency in brain neurons. *Biochem. Biophys. Res. Commun.* **2020**, *524*, 621–628. [[CrossRef](#)] [[PubMed](#)]
53. Maruyama, T.; Dougan, S.K.; Truttmann, M.C.; Bilate, A.M.; Ingram, J.R.; Ploegh, H.L. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **2015**, *33*, 538–542. [[CrossRef](#)] [[PubMed](#)]
54. Hu, Z.; Shi, Z.; Guo, X.; Jiang, B.; Wang, G.; Luo, D.; Chen, Y.; Zhu, Y.-S. Ligase IV inhibitor SCR7 enhances gene editing directed by CRISPR-Cas9 and ssODN in human cancer cells. *Cell Biosci.* **2018**, *8*, 12. [[CrossRef](#)] [[PubMed](#)]
55. Jayavaradhan, R.; Pillis, D.M.; Goodman, M.; Zhang, F.; Zhang, Y.; Andreassen, P.R.; Malik, P. CRISPR-Cas9 fusion to dominant-negative 53BP1 enhances HDR and inhibits NHEJ specifically at Cas9 target sites. *Nat. Commun.* **2019**, *10*, 2866. [[CrossRef](#)]
56. Wienert, B.; Nguyen, D.N.; Guenther, A.; Feng, S.J.; Locke, M.N.; Wyman, S.K.; Shin, J.; Kazane, K.R.; Gregory, G.L.; Carter, M.A.M.; et al. Timed inhibition of CDC7 increases CRISPR-Cas9 mediated templated repair. *Nat. Commun.* **2020**, *11*, 2109. [[CrossRef](#)]
57. Lin, S.; Staahl, B.T.; Alla, R.K.; Doudna, J.A. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* **2014**, *3*, e04766. [[CrossRef](#)]
58. Gu, B.; Posfai, E.; Rossant, J. Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat. Biotechnol.* **2018**, *36*, 632–637. [[CrossRef](#)]
59. Ranawakage, D.C.; Okada, K.; Sugio, K.; Kawaguchi, Y.; Kuninobu-Bonkohara, Y.; Takada, T.; Kamachi, Y. Efficient CRISPR-Cas9-mediated knock-in of composite tags in Zebrafish using long ssDNA as a donor. *Front. Cell. Dev. Biol.* **2020**, *8*, 598634. [[CrossRef](#)]
60. Mizuno, N.; Mizutani, E.; Sato, H.; Kasai, M.; Ogawa, A.; Suchy, F.; Yamaguchi, T.; Nakauchi, H. Intra-embryo gene cassette knockin by CRISPR/Cas9-mediated genome editing with adeno-associated viral vector. *iScience* **2018**, *9*, 286–297. [[CrossRef](#)]
61. Chen, S.; Sun, S.; Moonen, D.; Lee, C.; Lee, A.Y.; Schaffer, D.V.; He, L. CRISPR-READI: Efficient generation of knockin mice by CRISPR RNP electroporation and AAV donor infection. *Cell Rep.* **2019**, *27*, 3780–3789.e4. [[CrossRef](#)]
62. Yasuhara, T.; Kato, R.; Hagiwara, Y.; Shiotani, B.; Yamauchi, M.; Nakada, S.; Shibata, A.; Miyagawa, K. Human Rad52 promotes XPG-mediated R-loop processing to initiate transcription-associated homologous recombination repair. *Cell* **2018**, *175*, 558–570.e11. [[CrossRef](#)] [[PubMed](#)]
63. Xue, C.; Greene, E.C. New roles for RAD52 in DNA repair. *Cell Res.* **2018**, *28*, 1127–1128. [[CrossRef](#)] [[PubMed](#)]

64. Li, G.; Wang, H.; Zhang, X.; Wu, Z.; Yang, H. A Cas9-transcription factor fusion protein enhances homology-directed repair efficiency. *J. Biol. Chem.* **2021**, *296*, 100525. [[CrossRef](#)] [[PubMed](#)]
65. Lafuente-Barquero, J.; Garcia-Rubio, M.L.; Martin-Alonso, M.S.; Gomez-Gonzalez, B.; Aguilera, A. Harmful DNA:RNA hybrids are formed in cis and in a Rad51-independent manner. *eLife* **2020**, *9*, e56674. [[CrossRef](#)] [[PubMed](#)]
66. Bader, A.S.; Hawley, B.R.; Wilczynska, A.; Bushell, M. The roles of RNA in DNA double-strand break repair. *Br. J. Cancer* **2020**, *122*, 613–623. [[CrossRef](#)]
67. Lai, C.W.; Chen, H.L.; Tsai, T.C.; Chu, T.W.; Yang, S.H.; Chong, K.Y.; Chen, C.M. Sexually dimorphic expression of eGFP transgene in the Akr1A1 locus of mouse liver regulated by sex hormone-related epigenetic remodeling. *Sci. Rep.* **2016**, *6*, 24023. [[CrossRef](#)]
68. Yen, C.C.; Chang, W.H.; Tung, M.C.; Chen, H.L.; Liu, H.C.; Liao, C.H.; Lan, Y.W.; Chong, K.Y.; Yang, S.H.; Chen, C.M. Lactoferrin protects hyperoxia-induced lung and kidney systemic inflammation in an in vivo imaging model of NF-kappaB/luciferase transgenic mice. *Mol. Imaging Biol.* **2020**, *22*, 526–538. [[CrossRef](#)]
69. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **2003**, *31*, 3406–3415. [[CrossRef](#)]
70. Fischer, K.B.; Hannah, C.K.; Edward, C.M. Sources of off-target expression from recombinase-dependent AAV vectors and mitigation with cross-over insensitive ATG-out vectors. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 27001–27010. [[CrossRef](#)]