HHS Public Access

Author manuscript

Gene Ther. Author manuscript; available in PMC 2015 August 01.

Published in final edited form as:

Gene Ther. 2015 February ; 22(2): 190–195. doi:10.1038/gt.2014.109.

Adeno-associated virus inverted terminal repeats stimulate gene editing

ML Hirsch[1,2]

[1]Gene Therapy Center, University of North Carolina at Chapel, Chapel Hill, NC, USA

[2]Department of Ophthalmology, University of North Carolina, Chapel Hill, NC, USA

Abstract

Advancements in genome editing have relied on technologies to specifically damage DNA which, in turn, stimulates DNA repair including homologous recombination (HR). As off-target concerns complicate the therapeutic translation of site-specific DNA endonucleases, an alternative strategy to stimulate gene editing based on fragile DNA was investigated. To do this, an episomal gene-editing reporter was generated by a disruptive insertion of the adeno-associated virus (AAV) inverted terminal repeat (ITR) into the *egfp* gene. Compared with a non-structured DNA control sequence, the ITR induced DNA damage as evidenced by increased gamma-H2AX and Mre11 foci formation. As local DNA damage stimulates HR, ITR-mediated gene editing was investigated using DNA oligonucleotides as repair substrates. The AAV ITR stimulated gene editing >1000-fold in a replication-independent manner and was not biased by the polarity of the repair oligonucleotide. Analysis of additional human DNA sequences demonstrated stimulation of gene editing to varying degrees. In particular, inverted yet not direct, Alu repeats induced gene editing, suggesting a role for DNA structure in the repair event. Collectively, the results demonstrate that inverted DNA repeats stimulate gene editing via double-strand break repair in an episomal context and allude to efficient gene editing of the human chromosome using fragile DNA sequences.

INTRODUCTION

DNA-inverted repeat sequences (IRs) are abundant in prokaryotic and eukaryotic genomes and are considered ‘hot spots’ of genomic instability.[1,2] Such instability is thought to originate from intra-strand nucleotide (nt) interactions, resulting in structured extrusions, generally termed ‘hairpins’ herein. These hairpins, or remnants thereof, are often found near sequence deletions, mutations, duplications and chromosomal rearrangements, all of which contribute to the oncogenesis via chromosome instability.[3] In bacteria, yeast and mammals, IRs are processed to DNA double-strand breaks (DSB) in a manner that appears both dependent and independent of DNA replication.[2,4–6] Depending on the mechanism of DSB repair, these palindromic structures are often deleted or, to a lesser extent, can be repaired

Correspondence: M Hirsch, Department of Ophthalmology, Gene Therapy Center, University of North Carolina at Chapel Hill Campus Box 7352, Chapel Hill, NC 27599-7352, USA. mhirsch@email.unc.edu.

CONFLICT OF INTEREST

The author declares no conflict of interest.

via homologous recombination (HR). Consistently, reports in *Escherichia coli* and yeast have demonstrated the ability of IRs to stimulate HR.[2,7,8]

Recent genetic advancements have demonstrated the ability to modify the human genotype via HR. In general, the efficiency of HR is low (≈1 in a million),[9,10] however, a DSB generated at, or near, the target sequence induces HR by several orders of magnitude.[11–14] As such, much attention has focused on the generation of site-specific endonucleases to generate specific DSBs, or a single-strand nick, near the site to be modified.[15–17] However, off-target cleavage activity associated with these enzymes often generates unwanted mutations and toxicity, thus complicating the transition of this technology to therapeutic applications.[18–20] In addition to protein endonucleases, other methods of generating specific DSBs include triplex-forming oligonucleotides, the Cas9/CRISPR system and group II intron ribonucleic proteins, all of which demonstrate varied gene-editing efficiencies.[21–23]

It is well appreciated that chromosomal instability is associated with genetic rearrangements, oncogenesis, cell death and genetic diseases. Furthermore, repeated and structured DNA sequences are known elements of chromosomal instability.[1,2] This understanding generated the hypothesis that structured DNA sequences stimulate HR in the vicinity of the element. The work herein confirms this hypothesis using viral and human IRs in conjunction with a sensitive episomal reporter of gene editing. From a mechanistic standpoint, the adeno-associated virus serotype 2 (AAV2) inverted terminal repeat (ITR) sequence induced gamma-H2AX and MreII DNA damage signaling, was processed in the absence of a repair template and stimulated replication-independent HR. This work characterizes a unique format to optimize DNA structure-mediated HR and to gain a mechanistic understanding of DNA instability.

RESULTS AND DISCUSSION

Initially, I indirectly investigated whether the 165 nt AAV2-ITR is processed to a DSB in human 293 cells via gamma-H2AX staining. H2AX is a histone that is phosphorylated by a PI3 kinase on ser 139 (gamma-H2AX) in response to DSBs.[24] To do this, the AAV2-ITR (double D format),[25] or a size-matched control sequence, was inserted into the I-SceI site of a previously described defective *egfp* plasmid reporter (Figure 1).[12] Single-strand DNA secondary structure software (mfold) predicts a single T-shaped folding for the AAV2-ITR, whereas multiple less energetically favorable conformations were predicted for the control sequence, suggesting no preferred structure (termed No Structure). As a positive control for episomal DSBs, or perhaps off-target chromosome DSBs, a defective *egfp* reporter containing the I-SceI site was evaluated in the presence and absence of the I-SceI endonuclease (Figures 1 and 2).[12] Etoposide, which indirectly induces DSBs served as a positive control for host chromosomal damage.[26] Using human 293 cells, a significant twofold increase in gamma-H2AX staining was noted 8h post transfection of the AAV2-ITR plasmid compared with the No Structure control sequence as quantitated by flow cytometry (Figure 2a). This twofold increase was similar to that noted for the I-SceI target in the presence of the I-SceI endonuclease (Figure 2a). Greater than 70% of cells stained positive for gamma-H2AX following etoposide treatment (Figure 2a). These results demonstrate that

both I-SceI, in the presence of its target site, and the AAV2-ITR induce DNA DSB signaling.

The Mre11 protein is a endo- and exonuclease involved in HR and non-homologous end joining repair of DSBs, as well as in telomere maintenance.[27] Previous work reported that Mre11, in conjunction with Nbs1 and Rad50, acts as the primary DSB sensor, and an independent report demonstrated binding of the complex to the AAV2-ITR.[28–30] To confirm the later observation, Mre11 localization was monitored following recombinant AAV2 transduction by immunofluorescence. Discrete clusters of Mre11 foci were detected only in cells transduced by recombinant AAV2 (rAAV2), consistent with a previous report (Figure 2b).[29] Next, similar experiments were performed following the transfection of pAAV2-ITR or the No Structure control. Intense Mre11 staining of larger clusters was noted for pAAV2-ITR, but not in the No Structure-treated control cells (Figure 2c). The more intense staining observed in the pAAV2-ITR transfection, when compared with the smaller foci noted for rAAV infected cells, may indicate the greater intracellular DNA concentration of the double-strand circular plasmid compared with the initially single-strand linear transduced rAAV genomes, or perhaps different intra-nuclear distributions.

DNA structures are often processed by the host DNA machinery during transcription and/or replication. The above data of induced Mre11 foci formation and increased gamma-H2AX (Figure 2) demonstrate that in a non-viral context the AAV2-ITR sequence induces a DNA damage response, and perhaps, is processed similar to rAAV genomes post transduction.[29,31,32] To investigate this, polymerase chain reaction amplification was performed across the episomal insertion of either the No Structure or AAV2-ITR defective *egfp* reporter. As a template for these reactions, the respective plasmid prior to transfection or Hirt DNA recovered 24 h post transfection was utilized with a robust polymerase. A successful amplification reaction results in a 680 bp amplicon for both the No Structure and AAV2-ITR plasmid inputs, and theoretically all non-processed Hirt DNA templates (Figures 1 and 3a). Amplification of the AAV2-ITR input and AAV2-ITR Hirt templates gave two additional bands presumably consistent with the position of the amplification primers and a site within the AAV2-ITR (260 bp and 400 bp, respectively; Figures 1 and 3a). However, a unique polymerase chain reaction product was observed in a manner specific to the AAV2-ITR Hirt DNA amplification at an approximate size of 480 bp (Figure 3a). This product was cloned and recovered from SURE cells; however, despite repeated sequencing attempts it was not possible to obtain the entire sequence of the processed AAV2-ITR. To determine if the two *Sma*I sites within the ITR were maintained in the unique junction amplicon(s), *Sma*I digestion was performed. The results demonstrated a partial digest of the amplicons, suggesting that intracellular processing of the ITRs is not identical for all molecules (data not shown).

The experiments above demonstrate by two independent measures that the AAV2-ITR sequence in a plasmid induces DSB signaling (Figure 2). Furthermore, this sequence is targeted by cellular machinery and is processed to a smaller form perhaps by non-homologous end joining, as a homologous repair template was not provided (Figure 3). As local DSBs induce HR,[13] the ability of the AAV2-ITR to stimulate gene editing was investigated in the defective *egfp* plasmid reporter context (Figure 1). As a repair substrate, a

single-strand DNA oligonucleotide, composed of 40nts of *egfp* homology to both sides of the disruptive hairpin insertion, was used in sense (Fwd) or antisense orientations (Rev). A non-homologous oligonucleotide was used as a negative control.[33] Three days following co-transfection of the AAV2-ITR target plasmid and the *egfp* repair oligonucleotide, ~ 0.2% of cells were GFP+, indicating homologous repair of the defective reporter (Figure 3b). In contrast, no GFP+ cells were noted using the non-homologous repair substrate or the No Structure target with either repair substrate (Figure 3b). Furthermore, the sense and antisense repair substrates stimulated gene editing to similar levels, consistent with our report of no polarity bias for inter-molecular HR when targeting a region that is not replicated (Figure 3b).[33] To allow a comparison with endonuclease-mediated gene editing, the pA658 control plasmid containing the I-SceI site, in the presence of I-SceI, was investigated for *egfp* correction. The results demonstrate that HR stimulated by the nuclease induced break was fivefold more likely to occur when compared with either hairpin-induced break (Figure 4b).

Some previous reports of IR-induced DSBs invoke a replication-dependent mechanism,[34–36] however, the target plasmids used herein do not contain a known mammalian origin of replication. To determine whether the AAV2-ITR target plasmid undergoes replication in these experiments, extra-chromosomal (Hirt) DNA was harvested 3 days post transfection. To remove the transfected DNA, while preserving any replicated forms, samples were digested with DpnI, an endonuclease that recognizes a common bacterial-specific methylation motif. Then, the DNA was analyzed by non-denaturing Southern blot analysis using an *egfp* probe. In the presence of DpnI, complete digestion of the Hirt DNA was observed (plasmid is 9 kb, Figure 3c). This result demonstrates that significant target molecule replication does not occur, and therefore, suggests that the AAV2-ITR stimulates gene editing by a replication-independent mechanism.

Next, the ability of a similar DNA sequence derived from human chromosome 19 (Ch19), AAVS1, to stimulate gene editing was investigated.[37,38] AAVS1 shares 80% sequence identity to the AAV2-ITR, however, AAVS1 is predicted to fold into a hairpin having a perfect 67 nt stem with a 3 nt loop (Figure 4a). The 148 nt ‘DD’ AAVS1 sequence[38] was cloned into the gene editing reporter (Figure 1) and used as the episomal target in the presence of the *egfp* oligonucleotide repair substrate. GFP+ cells, indicative of targeted HR, were quantitated by flow cytometry 48 h post transfection. Consistent with the results using AAV2-ITR, no GFP+ cells were detected using the non-homologous oligonucleotide (Figure 4b). In the presence of the homologous repair oligonucleotide, gene editing stimulated by the AAVS1 hairpin was modestly elevated compared with AAV2-ITR, but not statistically different (Figure 4b). Assuming that secondary structure predictions are physiologically relevant, these results suggest that AAV2-ITR and AAVS1 sequence similarities, and not necessarily the predicted DNA secondary structure (T-shaped vs stem-loop), are important for IR-induced gene editing in these instances.

The work thus far demonstrates that a viral and a Ch19 sequence stimulate gene editing at levels approaching those reported for several first generation site-specific endonucleases, albeit in an episomal context with ~150 target copies per cell (Figure 4b; data not shown).[12,39] However, unlike endonuclease scaffolds that can be targeted towards a particular DNA sequence of interest, DNA structure-induced chromosomal gene editing

would require an endogenous IR at, or near, the site to be modified, such as Alu repeats. Over one million Alu repeats comprise over 10% of the human genome and are present in multi-copy on every arm of every human chromosome. Therefore, toward the possibility of IR-induced gene editing throughout the human genome, Alu DNA, which is associated with various types of chromosome instability, was investigated in an episomal context.[40] Previously reported Alu direct or indirect repeat sequences (652 nt in size with a 12 nt spacer)[40] in the gene editing reporter (Figure 1) were transfected along with an oligonucleotide repair substrate, and GFP+ cells were quantitated 3 days later. In the case of Alu direct repeats, GFP+ cells were observed at very low levels in the presence and absence of a homologous repair oligonucleotide (Figure 5). In contrast, a 20-fold increase in GFP+ cells was noted for the Alu indirect repeat sequences in the presence of homologous repair substrate, demonstrating the importance of DNA structure for stimulation of HR.[41]

It is well appreciated that local DNA damage can stimulate gene editing over 1000-fold.[12,15,29,30] One strategy to stimulate targeted HR, which has been reported in bacteria and yeast, is the use of DNA IRs.[33,41] In yeast, the human Alu indirect repeat sequence stimulated a 30-fold increase in gene editing in an Mre11-dependent reaction.[41,42] This dependence is consistent with the observations in a human context of pAAV2-ITR induced Mre11 foci formation (Figure 2c), which is not surprising given the diverse roles the Mre11 complex has in DNA repair.[29,30] Another report of IR-induced HR was our previous demonstration that the AAV2-ITRs stimulate inter-molecular recombination of transduced rAAV genomes.[33] In that report, an oligonucleotide was used to tether distinct viral genomes and also functioned as the repair substrate, thereby generating a larger DNA molecule. Interestingly, no significant difference in the repair frequency was observed based on oligonucleotide polarity when targeting a duplexed region of the transduced genome.[29,30,33] In contrast, a consistent bias was noted when targeting single-strand regions that undergo replication, and in that case, the oligonucleotide orientation that corresponded to the template strand was more efficient for the HR event.[29,30,33] The results obtained herein are consistent, in that no repair oligonucleotide polarity bias was observed for the DNA structure-induced editing events using an episomal template that is transcribed, but does not undergo observable replication (Figure 3b). As the episomal reporter of this work undergoes transcription, the role of this event in the DNA repair process remains uncharacterized.

Regarding the ‘natural’ processing of the AAV2-ITR in the absence of a repair substrate, the data allude to deleted, heterogeneous species likely generated by non-homologous end joining and/or intra- or inter-molecular HR (Figure 3a). Consistently, an early report using restriction analyses demonstrated that, host-cell processing of plasmid encoded AAV2-ITR sequences results in various ITR deletions, of which only some of these were competent for AAV production.[43] Importantly, that same work demonstrated that the alterations in the ITR sequence are not necessarily final, and in the case of the AAV genomes, ITR deletions could be ‘corrected’ back to wild type ITR sequence.[43] This observation calls into question the exact molecule, or molecules, mediating the AAV2-ITR-induced gene editing event and suggests the possibility that processed pAAV2-ITR (Figure 3) could be ‘corrected’ by non-processed ITR sequences.[43] Such heterogeneity could possibly account for our inability to obtain reproducible sequencing data for these molecules. Although not observed with the

negative control, it can be argued that polymerase chain reaction amplification of the pAAV-ITR2 Hirt DNA, either by the method used herein or for other approaches, such as bisulfite sequencing, is sequence biased and introduces artifacts further calling into question the precise sequences of the processed ITR sequences. Nonetheless, the results indicate that pAAV2-ITR is processed at multiple positions into smaller species of unknown structure/function (Figure 3).

In addition to the AAV2-ITR, the ability of other hairpins to stimulate gene editing was demonstrated, including an AAV2-ITR-like sequence found on Ch19 (AAVSI). AAVSI shares 80% sequence identity with the AAV2-ITR and, assuming that secondary structure predictions are physiologically relevant, a primary difference between these two sequences is that AAV2-ITR is T-shaped, whereas AAVSI forms a linear hairpin (Figure 4a). Currently, the reasons underlying the recombination potential of these tested sequences is unclear, and as DNA sequence and structure are not mutually exclusive and DNA secondary structure predictions may not be physiologically relevant, definitive conclusions are not forthcoming.

In this work a sensitive and quantitative method was used to evaluate the ability of different DNA sequences to induce targeted HR. This system relies on an episomal gene-editing reporter and takes advantage of the ability of structured DNA to induce DSBs and host-cell processing (Figures 2 and 3). The ability of AAV2-ITR, AAVS1, and Alu repeats to stimulate gene editing is demonstrated, the latter of which are scattered throughout the human genome. As previous work in yeast and in human cells has demonstrated gene editing at sites distant from a targeted DSB,[14,44] the collective work raises the intriguing possibility that existing fragile sequences, such as indirect Alu repeats, can serve to stimulate local gene editing throughout the human genome. Since the data herein is limited to episomal molecules, it is apparent that additional factors such as open or relaxed chromatin will influence the accessibility of chromosome regions to repair. Regarding episomal utility, the results of this work and ongoing experiments suggest that different ITRs stimulate HR at different efficiencies (Figures 4 and 5). As previous work has demonstrated that AAV large gene transduction is enhanced by directing concatemerization via ITRs from different serotypes,[45] it is possible that the system herein can also be used to screen wild type and synthetic ITR recombination potentials to enhance oversized AAV transduction (that is, identical ITRs with a high recombination potential could be used to direct transgene reconstruction following split vector transduction).

MATERIALS AND METHODS

Cell culture and manipulations

HEK-293 cells were maintained as previously reported.[46] Transfections were performed using poly(ethylenimine) on 50 000 293 cells per well (500 μl final volume) in a 24-well plate. The target plasmid was used at 0.5 μg per well and the repair oligonucleotide at 1.5 μg per well. The non-homologous oligonucleotide 80-mer was previously described,[33] whereas the GFP 80-mer corresponds to the 40 nt immediately flanking both sides of the disruptive insertion in pA658.[12] For the single-strand AAV2-CMV-eGFP vector infection experiment,

vector was produced following a standard protocol[47] and administered at 10 000 viral genomes per cell.

Amplification across the disruptive insertion in the gene editing *egfp* reporter relied on standard polymerase chain reaction using LA Taq polymerase (TaKaRa, Mountain View, CA, USA). The primer sequences within the *egfp* coding sequence used in those reactions contained *Hin*dIII and *Xba*I for junction cloning and were the following: Fwd CGCATAAGCTTGGACGGCGACGTAAA and Rev CGATTCTAGATACTCCAGCTTGTGCC. Plasmid DNA, and small molecular weight DNA (Hirt) recovered by a standard protocol,[32] were used as amplification templates. The replication experiments also relied on Hirt DNA recovery following plasmid transfection, DpnI digestion to remove delivered plasmid, non-denaturing gel electrophoresis and detection on a nitrocellulose membrane (Hybond XL, Amersham, England) with random radio-labeled probes generated from the *egfp* cDNA.[47]

Plasmid construction

The gene editing parent plasmid pA658 was reported previously.[12] Target plasmids were generated by blunt cloning into the unique I-SceI site of pA658. In general, *E. coli* Sure Cells (Agilent, Santa Clara, CA, USA) were used for transformation and plasmid recovery. The AAV2-ITR sequence was taken from the 165 nt EcoRI fragment of pDD,[25] the AAVSI sequence was acquired from the *Hin*dIII/*Kpn*I digestions of a reported plasmid,[38] and the Alu indirect and direct repeat sequences were provided by Dr K Lobachev.[40] Sequence verification was performed on the 'DD' AAV2-ITR and containing plasmid from both directions following *SmaI* digestion, however, the 5 nt in between the two *Sma*I sites were not confirmed. 'DD' AAVS1 sequencing followed a SfiI digestion. The No Structure control was generated using a non-coding sequence of pLKO.1 TRC (nts 152–296) and cloned into the I-SceI of pA658. In the case of the Alu repeats, the generated plasmids containing the indirect or direct variation were verified by digestion owing to difficulties in sequencing these elements.[40]

Cellular staining

Gamma-H2AX staining was performed using the phospho- (ser 139)-H2 A.X antibody (2 μg ml^{-1}; Millipore, Billerica, MA, USA) as previously described.[48] MreII cellular staining followed an established protocol.[29]

GFP+ cell detection

Flow cytometry was performed as described.[33] Approximately 250 000 cells were analyzed per replicate and more than nine different replicates were performed for each experimental group. Statistical significance was determined by a Student's *t*-test.

ACKNOWLEDGEMENTS

I thank Drs C Yin, C Hewitt, K Lobachev, and R J Samulski for molecular biology expertise, reagents and helpful discussions. This work was supported by the Northwest Genome Engineering Consortium, the Jain Foundation, and the NIH (RO1AI072176-06A1, RO1AR064369-01A1). Funding was also provided in part through a departmental Unrestricted Grant from Research to Prevent Blindness, New York, NY, USA.

REFERENCES

1. Collins J, Volckaert G, Nevers P. Precise and nearly-precise excision of the symmetrical inverted repeats of Tn5; common features of recA-independent deletion events in *Escherichia coli*. Gene. 1982; 19:139–146. [PubMed: 6292045]
2. Gordenin DA, Lobachev KS, Degtyareva NP, Malkova AL, Perkins E, Resnick MA. Inverted DNA repeats: a source of eukaryotic genomic instability. Mol Cell Biol. 1993; 13:5315–5322. [PubMed: 8395002]
3. Lobachev KS, Rattray A, Narayanan V. Hairpin- and cruciform-mediated chromosome breakage: causes and consequences in eukaryotic cells. Front Biosci. 2007; 12:4208–4220. [PubMed: 17485368]
4. DasGupta U, Weston-Hafer K, Berg DE. Local DNA sequence control of deletion formation in *Escherichia coli* plasmid pBR322. Genetics. 1987; 115:41–49. [PubMed: 3030883]
5. Akgun E, Zahn J, Baumes S, Brown G, Liang F, Romanienko PJ, et al. Palindrome resolution and recombination in the mammalian germ line. Mol Cell Biol. 1997; 17:5559–5570. [PubMed: 9271431]
6. Leach DR, Okely EA, Pinder DJ. Repair by recombination of DNA containing a palindromic sequence. Mol Microbiol. 1997; 26:597–606. [PubMed: 9402028]
7. Eykelenboom JK, Blackwood JK, Okely E, Leach DR. SbcCD causes a double-strand break at a DNA palindrome in the *Escherichia coli* chromosome. Mol Cell. 2008; 29:644–651. [PubMed: 18342610]
8. Farah JA, Hartsuiker E, Mizuno K, Ohta K, Smith GR. A 160-bp palindrome is a Rad50. Rad32-dependent mitotic recombination hotspot in Schizosaccharomyces pombe. Genetics. 2002; 161:461–468. [PubMed: 12019258]
9. Koller BH, Smithies O. Altering genes in animals by gene targeting. Ann Rev Immunol. 1992; 10:705–730. [PubMed: 1591000]
10. Capecchi MR. Altering the genome by homologous recombination. Science. 1989; 244:1288–1292. [PubMed: 2660260]
11. Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. Science. 2003; 300:763. [PubMed: 12730593]
12. Porteus MH, Cathomen T, Weitzman MD, Baltimore D. Efficient gene targeting mediated by adeno-associated virus and DNA double-strand breaks. Mol Cell Biol. 2003; 23:3558–3565. [PubMed: 12724414]
13. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. The double-strand-break repair model for recombination. Cell. 1983; 33:25–35. [PubMed: 6380756]
14. Stuckey S, Storici F. Genetic modification stimulated by the induction of a site-specific break distant from the locus of correction in haploid and diploid yeast *Saccharomyces cerevisiae*. Methods Mol Biol. 2014; 1114:309–324. [PubMed: 24557912]
15. Metzger MJ, McConnell-Smith A, Stoddard BL, Miller AD. Single-strand nicks induce homologous recombination with less toxicity than double-strand breaks using an AAV vector template. Nucleic Acids Res. 2011; 39:926–935. [PubMed: 20876694]
16. Certo MT, Gwiazda KS, Kuhar R, Sather B, Curinga G, Mandt T, et al. Coupling endonucleases with DNA end-processing enzymes to drive gene disruption. Nat Methods. 2012; 9:973–975. [PubMed: 22941364]
17. Smith J, Grizot S, Arnould S, Duclert A, Epinat JC, Chames P, et al. A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. Nucleic Acids Res. 2006; 34:e149. [PubMed: 17130168]
18. Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by *in vitro* selection. Nat Methods. 2011; 8:765–770. [PubMed: 21822273]
19. Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppel C, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nat Biotechnol. 2011; 29:816–823. [PubMed: 21822255]

20. Lin Y, Cradick TJ, Brown MT, Deshmukh H, Ranjan P, Sarode N, et al. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. Nucleic Acids Res. 2014; 42:7473–7485. [PubMed: 24838573]
21. Mastroianni M, Watanabe K, White TB, Zhuang F, Vernon J, Matsuura M, et al. Group II intron-based gene targeting reactions in eukaryotes. PLoS One. 2008; 3:e3121. [PubMed: 18769669]
22. Knauert MP, Glazer PM. Triplex forming oligonucleotides: sequence-specific tools for gene targeting. Hum Mol Genet. 2001; 10:2243–2251. [PubMed: 11673407]
23. Carroll D. A CRISPR approach to gene targeting. Mol Ther. 2012; 20:1658–1660. [PubMed: 22945229]
24. Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J Biol Chem. 1998; 273:5858–5868. [PubMed: 9488723]
25. Xiao X, Xiao W, Li J, Samulski RJ. A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle. J Virol. 1997; 71:941–948. [PubMed: 8995611]
26. Pommier Y, Leo E, Zhang H, Marchand C. DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. Chem Biol. 2010; 17:421–433. [PubMed: 20534341]
27. Stracker TH, Petrini JH. The MRE11 complex: starting from the ends. Nat Rev Mol Cell Biol. 2011; 12:90–103. [PubMed: 21252998]
28. Petrini JH, Stracker TH. The cellular response to DNA double-strand breaks: defining the sensors and mediators. Trends Cell Biol. 2003; 13:458–462. [PubMed: 12946624]
29. Schwartz RA, Carson CT, Schuberth C, Weitzman MD. Adeno-associated virus replication induces a DNA damage response coordinated by DNA-dependent protein kinase. J Virol. 2009; 83:6269–6278. [PubMed: 19339345]
30. Cervelli T, Palacios JA, Zentilin L, Mano M, Schwartz RA, Weitzman MD, et al. Processing of recombinant AAV genomes occurs in specific nuclear structures that overlap with foci of DNA-damage-response proteins. J Cell Sci. 2008; 121:349–357. [PubMed: 18216333]
31. Choi VW, McCarty DM, Samulski RJ. Host cell DNA repair pathways in adeno-associated viral genome processing. J Virol. 2006; 80:10346–10356. [PubMed: 17041215]
32. Inagaki K, Ma C, Storm TA, Kay MA, Nakai H. The role of DNA-PKcs and artemis in opening viral DNA hairpin termini in various tissues in mice. J Virol. 2007; 81:11304–11321. [PubMed: 17686847]
33. Hirsch ML, Storici F, Li C, Choi VW, Samulski RJ. AAV recombineering with single strand oligonucleotides. PLoS One. 2009; 4:e7705. [PubMed: 19888330]
34. Michel B, Grompone G, Flores MJ, Bidnenko V. Multiple pathways process stalled replication forks. Proc Natl Acad Sci USA. 2004; 101:12783–12788. [PubMed: 15328417]
35. Petermann E, Helleday T. Pathways of mammalian replication fork restart. Nat Rev Mol Cell Biol. 2010; 11:683–687. [PubMed: 20842177]
36. Weinert T, Kaochar S, Jones H, Paek A, Clark AJ. The replication fork's five degrees of freedom, their failure and genome rearrangements. Curr Opin Cell Biol. 2009; 21:778–784. [PubMed: 19913398]
37. Young SM Jr, Xiao W, Samulski RJ. Site-specific targeting of DNA plasmids to chromosome 19 using AAV cis and trans sequences. Methods Mol Biol. 2000; 133:111–126. [PubMed: 10561835]
38. Hewitt FC, Samulski RJ. Creating a novel origin of replication through modulating DNA-protein interfaces. PLoS One. 2010; 5:e8850. [PubMed: 20107513]
39. Porteus MH, Carroll D. Gene targeting using zinc finger nucleases. Nat Biotechnol. 2005; 23:967–973. [PubMed: 16082368]
40. Lobachev KS, Stenger JE, Kozyreva OG, Jurka J, Gordenin DA, Resnick MA. Inverted Alu repeats unstable in yeast are excluded from the human genome. EMBO J. 2000; 19:3822–3830. [PubMed: 10899135]
41. Storici F, Durham CL, Gordenin DA, Resnick MA. Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. Proc Natl Acad Sci USA. 2003; 100:14994–14999. [PubMed: 14630945]

42. Storici F, Lewis LK, Resnick MA. *In vivo* site-directed mutagenesis using oligonucleotides. Nat Biotechnol. 2001; 19:773–776. [PubMed: 11479573]

43. Samulski RJ, Srivastava A, Berns KI, Muzyczka N. Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats of AAV. Cell. 1983; 33:135–143. [PubMed: 6088052]

44. Katz SS, Gimble FS, Storici F. To nick or not to nick: comparison of I-SceI single-and double-strand break-induced recombination in yeast and human c ells. PLoS One. 2014; 9:e88840. [PubMed: 24558436]

45. Yan Z, Zak R, Zhang Y, Engelhardt JF. Inverted terminal repeat sequences are important for intermolecular recombination and circularization of adeno-associated virus genomes. J Virol. 2005; 79:364–379. [PubMed: 15596830]

46. Hirsch ML, Green L, Porteus MH, Samulski RJ. Self-complementary AAV mediates gene targeting and enhances endonuclease delivery for double-strand break repair. Gene Ther. 2010; 17:1175–1180. [PubMed: 20463753]

47. Grieger JC, Choi VW, Samulski RJ. Production and characterization of adeno-associated viral vectors. Nat Protoc. 2006; 1:1412–1428. [PubMed: 17406430]

48. Bower JJ, Karaca GF, Zhou Y, Simpson DA, Cordeiro-Stone M, Kaufmann WK. Topoisomerase IIalpha maintains genomic stability through decatenation G(2) checkpoint signaling. Oncogene. 2010; 29:4787–4799. [PubMed: 20562910]

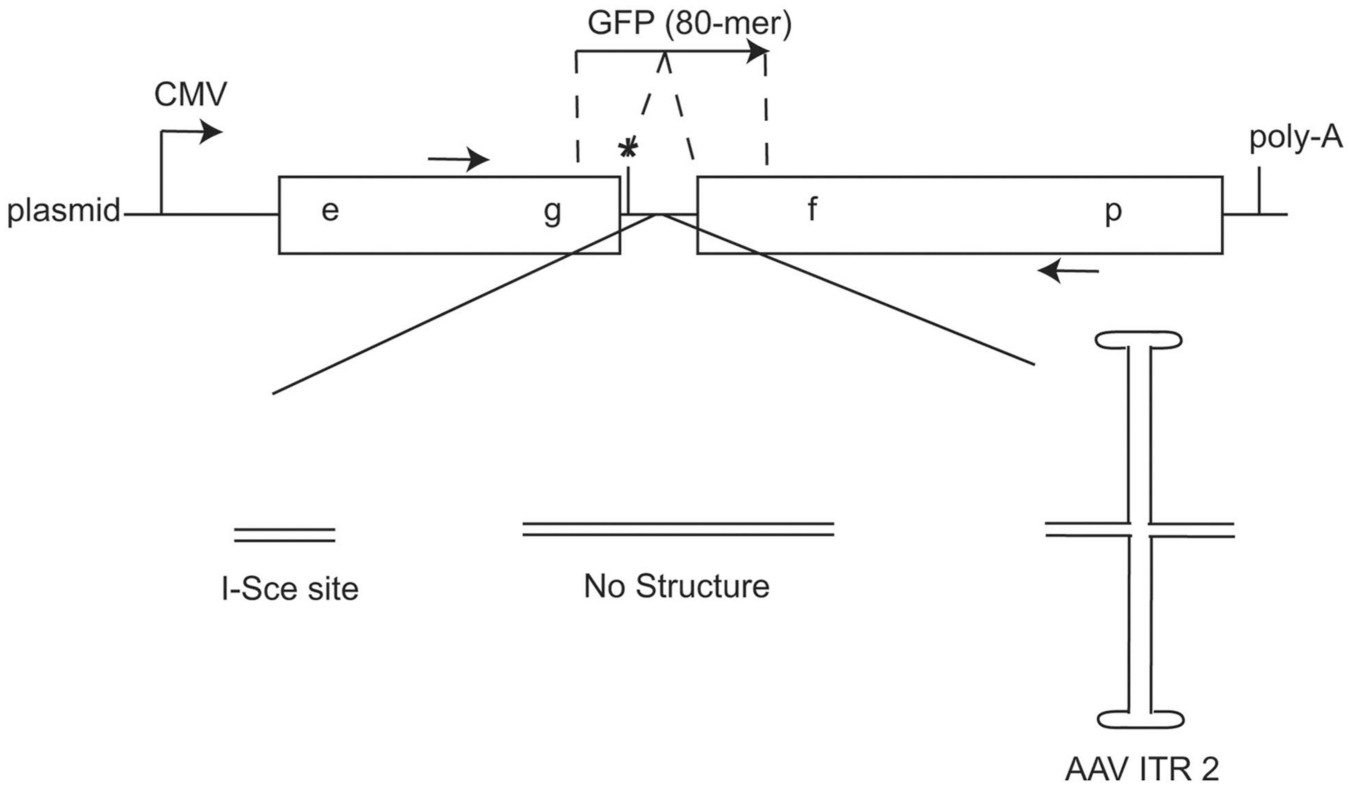

Figure 1.

Episomal reporter of gene editing. The episomal *egfp* reporter of gene editing contains a CMV promoter a defective *egfp* gene and a poly-adenylation sequence. The *egfp* gene is interrupted by out-of-frame stop codons (asterisk) and by the insertion of the depicted genetic elements: (i) I-Sce endonuclease site, (ii) a 165 nt sequence with no significant secondary structure and (iii) the 165 nt AAV2-ITR sequence. A sense single-strand DNA oligonucleotide having 40 nt of homology to either side of the insertion (dotted lines) is used as the repair substrate. A homologous recombination event involving the repair and target molecules results in a corrected *egfp* sequence and thus, a GFP+ cell phenotype.

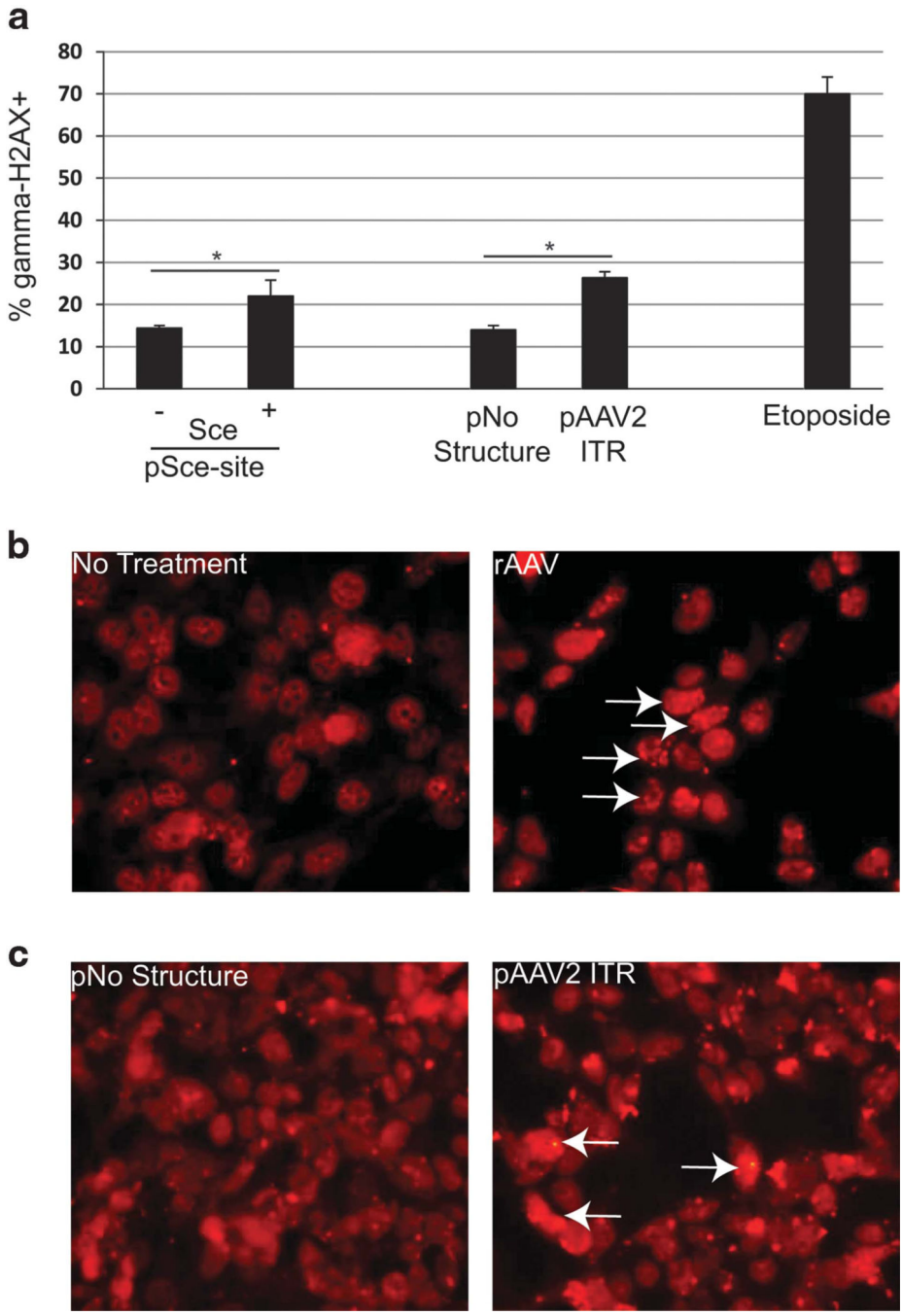

Figure 2.

The AAV2-ITR embedded in a plasmid context induces DNA damage signaling. (**a**) An episomal gene correction reporter containing one of the three insertions depicted in Figure 1 was transfected into 293 cells. Eight hours post transfection, cells were harvested and stained for gamma-H2 AX, which was then quantitated by flow cytometry. The I-SceI recognition sequence, in the presence of the I-SceI endonuclease, was used as a positive control for episomal double-strand breaks, whereas etoposide treatment was used as a positive control for host chromosome double-strand breaks. (**b**) 293 cells were given rAAV

and 12 h post infection, the cells were stained with an MreII antibody and observed by microscopy. (**c**) 293 cells were transfected with the indicated plasmids, stained with an MreII antibody 12 h later, and observed by microscopy. Arrows emphasize foci formation and the asterisk indicates a *P*-value < 0.05 using a student's *t*-test.

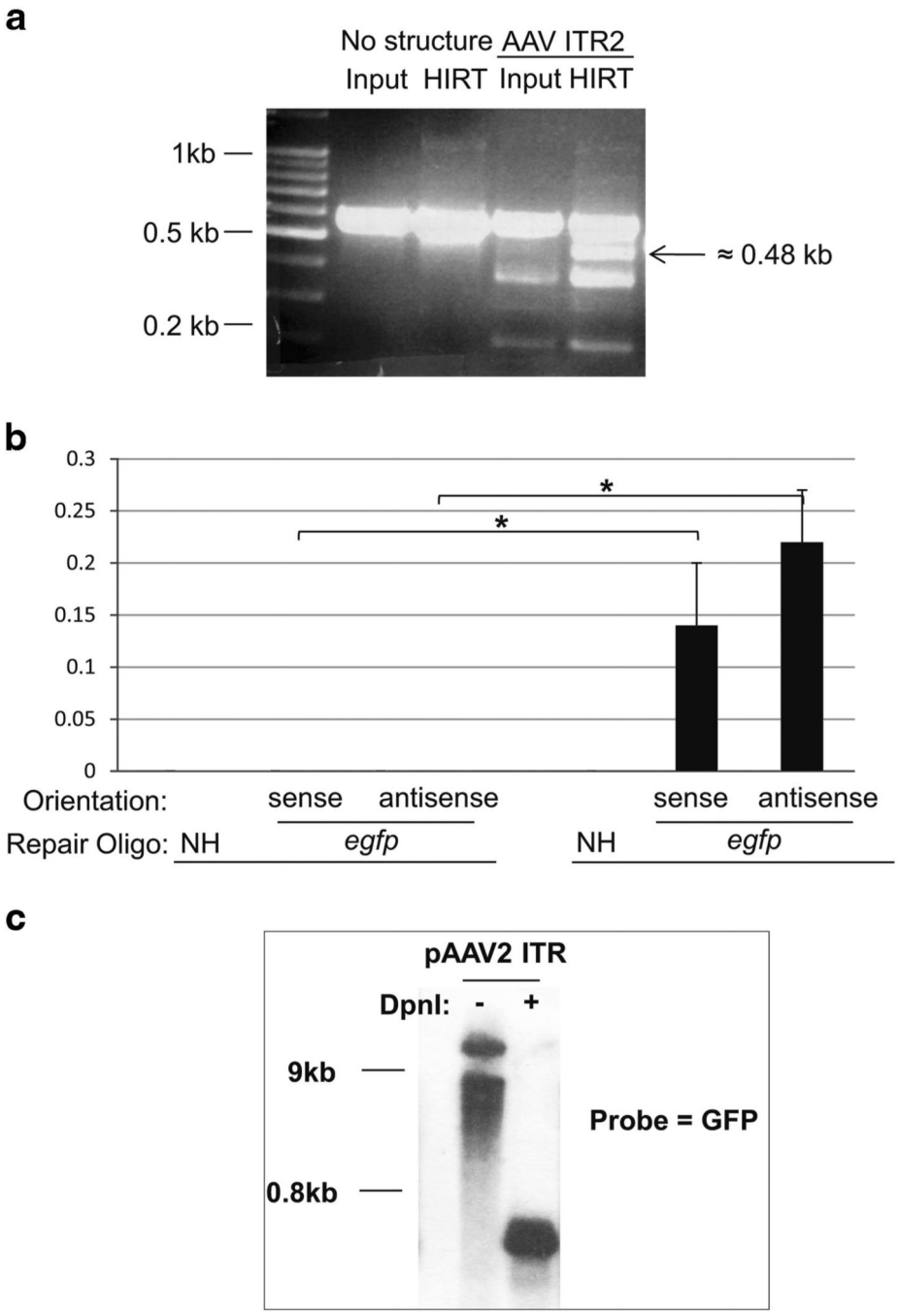

Figure 3.

The AAV2-ITR stimulates gene correction. (**a**) 293 cells were transfected with a defective *egfp* plasmid containing the No Structure or AAV2-ITR insertion sequences. Two days later, Hirt DNA was isolated and used as a PCR template with *egfp* primers designed to amplify across the insertion sequence. In a separate reaction, the input plasmid (used for the respective transfections) served as template. Amplicons were then separated on an agarose gel and visualized with ethidium bromide. (**b**) The gene correction reporter plasmid containing the indicated sequence insertion was used for 293 cell transfection. In addition, a

non-homologous (NH) or *egfp* repair oligonucleotide was included in the transfection as a repair substrate. Two days later, GFP+ cells were quantitated by flow cytometry. (**c**) 293 cells were transfected with the indicated plasmid and Hirt DNA was recovered 2 days later. The recovered DNA was analyzed by non-denaturing Southern blotting with or without prior digestion by DpnI. The asterisk indicates a *P*-value <0.05 using a student's *t*-test for the depicted samples.

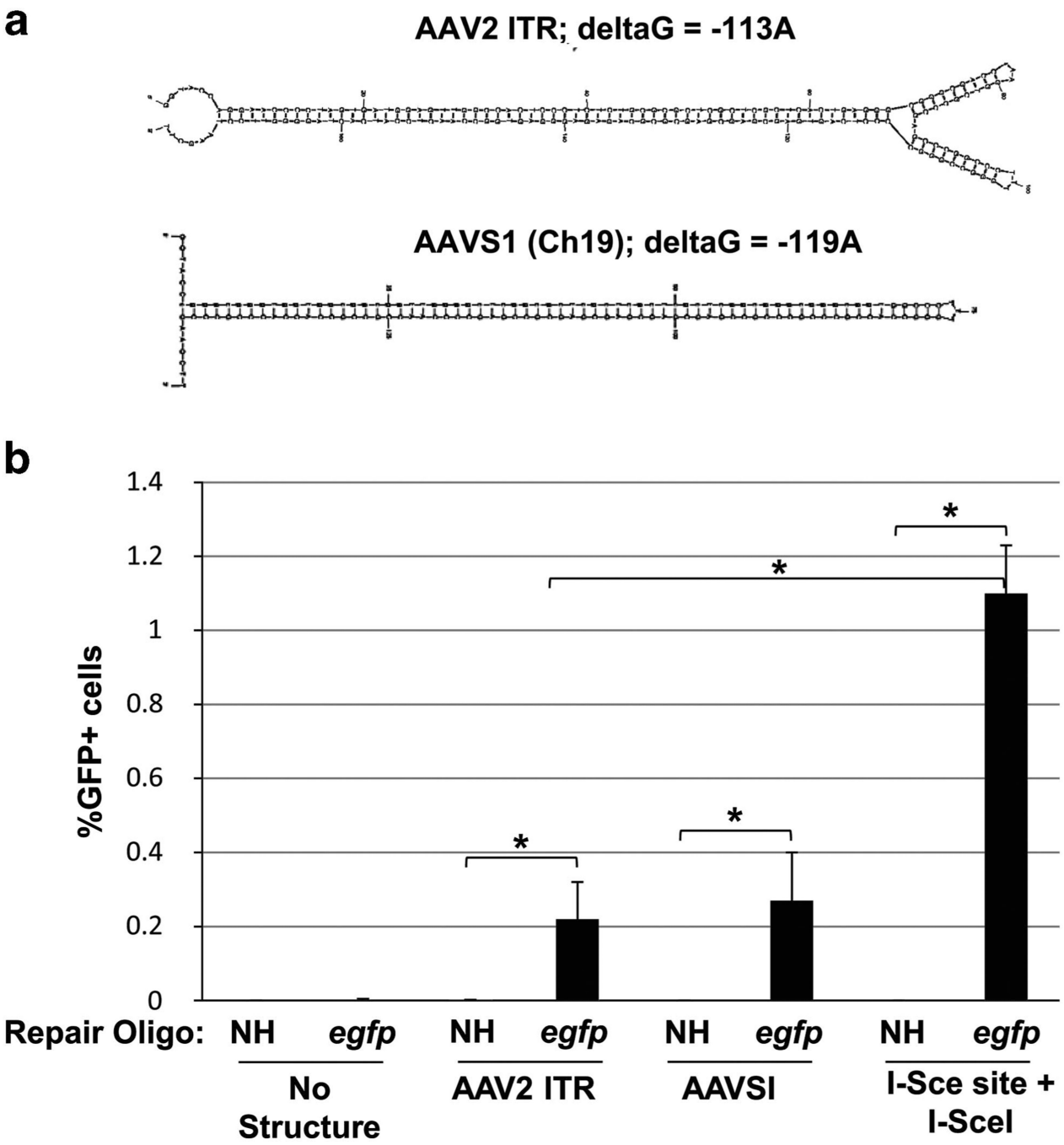

Figure 4.

Viral and human DNA hairpins stimulate gene correction. (**a**) DNA fold web server (http://mfold.rna.albany.edu/?q =mfold/dna-folding-form) was used to predict the secondary structures of the indicated hairpins. The most energetically favorable structure is depicted. (**b**) 293 cells were transfected with the gene editing reporter depicted in Figure 1, containing the indicated hairpin along with the indicated DNA oligonucleotide as a repair substrate (NH=non-homologous or *egfp*). Gene correction (GFP+ cells) was quantitated 72 h post

infection via flow cytometry. The asterisk indicates a *P*-value <0.05 using a student's *t*-test for the depicted samples.

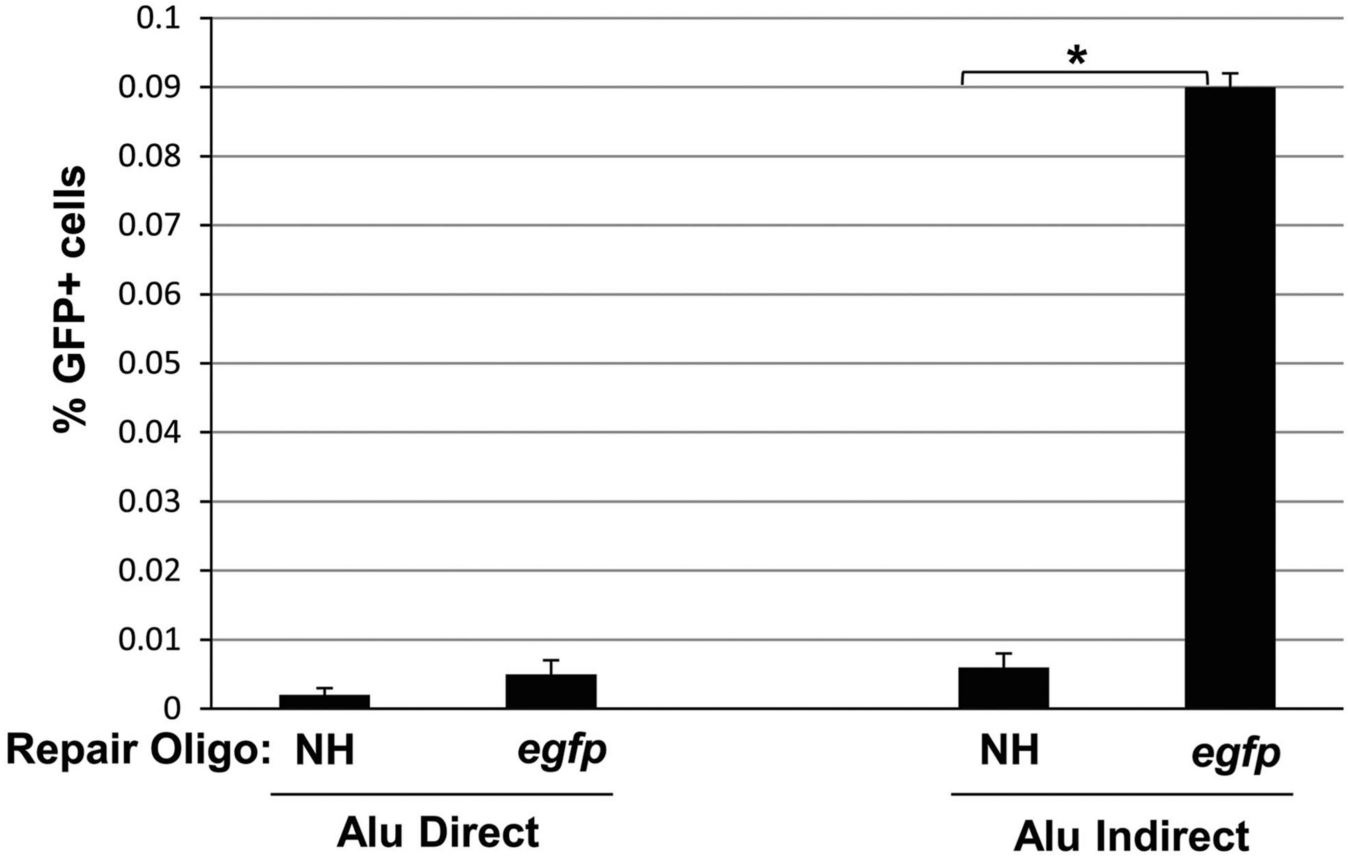

Figure 5.

Alu inverted repeats stimulate gene correction. Alu direct or indirect repeats were cloned into the gene editing reporter depicted in Figure 1. These resultant 'target' plasmids were then independently transfected into 293 cells along with the indicated repair oligonucleotide (NH=non-homologous or *egfp*). Three days post transfection GFP+ cells were quantitated using flow cytometry. The asterisk indicates a *P*-value <0.05 using a student's *t*-test.