Increasing the efficiency of targeted genome engineering by directly manipulating DNA double strand break repair pathways

Kamila S Gwiazda

A dissertation
submitted in partial fulfillment of the
requirments for the degree of

Doctor of Philosophy

University of Washington
2014

Reading Committee:
Andrew Scharenberg, Chair
David Rawlings, Co-Chair
Keith Jerome
Michael Lagunoff

Program Authorized to Offer Degree:
Molecular and Cellular Biology

University of Washington

Abstract

Targeted genome engineering is a technology that has the potential to vastly improve clinical medicine, as well as biological research. With the ability to design enzymes targeting any given DNA locus of interest, and mutate, delete, or repair the sequence, we can manipulate the genome to a degree that was not possible even a decade ago. However, both the rate at which these desired modifications occur following a nuclease-induced double strand break, as well as the type of repair that occurs, are variables we do not yet have full control over. In this work, we have found that introducing exogenous DNA end-modifying enzymes such as Trex2 and Artemis has a significant impact on promoting one type of repair, NHEJ. In contrast, altering the structure of the donor DNA by linearizing it *in vivo* promotes repair by HR. Additionally, we present evidence that not all designer endonuclease breaks behave similarly: those that leave 3' overhangs are much more amenable to repair by HR and SSA than those with 5' overhangs.

Increasing the efficiency of targeted genome engineering by directly manipulating DNA double strand break repair pathways

Kamila S Gwiazda

Chair of the Supervisory Committee:
Andrew M Scharenberg, MD &
Co-Chair David J Rawlings, MD
Department of Immunology

Abstract

Targeted genome engineering is a technology that has the potential to vastly improve clinical medicine, as well as biological research. With the ability to design enzymes targeting any given DNA locus of interest, and mutate, delete, or repair the sequence, we can manipulate the genome to a degree that was not possible even a decade ago. However, both the rate at which these desired modifications occur following a nuclease-induced double strand break, as well as the type of repair that occurs, are variables we do not yet have full control over. In this work, we have found that introducing exogenous DNA end-modifying enzymes such as Trex2 and Artemis has a significant impact on promoting one type of repair, NHEJ. In contrast, altering the structure of the donor DNA by linearizing it *in vivo* promotes repair by HR. Additionally, we present evidence that not all designer endonuclease breaks behave similarly: those that leave 3' overhangs are much more amenable to repair by HR and SSA than those with 5' overhangs.

Chapter 1

Introduction

Genetic engineering has been a tool used by humanity long before scientists began doing molecular biology: farmers have long been selecting the best crops and livestock, and animals have been bred for various desired traits. Not until Mendel discovered heritability in his famous pea plant experiments in 1866 did scientists begin to grasp the potential for a cellular basis of phenotype – elements later termed genes. Watson and Crick's famed discovery nearly one hundred years later, wherein they determined the structure of the DNA molecule, paved the way for the field of molecular biology, and allowed scientists to begin studying – and manipulating - the human genetic code in earnest.

1.1 A history of gene therapy

Incorporation of foreign DNA into the genome of cells, both for therapeutic and research purposes, is a desirable goal for the study of many biological systems and a possible treatment and cure for a myriad of debilitating diseases, congenital or otherwise. Following the discovery of recombinant DNA technology and the ability to deliver DNA into human cells, via viral vectors, the notion of gene addition to cure monogenic diseases was first suggested in the 1970's, and after a number of years and experimental models (1), was accepted as a viable clinical possibility.

The first human clinical therapeutic gene therapy trial was conducted on two young girls who had adenine deaminase deficient severe combined immunodeficiency

(ADA-SCID) in 1990. Their T-cells were transduced with the correct ADA gene by a retroviral vector *ex vivo* (2); while the response was moderate, the safety of the treatment was established, and a second ADA-SCID trial began in Europe (3). A number of gene transfer trials for different diseases followed, including the first showing *in vivo* adenoviral or retroviral gene delivery could successfully transfer a gene into the human brain, in this case the beta-galactosidease gene to image malignant glioblastoma (4). This began the work to use gene transfer as a method to not only treat congenital monogenic diseases, but also to treat various forms of cancer, which are currently the majority of the focus of gene therapy trials (65%) (5).

1.2 The need for targeted genome engineering

The first direct failure of a gene therapy trial came in 1998, with the death of a young patient treated with a very high dose of adenovirus for a deficiency of ornithine transcarbamylase; his death, due to multi-organ failure, was directly attributed to the vector used in the treatment (5). Two years later, a notable trial began whose long-term results highlighted the need for a different approach: of the ten children treated for X-linked SCID with retroviral vectors driving gC expression from the viral LTR, nine responded positively but four developed leukemia as a result of semi-random vector integration near the proto-oncogene LMO2, among others (6). Analysis of viral integration profiles revealed that the gc retroviral vector preferentially integrated within transcriptionally active genes or gene regulatory areas, leading to insertional mutagenesis or oncogenesis (7, 8). Another trial for chronic granulomatous disease resulted in similar problems using a retroviral vector, where transgene-expressing cells

exhibited clonal dominance due to insertional mutagenesis and a consequent survival advantage (9). While the development of safer vectors, the emergence of 'safe harbor' genomic integration sites, and numerous subsequent trials that have not led to any adverse events, the fear of insertional mutagenesis and oncogenesis remains an important consideration for the use of traditional gene therapy.

Meanwhile, studies in yeast, later confirmed in mammalian cells (10) discovered that creating a double-strand break increased the rate of homologous recombination repair 2-3 orders of magnitude at the site of the double strand break (11) . Concurrently, endonuclease enzymes were discovered in yeast that could create targeted double-strand breaks in specific areas in the genome, based on sequence-specific protein-DNA interactions (12). This led to the suggestion that rather than relying on gene addition by viral vectors (for which random (or semi-random) insertion of a transgene could induce insertional mutagenesis (13), and which lack endogenous regulatory sequence and inherited epigenetic marks), direct editing of the endogenous mutation could be achieved, built off the platform of these endonucleases. Thus targeted genome engineering, both for targeted gene repair, as well as gene disruption, began to take shape (14, 15).

.

1.3 Development of designer endonuclease technology

The discovery of the yeast I-SceI and HO homing endonucleases (HE's) has spurred the discovery (16, 17), as well as engineering (18, 19, 20) of many more HEs; as well as the development of synthetic designer endonucleases: fusions of a specific DNA binding domain to a DNA cleavage domain. Presently, there are four main "flavors"

of designer engineered enzymes: homing endonucleases (HE's), zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and CRISPR RNA-guided technology. Additionally, a newly developed fusion between homing endonucleases and TALENs, termed megaTALs (21) is a promising hybrid technology that addresses the main drawbacks of the HE and TALEN platforms.

All designer endonucleases function by binding their cognate DNA targets and creating a double-strand break, though they differ in their biochemistry of cleavage: HE's and megaTALs leave 3' DNA overhangs, while the Fok1 nuclease-based TALENs and ZFNs leave 5' DNA overhangs. CRISPR, which is based on RNA guides to bind target chromosomal loci and the Cas9 endonuclease cleavage, is a blunt cutting enzyme, though the creative design of RNA guides and Cas9 nickases (22) has enabled myriad opportunities for creating different types of DSBs made by this system – although it is not yet known to what extent coordinated nicking is able to occur, particularly for closely spaced nicks which would be predicted to create steric hindrance between adjacently bound CRISPR Cas9/guide RNA complexes.

1.3a Homing endonucleases

Homing endonucleases are naturally occurring selfish genetic elements, existing only to replicate themselves. They are mostly found in the fungal and archael kingdoms, and to date, only a fraction of the potential existing HE's have been discovered and characterized (17). The type of HE used in most genome engineering studies is the LAGLIDADG family, which consists of DNA binding and cleavage domains that act in a concerted manner to bind and cleave their DNA targets. HE DNA targets are, on average, 18-20bp long; the minimum length for a unique target in the human genome,

statistically, is 14bp. Importantly, and uniquely to HEs, because their DNA binding and cleavage are concerted, there is less chance for off-target cleavage compared to dimeric enzymes where the two activities are unlinked, even though the dimeric enzyme DNA targets might be longer.

The cleavage activity, DNA binding kinetics, and enzyme turnover are, like in any enzyme family, unique for each individual HE, and in the past the difficulty in engineering HEs towards new targets has impeded their use. However, with structure-based rational design and directed evolution, the activity of an enzyme towards a specific target can be greatly improved (23) or altogether re-engineered towards a new target (20, 24). Additionally, new technologies are drastically reducing the time required for engineering of new HE targets from existing scaffolds, which to date has been a laborious and time-intensive process (19, J. Jarjour, unpublished). Together with discovering new scaffolds, or creating new ones by making chimeric HE's (25), creating HEs for a given genomic target is now a very realistic and achievable possibility.

There are a number of reasons that HEs remain some investigators' designer nuclease of choice for genome engineering endeavors. Compared to other nucleases, they are small, compact, single-protein enzymes, making them simpler to deliver to cells of interest via vectors with limited packaging capacity (21). Further, because they are monomeric proteins with concerted binding and cleavage, multiplexing is possible: multiple different HEs can be delivered at the same time to hit many different targets, without the worry of cross-reactivity. Perhaps one of the more important benefits, however, is their biochemistry: they leave double strand breaks with 4bp 3' overhangs, to which the DNA repair machinery responds differently than to 5' overhangs (26).

Indeed, these 3' overhangs can be leveraged to both specifically increase either gene disruption or gene repair to a greater extent than 5' overhangs, a concept which will be explored further in this thesis.

1.3b Zinc Finger Nucleases (ZFNs) and Transcription like effector nucleases (TALENs)

ZFNs and TALENs are both modular Fok1-based hybrid nucleases. ZFNs are made up of three or four zinc finger proteins, which each make contact with 3bp of DNA, bound to the non-specific Type II-S endonuclease Fok1. TAL effectors are made up of repetitive protein arrays (33-35 amino acids each), which contain DNA binding specificity within a specific location of each array, termed the Repeat Variable Diresidue (RVD). Both enzyme platforms are based on Fok1 DNA cleavage, and since this endonuclease must dimerize, both ZFNs and TALENs are delivered as pairs that dimerize *in vivo*. Wild-type Fok1 has since been engineered to be more active (27), as well as act as an obligate heterodimer, thus only dimerizing with its obligate pair to reduce off-target cleavage and reduce genotoxicity (28).

TALEs were originally discovered as plant pathogens secreted by Xanthomonas bacteria before they became TALENs (29), and they are highly modular and easily engineerable: each RVD corresponds a given nucleotide sequence, so building a TALEN target is relatively simple compared to other platforms. Several models have been built to make TALEN construction fluid and high-throughput (eg. 30, 31). ZFNs, which are also somewhat modular, are not as simple to engineer as initially anticipated (32), although computational platforms (33), as well as company programs (34) have been built to address some of these issues. However, the ease of TALEN construction

and targeting is a considerable advantage to this platform, although the size and repetitive nature of TALENs makes them difficult to package in some viral vectors (21).

The safety profile of ZFNs and TALENs varies depending on each enzyme; the rates of off-target cleavage depend on the specificity of zinc finger/TALE binding, among other considerations. Both computational (35) and high-throughput *in vitro* selection (36) methods have been designed to predict ZFN and TALEN off-target cleavage. Interestingly, despite the two platforms sharing the same DNA cleavage enzyme, ZFNs and TALENs leave distinct mutagenic signatures, at both on- and off-target sites (37). While TALENs are relatively new, only having been discovered five years ago, ZFNs have been systematically developed over the last two decades, (38) and have been successfully used in a large number of animal models to knock out various genes (see 1.4). Indeed, a pair of carefully characterized ZFNs is currently in clinical trial for treatment of HIV by disrupting the CCR5 gene in the CD4+ T-cell compartment (see 1.4).

1.3c Mega-TALs

MegaTALs are a fusion between a TALE array and a homing endonuclease, joined by a flexible linker on the C-terminal half of the TALE and N-terminal end of the HE (21). This platform was developed to harness the advantages of HEs and TALENs, while minimizing their detriments. MegaTALs are single-chain proteins, and because they contain two intrinsic DNA binding domains (TALE and HE), the TALE array number can be shortened without significantly affecting target specificity, thereby lessening the number of TALE repeats and making it more amenable toward lentiviral packaging and delivery (21). By extension, multiplexing is also more possible with megaTALs than with

ZFNs or TALENs: because HEs do not need to dimerize in a megaTAL, multiple enzymes can be delivered to target various targets at once, without concerns about cross-reactivity as encountered with a Fok1-restricted dimerizing nuclease.

Beyond easing nuclease delivery, megaTALs significantly lower the engineering bar for HEs. Due to the additional binding affinity that the TALE domain provides, lower affinity HEs can be used in megaTAL fusions and be sufficiently active enzymes, whereas on their own the cleavage activity of a low affinity enzyme – which is intrinsically tied to their binding affinity – would be quite poor (21, 20).

1.4 Genome engineering at work: animal and human disease models

The previous decade has seen many different laboratories showing that genome engineering (both gene repair and gene disruption) is possible, in many different animal and plant models. Using ZFNs, HEs, and most recently TALENs, a collection of labs has made significant progress establishing both the safety and efficacy of these tools in mouse, rat, zebrafish, *Drosophila*, *Xenopus*, *C. elegans*, *Bombyx mori*, a number of different plant species, along with cultured and primary human cells (reviewed in (39, 40). With the aforementioned technological advances in HE structural design, enzymes have been engineered towards various human loci: Btk, XPC, RAG, (20, 24, 41). And importantly, all four platforms (megaTALs, HEs, ZFNs and TALENs) have now been successfully used to modify not only cultured human cells, but now ES and iPS cells as well (42, 43, 21, 44).

Of active translational interest are a number of projects using genome engineering tools: various cancers, in so far as studies are being conducted to knock-

out the T-cell receptor in CD-19+ chimeric antigen receptor (CAR) T-cells (45) or in cytotoxic T lmphocytes (CTLs) (46), as well as knocking out latently infected cells containing viral DNA (47), such as with HSV (48) or HBV (49) using HE's or ZFNs.

One of the first clinically relevant proof-of-principle genome engineering studies successfully treated a murine model of hemophilia B using ZFNs targeted to F9 (factor IX) with an AAV based targeted vector (50). The first human clinical trial using a designer endonuclease is currently underway using a CCR5-targeted ZFN (51, 52). The rationale is based on the 'Berlin patient' (53), an HIV positive patient who was treated for acute myeloid leukemia using a HSC transplant from a CCR5 Δ32 donor, and whose HIV viral loads have been subsequently eradicated (54). Consequently, ZFNs were developed to disrupt the CCR5 gene near the Δ32 locus, and when delivered *ex vivo* to primary human T-cells (55) and HSCs (56) were successfully expressed and able to create indels at the target locus. Most importantly, these disrupted cells engrafted into HIV-infected mice led to lower viral loads (56).

1.5 Key obstacles in the field

Much progress has been made in developing a large variety of designer endonucleases, as outlined above. Despite these promising advances, the efficiency of targeted repair and bi-allelic targeted disruption with engineered nucleases in primary cells remains low (55, 52). Further, occurrences of deleterious translocations resulting from persistent and/or off-target DNA breaks have been observed [2,3,4]. Importantly, we do not yet have a concrete understanding of DNA end processing that occurs at the

DSB site, a means of precisely controlling the outcome, nor a way to bias repair toward a certain pathway, although advances have been made towards these aims (44).

Among the main hurdles to overcome is the low frequency of homologous recombination in primary mammalian cells, even with the inclusion of a double-strand break. There is evidence that in post-mitotic somatic cells, HR no longer occurs, since it is dependent on cells entering S or G2 phase of the cell cycle (57). Depending on the context of the experiment, selection – both positive or negative – can be applied, such as in the case of mouse ES targeting, but when scarless gene repair is desired, a selectable marker cannot be inserted to select for cells that have undergone an HR event. In some cases, low frequency HR gene repair could be sufficient, such as when the repaired cells have a strong survival advantage over non-repaired cells (E. Curinga, unpublished). However, this scenario is not representative of all cell compartments, nor is it ideal, especially when cells are being repaired *ex vivo* and so cannot necessarily be expanded before infusion.

Another key consideration is delivery of nuclease and donor to the cell type of interest. Unlike gene therapy, this is based on transient expression, and thus not suited for integrating lentiviral vectors. Non-integrating viral vectors have been developed, including non-integrating lentivirus and AAV (eg. (58, 59). Although some of these vectors have had success transducing many different cell types of interest, no vector can transduce all human cells. Promising new methods involving transient delivery are also being developed, including mRNA (60, 61), minicircles (62), and nanoparticles (63).

Taken together, it can be seen that the field of genome engineering has made incredible progress in the past decade, but significant challenges remain. The aim of

this thesis is to define molecular mechanisms at work in manipulation of DNA end processing and donor DNA structure, with the goal of increasing the efficiency of both targeted genome disruption and repair.

1.6. Double-strand break repair pathways

1.6a Non-homologous end-joining (NHEJ) and homology-directed repair (HDR)

Double-strand breaks (DSBs) are the most dangerous lesions a genome can encounter, and thus there are a number of ways cells can repair the damage, which are highly conserved among eukaryotic cells. They can be grouped into two main pathways (**Fig. 1**): classic non-homologous end-joining (cNHEJ), which can seamlessly ligate breaks or introduce small insertions/deletions (indels); or homology-directed repair (HDR), a 5' resection-dependent pathway that uses varying amounts of homologous

DNA to copy sequence information to repair the break.

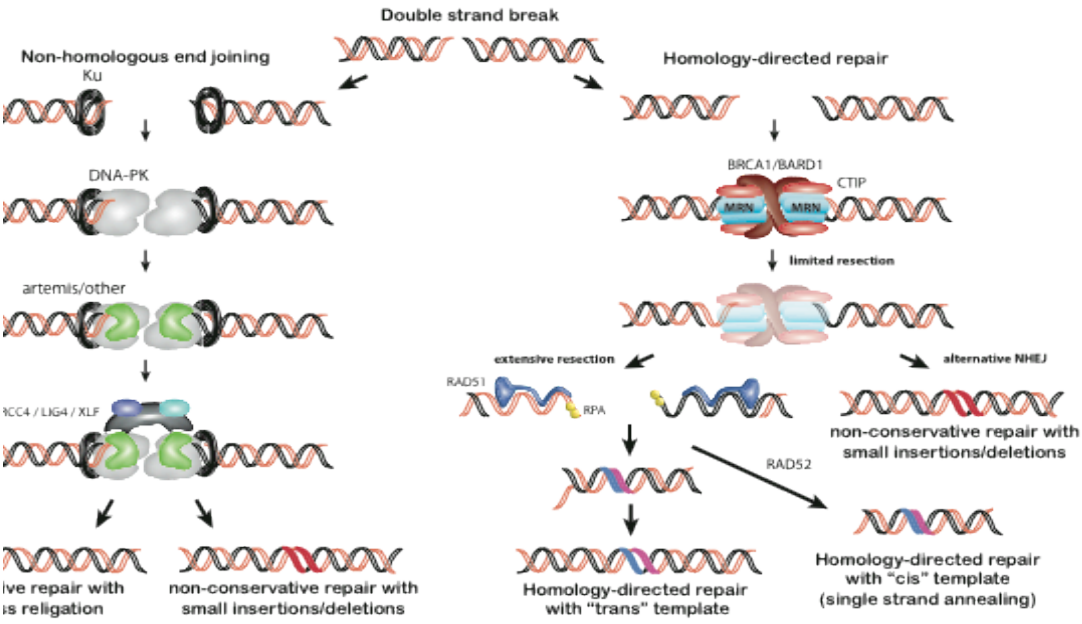


Figure 1. **Double strand break repair pathway choice**. The two main pathway arms, non-homologous end-joining (NHEJ) and homology-directed repair (HDR) first diverge at the step of end-resection, by MRN, or DNA-binding by Ku.

Following a DSB, the damage is sensed by key signaling proteins that can, in part, determine which pathway the DSB will be repaired by. To initiate cNHEJ, the protein heterodimer Ku binds the free DNA ends, which recruits the kinase DNA-PKcs, promoting its auto-phosphorylation and phosphorylation of other targets. This complex prevents resection by MRN, but recruits processing factors such as Artemis and others (nucleotide kinase, gap filling polymerases, if required). Ligation is performed by DNA ligase IV, scaffolded by XRCC4 and XLF. In contrast, HDR begins by MRN binding to DNA, together with auto-phosphorylation of the kinase ATM, to begin resection and arrest of the cell cycle (ATM-dependent signals). MRN, a complex made up of three proteins – one of which, Mre11, is an exonuclease – beings the initial resection step. However, for both homologous recombination (HR) and single-strand annealing (SSA,

detailed below), long range (secondary) resection is required, and this is carried out by different nuclease/helicase pairs: either Exo1/Bloom, or Dna2/Bloom (64). If secondary resection does not occur, the MRN-resected DNA can be ligated by microhomology-mediated end-joining (MMEJ, detailed below). HR is the only HDR pathway that can be error-free, as it uses a homologous strand of donor DNA as the template to repair the break. This is most often the sister chromatid, but can also be a homologous chromosome or exogenously delivered donor template (eg. plasmid, viral vector). Importantly, the type of DNA break can influence its repair: designer endonucleases leave chemically clean breaks, which require minimal processing, and thus can be seamlessly ligated. This can result in the re-creation of the target site (65), which leads to a cycle of cleavage and precise repair, creating a 'persistent DNA break', where only a small subset of breaks escape the cycle by being resolved in a mutagenic way (66). Indeed, it is estimated that of double strand breaks repaired by NHEJ, ~15% require the nuclease processing activity of Artemis (67), suggesting that this small subset needs to be modified before the ends can be ligated. This subset of breaks is thought to include DNA modifications caused by damaging agents (68) that need to be destroyed before the DNA can be enzymatically ligated.

1.6b Microhomology mediated end-joining (MMEJ) and single strand annealing (SSA)

As noted above, HDR includes strict homologous recombination, as well as single strand annealing (SSA) and microhomology mediated end-joining (MMEJ). All three are resection-dependent, and thus are thought to be downstream of the MRN arm of the DSB repair pathway. MMEJ and SSA differ from HR in that they are not strictly

template dependent and are always mutagenic; they are thought to be back-up pathways when cNHEJ or HR fail (69, 70, 71).

MMEJ, also called alternative non-homologous end-joining (aNHEJ), is less-well understood than other repair pathways. However, it is known to compensate for repair when key cNHEJ factors are lacking (eg. LigIV, Ku: 71). Rather than relying on DNA Ligase IV/XRCC4 for ligation, MMEJ uses Ligase III and XRCC1 (72). After resection, the DNA ends are typically joined using micro-homologies (less than 10nt) found at or near the ends.

Single strand annealing (SSA) is arguably the most mutagenic repair pathway, as it always involves considerable loss of sequence. It is an active pathway in yeast, and recently was discovered to be active in human cells as well (72, 26), though the extent to which it is used may depend on cell type and the context of the break (72, 26, 71). Like HR, SSA involves two-step resection, but instead of using an extrachromosomal DNA donor, SSA depends on finding intra-chromosomal sequence repeats, followed by deletion of the intervening sequence. As the human genome consists of a considerable amount of repetitive DNA (73), there is considerable chance that SSA could cause extensive genomic damage; the aspect of how SSA is regulated is less well understood.

1.6c Assessing DNA repair via flow-cytometry

1. Traffic Light Reporter (TLR)

To study DNA repair of an endonuclease-mediated double-strand break, the Traffic Light Reporter (TLR) assay was developed in the Scharenberg lab (**Fig. 2**) (74). Briefly, this assay is based upon stable genomic integration and expression of the TLR cassette,

which can be delivered by lentivirus or plasmid electroporation, in a cell line of choice. The assay measures DSB repair at a nuclease target site, built within the reporter, by expression of fluorescent proteins: GFP, which indicates repair by homologous recombination, and mCherry, which indicates a mutagenic NHEJ event resulting in a +3 frame-shift. The TLR reporter thus allows the measure of DSB repair following treatment with a designer endonuclease using flow cytometry, as well as the ability to sort cells that are known have undergone certain repair events, which can then be further analyzed at the genomic level.

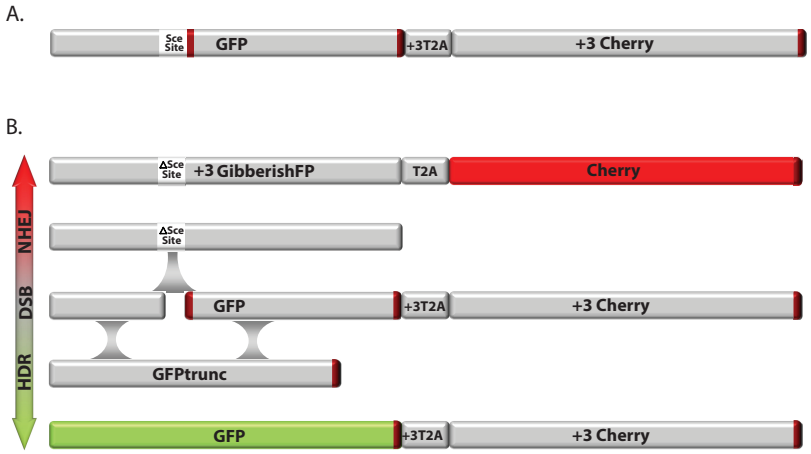


Figure 2. Diagram of the TLR. Schematic depicting different engineering outcomes following the induction of a site specific double stranded break (DSB). If the break is resolved through the HDR pathway the full GFP sequence will be reconstituted and cells will fluoresce green; if the break undergoes mutagenic non-homologous end joining, GFP will be translated out of frame and the T2A and mCherry sequences are rendered in frame to produce red fluorescent cells.

2. Single-strand Annealing Traffic Light Reporter (SSA-TLR)

A modified TLR reporter was built to measure SSA in addition to HR and NHEJ: the SSA-TLR (**Fig. 3**) (26). This reporter cassette is flanked by an interrupted iRFP open reading frame, immediately downstream of the promoter, that shares 762bp of

homology on either end, separated by 1.8kb of [GFP.T2A.mCherry] coding sequence. A nuclease target site is embedded within the GFP ORF, as in the standard TLR; following treatment with designer endonuclease, repair by SSA is possible if enough resection occurs to expose the homologous iRFP arms on either end of the reporter cassette (26). The advantage of the SSA-TLR, over similar reporters designed in the past (66, 72), is that it allows simultaneous measurement of all three DSB repair pathways and can provide unbiased data with respect to cellular repair pathway choice, within a given context of cell and DSB break type.

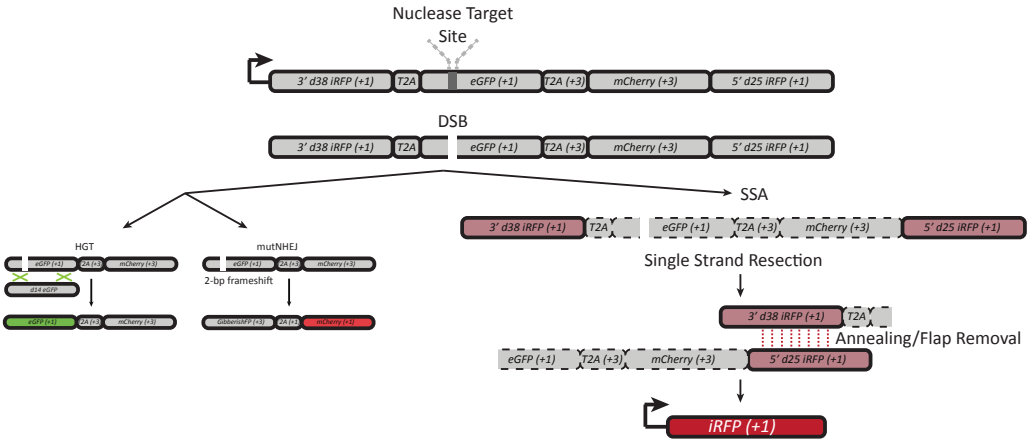


Figure 3. Diagram of the SSA-TLR. Schematic showing the different genome editing outcomes after a DSB is made. The previously depicted NHEJ and HDR pathways remain the same as above and result in GFP and mCherry expression respectively. If the break undergoes repair via SSA, single-strand resection will reveal the homology between the two arms, which will subsequently be processed to result in a fully functional *iRFP* ORF.

Chapter 2

Pairing Exonucleases with Designer Endonucleases to Increase Targeted Gene Disruption

2.1 Introduction

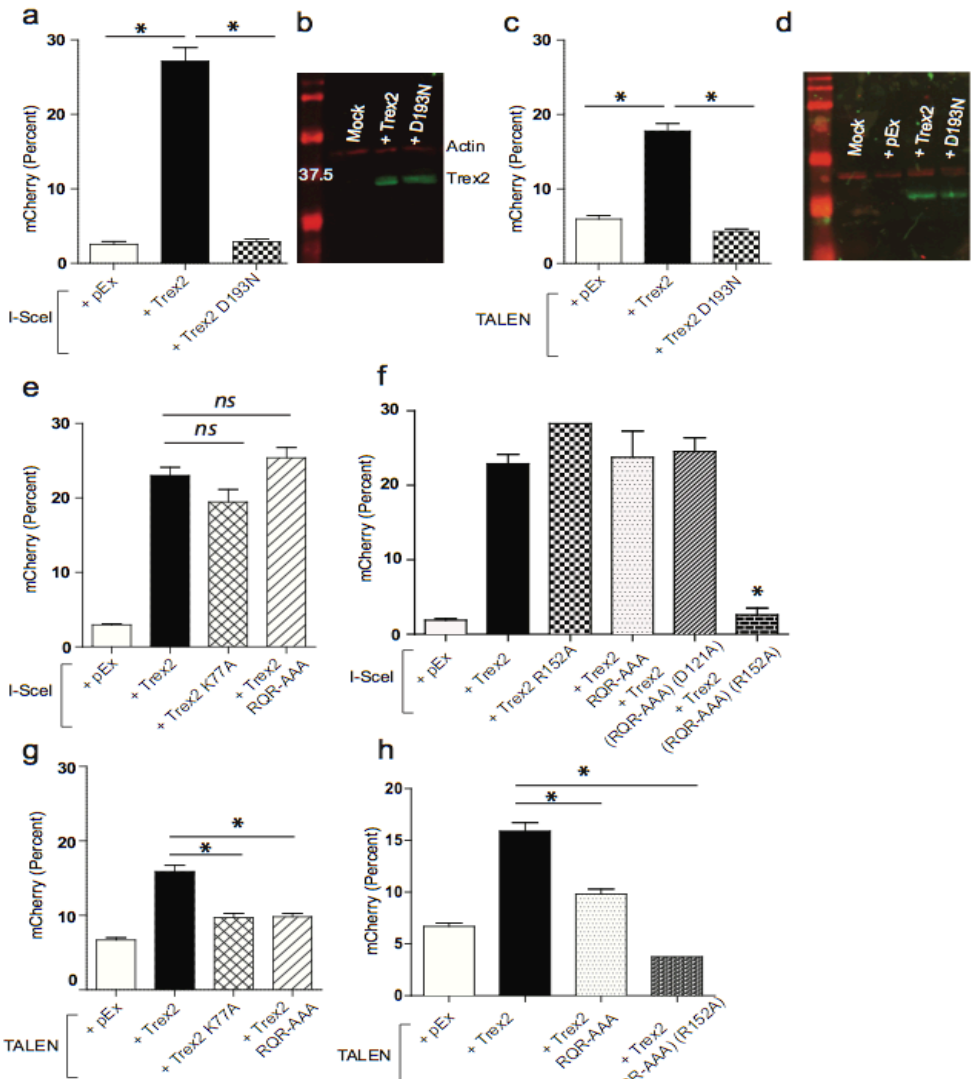
Targeted gene disruption has been relatively easier to achieve than targeted repair: mammalian NHEJ is the most active pathway in most cells types, can be engaged at all points of the cell cycle, and is the main repair pathway in post-mitotic cells. However, since designer endonucleases leave chemically clean breaks that require little processing, and thus engagement of the classical (Ku-dependent) NHEJ pathway frequently leads to seamless repair, a cycle of ligation and re-cleavage can occur, leading to what has been termed a 'persistent break' (66). This cycle is broken only when a mutagenic repair event occurs and the nuclease site is destroyed: either when a processing enzyme is recruited (such as Artemis or a gap filling polymerase) to enzymatically modify the broken DNA ends, or if mutagenic microhomology-mediated end-joining (MMEJ) is stimulated rather than cNHEJ. Importantly, the initiation step of MMEJ is thought to be different than cNHEJ, in that MMEJ begins with DNA end resection rather than Ku and DNA-PKcs binding; thus one possible way to bias pathway choice towards mutagenic repair is at this initial pathway divergence.

The activity (DNA binding and cleavage) of a given endonuclease can determine its efficacy *in vivo*, and factors such as epigenetics and chromatin compaction (26) can influence its access to the DNA. Together, the combined affect of these factors is broad, and this has lead to a wide range of mutagenic frequencies seen with current designer

endonuclease enzymes. Ultimately, the goal is to increase the frequency of targeted gene disruption, especially as bi-allelic disruption is frequently desired, by a general mechanism applicable to all designer nuclease platforms.

Previous attempts at this have had moderate success. One of the first was cold shock – incubation of cells at 30°C rather than 37°C – has been shown to increase the frequency of mutNHEJ repair with ZFNs in two different mammalian cell lines (75). Mild hypothermia treatment slows cell growth and increases steady-state levels of ZFN protein; thus the authors suggest that the increase in ZFN accumulation allows for more opportunity for on-target DNA cleavage. This method has since been effectively applied towards primary T cells (45), as well as TALENs (A. Grier, unpublished). Other approaches include attempts to steer DSB repair pathway choice towards mutagenic NHEJ by pharmacological means to inhibit HR proteins (76). However, direct inhibition of the DSB response can put cells at risk for apoptosis (for example, inhibition of HR affects cells' ability to fix stalled replication forks), and so is not an ideal solution.

Previous work has suggested that the 4 base-pair (bp) 3' DNA overhangs left by the HE I-SceI can be processed by exogenous delivery of the non-processive 3-5' exonuclease Trex2, thereby limiting cycles of seamless re-ligation (66). Trex2 is a recently discovered homolog of Trex1, the predominant 3' mammalian exonuclease responsible for processing foreign DNA (77, 78), but has no known function in mammalian cells. We hypothesized that expressing Trex2, and perhaps other enzymes that directly modify DNA ends, following an endonuclease-induced DSB would be an effective and safe means to increase mutNHEJ.

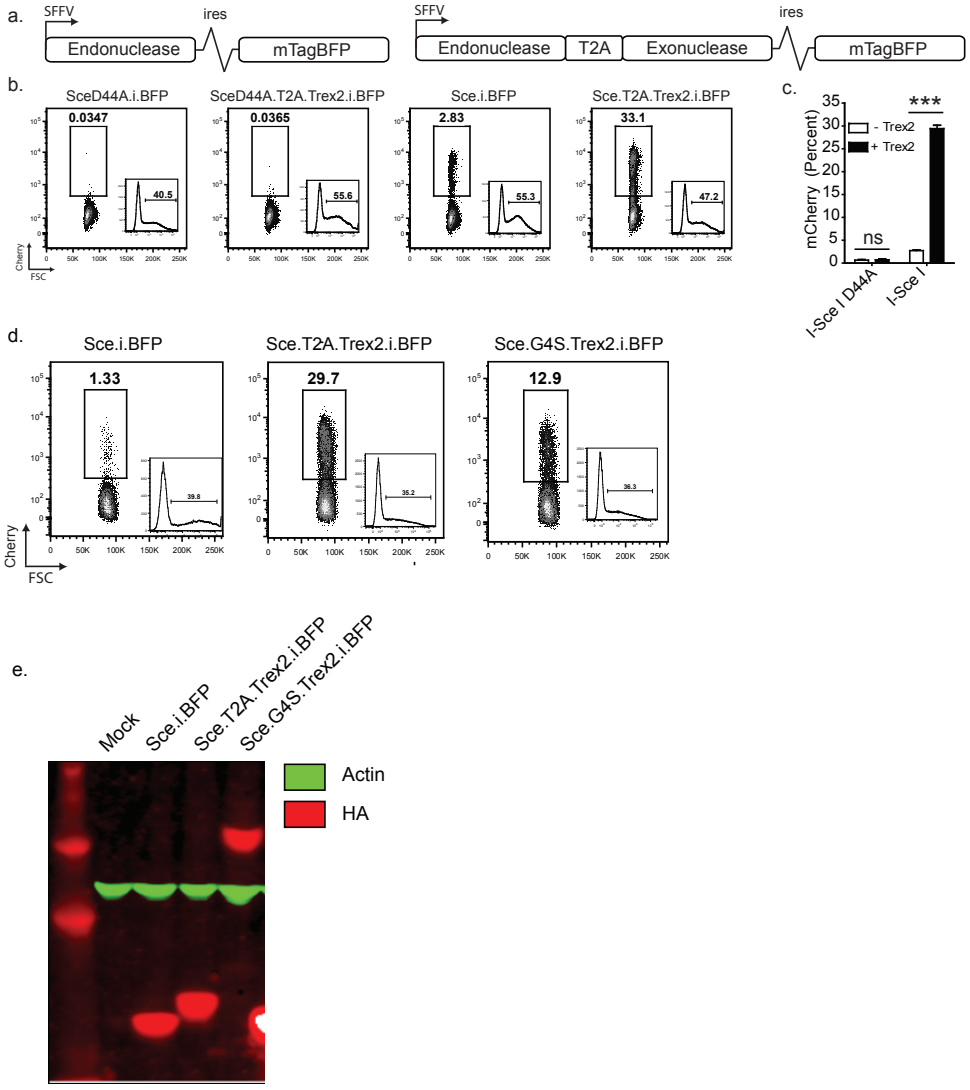


2.2 Using the exonuclease Trex2 to process 3' overhangs left by HE DSBs

Given the earlier Trex2 *in vitro* data (66), we hypothesized that we could extend this finding to homing endonucleases *in vivo* to increase mutagenic NHEJ events, thereby increasing targeted gene disruption. Thus, we co-expressed Trex2 with the model HE I-SceI in HEK293T cells that stably expressed the Traffic Light Reporter (with I-SceI target), and found that over-expression of Trex2 significantly increased the proportion of mutagenic NHEJ (mutNHEJ) outcomes, as quantified by TLR mCherry

expression (**Fig. 4a, b**). This reporter generates mCherry fluorescent positive cells in response to a frameshift into the +3 reading frame at the target site, indicating targeted disruption. Two different constructs were assessed to co-express Trex2 with I-SceI: a flexible linkage that directly couples the two enzymes, termed G4S, or a peptide linker that enables two separate proteins to be translated, termed T2A, a viral 'peptide skip' motif. All constructs included an mTagBFP fluorescent protein to track transfection efficiency and protein expression, downstream of an internal ribosome entry site (IRES). Driven by a viral SFFV promoter, the construct coupling I-SceI with Trex2 by T2A significantly increased mutagenic NHEJ 11-fold over I-SceI alone (**Fig. 4c**). However, the construct that directly linked I-SceI to Trex2 with a G4S linker gave a significantly lower mutNHEJ frequency compared to the T2A linker (**Fig. 4d**), although no degradation of I-SceI was seen using a western blot (**Fig. 4e**). We reasoned that since Trex2 must dimerize to be fully catalytically active (79), perhaps being directly linked to a homing endonuclease may inhibit some of its catalytic activity. We thus used T2A constructs for all subsequent experiments involving Trex2 co-delivery. Additionally, the enzymatically dead I-SceI^{D44A} construct, with or without Trex2, had no effect on mutNHEJ, signifying that a double-strand break must be made by the HE in order for Trex2 to exert its effect.

Figure 4. Coupling endonucleases to exonucleases increases gene disruption. (**a**) Schematic of Trex2 exonuclease expression vectors. SFFV, spleen focus-forming virus promoter/enhancer; IRES, internal ribosomal entry site. (**b**) Representative flow plots of mCherry (mutNHEJ) of HEK293T cells transfected with the indicated vectors, with BFP expression in the inset. (**c**) Quantification of mCherry expression in BFP-positive HEK293T cells transfected with the indicated vectors and analyzed 72 h after transfection (*n* = 3). (**d**) Representative flow plots of mCherry (mutNHEJ) of HEK293T cells transfected with the indicated vectors, with BFP expression in the inset. (**e**). Western blot of HEK293T cells transfected with the indicated vectors.



To extend the Trex2 observations to other homing endonucleases, we co-expressed it with I-AniI, both with the native less active enzyme, as well as the more active Y2 variant (23) in HEK293T TLR cells (I-AniI target). Trex2 expression significantly increased mutNHEJ with both enzyme variants, and impressively so with the relatively inactive I-AniI (**Fig. 5a**), suggesting that once a double-strand break is

made, even if rarely, Trex2 can capture the DNA ends and process them. Although the affinity of native I-AniI for DNA is 100nM and Trex2 is 97nM (80), structural data for Trex2 suggests that it has an ideal binding pocket for HE products: a 4bp long 3' DNA overhang (80). Additionally, despite Y2 expression resulting in a high level of baseline mutNHEJ, the addition of Trex2 increased mutNHEJ by 2.5-fold (**Fig. 5b**), likely due to Trex2's ability to promote escape from the cycle of the 'persistent break'.

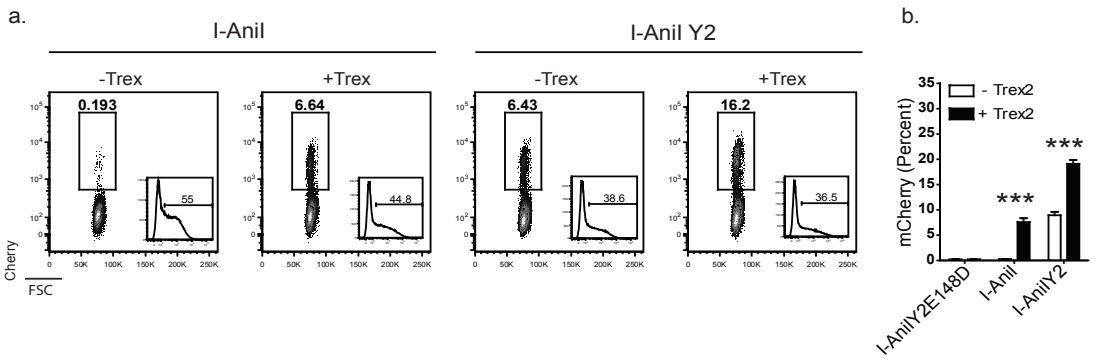


Figure 5. Trex2 increases mutNHEJ with the I-AniI HE. (**a**) Representative flow plots of mCherry and BFP of HEK293T cells transfected with the indicated vectors, quantified in (**b**).

2.3 Trex2 co-expression causes near-complete disruption even at low HE expression levels

Using the TLR to assess disruption is a quick and efficient method, but because mCherry expression results from a +3 frameshift event, the TLR may be under-reporting up to 60% of disruption events (74). Thus, to assess the total gene disruption rate, in addition to assessing it at different levels of HE expression, we sorted I-SceI and I-SceI-T2A-Trex2 transfected HEK293T TLR cells based on three levels of BFP expression. Genomic DNA was obtained and using nested PCR, the area flanking the TLR I-SceI target site was amplified from each of the populations. We then digested the PCR

product with recombinant I-SceI enzyme to look for a resistant band, indicative of a mutagenic event at the genomic TLR locus that destroyed the I-SceI target site.

At low endonuclease expression levels, we observed a 25-fold increase in total gene disruption between I-SceI and I-SceI coupled to Trex2 (2.2 to 50.2% respectively), and nearly 100% of targets were disrupted in the medium and high expression gates of I-SceI T2A Trex2 (90.3, and 97.1% respectively) (**Fig 6**). These results indicate that while I-SceI exhibits a dose dependent increase in gene disruption, I-SceI coupled to Trex2 quickly becomes saturated. Further, while we see between 27%-40% mCherry signal with Trex2 by the TLR, a more unbiased approach by genomic analysis indicates that indeed this is an under-estimation of the total levels of disruption. Importantly, there is little difference in total disruption between the mid and high levels of HE expression with Trex2. Off-target cleavage with designer endonucleases is a significant concern, so if Trex2 expression can effectively lower the dose at which an HE can be delivered without adversely affecting mutNHEJ, it would be a very beneficial technology.

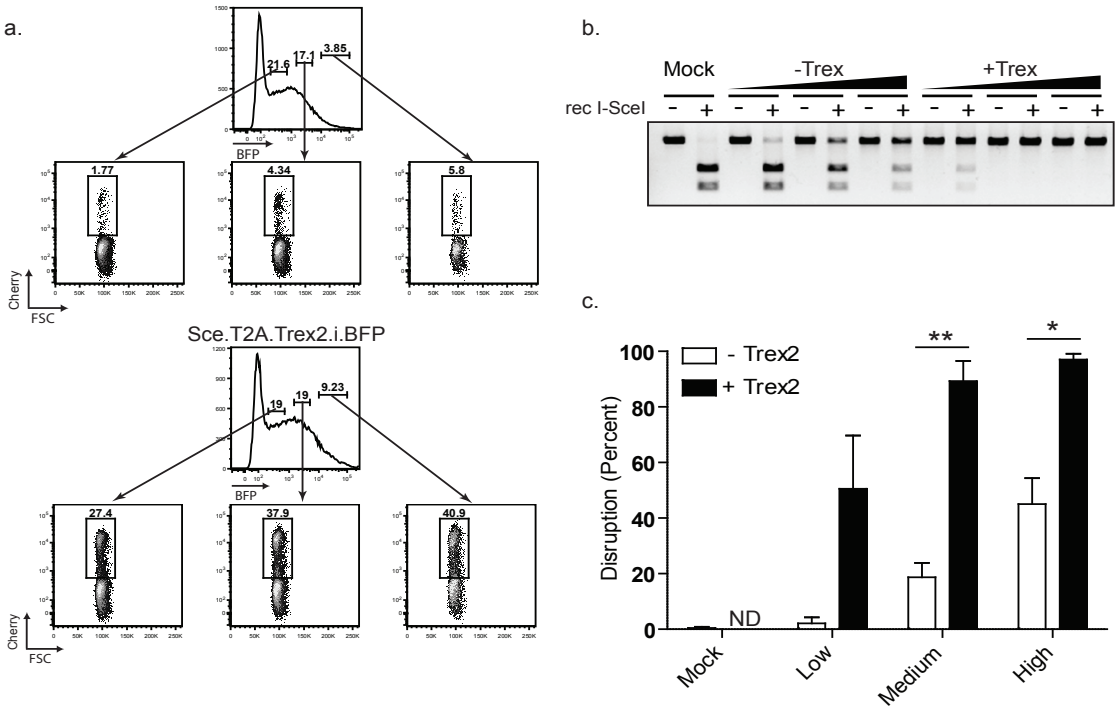


Figure 6. Trex2 expression increases mutNHEJ at levels of I-SceI expression (**a**) Representative flow plots of sorting gates for I-SceI with or without Trex2 (top and bottom, respectively (**b**) Sce-digestion NHEJ assay to measure total gene disruption. Cells were sorted based on the gates indicated in **a**. rec, recombinant. (**c**) Quantification of band intensity (undigested out of total) of three independent experiments of the Sce digestion assay as performed in **b**. ND, not determined.

2.4 Trex2 promotes a shift in mutagenic spectra towards small deletions

Further sequencing analysis of Trex2's activity on HE-mediated 3' DNA overhangs suggests that it excises nucleotides to preferentially leave small deletions: a comparison of the mutation spectra between I-SceI alone and I-SceI.T2A.Trex2 showed a trend towards small deletion events in the exonuclease treated cells (**Fig. 7**). While mutNHEJ events from I-SceI alone can lead to insertions, large deletions, or small deletions, Trex2 predominantly leads to 2-4bp deletions. As noted above, Trex2 structural analysis has suggested that Trex2's catalytic site forms a binding pocket to process up to a 4bp 3' overhang (81), which dovetails with these results.

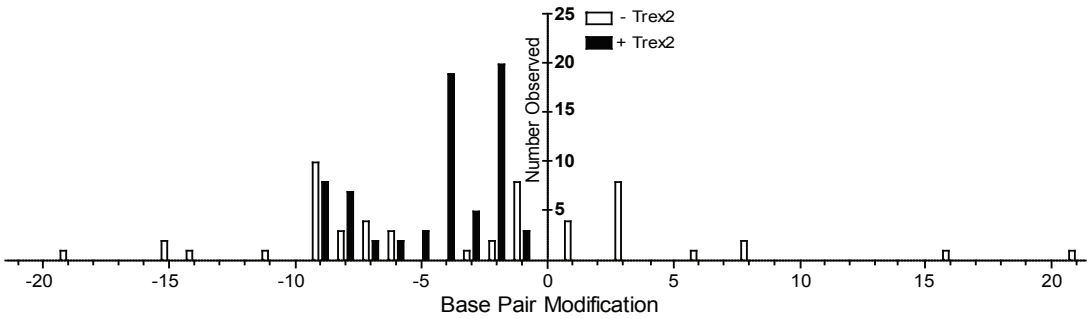


Figure 7. Trex2 promotes increase toward small deletions when co-expressed with I-SceI. Sequence analysis of HEK293T cells transfected with I-SceI alone or I-SceI with Trex2, PCR amplified, and TLR region analyzed for in-dels.

An important implication from this sequencing analysis, but one which has not been directly tested, is that the action of Trex2 may inhibit deleterious translocations. If Trex2 captures DNA ends immediately following an HE-induced DSB, processes a small amount of sequence, presumably the two broken DNA ends would be held in close enough proximity to be re-ligated back to one another before diffusing farther away. Further, the Trex2 dimer structure may even allow for simulateous processing of both DNA ends. When there is significantly more DNA end-processing, it may allow for greater physical separation of the broken DNA strands, and with DNA damage occurring elsewhere in the genome, or if there is off-target cleavage occurring by the HE, a dangerous translocation may be more likely.

2.5 Trex2 promotes the earlier onset of mutNHEJ

Given our results that Trex2 can promote mutNHEJ even at the lowest I-SceI expression levels, and we suggest that it enables the escape from the persistent break cycle, we were curious whether Trex2 expression also promotes an earlier onset of

mutNHEJ events. This would be advantageous in that the eventual goal is to minimize the exposure time of cells to designer endonucleases, in order to minimize the potential for off-target cleavage. To address this, we performed a kinetic analysis where cells were transfected with either Sce.ires.BFP or Sce.T2A.Trex2.ires.BFP, or the corresponding Sce^{D44A} mutants, and tracked BFP expression and mutNHEJ events (mCherry expression) over seven days (**Fig 8**). While all constructs exhibited similar expression patterns (**Fig. 8a**), we found that Trex2 expression coincided with the appearance of disruption events at an earlier time-point (24 hours), when I-SceI expression alone caused negligible mutNHEJ (**Fig. 8b**). Importantly, at 48hrs when BFP expression was maximal, mutNHEJ with Trex2 was 78% of max, while I-SceI alone was only 41% of max. This data further confirms the utility of Trex2 as a tool for use in targeted gene disruption with HE's: it significantly promotes mutNHEJ at low HE expression levels, as well as at lower and earlier levels of HE accumulation.

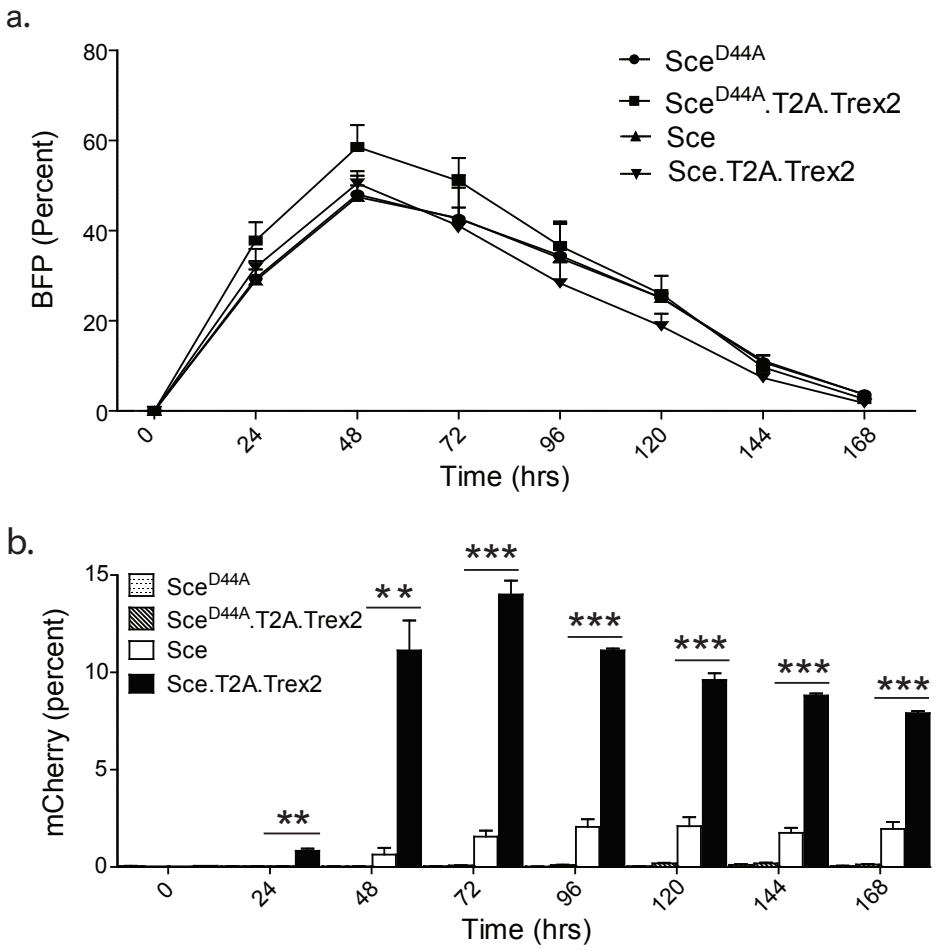


Figure 8. Trex2 promotes the earlier onset of mutNHEJ (**a**). Analysis of BFP expression at the time-points indicated following transfection with the indicated vectors. (**b**) Quantification of 3 experiments of HEK293T cells transfected with the indicated vectors, analyzed at the indicated time-points.

2.6 Trex2 is effective at 3' recessed ends left by a TALEN and ZFN

We next evaluated whether Trex2-enhanced disruption is applicable to different nuclease platforms. Namely, whether Trex2 can process 3' ends when they are recessed, which occurs at the 4bp 5' overhangs left by the Fok1-based TALENs and ZFNs. We made HEK293T TLR stable cell lines with targets for the CCR5 TALEN (82) and the VegF 2468 ZFN (83) enzymes, and transfected cells with either the CCR5 TALEN or VegF ZFN with or without Trex2. In both cases, we found that Trex2

significantly increased mutNHEJ above either enzyme alone (**Fig 9**). Importantly, this indicates that the 3' recessed ends left by the TALEN- and ZFN-based Fok1 endonuclease are accessible to Trex2 activity, despite not being either single-stranded or overhanging. There is previous *in vitro* data that Trex2 can indeed resect recessed ends (77), or access 3' ends following 5' resection; and our results confirm that the enzyme acts in a similar fashion on genomic DNA *in vivo*. Critically, it also implies that Trex2 can be used as a general tool for all three nuclease platforms to increase targeted gene disruption.

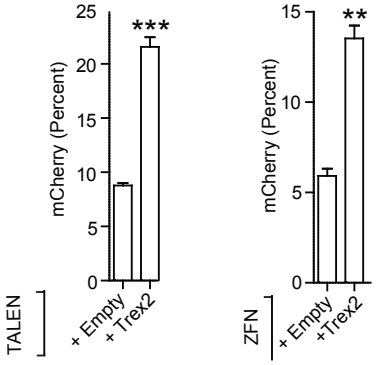


Figure 9. Trex2 increases mutNHEJ with a TALEN and ZFN. HEK293T TLR cells transfected with the CCR5 TALEN or VegF 2468 with empty vector (pEx) or Trex2 and assayed for mCherry.

2.7 Genomic analysis of Trex2 with VegF ZFN and CCR5 TALEN

As above, to obtain a less biased estimate of mutNHEJ caused by Trex2, we performed genomic analysis on both the TLR as well as the endogenous loci of cells treated with either CCR5 TALEN or VegF ZFN with or without Trex2. We sorted cells transfected with TALEN or ZFN, with or without Trex2, based on overall BFP expression. Genomic DNA was obtained and using nested PCR, either the area flanking the TLR target sites or the endogenous CCR5 and VegF genes was amplified from each of the populations (**Fig. 10**). Unfortunately, data with the CCR5 TALEN +/- Trex2 at its

endogenous locus could not be obtained, as it seems that the HEK293T cell line naturally has the CCR5Δ32 mutation, towards which this enzyme is targeted. However, at the TLR locus, Trex2 caused an increase in mutNHEJ to 48%, a 2-fold increase from TLR mCherry levels (**Fig. 10a, b**). Sequence analysis indicates that mutation spectra are biased towards mid-sized deletions (8-9bp), while TALEN-treated cells alone exhibit a wide range of deletions, which is commonly seen with different TALENs (37).

In contrast, at both the VegF TLR and endogenous locus, Trex2 treatment resulted in a sequence bias towards small deletions (2-4bp), similar to that seen with I-SceI, but the ZFN itself tended to result in insertions rather than deletions (**Fig 10 c-f**). Overall baseline mutation rates with the ZFN itself were higher at the endogenous locus (71%), where the addition of Trex2 led to almost complete disruption, whereas at the TLR locus the VegF ZFN alone caused 37% disruption and Trex2 addition increased this almost two-fold.

This difference in mutation specra is interesting, and agrees with a report showing the distinctions between ZFN and TALEN mutation signatures (37). We suggest that the difference is based on the size of the spacer sequence – which creates the Fok1 target site – the TALEN's is 18bp, while the ZFN's is 7bp. As Fok1 cleavage is independent of DNA sequence, we speculate that Trex2 may be able to access DNA when TALE/zinc fingers are still bound (perhaps due to malleable nature of TAL binding at 5' end of target sequences), but the ZFN's limited spacer and perhaps higher 5' end affinity restricts the extent to which Trex2 can interact with or resect the resulting end.

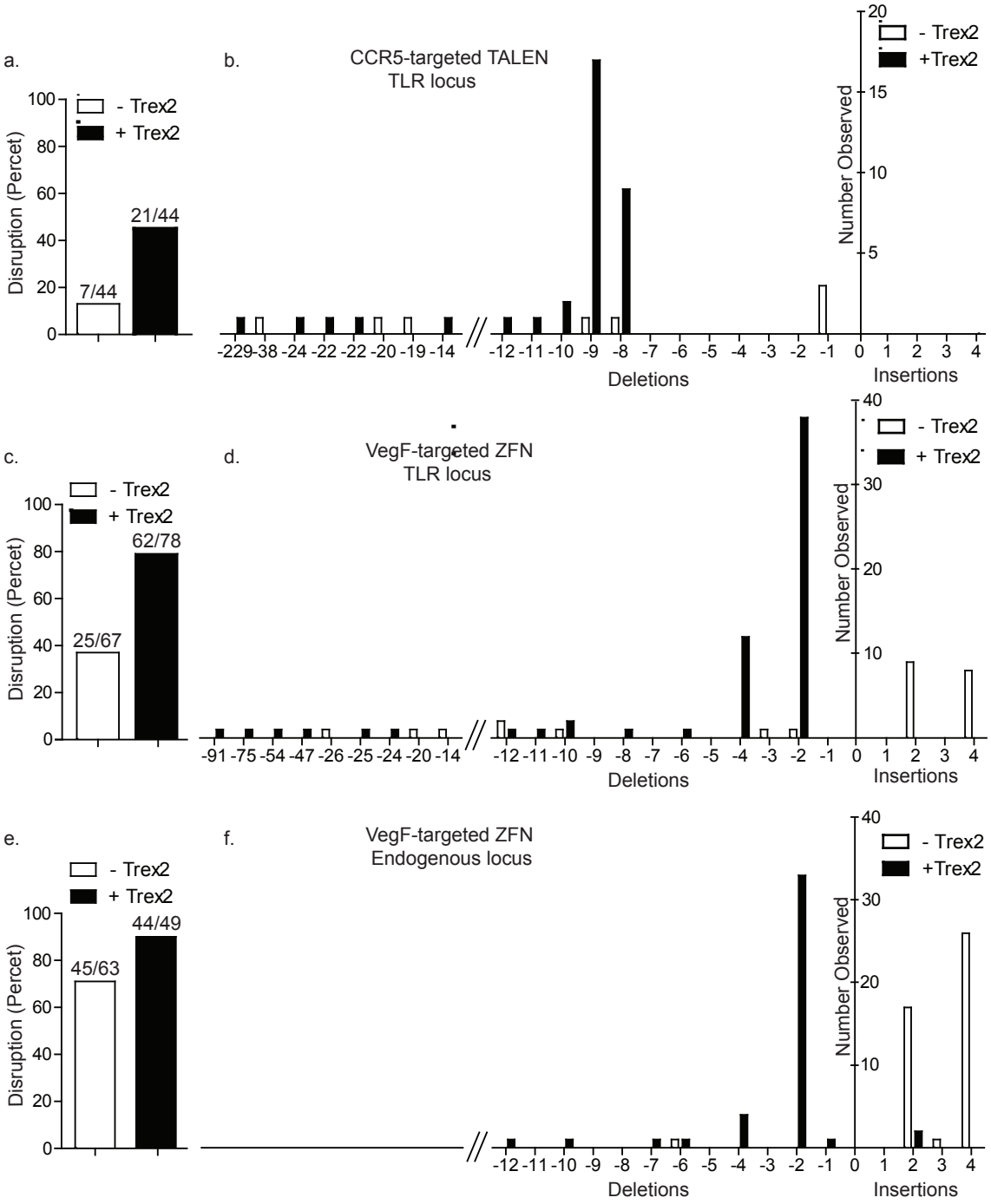


Figure 10. Modification rates and mutation spectra of TALEN or ZFN with and without Trex2 by sequencing analysis. (**a, c, e**). Quantification of overall disruption rates by CCR5 TALEN at TLR locus (a), VegF ZFN at TLR locus (c) or endogenous locus (e) with or without Trex2. Numbers above bars represent mutated sequences over total sequences analyzed. (**b, d, f**). Base pair modifications induced by CCR5 TALEN (b), or VegF ZFN (d, f) with or without Trex2.

2.8 An endo/exonuclease screen to find other DNA end-modifying enzymes

We were curious if we could extend the results of coupling designer nucleases with Trex2 to other DNA end-processing enzymes, especially those with 5' exonuclease activity. We therefore created a candidate library of 13 enzymes, including 5-3' exonucleases, possessing an array of biochemical end-processing activities derived from mammalian, bacterial or viral origin (**Table 1**). We then screened this library by co-expressing each enzyme with either I-SceI, the CCR5 TALEN, or the VegF 2468 ZFN in the respective HEK293T TLR cells (**Fig. 11**). In a manner similar to Trex2, five of these enzymes (Artemis, Tdt, Apollo, Rad2, and Exo1) robustly increased the gene disruption efficiency of I-SceI (**Fig. 11a**). Additionally, we analyzed the gene disruption activity of these five enzymes at three levels of I-SceI expression (quantified by the mean fluorescence intensity, MFI, of the BFP fluorophore). We found that expression of most of these enzymes with I-SceI increased its efficiency at low levels of the HE, similar to effect of Trex2 (**Fig. 11 d, e**). In contrast, we did not observe a significant effect of any enzyme on increasing the gene disruption efficiency of the VegF ZFN, while only the endonuclease Artemis increased mutNHEJ events with the CCR5 TALEN (**Fig. 11b, c**). This result was surprising, given that several of these proteins possess processive 5' exonuclease activity. We speculate that DSB's with Fok1 5' overhangs are more amenable to gap filling, as they have a 'primable' recessed 3' OH, and are thus less amenable to exonuclease processing.

Table 1: Library of DNA end-processing enzymes.

Enzyme	Gene name	Activity	Species of origin	NLS added
Apollo	SNM1B	5-3' exonuclease	Human	No
Artemis	Artemis	5-3' exonuclease	Human	No
Dna2	DNA2	5-3' exonuclease, helicase	Human	No
Exo1	EXO1	5-3' exonuclease	Human	No
Fen1	FEN1	5' flap endonuclease	Human	No
Mre11	MRE11	5-3' and 3-5' exonuclease	Human	No
Rad2	catalytic domain of Exo1	5-3' exonuclease (Exo1 catalytic domain)	Human	No
TdT	TdT	Single-stranded template-independent DNA	Human	No
RecE	RecE	5-3' exonuclease	*E. coli*	Yes
Lambda exonuclea	exonuclease	5-3' exonuclease	Bacteriophage l	Yes
Sox (T24I mutation)	SOX	5-3' alkaline exonuclease	Kaposi's sarcoma-associated herpesvirus	Yes
Vaccinia DNA polymeras	E9L	3-5' exonuclease	Vaccinia poxvirus	Yes
UL-12	UL12	5-3' alkaline exonuclease	Herpes simplex virus (HSV)-1	Yes

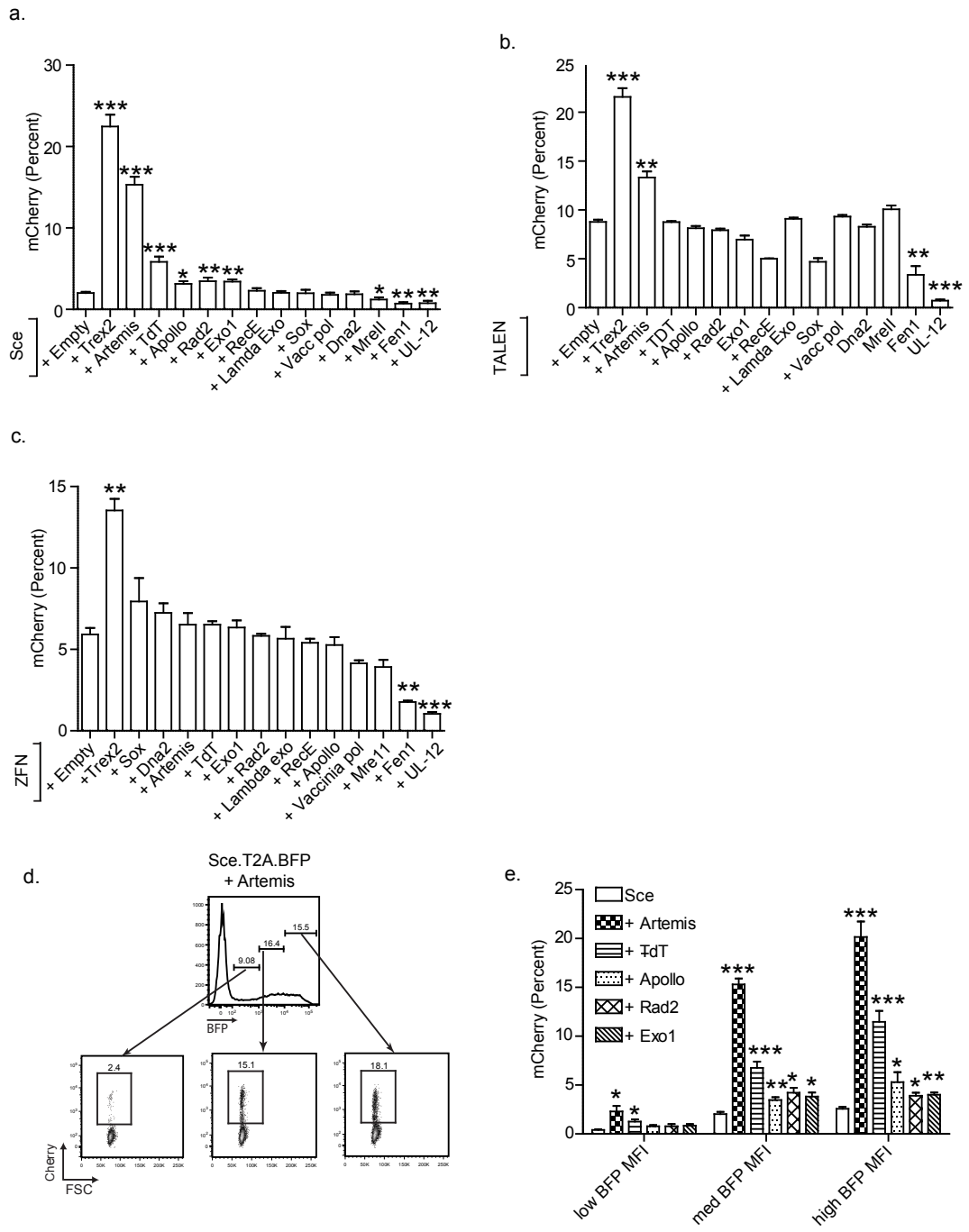


Figure 11. DNA end-processing enzymes library screen. (**a**–**c**) Quantification of mCherry expression in BFP-positive HEK293T cells harboring the Sce-TLR, the *CCR5* TALEN-TLR or the VegF ZFN TLR. (**d**) Representative flow plots indicating different expression levels of I-SceI co-expressed with Artemis, leading to increasing mutNHEJ. (**e**) Quantification of mutNHEJ at different expression levels I-SceI (quantified by BFP MFI) with the top five enzymes in the library screen.

2.9 Genomic analysis of Artemis co-expression with HE and TALEN

To look more closely at the activity of Artemis at the genomic level, we assessed the total gene disruption rate rather than just that reported by the TLR with I-SceI and the TALEN. We sorted HEK293T TLR cells transfected with I-SceI or CCR5 TALEN with or without Artemis based nuclease BFP expression. Genomic DNA was obtained and using nested PCR, the area flanking the TLR target sites were amplified from each of the populations. We then digested the PCR product with recombinant I-SceI enzyme to look for a resistant band, indicative of a mutagenic event at the genomic TLR locus that destroyed the I-SceI target site. Because the CCR5 TALEN TLR cell line was built with the spacer containing the 18bp I-SceI target site, this experiment could be done with both the HE as well as the TALEN-transfected cells. We found that expression of Artemis increases mutNHEJ frequency to a similar extent as Trex2 at TALEN-induced 5' overhang breaks (39% and 45%, respectively), though it is less effective than Trex2 at 3' HE-induced 3' overhang DSB's (32% and 79%) (**Fig. 12 a, b**).

Sequence analysis indicates that mutation spectra are biased towards mid-sized deletions (8-9bp) when co-expressed with the CCR5 TALEN (**Fig. 12d**), similar to Trex2. In contrast, when co-expressed with the I-SceI HE it leaves a broader mutagenic deletion spectrum (**Fig. 12c**). Additionally, Artemis co-expression with I-SceI causes larger deletions than the expression of Trex2, suggesting that Artemis, directly or indirectly, leads to more extensive and heterogeneous 3' end processing following an I-SceI induced break.

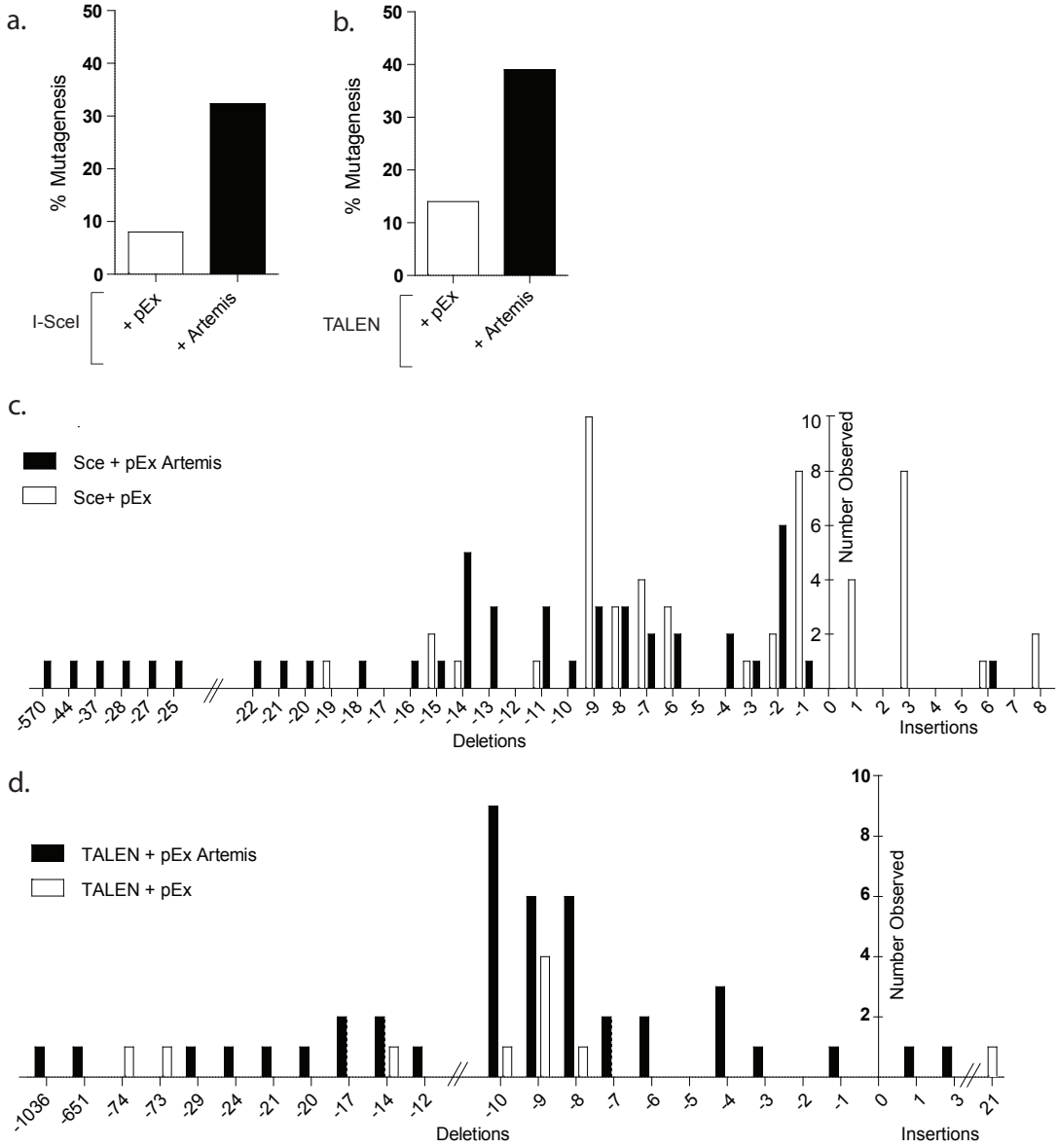


Figure 12. Genomic analysis of Artemis co-expression with I-SceI and CCR5 TALEN. (**a,b**) Quantification of total disruption rates of Artemis co-expression with I-SceI (**a**) or CCR5 TALEN (**b**) in HEK293T TLR cells. (**c, d**). Sequence analysis at the TLR locus in HEK293T cells.

2.10 Trex2 and Artemis significantly increase mutNHEJ at silenced loci

Artemis has been implicated in repair of DSBs at heterochromatic loci (84, 85), including resection of tightly packed nucleosomes (85). Given the importance of

designer nucleases in accessing a multitude of genomic loci, including silenced and inactive chromatin, we were curious whether expression of Artemis would have a similar effect in promoting access to silenced loci. Additionally, we hypothesized that expressing Trex2 would also increase disruption rate at silenced loci due to its ability to access most breaks; even a rare DSB event could be captured and mutagenized.

We recently published a TLR reporter that measures both silenced and active loci, via the expression of iRFP (26), and we used this system to assess whether co-expression of Artemis or Trex2 with I-SceI affects mutNHEJ. We co-expressed I-SceI with Artemis or Trex2 in silenced (iRFP-) or active (iRRP+) cells, sorted for I-SceI expression (BFP+ cells), isolated genomic DNA and PCR-amplified across the locus, and assessed NHEJ frequency by recombinant I-SceI digest assay. We found that both enzymes significantly increased the efficiency of I-SceI to increase mutagenic NHEJ, from 3% with I-SceI alone, to 29 % with Artemis and 74% with Trex2 (**Fig. 13 a-c**). These results suggest that I-SceI is able to access and cleave silenced loci and that using Trex2 or Artemis may be advantageous for the success of gene editing at silenced loci.

Earlier work has suggested that Artemis promotes large deletions at silenced loci, perhaps due to resection of tightly packed nucleosomes (85). We reasoned that expression of Artemis at our silenced locus would promote a similar effect, given the extent of silencing around the promoter surrounding the TLR reporter (26). We analyzed the sequences from the PCR amplicons obtained from genomic DNA from the silenced cells, analyzing 82 sequences and plotting the mutation spectra comparing cells expressing I-SceI with and without Artemis (**Fig. 13d**). We did not find any

amplicons with deletions greater 16bp, similar to what we observed at relaxed, open loci (see **Fig. 12c**). With resected nucleosomes, we would expect to find much larger, discrete deletions (nucleosomes are ~ 150bp), suggesting that in this model, within this reporter locus at least, Artemis does not promote nucleosome resection.

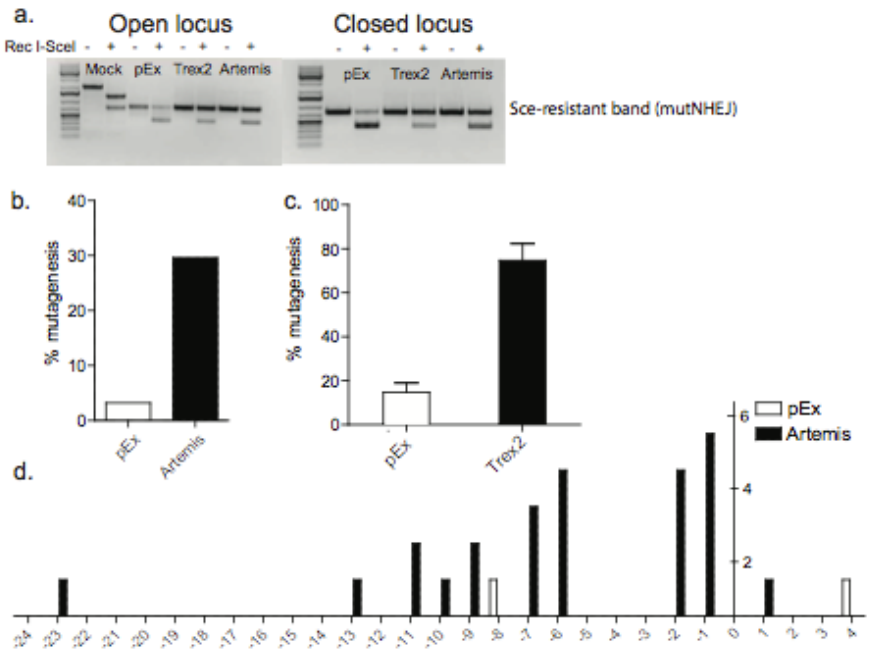


Figure 13. Trex2 and Artemis increase mutNHEJ at silenced loci (AR-TLR) (**a**) Sce digestion NHEJ assay to measure total disruption at both open and closed loci with the indicated vectors, quantifed in (**b, c**) for the closed locus. (**d**) AR-TLR sequence analysis of deletions induced by co-expression of I-SceI with Artemis at the closed AR-TLR locus to determine whether Artemis affects nucleosome resection.

2.11 Summary

Further work has since been done with primary cells to show both that Trex2 expression enhances mutNHEJ, and that its does not sensitize cells to DNA damage. Primary human CD34+ HSCs expressing a CCR5-targeted HE saw in increase in mutNHEJ from 5% to 37% with Trex2 addition (44). Treatment with DNA damaging agents in primary MEF cells and human CD34+ HSCs in the presence of Trex2 yielded no difference in the rate of cell death over cells that did not express Trex2 (44). Thus

Trex2 expression is both effective, and does not increase DNA damage sensitivity, in primary human cells of genome engineering interest.

Overall, we have demonstrated that coupling designer endonucleases with DNA end-processing enzymes is a robust way to improve the rates of targeted disruption in a variety of cell types and species, without associated toxicity to the host. This is an important advance for several reasons: 1) DSBs trigger cell cycle checkpoints to arrest division until the break has been resolved; in the case of a 'persistent break', cells may arrest indefinitely, leading to apoptosis. The use of end-processor coupling is an effective approach to interrupt such a 'persistent break' cycle. 2) While expression of an endonuclease over time should in theory lead to a fixation of mutagenic events, engineering applications often necessitate transient delivery of the endonuclease, providing only a short window in which enzyme concentration is sufficient to achieve breaks. Dividing cells quickly dilute transient expression vectors, and thus the time available for the intended genetic changes to occur is reduced. 3) Persistent breaks can be a source of translocations. Because coupling endonucleases to exonucleases reduces break persistence, and so the chance of gross chromosomal rearrangements, it may improve the safety of endonuclease-induced targeted disruption for therapeutic applications. 4) With the demand for more complex changes to genomes comes the desire to make multiple changes in a single round of mutagenesis, for example multi-allelic knockouts and multiplexing. Our data suggest that "exo-plexing" could provide an efficient means to this end. Lastly, the demonstration that endo/end-processor coupling can be used to drive gene disruption via mutagenic NHEJ suggests that similar manipulations aimed at enhancing homologous recombination may also be possible.

Additionally, the use of end-processing enzymes with different activities provides another level of control for a desired engineering outcome, and also suggests the possibility of synergy between different enzymes to achieve maximal or unique types of knockout effects.

In conclusion, we show that DNA double strand breaks generated in living cells can be efficiently modified by a diverse array of DNA end-processing enzymes. While the current endonuclease-induced genome engineering paradigm relies on the native resolution of DNA breaks, the development of molecular strategies to bias repair towards a desired outcome is an important next step towards enabling rapid, precise and efficient site-specific genome modification required for routine large scale genome engineering.

2.10 Materials and Methods

Construct Assembly

The library of DNA processing enzymes was cloned into the ‘pExodus’ vector with genes synthesized by Genscript (Piscataway, NJ) as cDNA codon-optimized for human expression. All other constructs were cloned into the “RRL”(addgene #12252) or CVL lentiviral backbones using standard molecular biology techniques.

Cell Culture

HEK293T cells were cultured in glutamine-free Dulbecco’s modified Eagle’s medium supplemented with 2mM L-glutamine, 10% Fetal Bovine Serum (FBS) and 5% penicillin/streptomycin.

Transfection

1.0x 10^5 HEK293T cells were plated 24 hours prior to transfection in 24-well plates. 0.5μg DNA was used for each expression vector, and transfected using Fugene6 or XtremeGene9 (Roche) according to the manufacture's protocol. Cells were analyzed on a flow cytometer 72 hours post-transfection, unless otherwise indicated

Traffic Light Reporter (TLR) Assay

TLR cell lines for HE, TALEN, and ZFN targets were generated as previously described (74). Briefly, the TLR assay consists of a GFP open reading frame with an embedded designer endonuclease target site followed by a T2A.mCherry in the +3 frame. No fluorescent protein is made under basal conditions. If a nuclease is delivered, and a mutagenic NHEJ event occurs repairing the TLR in the +3 frame, mCherry is expressed. If a nuclease is delivered with donor (C-terminally truncated GFP), and the TLR is repaired using HR, GFP is expressed. The Epigenetic TLR is as described (26). Briefly, this TLR is the same as the normal TLR, except that rather than being selected using puromycin, this construct has an iRFP fluorophore immediately downstream of the SFFV promoter, allowing for the visualization and sorting of cells that have been silenced (iRFP-) or are active (iRFP+).

Lentivirus Generation

Recombinant lentivirus (LV) was produced by transient co-transfection of HEK293T cells as previously described (74). Briefly, cells were transfected in 10cm dishes using PEI (Polysciences, Inc.) with 6 μg vector, 1.5 μg pMD2G envelope plasmid (VSV-G), and 3 μg psPAX2 per plate. Virus was titered by Lenti-x p24 rapid ELISA kit (Clontech). MOI was calculated for fluorescently tagged virus experiments by dividing the amount of infectious units added by the number of cells plated. Infectious units were calculated

by transducing 2.0 x10^5 HEK293T cells with increasing amounts of viral stocks and analyzing them on a flow cytometer 72 hours post transduction and applying the following formula for volumes of virus that yielded between 5-20% fluorescent positive cells (($2.0x10^5$ x percent fluorescent positive cells)/100)/ volume viral stock added).

Flow Cytometry

Cells were analyzed on a BD LSRII or BD FACS ARIAII. The mCherry fluorophore was excited using a 561nm laser and acquired with a 610/20 filter. The mTagBFP fluorophore was excited on a 405nm laser with a 450/50 filter. Data was analyzed using FlowJo software (FlowJo, Ashland OR).

Sce Digestion NHEJ Assay

Genomic DNA was isolated using Qiagen's DNA easy kit, and TLR target sites were PCR-amplified as previously described (74). 100ng of each PCR product was digested *in vitro* with recombinant I-SceI (New England Biolabs) for 6 hours at 37^0C. DNA was separated using a 1% agarose gel stained with ethidium bromide. Percent disruption was calculated by quantifying band intensity using Image J software, and dividing the intensity of the undigested band by the total.

NHEJ Sequencing

Genomic DNA was isolated from cells using Qiagen DNAeasy kit (Qiagen) and the HE target region was amplified as previously described (74). PCR products were cloned using a CloneJET PCR cloning kit (Fermentas) according to the manufactures protocol, followed by transformation into chemically competent DH5α *E.coli* bacteria. Bacterial colonies were directly sequenced using a standard colony sequencing protocol. Sequences were analyzed using the "Contig Express" software provided in Invitrogen's

Vector NTI software suite (Invitrogen).

TALEN engineering

A Xanthomonas pthXo1 golden gate destination ORF and an RVD plasmid library were a kind gift from the Voytas lab. A CCR5 specific TALEN pair were assembled with a Golden Gate cloning strategy as described (30) into a pthXO1 scaffold truncated at positions N154 and C63 as previously described. (82).

Statistical Analysis

Error bars represent standard error of the mean (SEM), and *p*-values (with * representing $p<0.05$, ** $p<0.005$, and *** $p<0.0005$) were calculated using the Student's two-tailed unpaired t-test to compare the samples indicated in this and all subsequent figures.

Chapter 3

Probing the mechanistic interactions of Trex2 and Artemis with NHEJ factors following designer endonuclease cleavage

3.1 Introduction

In an effort to increase the efficiency of targeted disruption, we show that coupling the exonuclease Trex2, or the endonuclease Artemis, to both the HEs and Fok1-based TALENs or ZFNs significantly increases the frequency of mutagenic NHEJ repair (mutNHEJ) ((44), and **Ch. 2**). However, before these tools can be used clinically, they must be carefully characterized. Indeed, the Trex2 gene was discovered relatively recently and its endogenous cellular function remains unknown. The role of Artemis as a cNHEJ enzyme is understood, but the consequences of coupling these DNA end-modifiers with designer endonucleases must be thoroughly assessed. Taken together, our characterization of these new technologies may provide researchers further insight into which designer nuclease and DNA processing enzyme combination can be best for a desired mutagenic outcome, and may help explain the ability of these nucleases to affect DSB repair pathway choice.

3.2 Trex2 and Artemis divert repair events from HR to NHEJ

Both Trex2 and Artemis significantly increase mutNHEJ, so we wanted to determine whether this increase comes at the expense of diverting events from other pathways. Using HEK293T TLR cells, we transfected I-SceI or VegF ZFN and d14 GFP

donor with or without Trex2 or Artemis, and found a slight but significant decrease in HR with both nucleases (**Fig. 14 a, b**). Because NHEJ is active during all stages of the cell cycle, these results suggest that the activity of Trex2 and Artemis divert DSB repair down the NHEJ pathway rather than HR, despite the presence of exogenous donor. The ability to control, or at least better understand, break repair pathway choice is an attractive one when undertaking a genome engineering project, so given this result we wanted to characterize the mechanistic facets of Trex2 and Artemis that could shed light on how they create this pathway diversion.

Figure 14 (Chapter 3)

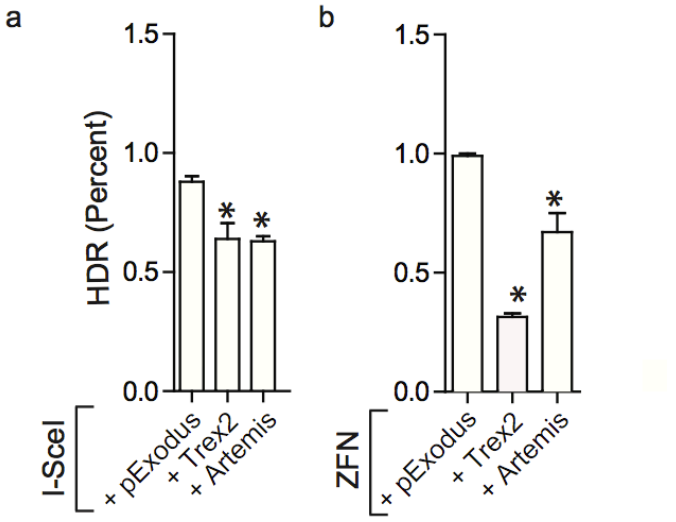


Figure 14. Trex2 and Artemis inhibit HR. Quantitation of homology-directed repair (HDR) of HEK293T I-SceI (**a**) or CCR5 (**b**) TLR cells expressing I-SceI or CCR5 TALEN and either pEx (empty vector), Trex2, or Artemis. Measurement of HDR is based on GFP from the TLR assay, 72 hrs post-transfection.

3.3 Trex2 directly catalyzes DNA resection to increase mutNHEJ

Data from *in vitro* assays has shown Trex2 can resect a variety of DNA ends, including overhanging and recessed 3' ends (77). We were interested in directly testing whether Trex2 acts directly on genomic DNA following a endonuclease-induced DSB *in*

vivo. We tested a catalytic mutant of Trex2, D193N (79), in the TLR assay: Trex2^{D193N} expression with I-SceI or CCR5 TALEN completely abolished the effect of promoting mutagenic NHEJ (**Fig. 15 a, c**). The expression level of catalytically dead Trex2 was equal to wild type protein, measured by western-blot (**Fig 15 b,d**). To further confirm the involvement of Trex2 catalysis, we tested further catalytic mutants, D14A E16A, which coordinates calcium ion binding (86). Transfection of HEK293T TLR cells with I-SceI and Trex2^{wt} or Trex2^{D14AE16A} gave a similar result to the D193N mutant: mutNHEJ with the mutant Trex2 is no different than baseline (data not shown).

Together, these results suggest that Trex2 directly catalyzes the resection of DNA after the DSB created by designer nuclease, and can access the DNA ends left by the Fok1 nuclease *in vivo* (either recessed 3' ends, or 3' overhangs created by extensive 5' resection).

Figure 15. Trex2 directly catalyzes DNA. (**a, c**) TLR assay of HEK293T cells transfected with I-SceI (**a**) or CCR5 TALEN (**c**) and either pEx, Trex2^{wt} or catalytically inactive Trex2^{D193N}. (**b, d**) Western blot showing equal expression of both forms of Trex2. (**e-h**) TLR assay of HEK293T cells transfected with I-SceI (**e, f**) or CCR5 TALEN (**g, h**) and the indicated Trex2 DNA binding mutants and dimerization mutant. *ns*, not significant.

3.4 Differential requirements for Trex2 activity on DNA *in vivo*

To further characterize the mechanistic action of coupling Trex2 with designer nuclease, we tested whether dimerization and DNA binding were also required for its ability to promote mutNHEJ. As described by Chen et al, (79) the point mutant K59A (K77 equivalent in mouse) abolishes dimerization of Trex2, while the triple mutant R163A/R165A/R167A (R/Q/R equivalent in mouse) destroys Trex2's ability to bind DNA. Surprisingly, we found that when co-expressed with I-SceI, these mutant enzymes were

not significantly less effective than wild type Trex2 at promoting mutNHEJ (**Fig. 15e**). However, when we added an additional binding mutant, R152A (86), Trex2 activity was abolished (**Fig. 15f**). Curiously, the other residue suggested to be involved in DNA binding in this structural analysis, D121 (86), had no effect when coupled to the triple mutant.

Surprisingly, Trex2 behaved differently at Fok1-induced 5' overhang DSBs. When we co-expressed Trex2 with the CCR5 TALEN in HEK293T TLR cells, we found that dimerization ($Trex^{K77A}$) and DNA binding (just the triple mutant) were required for full mutNHEJ activity, despite equal protein expression (Fig. 15 g, h, data not shown). When the fourth DNA binding mutant was added (R152A with R163A/Q165A/R167A), mCherry levels were the same as baseline.

These results suggest that at 5' breaks – where 3' ends are recessed – in the context of TALEN co-expression, Trex2 perhaps cannot access the 3' ends as well as at a 3' overhang break left by a homing endonuclease.

3.5 Mutating Trex2's β-hairpin does not impair its activity

We were curious whether we could discover a potential protein-interaction domain in Trex2, given that most NHEJ enzymes require protein scaffolding to be active at DNA DSBs. One previous report suggested Trex2 interacts with pol δ, presumably acting a proof-reading exonuclease (87), though there has been no follow-up evidence since. In comparing Trex2's structure, and the structure of its closest relatives, its protein fold most closely resembles that of the E. coli RecQ proteins, epsilon and exo1 (**Fig. 16a**) (81). Based on sequence alignment, Trex2 and the bacterial proteins share

most catalytic sites, but one feature stands out: Trex2 contains a β-harpin, which neither bacterial proteins have sequence for (**Fig. 16a**). Additionally, this β-harpin sequence aligns with Trex1's polyproline domain, which may be involved in Trex1's protein-protein interactions (86), and this hairpin is located on the outside of the enzyme, well-positioned to interact with other proteins (**Fig. 16b**). We thus hypothesized that the hairpin may be involved in Trex2 protein-protein interactions with NHEJ factors, so we created β-harpin mutants to test this. We disrupted the negatively charged residues in the middle of the hairpin, D42 D43, using alanine or lysine, and also mutated the whole portion of the hairpin that is exposed and potentially interacting to alanine (H-D-D-S-G) (**Fig. 16b**). After transfecting HEK293T TLR cells with I-SceI or CCR5 TALEN and the corresponding mutants, we found that none of the mutants had any effect on mutNHEJ that was different from $Trex2^{wt}$ (**Fig. 16 c, d**). This data indicates that the β-harpin has no effect on Trex2's ability to resect DNA following a DSB.

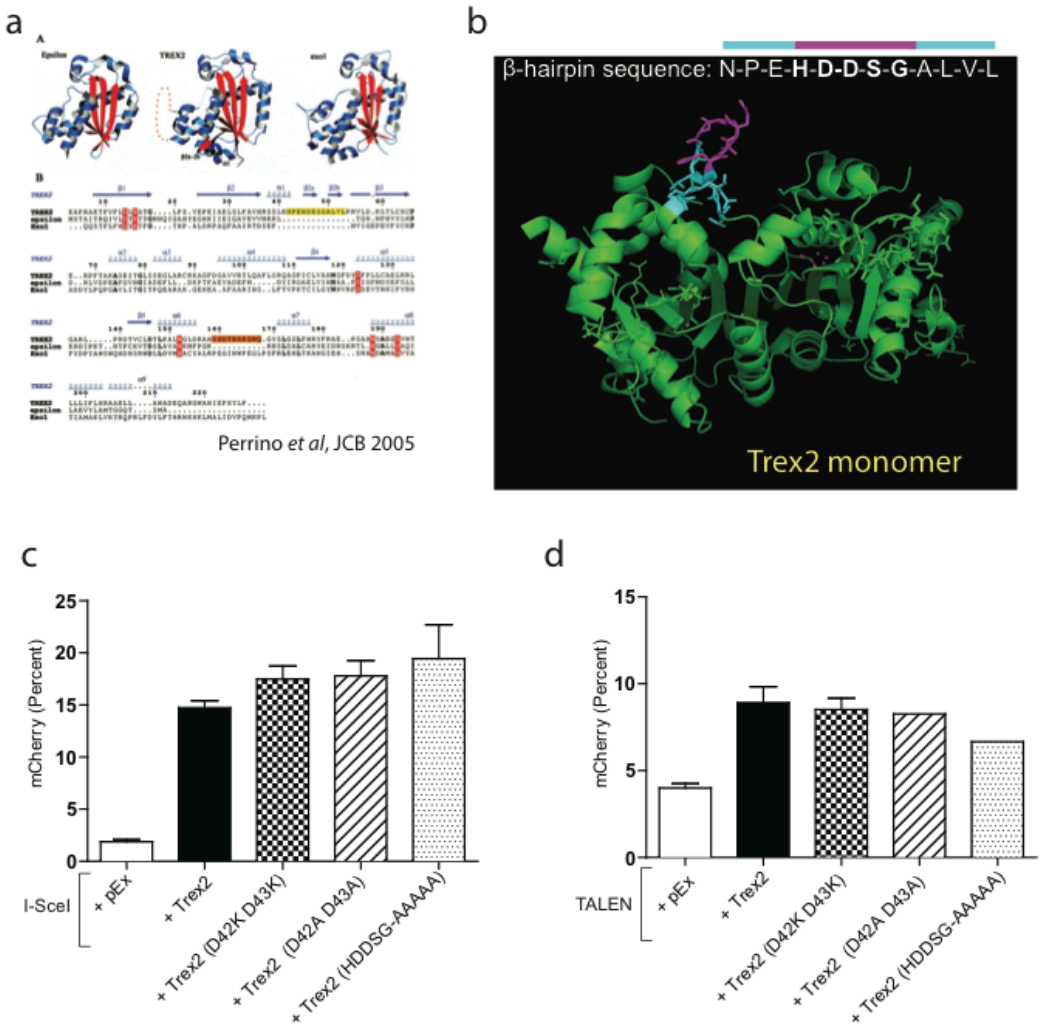


Figure 16. Trex2's β-harpin has no effect on its activity. (**a**) Alignment of Trex2 sequence with its closest bacterial family members, the RecQ enzymes, Exo1 and epsilon. The three proteins share homology, catalytic residues, and overall protein fold, but only Trex2 has a β-harpin (from Perrino et al, 2005). (**b**) A pymol representation of the β-harpin, with positively charged residues outlined in pink. (**c, d**) The TLR mCherry assay of HEK293T cells transfected with the indicated mutants of Trex2 with I-SceI (**c**) or CCR5 TALEN (**d**).

3.6 Artemis directly catalyzes DNA ends

Similar to our characterization of Trex2, we were interested in whether Artemis directly catalyzes resection of DNA ends following a designer endonuclease-induced DSB. We thus made a point mutation in the β-CASP domain of Artemis, H254A: a catalytic histidine which when mutated to alanine abolishes endonuclease activity (88).

When wild type or mutant Artemis were co-expressed with I-SceI or CCR5 TALEN in the TLR assay, we found that in both cases Artemis' endonuclease activity was required to promote mutagenic NHEJ (**Fig 17 a, b**). Thus, this data suggests that Artemis promotes mutNHEJ by directly processing DNA.

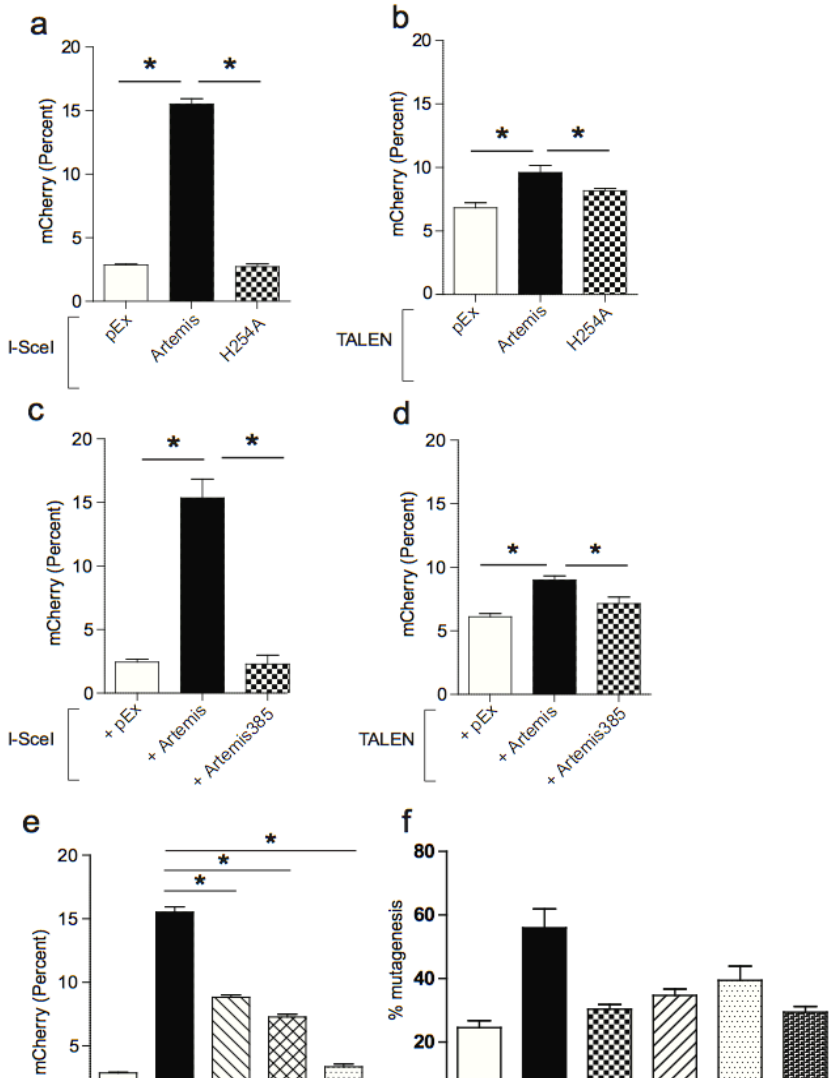


Figure 17. Artemis requires catalytic and signaling domains to promote mutNHEJ. (**a, b**). TLR assay of HEK293T cells transfected with I-SceI (**a**) or CCR5 TALEN (**b**) and Artemis^{wt} or catalytically inactive Artemis^{H254A}. (c-f) TLR assay of HEK293T cells

transfected with I-SceI (**c, e**) or CCR5 TALEN (**d, f**) and the indicated signaling domain mutants of Artemis.

3.7 Artemis requires interactions with cNHEJ proteins to promote mutNHEJ

We were next interested in whether we could separate the catalytic and signaling domains of Artemis for the purposes of genome engineering. Artemis consists of an N terminal catalytic half, comprised of metallo β-lactamase and β-CASP domains; and a C terminal signaling domain, through which phosphorlyation signals and interactions with DNA repair factors occur (89). Although models exist in which the signaling domain is required for full activation of Artemis, we were curious whether in our over-expression system, only the catalytic core could act to promote NHEJ. We found that at both 3' and 5' overhang breaks, the N-terminal catalytic core of Artemis (termed Artemis^{385}, for the first 385 amino acids), was not sufficient to cause an increase in mutagenic NHEJ (**Fig. 17 c, d**). This suggests that although Artemis directly processes DNA, it requires its signaling domain to do so.

Thus, we sought to determine whether in our over-expression system, Artemis interacts with its cNHEJ partners, and whether it requires these interactions to promote mutNHEJ. We created three Artemis mutants: W489A, which destroys its interaction with DNA Ligase IV (90); L401G/R402N, which disrupts its interaction with DNA-PKcs (91), and the triple mutant, which disrupts both (90). We found that when co-expressed with I-SceI in the TLR assay, the individual mutants inhibited mutNHEJ (**Fig 17e**). When we co-expressed with triple mutant with I-SceI, there was no increase in mutNHEJ signal over baseline, suggesting that Artemis requires interaction with both DNA-PKcs and LigIV to fully promote mutagenic NHEJ (**Fig. 17e**). Similarly, we found

that co-expressing the individual mutants with the CCR5 TALEN completely disrupted mutNHEJ (**Fig. 17f**).

3.8 Trex2 and Artemis prevent deleterious SSA by catalytic processing of DNA

Using the recently developed SSA-TLR, which concurrently measures mutNHEJ, homologous recombination, and single strand annealing (SSA) (see Figure 3, (26), we were curious whether our DNA end-modifying enzymes affected long-range resection and influenced SSA rates. This is a relevant concern, given that we recently discovered that following a HE-induced break, SSA is a major repair pathway, in both HEK293T SSA-TLR cells (26) and primary cells (72). However, we found that transfection of I-SceI with Trex2 or Artemis into SSA-TLR HEK293T cells significantly decreased the frequency of repair by SSA (iRFP+ cells), while concomitantly increasing mutNHEJ (**Fig 18a**). Of note, Trex2 expression impaired repair by SSA to a greater extent than Artemis, and was more effective at promoting mutNHEJ in this model cell line. These results support the translational/clinical use of co-expressing DNA processing enzymes with HEs, as they decrease the endogenous rates of deleterious SSA induced by the HE alone.

We hypothesized that Trex2 and Artemis were preventing resection due to direct catalysis of DNA ends, promoting I-SceI-induced breaks down NHEJ rather than SSA. We reasoned that since the initial DNA capture step is thought to be deterministic to repair pathway choice – whether NHEJ by Ku/DNA-PKcs, or resection and SSA/HR by MRN and CtIP, perhaps expression of Trex2 and Artemis places these enzymes at the critical point to access the DNA ends to catalyze the overhang and promote NHEJ.

Indeed, we find that when the catalytic residues of both Trex2 (D193N) and Artemis (H254A) are mutated, NHEJ is inhibited while SSA is enhanced relative to wild type enzyme (**Fig. 18 b, c**). Interestingly, the Trex2 catalytic mutant only partially inhibits SSA compared to control; this may be due to a predicted dominant negative effect of D193N, where it binds DNA but cannot catalyze resection. Surprisingly, the K77A mutant can fully promote mutNHEJ, but its ability to prevent SSA is similar to catalytically dead Trex2. Perhaps dimerization of Trex2 is required to efficiently access both DNA strands at the same time, otherwise SSA can occur; mutNHEJ is not affected, because the monomer is still catalytically active. In contrast, while the Artemis signaling mutants prevent NHEJ, they do not significantly inhibit SSA, suggesting that Artemis must directly catalyze the DNA ends to prevent long-range resection.

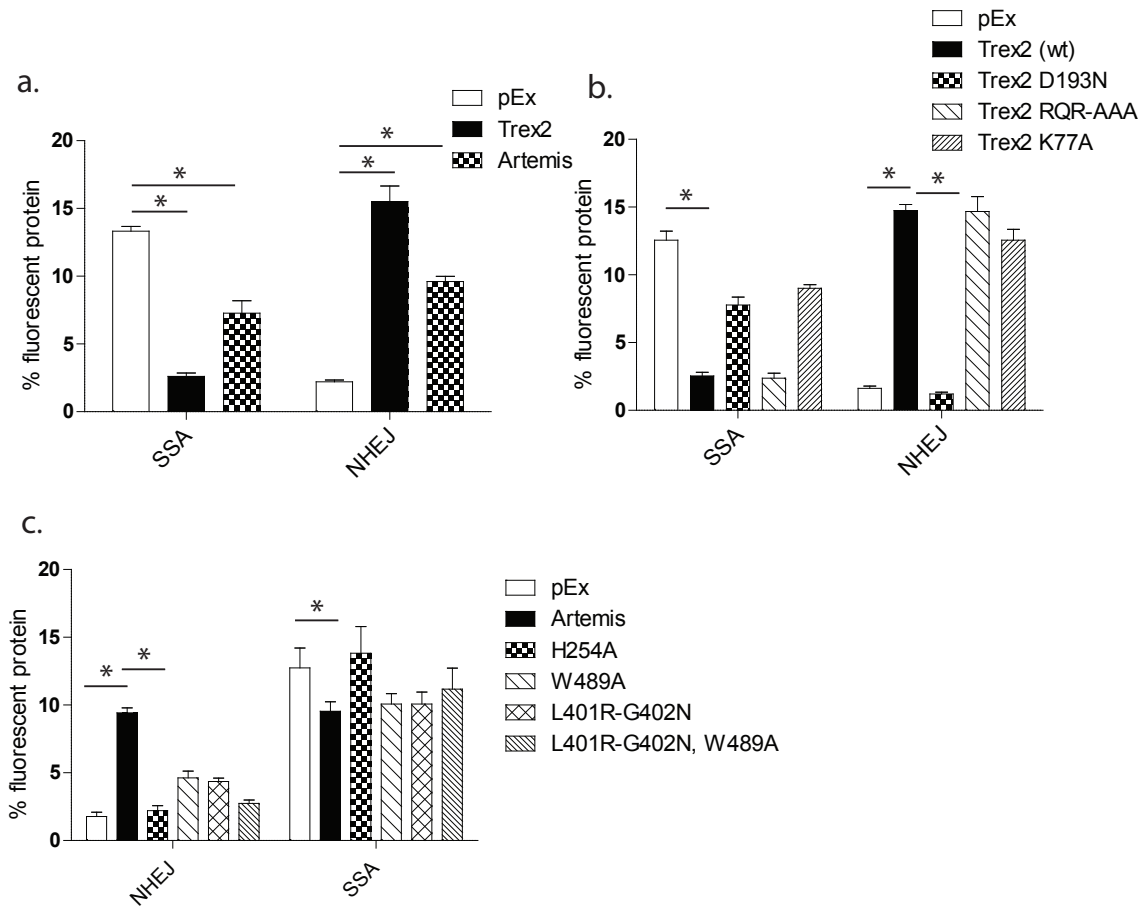


Figure 18. Trex2 and Artemis inhibit SSA. (**a-c**) SSA-TLR HEK293T cells transfected with I-SceI with the indicated vectors, and assayed for mutNHEJ (mCherry) and SSA (iRFP).

3.9 Summary

We have shown that both Trex2 and Artemis act directly on DNA to process ends and promote mutagenic NHEJ through creating small deletions with both platforms of designer endonuclease. They both divert DSB events from HR and SSA towards NHEJ, though Trex2 is more effective at SSA inhibition. Through mutational analysis, we show that Trex2 must directly contact DNA in order to be fully active, and show evidence that deleting a potential protein-ineration domain, its β-hairpin, has no effect, suggesting it may act autonomously of NHEJ factors. In contrast, we show that Artemis expression

promotes its catalytic activity at DSBs but that this depends on its interaction with key cNHEJ factors LigIV and DNA-PKcs.

Based on this data, we have created a model of Trex2 and Artemis activity when co-expressed with a designer endonuclease. We suggest that immediately following a break, Trex2 can localize to the DNA and resect the ends, with or without other factors present; however, its presence prevents long-range resection, MRN/CtIP binding, and thus HR/SSA. This is dependent on Trex2 being catalytically active and able to dimerize, perhaps to bridge across the DNA break and capture both ends. Artemis, however, is only effective as part of an active cNHEJ complex, but its presence increases overall processing when cNHEJ is active. Because cNHEJ already the major mammalian repair pathway, and endogenous Artemis is rarely used in chemically clean breaks such as those left by designer nucleases, we suggest that its over-expression promotes its natural localization to the DSB scaffold and mutagenic processing of the clean DSB.

In summary, we have characterized two DNA modifying enzymes, Trex2 and Artemis, which have a significant effect on increasing gene disruption with at least two endonuclease platforms, by assessing their mechanisms of action. We suggest these enzyme-coupling technologies have the potential to safely and dramatically increase the utility of designer endonucleases for genetic knockout applications.

3.10 Materials and Methods

Cell Culture

HEK293T cells were cultured in glutamine-free Dulbecco's modified Eagle's medium

supplemented with 2mM L-glutamine, 10% Fetal Bovine Serum (FBS) and 5% penicillin/streptomycin.

Transfection

1.0x 10^5 HEK293T cells were plated 24 hours prior to transfection in 24-well plates. 0.5μg DNA was used for each expression vector, and transfected using Fugene6 or XtremeGene9 (Roche) according to the manufacture's protocol. Cells were analyzed on a flow cytometer 72 hours post-transfection, unless otherwise indicated

Traffic Light Reporter (TLR) Assay

TLR cell lines for HE, and TALEN targets were generated as previously described(74). Briefly, the TLR assay consists of a GFP open reading frame with an embedded designer endonuclease target site followed by a T2A.mCherry in the +3 frame. No fluorescent protein is made under basal conditions. If a nuclease is delivered, and a mutagenic NHEJ event occurs repairing the TLR in the +3 frame, mCherry is expressed. If a nuclease is delivered with donor (C-terminally truncated GFP), and the TLR is repaired using HR, GFP is expressed. The SSA TLR is as described (26). Briefly, this TLR construct is the same as the regular TLR except that it has an interrupted iRFP open reading frame upstream of the GFP and downstream of mCherry, separated by 1.8kb, which shares 765bp of homology. Following long-range resection, the iRFP homology arms are revealed, allowing SSA to occur and an intact iRFP protein to be expressed and visualized.

Flow Cytometry

Cells were analyzed on a BD LSRII or BD FACS ARIAII. The mCherry fluorophore was excited using a 561nm laser and acquired with a 610/20 filter. The mTagBFP

fluorophore was excited on a 405nm laser with a 450/50 filter. Data was analyzed using FlowJo software (FlowJo, Ashland OR).

Sce Digestion NHEJ Assay

Genomic DNA was isolated using Qiagen's DNA easy kit, and TLR target sites were PCR-amplified as previously described (74). 100ng of each PCR product was digested *in vitro* with recombinant I-SceI (New England Biolabs) for 6 hours at 37^0C. DNA was separated using a 1% agarose gel stained with ethidium bromide. Percent disruption was calculated by quantifying band intensity using Image J software, and dividing the intensity of the undigested band by the total.

NHEJ Sequencing

Genomic DNA was isolated from cells using Qiagen DNAeasy kit (Qiagen) and the HE target region was amplified as previously described (74). PCR products were cloned using a CloneJET PCR cloning kit (Fermentas) according to the manufactures protocol, followed by transformation into chemically competent DH5α *E.coli* bacteria. Bacterial colonies were directly sequenced using a standard colony sequencing protocol. Sequences were analyzed using the "Contig Express" software provided in Invitrogen's Vector NTI software suite (Invitrogen).

Statistical Analysis

Error bars represent standard error of the mean (SEM), and *p*-values (with * representing $p<0.05$, ** $p<0.005$, and *** $p<0.0005$) were calculated using the Student's two-tailed unpaired t-test to compare the samples indicated in this and all subsequent figures.

Chapter 4

Increasing HR by altering the structure of donor DNA

4.1 Introduction

Thus far we have successfully achieved significant increases in targeted gene disruption by co-expressing designer endonucleases with DNA end-modifying enzymes. However, none of these exonucleases, despite some having processive 5' exonuclease activity, were able to promote significantly meaningful increases in targeted gene repair (44). This repair arm of genome engineering is arguably of greater clinical – and overall biological – importance, as it has been incredibly difficult to seamlessly incorporate foreign DNA into primary cells, and many of the most translationally relevant applications require homologous-recombination-mediated repair of a targeted break using an exogenous template, rather than the simpler targeted gene disruption through NHEJ. One potential limitation of this pathway is the template DNA, which may be subject to epigenetic modification or processing by endogenous DNA repair mechanisms. Thus, to explore alternative means to enhance recombination-mediated repair with an exogenous template, we decided to change our approach from modifying the DNA ends post-DSB, to modifying the structure of the DNA template donor itself.

Currently the best balance between efficient donor delivery (into primary cells), safety, and engagement of HR, is through the use of various non-integrating viral vectors. However, directly altering the structure of the donor DNA itself is possible. This was first shown in Xenopus oocytes: donor DNA was delivered as a naked circular plasmid and linearized *in vivo* by I-SceI cleavage; HR between the linear donor and

chromosome was significantly increased over that using circular plasmid (92). In parallel, the observation was made in plants that cleaving the donor (but not the target) using I-SceI enhanced HDR frequency, in an extrachromosomal assay (93). Subsequently, this method has been applied using a combination of ZFNs and I-SceI in *Drosophila* (94), as well as AAV-specific nickases (95) and ZFNs (96, 97) in human cell lines. However, no one has directly compared types of double strand breaks (3' vs. 5' overhang) on the efficiency of using a linearized donor to promote HDR, or done a comparative analysis indicating which designer endonuclease platform is most efficient with this technology. Indeed, recent work has suggested that 5' ends may be more suitable for utilization of ssODN donors (22), but these have not been tested with HE's. Thus our aim was to broaden our understanding of providing an *in vivo* cleavable donor to human cells, and compare all four designer nuclease types: HEs, ZFNs, TALENs, and the recently developed MegaTALs.

4.2 *In vivo* donor cleavage with HE's significantly increases HDR

Using the TLR assay, we used multiple HEK293T cell lines to test three different HE's with an *in vivo* linearizable donor: I-SceI, I-AniI-Y2, and a chimeric HE fusion of I-Onu and I-Ltr, Onu-Ltr (25). We designed donor constructs with a cleavage site for the corresponding HE downstream of the SFFV promoter, upstream of the GFP coding ORF (**Fig. 19a**), designated as cleavage site **1**. The position of this cleavage site is within the homology tract of the chromosomal TLR, but outside of the GFP coding sequence. Notably, the donor construct has a GFP C-terminal truncation (14 amino acids, see (74)), such that repair by HDR using either the circular or *in vivo* linearized

donor that results in a repaired GFP+ signal can only occur through invasion of the antisense chromosomal 3' tail and non crossover-mediated synthesis-dependent strand annealing (SDSA).

In all three cases, treatment of HEK-293T TLR cells with cleavable donor and HE resulted in a significant increase in GFP+ cells over treatment with non-cleavable donor (Fig. 19 b, e, g). Further, with all three HE's, we found that treatment with cleavable donor significantly increased the ratio of HDR:NHEJ events, suggesting that DSB repair choice may be biased toward HDR by the use of *in vivo* linearized donor (**Fig. 19 c, f, h**).

We also created a donor construct in which we cloned two I-SceI target sites in tandem at position **1** (**Fig. 20a**). We found that using this cleavable donor construct increases the efficiency of HDR, while maintaining NHEJ rates constant (**Fig. 20 b, c**). This suggests that although the dose of donor DNA is kept constant, by increasing the number of nuclease target sites we can increase the 'apparent dose' of donor molecules, without overwhelming the cell with donor DNA.

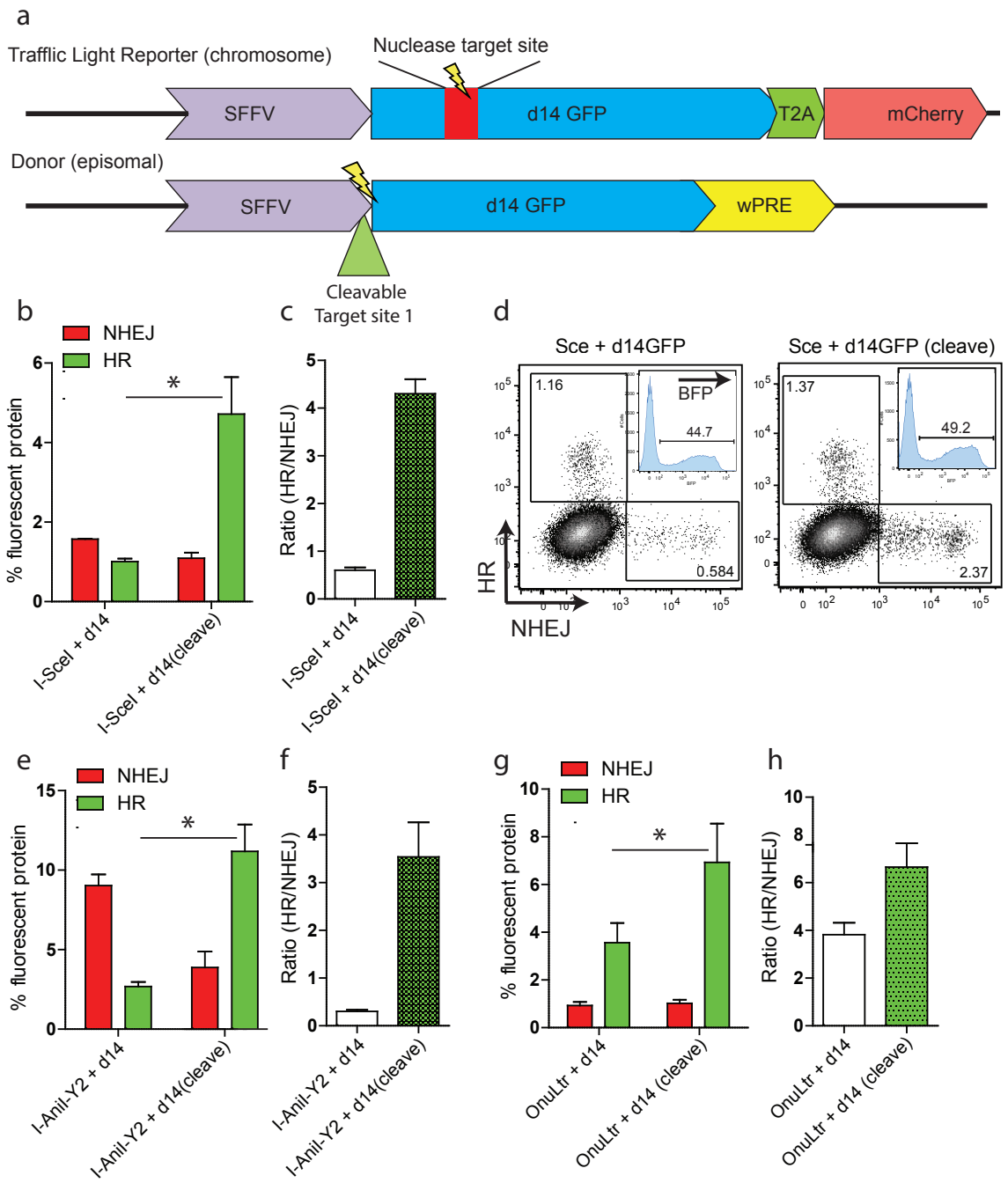


Figure 19. ***In vivo*** **cleavage of donor with homing endonucleases increases HR.** (**a**). Schematic of chromosomally integrated TLR (top) and episomal donor (below), with nuclease target sites indicated. (**b, e, h**) TLR assay of HEK293T cells transfected with I-SceI (**b**), I-AniI-Y2 (**e**) or OnuLtr (**g**) and either non-cleavable donor or cleavable donor. (**c, f, h**) Quantitation of the ratio of HR:NHEJ. (**d**) Representative flow plot of a TLR experiment indicating HR (GFP) and NHEJ (mCherry) following transfection of the indicating vectors, measured 72hrs post transfection.

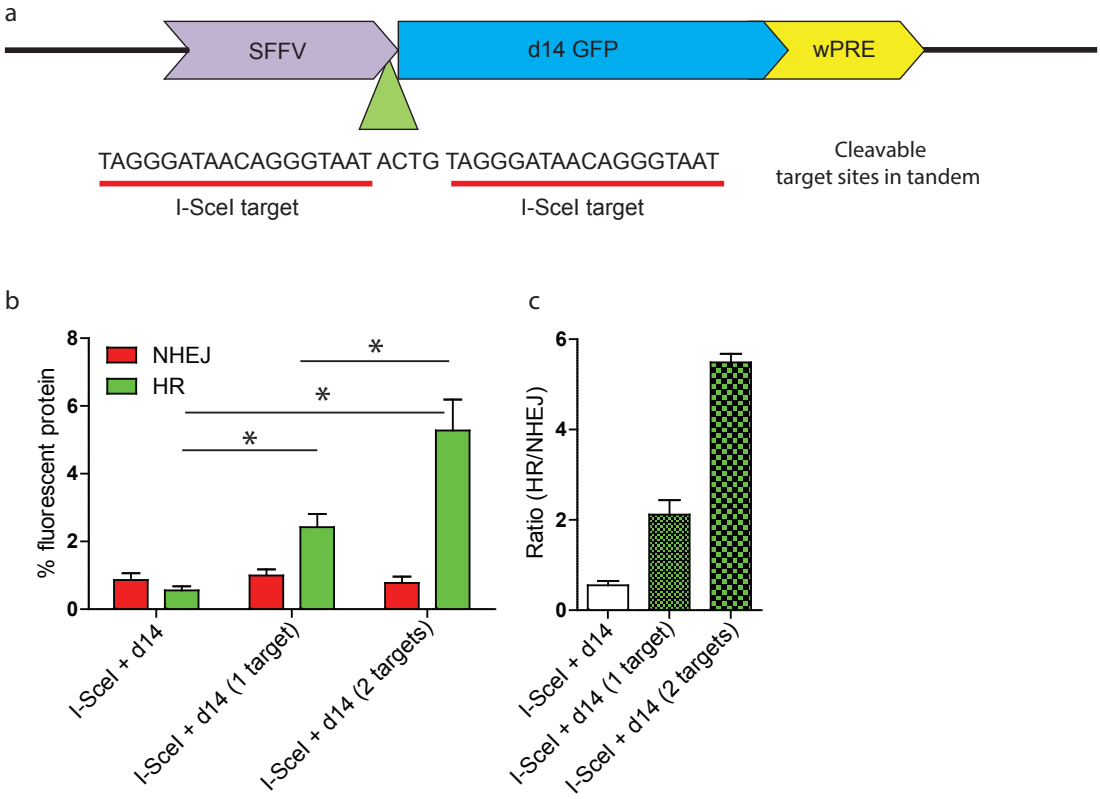


Figure 20. Two cleavable sites in tandem increase HR. (**a**) Schematic of the donor with two tandem I-SceI cleavable sites, upstream of the GFP (C-terminally truncated). (**b**) Quantification of TLR assay of HEK293T cells transfected with indicated vectors. (**c**) Quantification of the ratio of HR:NHEJ following TLR experiments in HEK293T cells transfected with the indicated vectors.

4.3 *In vivo* donor cleavage with MegaTALs significantly increases HDR

Similarly, we tested a novel platform with cleavable donor, the megaTAL nuclease, which was recently developed in our laboratory (21): a fusion of a homing endonuclease (in this case I-AniI) and one half of a TALEN pair (the CCR5 L538 TALE (82)). Because megaTALs use HEs as the cleavage nuclease, which leave 3' overhang DSBs, we hypothesized that they would behave similarly to HE's in response to a cleavable donor. We tested three RVD lengths of megaTAL: 5.5, 6.5, and 16.5, with TLRs specific for 6.5 and 16.5 (the target is shared between the last 5.5 and 16.5 TALE

RVDs so was used for both enzymes), and cleavable donor constructs were cloned that had megaTAL target sites at the **1** position. For each megaTAL, we found that treating cells with enzyme and cleavable donor significantly increased HDR over those containing non-cleavable donor (**Fig 21 a, b** and data not shown for 6.5 RVD MT). However, we found that this was only the case when the megaTAL target was cloned into the donor in the forward direction (L538-linker-I-AniI, **Fig. 21d**); when it was cloned in the reverse direction at the **1** position, treatment of cells with 5.5 and 16.5 megaTAL and cleavable donor abrogated HDR (**Fig 21 a,b**). Somewhat surprisingly, this was not limited to megaTALs: the same experiment with I-AniI-Y2 using the same donor constructs (megaTAL target in forward and reverse orientations) gave the same result (**Fig 21 c**). This suggests that this effect is independent of TALE-DNA binding and instead likely sequence-dependent at this position. Further, there was no significant difference with respect to HDR if a cleavable donor with megaTAL target (forward orientation) was used or donor with just the I-AniI target, indicating that only the sequence downstream of the cleavage site is necessary for HDR (**Fig. 21c**).

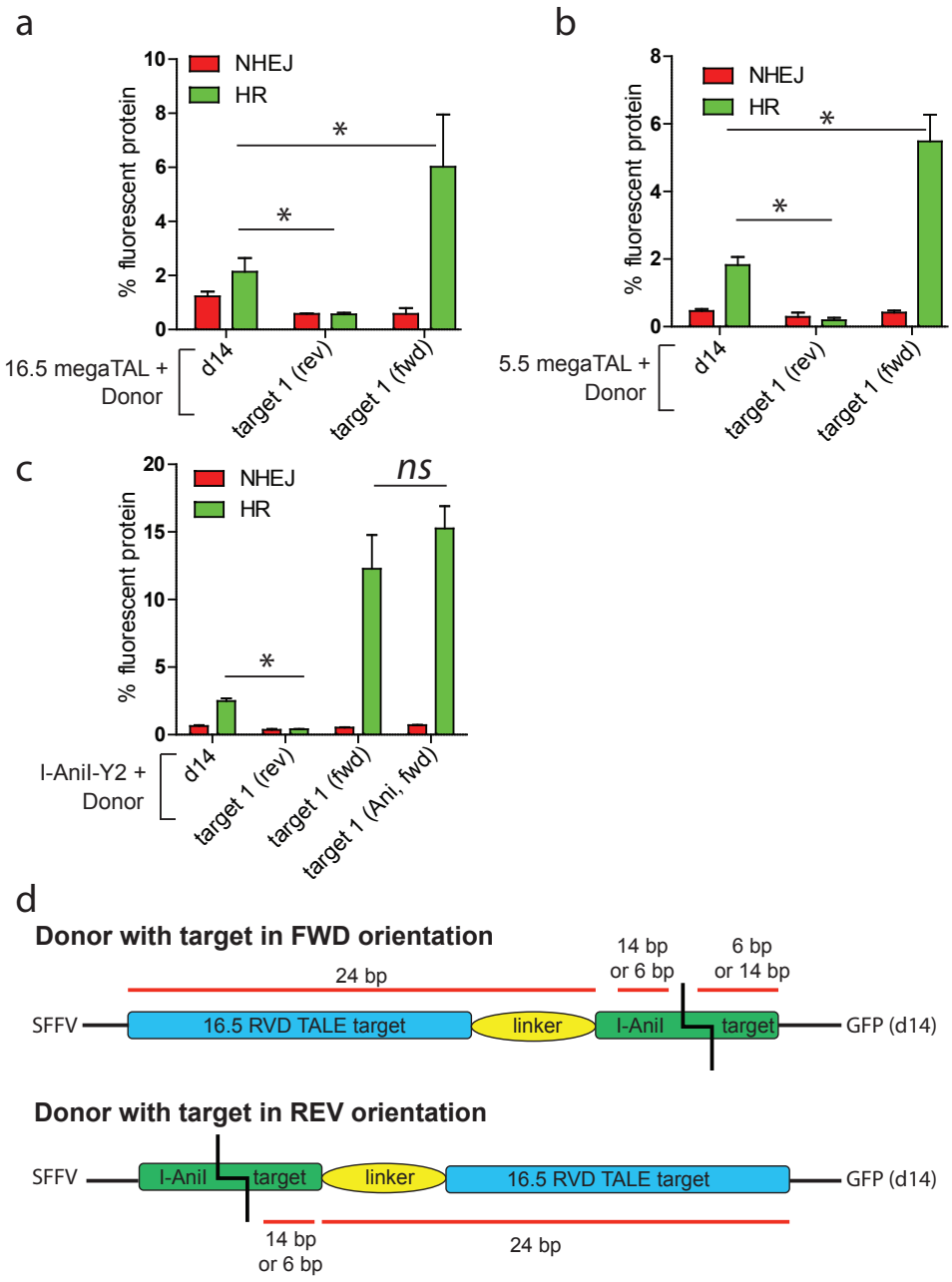


Figure 21. In vivo donor cleavage with megaTALs increases HDR. (**a-c**) TLR assay with 16.5 RVD megaTAL (**a**), 5.5 RVD megaTAL (**b**), and I-AniI-Y2 HE (**c**) in the same HEK293T cell line transfected with the indicated vectors. All donors had the megaTAL target (corresponding to 16.5 or 5.5 RVDs), in either the forward (fwd) or reverse (rev) orientations, unless noted otherwise. (**d**) Schematic of the donor with megaTAL targets in the forward (FWD) and reverse (REV) orientations.

4.4 Effects of position and orientation of the cleavable site on HDR

Given our results above, we speculated that the position of the cleavable site within the donor may affect HDR efficiency. We thus designed two new positions for a cleavage site, **2** and **3** (**Fig. 22a**). The **2** position, similar to **1**, is 1.4 kb upstream of the GFP ORF and completely homologous to corresponding TLR sequence. In contrast, the **3** position is 1.3 kb downstream of the GFP ORF, and contains ~ 600bp of TLR sequence non-homology, with the cleavage site at the end of this region. We transfected HEK293T cells with I-SceI and non-cleavable donor or these new cleavable donors, and found that there was no significant difference in HDR between a cleavable site at positions **1** or **2**; however, cleavage at the **3** position yielded HDR frequencies 2-fold higher than **1** or **2** (**Fig.22b**).

To determine whether the orientation of the target site is affected at different positions, we cloned forward and reverse 16.5 megaTAL target at the **3** position and treated cells with 5.5 and 16.5 megaTALs, as well as I-AniI-Y2, with both sets of cleavable donors and compared HR to non-cleavable donor. In contrast to the effect of target orientation at position **1**, there was no significant difference between having the megaTAL target in the forward or reverse orientation when cells were treated with I-AniI-Y2, or with either megaTAL, at position **3** (**Fig. 22 c-e**).

Given the success of using short linear DNA fragments, ssODNs (98) and work done in Drosophila where a homologous donor was excised from the genome (94), we were curious if we could further stimulate HDR by creating two nuclease target sites within the donor, at positions **1** and **3** to fully excise the donor DNA. However, we found that in the case of megaTALs, the presence of a short dsDNA fragment did not increase

HDR over an *in vivo* linearizable donor (**Fig. 22 d, e**), and using I-AniI-Y2 with a short dsDNA fragment tended to decrease HDR frequency over using a linearized (single cut) donor (**Fig. 22c**). We cannot rule out that this lack of effect is due to incomplete cleavage of the donor by the nucleases, although the appreciable level of NHEJ observed (mCherry+ signal) suggests this is unlikely. Another consideration to note is that ssODN donors, as well as chromosomally-based dsDNA fragments have only been shown to increase HDR with TALENs or ZFNs, whose Fok1 nuclease leaves 5' DNA overhangs. To our knowledge, this technology has not been successfully used with HE's. Both megaTALs and the I-AniI-Y2 HE leave 3' overhangs; perhaps dsDNA fragments are not effective donors with this type of DSB.

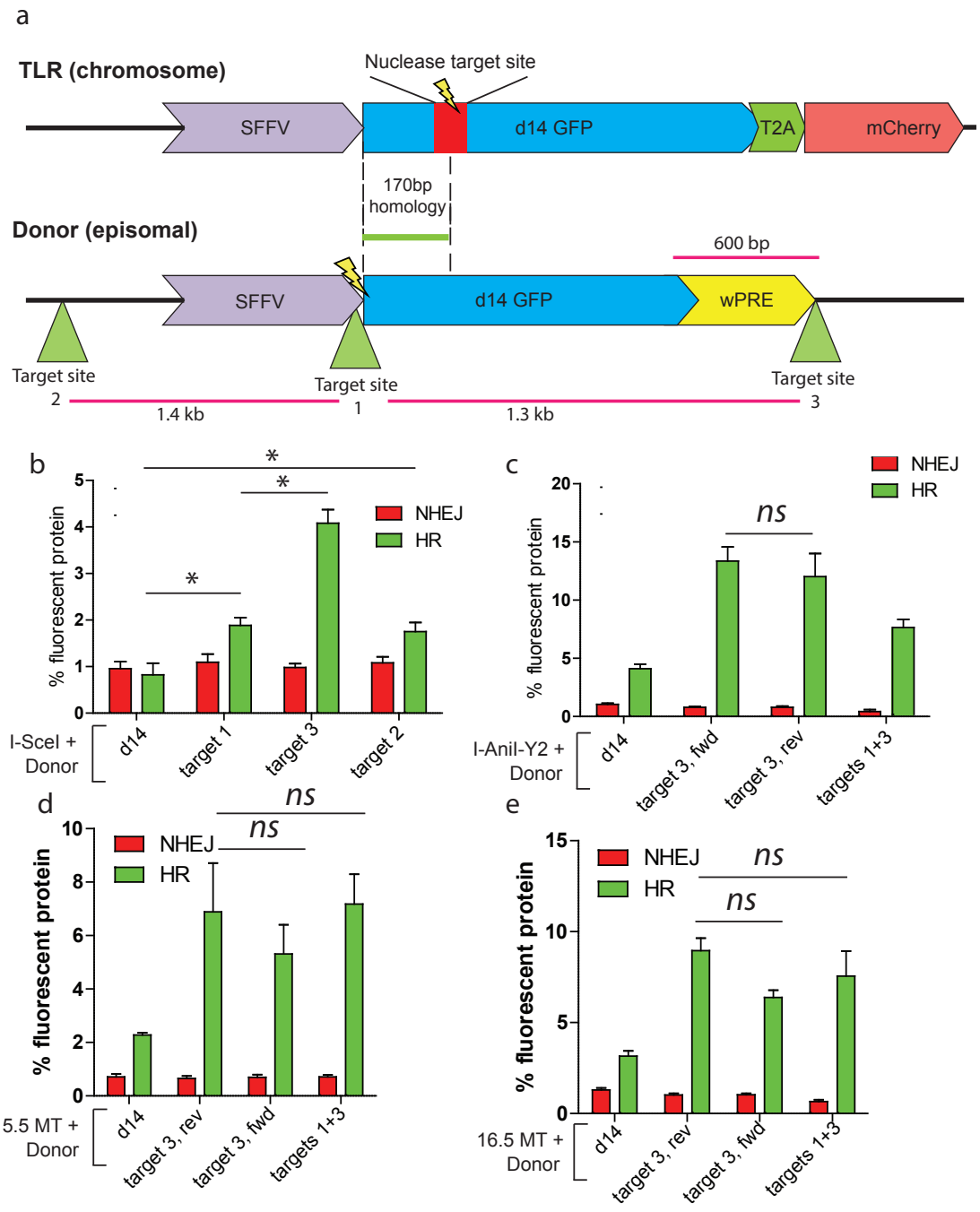


Figure 22. Orientation and positional effects of cleavable donor. (**a**) Schematic of TLR and donor, with two new cleavable positions. (**b-e**) TLR assay of HEK293T cells transfected with the indicated vectors. (**b**) Cells expressing I-SceI and donor with cleavable sites at different positions. (**c**) Cells expressing I-AniI-Y2 and cleavable donor (position 3) in different orientation. (**d**, **e**) Cells expressing megaTALs (5.5 RVD, **d**; 16.5 RVD, **e**) and cleavable donor (position **3**) in different orientations and dual-cleaving donor (positions 1 and 3).

4.5 *In vivo* cleavable donor has no effect on increasing HDR with TALENs but is effective with a ZFN

Next, we were interested if we could extend the results we obtained with cleavable donor with HE-based nucleases to the versatile Fok1-based nucleases: TALENs and ZFNs. We used two well-characterized enzymes, the CCR5 TALEN (82) and the VegF 2468 ZFN (83), as well as another TALEN developed in our laboratory against the T-cell MagT1 gene, and tested the effectiveness of cleavable donor in the TLR assay. Surprisingly, we found that for both TALENs, using a cleavable donor with target site at position **1** had no effect on HDR frequency (**Fig. 23 a, b**). Although it is possible that the lack of change in HDR could be due to inefficient donor cleavage by the TALENs, because we observed appreciable NHEJ levels (mCherry signal), this is unlikely. This gives us confidence that the TALENs are efficiently cleaving the chromosomal TLR locus, and thus likely also the episomal donor construct.

In contrast, when a cleavable donor with target site at position **1** was introduced with the VegF 2468 ZFN in the ZFN-target TLR, HDR was significantly increased (**Fig. 23c**). This result agrees with what has been previously reported with ZFNs (94, 97). The fold change increase in HDR with ZFN using cleavable donor (at position **1**) over non-cleavable was 2.9±0.2, similar to that seen with the 5.5 megaTAL (3.1±0.5), the 16.5 megaTAL (2.56±0.03), and I-AniI-Y2 (4±1). Importantly, however, the actual overall rate of HDR with the ZFN was much lower than with any HE-based enzyme.

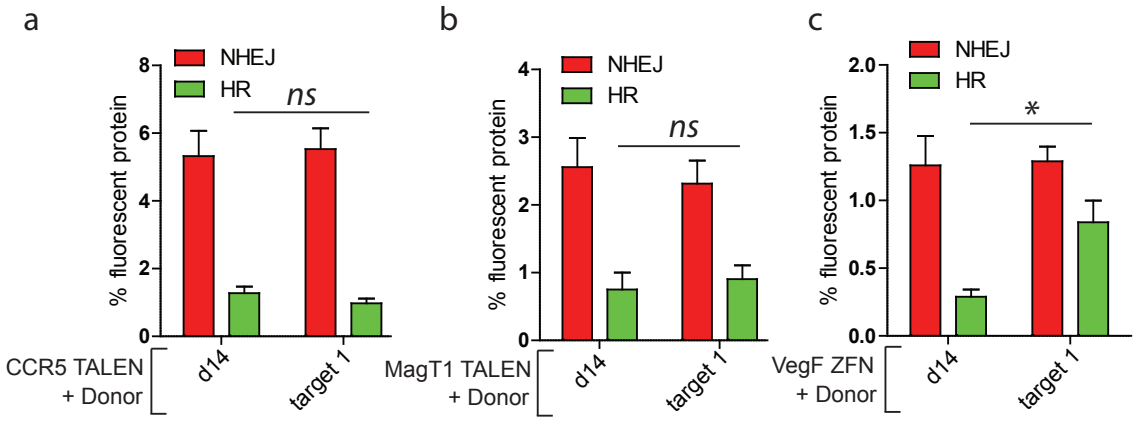


Figure 23. Cleaving donor in vivo has no effect on TALEN-induced breaks, but does increase HR with ZFN-induced breaks. (**a-c**) TLR assay of HEK293T cells transfected with the indicated designer nuclease and donors. *ns*, non-significant.

4.6 Cleaving the donor with an HE can increase HR following Fok1-induced chromosomal breaks

Given the disparate results between Fok1- and HE-based nucleases, we were curious whether the type of DNA break (3' vs. 5' overhang) of the chromosome or donor – or both – was critical for an *in vivo* linearized donor to increase HDR. In order to test this, we compared the effectiveness of cleaving the chromosome with an HE and the donor with Fok1-based nuclease, and visa versa. We used the I-SceI TLR cell line, which only has the target for the I-SceI HE, and co-transfected HEK293T cells with I-SceI and either non-cleavable donor or I-SceI position **3** donor, and compared this to a variety of other donors with their nucleases: MagT1 TALENs + MagT1position **1** donor; VegF + VegF position **1** donor; I-AniI Y2 + I-AniI position **1** donor; and 16.5 megaTAL + 16.5 position **3** donor (**Fig. 24**). We found that only those cells that had both the chromosome and donor cut with HE-based nucleases showed a significant increase in HDR. When the experiment was done in reverse, using either the VegF ZFN, MagT1 TALEN, or CCR5 TALEN TLR cell lines, the only increase in HDR was seen when cells

were treated with Fok1 nuclease to cut the chromosome and HE enzyme with its matched cleavable donor (**Fig. 24**).

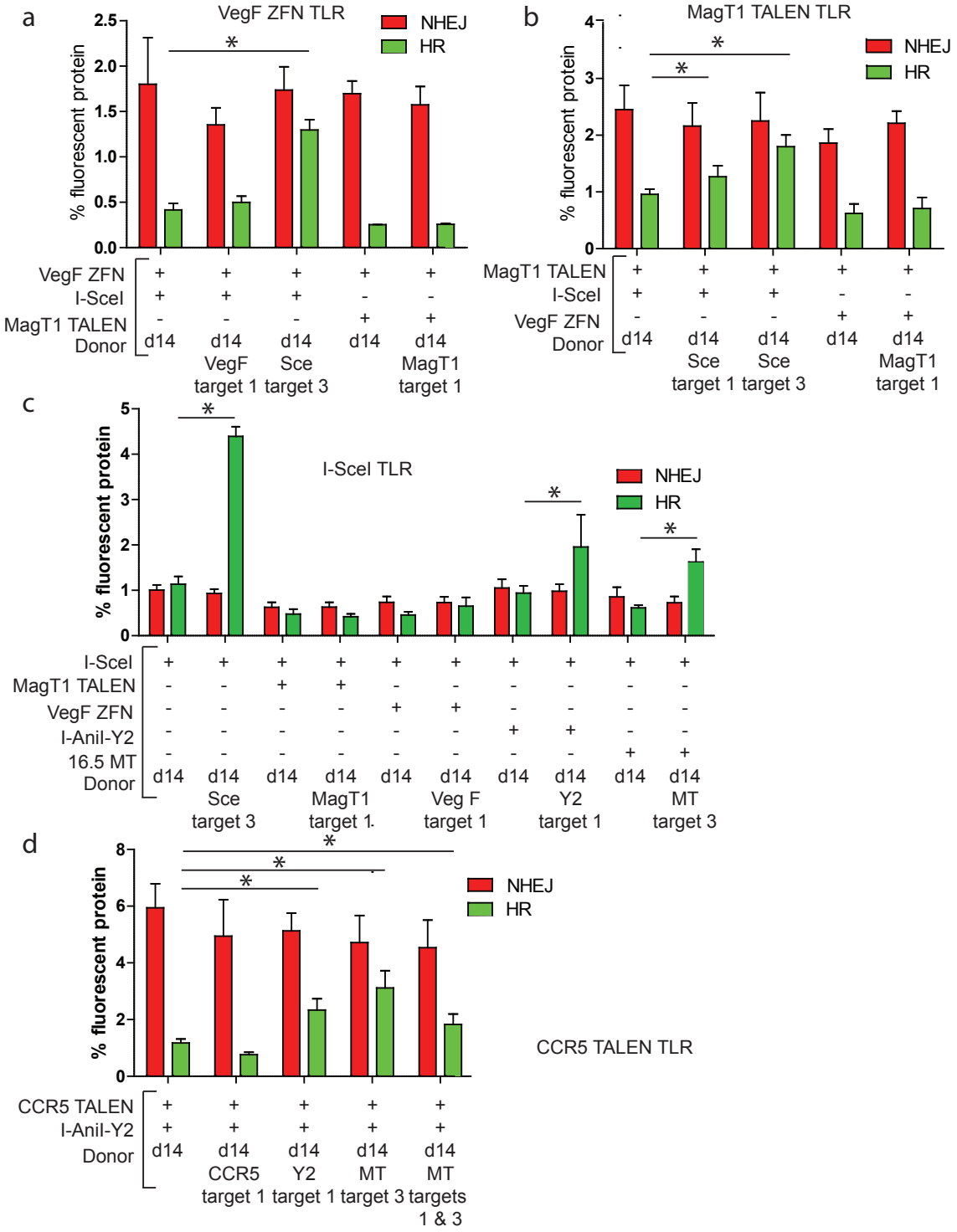


Figure 24. Only donors with 3' DSB overhangs increase HDR. (**a-d**) TLR assay experiments with mixed polarity of donor and chromosomal DSBs. (**a**) VegF ZFN TLR cells were treated with VegF nuclease to cleave the chromosome, and either the MagT1 TALEN 5' or I-SceI 3' nuclease/donor pairs, as indicated. (**b**) MagT1 TALEN TLR cells were treated with the MagT1 TALEN to cleave the chromosome, and and either the 3' I-SceI or 5' VegF ZFN nuclease/donor pairs. (**c**) I-SceI TLR cells were treated with the I-SceI HE to cleave the chromosome, and a variety of 3' or 5' nuclease/donor pairs, as indicated. (**d**) CCR5 TALEN TLR cells were treated with the CCR5 TALEN to cleave the chromosome and the I-AniI-Y2 HE to cleave its donor, with appropriate controls as indicated.

4.7 Cleavable donor does not act through the SSA pathway

As we have previously shown, the SSA pathway is very active in response to HE-induced 3' overhang DSBs, but not CCR5 TALEN 5' overhangs in HEK293T cells (26). Here we wanted to expand this observation by looking at the SSA response to megaTALs, as well as determine whether treatment with a cleavable donor changes the SSA/HDR repair ratio. First, we made a SSA-TLR line specific for the 16.5 megaTAL TLR, which also contains the 5.5 RVD megaTAL (and I-AniI alone) targets. We designed this SSA-TLR to be mCherry+ (a permanent +3 frame-shift in the TLR), to better track stable insertion of the reporter in our target cells. After treatment of these SSA-TLR HEK293T cells with 16.5 MT, 5.5 MT, or I-AniI-Y2, we found that the repair by SSA is dependent on RVD number (corresponding to TALE length) (**Fig. 25**). I-AniI-Y2 treatment resulted in the strongest level of SSA, followed by the shorter megaTAL (5.5 RVDs), with the longest megaTAL (16.5 RVDs) showed the weakest SSA response. I-AniI stimulates very little repair: 1.3% SSA, and negligible HR (data not shown). Interestingly, there was no difference between SSA levels between samples treated with normal donor or cleavable donor (target site at position **3**), although SSA levels were significantly lower than when no donor was present at all (total amounts of DNA in all transfections were kept equal). In all cases, treatment with cleavable donor was able

to stimulate HDR to similar extents as in the non-SSA TLR: there was no significant differences in fold change HDR between non-cleavable and cleavable donor between the two cell lines (data not shown). In addition, we used the cleavable donor with two target sites (positions **1** and **3**), which led to an increase in SSA and concomitant decrease in HR (**Fig. 25**). This donor most closely resembles a SSA substrate (when dsDNA is exised and resembles a homologous repeat) but rather than stealing SSA events and promoting HDR, its presence has the opposite effect with the 5.5 and 16.5 megaTALs, and decreases HR with I-AniI-Y2.

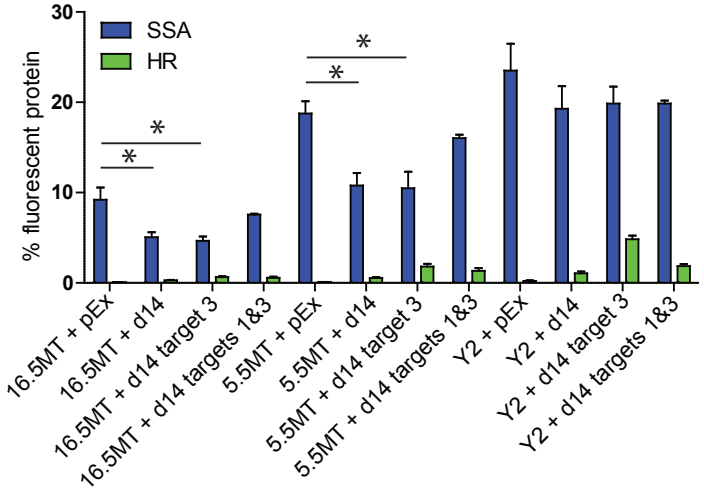


Figure 25. Cleavable donor does not act through the SSA pathway. SSA-TLR HEK293T cells transfected with vectors as indicated, and assayed for HR (GFP) and SSA (iRFP).

Given these results, and specifically that we do not see a decrease in SSA when cleavable donors are used over circular, we suggest that the cleavable donor does not act through the SSA pathway. Neither linearized nor fragmented donors do not steal repair events from SSA, since the same proportion of events are diverted down HDR if either circular or cleavable donor is present.

4.8 Determining whether a cleavable donor can work in a viral platform

Thus far we have tested cleavable donors within a plasmid delivery system, which is not an ideal platform for delivering DNA to primary cells due to their inflammatory response to foreign DNA. Most genome engineering applications are using non-integrating viral vectors to delivery sequence information to primary cells, including IDLV (integrase-deficient lentivirus), adenovirus (Ad), and AAV (adeno-associated virus). A recently published work using ZFNs cleaving donors *in vivo* showed that HR was significantly increased over non-cleaved donor in their model using adenovirus (97). Although Ad is a large and already linear virus, we hypothesized that we could use the much smaller AAV vector with our cleavable donor technology. AAV has a complex biology, and the vector can take on many forms in the cell - including circular and concatermeric - especially depending on whether it is packaged as single stranded (ss) or self-complementary (sc).

However, when we transfected HEK293T TLR cells with I-AniI-Y2 and transduced with either non-cleavable ssAAV, ssAAV with one cleavage site (position **3**), or ssAAV with two cleavage sites (positions **1** and **3**), there was no significant difference between treatments on HR (**Fig. 26a**). When we repeated the experiment using scAAV, we saw the same results (**Fig. 26b**). We went on to attempt to analyze the insertion rates of these constructs, to determine whether having a cleavable AAV donor – especially a dual-cleavable donor – can decrease the rate of AAV insertion, which a current concern with this vector platform (99). However, despite several efforts to ascertain insertion rates using PCR-amplification of genomic DNA from cells transfected

with Y2 and ssAAV donors, insertion could not be detected (data not shown). We hypothesized that this was due to AAV insertion at rates below PCR detection; to estimate this, we measured the insertion rate of an AAV6 MND.BFP construct into HEK293Ts following co-delivery with Y2. 3 days following transduction/transfection, AAV6 (BFP+) was expressed in 90% of cells, transduced at an MOI of 50k (**Fig. 26c**). We measured I-AniI-Y2 expression, and found it was present in 25% of cells (GFP+); it yielded mCherry mutNHEJ rates of 0.9% (Fig. 26c). After 8 days, the difference between BFP signal from AAV delivered alone (5.6% BFP+, due to remaining episomal AAV expression) and AAV co-delivered with Y2 (10% BFP+, due to episomal expression and insertion) was 4.4% (**Fig. 26d**). Thus an estimate of the upper limit of insertion in HEK293T cells, with an AAV6 construct transduced at an MOI of 50k, is likely too low to detect by standard PCR methods, especially when trying to look for quantitative differences between insertion rates.

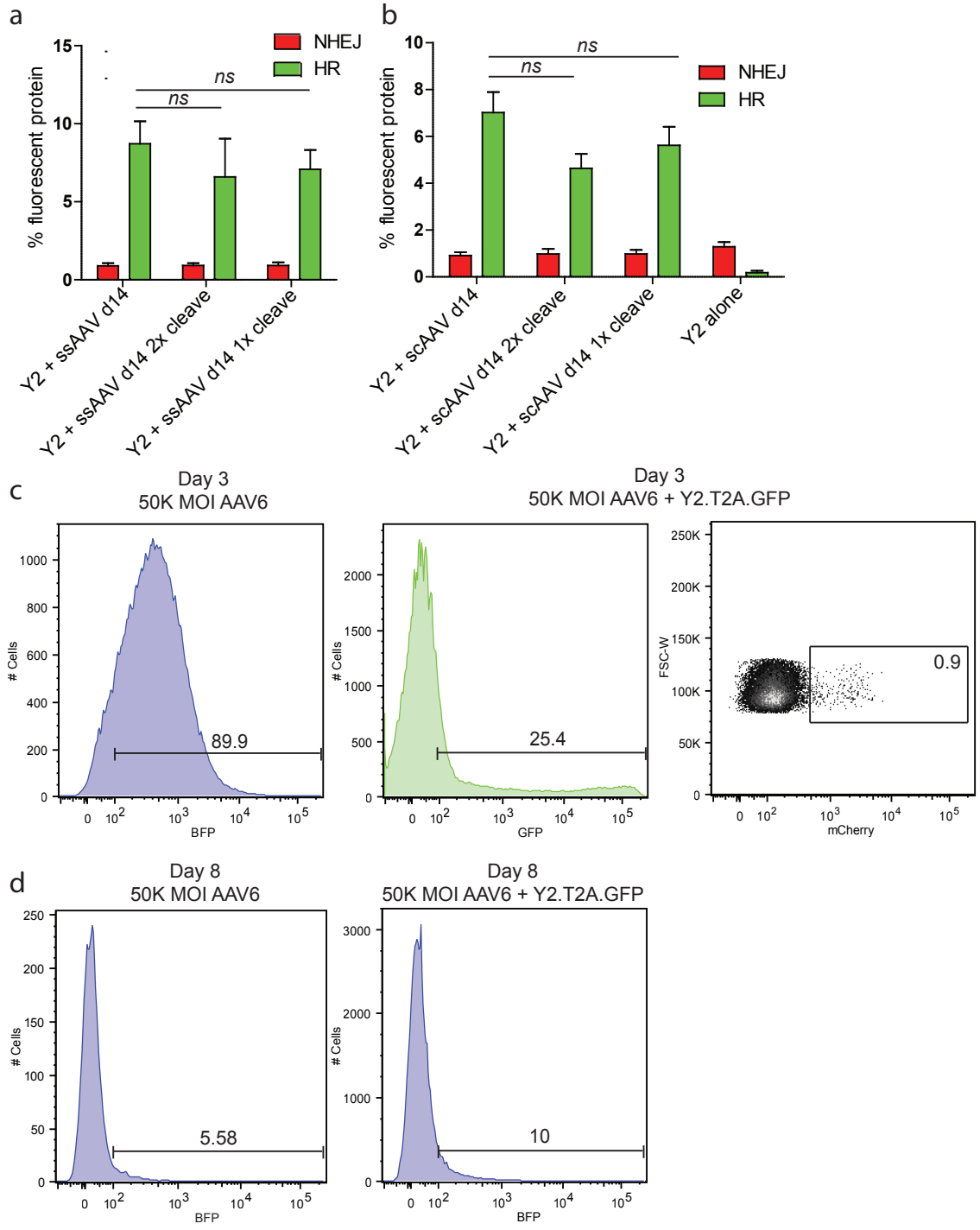


Figure 26. Cleavable donor does not increase the rate of HR with AAV. (**a, b**) Quantification of TLR experiments using single stranded (**a**) or self-complimentary (**b**) AAV vectors, with one, two, or no cleavable sites as indicated. (**c**) Flow plots following transduction/transfection of AA6 MND.BFP and Y2.t2a.GFP on day 3 and day 8 into HEK293T TLR cells harboring an I-AniI target site. Day 3 represents episomal expression of AAV, and mCherry represents mutNHEJ from successful Y2 expression. BFP expression in cells transduced with AAV alone represents random AAV integration

and residual episomal expression (5.58%); BFP expression in the population that received both AAV with Y2 represents random integration, integration at the TLR target site, as well as residual episomal AAV expression (10%).

4.9 Summary

Here we show, for the first time in mammalian cells, that specifically linearizing a DNA donor *in vivo* leads to significant increases in HDR frequency, when both donor and chromosome are cut with HE-based enzymes. Previous work has shown that using linear DNA, whether it be ssODN, excised dsDNA fragment, PCR product, or nicked vector (92, 94, 95, 97) is an effective way to stimulate HDR. We have expanded on this work to show that this strategy is more effective when 3' DSB overhangs are created rather than 5' overhangs, critically in the donor.

Our model for why a linearized vector is efficient at promoting HDR with HE-based nucleases is based on our observation that DSBs with 3' overhangs are significantly more amenable to resection than those created by a TALEN, measured by SSA repair(26). Previous work has indicated that there are differences in DSB repair based on both overhang length and polarity (100). Thus we suggest that following a HE-induced DSB, but not a TALEN-induced DSB, the recessed 5' strand on both the chromosome and donor is resected more efficiently, thereby exposing ssDNA 3' tails available for HR strand exchange. We expand on this observation by showing that megaTALs stimulate SSA in a manner dependent on the amount of TALE binding: the more TALE repeats are present, the higher the inhibition on resection and SSA. We suggest that TALENs, which have both the aspect of TALE-DNA binding and 5' overhangs, deliver two strikes against promoting resection.

A DSB within an episomal plasmid would localize this broken donor DNA to repair foci (101), possibly where the chromosomal DSB is also being repaired (some evidence suggests this may occur at the nuclear periphery, (102). Linearization also causes the duplex donor DNA to relax (compared to the native supercoiled vector), requiring less energy for strand displacement, second-end capture, and displacement loop (D-loop) formation. Our results show that a DSB downstream of the homology tract (at the **3** position) is more effective than those within homology tracts; we hypothesize that this is because DNA can be resected from this break while leaving the upstream homology unaffected and visible to the invading chromosomal strand. This helps explain why the orientation of the megaTAL target at the **3** position does not affect HDR rates, but strongly affects them at the **1** position. At **3**, the target in the donor is not part of the TLR homology tract, and as such the break can be made in either orientation: the extra megaTAL target sequence does not matter. However, at the **1** position, the megaTAL target is 170bp upstream of the chromosomal DSB; single molecule analysis of D-loops *in vitro* using magnetic tweezers estimate that ~79bp of DNA are involved in the homology pairing at a time (103), though longer D-loops (289bp) have also been measured (104). Since D-loops are not much longer than the distance between this break and the chromosomal (TLR) break, a 30bp mismatch at the beginning may present an energy barrier too high to overcome for strand pairing. Indeed, previous studies have shown that DNA mismatches within homology tracts can inhibit HDR (105, 106) and D-loop formation (104). In contrast, a 10bp mismatch at the ends does not affect strand pairing (104). If the megaTAL target is in the forward

orientation, the broken strand downstream only has 6bp of TLR non-homology; if the target is in the reverse orientation, it has 30bp.

We propose that if gene repair is the desired outcome, the optimal enzyme to use is a HE-based nuclease. With the availability of new methods that significantly decrease the time and effort needed to redesign the specificity of HE's (19, 20), it has become much more accessible to create good HE enzymes, and/or megaTALs, for theoretically any given genomic target.

4.10 Materials and Methods

Cell Culture

HEK293T cells were cultured in glutamine-free Dulbecco's modified Eagle's medium supplemented with 2mM L-glutamine, 10% Fetal Bovine Serum (FBS) and 5% penicillin/streptomycin.

Transfection

1.0x 10^5 HEK293T cells were plated 24 hours prior to transfection in 24-well plates. 0.5μg DNA was used for each expression vector, and transfected using Fugene6 or XtremeGene9 (Roche) according to the manufacture's protocol. Cells were analyzed on a flow cytometer 72 hours post-transfection, unless otherwise indicated

Traffic Light Reporter (TLR) Assay

TLR cell lines for HE, and TALEN targets were generated as previously described(44). Briefly, the TLR assay consists of a GFP open reading frame with an embedded designer endonuclease target site followed by a T2A.mCherry in the +3 frame. No fluorescent protein is made under basal conditions. If a nuclease is delivered, and a

mutagenic NHEJ event occurs repairing the TLR in the +3 frame, mCherry is expressed. If a nuclease is delivered with donor (C-terminally truncated GFP), and the TLR is repaired using HR, GFP is expressed. The SSA TLR is as described (26). Briefly, this TLR construct is the same as the regular TLR except that it has an interrupted iRFP open reading frame upstream of the GFP and downstream of mCherry, separated by 2.5kb, which shares ~750bp of homology. Following long-range resection, the iRFP homology arms are revealed, allowing SSA to occur and an intact iRFP protein to be expressed and visualized.

Flow Cytometry

Cells were analyzed on a BD LSRII or BD FACS ARIAII. The mCherry fluorophore was excited using a 561nm laser and acquired with a 610/20 filter. The mTagBFP fluorophore was excited on a 405nm laser with a 450/50 filter. Data was analyzed using FlowJo software (FlowJo, Ashland OR).

Statistical Analysis

Error bars represent standard error of the mean (SEM), and *p*-values (with * representing $p<0.05$, ** $p<0.005$, and *** $p<0.0005$) were calculated using the Student's two-tailed unpaired t-test to compare the samples indicated in this and all subsequent figures.

Chapter 5

The Future of Genome Engineering

5.1 Improving the Trex2 platform

Trex2 is a fantastic tool for genome engineering, and has already been used as a tool to improve mutagenesis rates (48). Indeed, another group has since engineered a single-chain monomeric version of Trex2 (107), obviating the need for dimerization.

To further improve the specificity and on-target rate of Trex2 DNA processing, while increasing its safety profile, it may be advantageous to directly couple (using a flexible protein linker) a DNA binding-defective Trex2 to a HE or megaTAL. This would ensure that this endo/exo enzyme would only resect designer endonuclease-induced breaks (those that only the HE binds to), rather than any randomly occurring DSBs throughout the genome. We hypothesize that Trex2's DNA binding deficiency would be made up for by the DNA binding affinities of the HE/TALE.

5.2 Towards Targeted Repair

5.2a RGENs with cleavable donor

One designer endonuclease platform we have not yet tested with cleavable donor is the new RNA-guided endonuclease system (RGEN), based on RNA-guided Cas9 cleavage. The Cas9 endonuclease cleaves DNA to create blunt ends, and recently a nickase version has been developed, allowing overhangs of almost any size and polarity to be created using two Cas9 nickases, (22). It would be interesting to test

the effect of a cleavable donor containing a blunt end DSB, as well as donors with longer 5' and 3' overhangs, on promoting HR.

5.2b Cleavable donor as a safety mechanism: decreasing AAV insertion

AAV vectors promote HR to a significant extent, and thus perhaps having a cleavable AAV will not further increase the possibly saturated rate of HR. Further, while AAV can become circular or concatameric within the cell (108, 109), it is inherently linear; perhaps creating an *in vivo* cleavage event may not have the desired effect of linearizing it per se.

However, the use of AAV vectors poses a small but significant threat of unwanted, and potentially deleterious, AAV genome insertion at genomic DSBs. The mechanism by which is occurs is not well understood, but it is thought that it occurs through NHEJ, and the AAV ITRs are critical for viral insertion. Thus, we hypothesize that perhaps by cleaving the AAV vector, especially a dual-cleavage event that excises the homologous template and deletes the ITR's, would prevent this insertion event. Or, alternatively, the cleavage could occur post-hoc, after an insertion event, to delete it via an SSA type repair event.

We attempted to test whether our cleavable AAV constructs were affecting the rate of AAV insertion, but the insertion rates were below the limit of detection with our PCR assay (data not shown). One could perform amplicon sequencing on genomic DNA obtained from PCR amplification, though this would be laborious and potentially biased due to PCR amplifying smaller amplicons more readily than larger ones. A superior method would be to design an insertion reporter to determine whether AAV

insertion or HR occurs following a DSB: for example, by interrupting or turning on different fluorescent protein coding sequences.

5.2c Cleavable viruses and minicircles

While our AAV results indicated that using a cleavable donor – both with one or two cleavage sites – was ineffective in promoting HR beyond the already high rate induced by AAV, there are other platforms that we did not test. Others have recently shown that *in vivo* ZFN cleavage within the (already linear) adenovirus increases its ability to promote HR (97).

There is a technology that, although difficult to work with, may prove the most effective with cleavable donors. Minicircles, essentially plasmids but without any residual baceterial sequence, have shown to be efficiently electroporated into many different primary cell populations (110). They remain episomal and do not carry the risk of insertional mutagenesis, and so their use carries distinct advantages over certain viral vectors. Given their circular nature, and the fact that they are delivered to the cell as DNA, we hypothesize that using *in vivo* cleavage with a minicircle would significantly enhance HR.

5.2d Harnessing the SSA pathway to promote HR

Post-mitotic cells, cells that are in G1 arrest, or cells that divide slowly, undergo HR minimally. When it is these cell populations we want to repair, it becomes a monumentally difficult task. In the case of stem cells, too many cell divisions - especially *ex vivo,* where the stem cell niche cannot be perfectly replicated - can lead to

differentiation or senescence, and thus limiting cell cycle progressions would be advantageous. In all these cases, a way to promote HR without requiring entry into S/G2 phase would be of great benefit.

We were unable to stimulate HR with any of the 5-3' processive exonucleases in our screen under normal conditions. However, when paired with a cleavable donor and I-SceI, two of these exonucleases, the viral proteins UL12 and Sox (from HSV-1 and KSHV, respectively), significantly increased HR over that stimulated by cleavable donor alone. We went on to show that both UL12 and Sox increase endogenous mammalian SSA with I-SceI in the SSA-TLR assay. This preliminary data suggests that if we could harness the SSA pathway, which is both active during all phases of the cell cycle, as well as in response to HE- and megaTAL-induced DSBs, it could be an alternative but powerful way to promote HR with linear DNA and processive 5' viral exonucleases. The cleavable donor mimics an SSA substrate, and perhaps a dual-cleavable donor, that has an excisable homologous dsDNA fragment, would be even more effective.

Expression of UL12, however, did not universally promote HR with cleavable donor. It caused notable toxicity when co-expressed with I-Ani-Y2, and did not affect HR with megaTALs or TALENs, owing perhaps to the restrictive TALE-DNA binding of these enzymes, which may preclude UL12's 5' exonuclease activity. In order to move this technology forward, UL12 needs to be more carefully characterized, and perhaps mutagenically modified in order to be less active – and thus less toxic – to human cells. It should also only be paired with a designer endonuclease that has low off-target rates – an enzyme like I-Ani-Y2, which has many known off-target sites, likely created many different genomic cleavage events that promoted UL12 resection and SSA, leading to

genotoxicity and cell death. There is also evidence that UL12 interacts with endogenous MRN DNA repair machinery (111, 112) while this is important for viral replication, it may not be required for promotion of SSA and indeed may participate in cellular toxicity. The UL12-MRN interaction domain has been mapped, and perhaps mutating this interaction could reduce UL12-mediated toxicity without affecting its SSA-promoting nuclease activity.

Another option is to look elsewhere for homologous exonucleases to UL12. HSV-1 is a pathogenic virus in humans, and as such has evolved to cause considerable, albeit not lethal, damage to its host cells. However, other alkaline exonucleases related to UL12 exist in the herpesvirus family, although HSV-1's UL12 is the best characterized. Perhaps using an exonuclease less suited to human cells – such as the alkaline exonuclease from the chimpanzee alpha-1 herpesvirus, might prevent unwanted interactions with human DNA repair proteins but retain the more conserved exonuclease activity, as it shares 80% amino acid identity with HSV-1 UL12, based on NCBI BLAST search.

References

1. Friedmann T. A brief history of gene therapy. Nat Genet. 1992 Oct;2(2):93-8.
2. Blaese RM, Culver KW, Miller AD, Carter CS, Fleisher T, Clerici M, et al. T lymphocyte-directed gene therapy for ADA- SCID: Initial trial results after 4 years. Science. 1995 Oct 20;270(5235):475-80.
3. Bordignon C, Mavilio F, Ferrari G, Servida P, Ugazio AG, Notarangelo LD, et al. Transfer of the ADA gene into bone marrow cells and peripheral blood lymphocytes for

the treatment of patients affected by ADA-deficient SCID. Hum Gene Ther. 1993 Aug;4(4):513-20.
4. Puumalainen AM, Vapalahti M, Agrawal RS, Kossila M, Laukkanen J, Lehtolainen P, et al. Beta-galactosidase gene transfer to human malignant glioma in vivo using replication-deficient retroviruses and adenoviruses. Hum Gene Ther. 1998 Aug 10;9(12):1769-74.
5. Wirth T, Parker N, Yla-Herttuala S. History of gene therapy. Gene. 2013 Aug 10;525(2):162-9.
6. Hacein-Bey-Abina S, Garrigue A, Wang GP, Soulier J, Lim A, Morillon E, et al. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J Clin Invest. 2008 Sep;118(9):3132-42.
7. Lewinski MK, Yamashita M, Emerman M, Ciuffi A, Marshall H, Crawford G, et al. Retroviral DNA integration: Viral and cellular determinants of target-site selection. PLoS Pathog. 2006 Jun;2(6):e60.
8. Bushman F, Lewinski M, Ciuffi A, Barr S, Leipzig J, Hannenhalli S, et al. Genome-wide analysis of retroviral DNA integration. Nat Rev Microbiol. 2005 Nov;3(11):848-58.
9. Fehse B, Roeder I. Insertional mutagenesis and clonal dominance: Biological and statistical considerations. Gene Ther. 2007 11/01;15(2):143-53.
10. Jasin M. Genetic manipulation of genomes with rare-cutting endonucleases. Trends Genet. 1996 Jun;12(6):224-8.
11. Taghian DG, Nickoloff JA. Chromosomal double-strand breaks induce gene conversion at high frequency in mammalian cells. Mol Cell Biol. 1997 Nov;17(11):6386-93.
12. Perlman PS, Butow RA. Mobile introns and intron-encoded proteins. Science. 1989 Dec 1;246(4934):1106-9.
13. Wu C, Dunbar CE. Stem cell gene therapy: The risks of insertional mutagenesis and approaches to minimize genotoxicity. Front Med. 2011 Dec;5(4):356-71.
14. Silva G, Poirot L, Galetto R, Smith J, Montoya G, Duchateau P, et al. Meganucleases and other tools for targeted genome engineering: Perspectives and challenges for gene therapy. Curr Gene Ther. 2011 Feb;11(1):11-27.
15. McMahon MA, Rahdar M, Porteus M. Gene editing: Not just for translation anymore. Nat Methods. 2011 Dec 28;9(1):28-31.
16. Jacoby K, Metzger M, Shen BW, Certo MT, Jarjour J, Stoddard BL, et al. Expanding LAGLIDADG endonuclease scaffold diversity by rapidly surveying evolutionary sequence space. Nucleic Acids Res. 2012 Jun;40(11):4954-64.
17. Takeuchi R, Lambert AR, Mak AN, Jacoby K, Dickson RJ, Gloor GB, et al. Tapping natural reservoirs of homing endonucleases for targeted gene modification. Proc Natl Acad Sci U S A. 2011 Aug 9;108(32):13077-82.
18. Baxter SK, Scharenberg AM, Lambert AR. Engineering and flow-cytometric analysis of chimeric LAGLIDADG homing endonucleases from homologous I-OnuI-family enzymes. Methods Mol Biol. 2014;1123:191-221.
19. Takeuchi R, Choi M, Stoddard BL. Redesign of extensive protein-DNA interfaces of meganucleases using iterative cycles of in vitro compartmentalization. Proc Natl Acad Sci U S A. 2014 Mar 18;111(11):4061-6.

20. Wang Y, Khan IF, Boissel S, Jarjour J, Pangallo J, Thyme S, et al. Progressive engineering of a homing endonuclease genome editing reagent for the murine X-linked immunodeficiency locus. Nucleic Acids Res. 2014 Mar 25.
21. Boissel S, Jarjour J, Astrakhan A, Adey A, Gouble A, Duchateau P, et al. megaTALs: A rare-cleaving nuclease architecture for therapeutic genome engineering. Nucleic Acids Res. 2014 Feb;42(4):2591-601.
22. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell. 2013 Sep 12;154(6):1380-9.
23. Takeuchi R, Certo M, Caprara MG, Scharenberg AM, Stoddard BL. Optimization of in vivo activity of a bifunctional homing endonuclease and maturase reverses evolutionary degradation. Nucleic Acids Res. 2009 Feb;37(3):877-90.
24. Arnould S, Perez C, Cabaniols JP, Smith J, Gouble A, Grizot S, et al. Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. J Mol Biol. 2007 Aug 3;371(1):49-65.
25. Baxter S, Lambert AR, Kuhar R, Jarjour J, Kulshina N, Parmeggiani F, et al. Engineering domain fusion chimeras from I-OnuI family LAGLIDADG homing endonucleases. Nucleic Acids Res. 2012 Sep;40(16):7985-8000.
26. Kuhar R, Gwiazda KS, Humbert O, Mandt T, Pangallo J, Brault M, et al. Novel fluorescent genome editing reporters for monitoring DNA repair pathway utilization at endonuclease-induced breaks. Nucleic Acids Res. 2014 Jan;42(1):e4.
27. Guo J, Gaj T, Barbas CF,3rd. Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. J Mol Biol. 2010 Jul 2;400(1):96-107.
28. Szczepek M, Brondani V, Buchel J, Serrano L, Segal DJ, Cathomen T. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat Biotechnol. 2007 Jul;25(7):786-93.
29. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, et al. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics. 2010 Oct;186(2):757-61.
30. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 2011 Jul;39(12):e82.
31. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol. 2012 May;30(5):460-5.
32. Ramirez CL, Foley JE, Wright DA, Muller-Lerch F, Rahman SH, Cornu TI, et al. Unexpected failure rates for modular assembly of engineered zinc fingers. Nat Methods. 2008 May;5(5):374-5.
33. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, et al. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell. 2008 Jul 25;31(2):294-301.
34. Hansen K, Coussens MJ, Sago J, Subramanian S, Gjoka M, Briner D. Genome editing with CompoZr custom zinc finger nucleases (ZFNs). J Vis Exp. 2012 Jun 14;(64):e3304. doi(64):e3304.
35. Fine EJ, Cradick TJ, Zhao CL, Lin Y, Bao G. An online bioinformatics tool predicts zinc finger and TALE nuclease off-target cleavage. Nucleic Acids Res. 2014 Apr 1;42(6):e42.

36. Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. Nat Methods. 2011 Aug 7;8(9):765-70.
37. Kim Y, Kweon J, Kim JS. TALENs and ZFNs are associated with different mutation signatures. Nat Methods. 2013 Mar;10(3):185.
38. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: Zinc finger fusions to fok I cleavage domain. Proc Natl Acad Sci U S A. 1996 Feb 6;93(3):1156-60.
39. Carroll D. Genome engineering with zinc-finger nucleases. Genetics. 2011 Aug;188(4):773-82.
40. Gaj T, Gersbach CA, Barbas CF,3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 2013 Jul;31(7):397-405.
41. Grizot S, Smith J, Daboussi F, Prieto J, Redondo P, Merino N, et al. Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease. Nucleic Acids Res. 2009 Sep;37(16):5405-19.
42. Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol. 2009 Sep;27(9):851-7.
43. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, et al. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol. 2011 Jul 7;29(8):731-4.
44. 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 Oct;9(10):973-5.
45. Torikai H, Reik A, Liu PQ, Zhou Y, Zhang L, Maiti S, et al. A foundation for universal T-cell based immunotherapy: T cells engineered to express a CD19-specific chimeric-antigen-receptor and eliminate expression of endogenous TCR. Blood. 2012 Jun 14;119(24):5697-705.
46. Provasi E, Genovese P, Lombardo A, Magnani Z, Liu PQ, Reik A, et al. Editing T cell specificity towards leukemia by zinc finger nucleases and lentiviral gene transfer. Nat Med. 2012 May;18(5):807-15.
47. Weber ND, Aubert M, Dang CH, Stone D, Jerome KR. DNA cleavage enzymes for treatment of persistent viral infections: Recent advances and the pathway forward. Virology. 2014 Apr;454-455C:353-61.
48. Aubert M, Boyle NM, Stone D, Stensland L, Huang ML, Magaret AS, et al. In vitro inactivation of latent HSV by targeted mutagenesis using an HSV-specific homing endonuclease. Mol Ther Nucleic Acids. 2014 Feb 4;3:e146.
49. Cradick TJ, Keck K, Bradshaw S, Jamieson AC, McCaffrey AP. Zinc-finger nucleases as a novel therapeutic strategy for targeting hepatitis B virus DNAs. Mol Ther. 2010 May;18(5):947-54.
50. Li H, Haurigot V, Doyon Y, Li T, Wong SY, Bhagwat AS, et al. In vivo genome editing restores haemostasis in a mouse model of haemophilia. Nature. 2011 Jun 26;475(7355):217-21.
51. Cannon P, June C. Chemokine receptor 5 knockout strategies. Curr Opin HIV AIDS. 2011 Jan;6(1):74-9.

52. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med. 2014 Mar 6;370(10):901-10.
53. Hutter G, Nowak D, Mossner M, Ganepola S, Mussig A, Allers K, et al. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. N Engl J Med. 2009 Feb 12;360(7):692-8.
54. Allers K, Hutter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, et al. Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation. Blood. 2011 Mar 10;117(10):2791-9.
55. Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol. 2008 Jul;26(7):808-16.
56. Holt N, Wang J, Kim K, Friedman G, Wang X, Taupin V, et al. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. Nat Biotechnol. 2010 Aug;28(8):839-47.
57. Iyama T, Wilson DM,3rd. DNA repair mechanisms in dividing and non-dividing cells. DNA Repair (Amst). 2013 Aug;12(8):620-36.
58. Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol. 2007 Nov;25(11):1298-306.
59. Hirsch ML, Samulski RJ. AAV-mediated gene editing via double-strand break repair. Methods Mol Biol. 2014;1114:291-307.
60. Van Tendeloo VF, Ponsaerts P, Lardon F, Nijs G, Lenjou M, Van Broeckhoven C, et al. Highly efficient gene delivery by mRNA electroporation in human hematopoietic cells: Superiority to lipofection and passive pulsing of mRNA and to electroporation of plasmid cDNA for tumor antigen loading of dendritic cells. Blood. 2001 Jul 1;98(1):49-56.
61. Beatty GL, Haas AR, Maus MV, Torigian DA, Soulen MC, Plesa G, et al. Mesothelin-specific chimeric antigen receptor mRNA-engineered T cells induce anti-tumor activity in solid malignancies. Cancer Immunol Res. 2014 Feb 1;2(2):112-20.
62. Gracey Maniar LE, Maniar JM, Chen ZY, Lu J, Fire AZ, Kay MA. Minicircle DNA vectors achieve sustained expression reflected by active chromatin and transcriptional level. Mol Ther. 2013 Jan;21(1):131-8.
63. Vitor MT, Bergami-Santos PC, Barbuto JA, de la Torre LG. Cationic liposomes as non-viral vector for RNA delivery in cancer immunotherapy. Recent Pat Drug Deliv Formul. 2013 Aug;7(2):99-110.
64. Nimonkar AV, Genschel J, Kinoshita E, Polaczek P, Campbell JL, Wyman C, et al. BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. Genes Dev. 2011 Feb 15;25(4):350-62.
65. Shrivastav M, De Haro LP, Nickoloff JA. Regulation of DNA double-strand break repair pathway choice. Cell Res. 2008 Jan;18(1):134-47.
66. Bennardo N, Gunn A, Cheng A, Hasty P, Stark JM. Limiting the persistence of a chromosome break diminishes its mutagenic potential. PLoS Genet. 2009 Oct;5(10):e1000683.
67. Riballo E, Kuhne M, Rief N, Doherty A, Smith GC, Recio MJ, et al. A pathway of double-strand break rejoining dependent upon ATM, artemis, and proteins locating to gamma-H2AX foci. Mol Cell. 2004 Dec 3;16(5):715-24.

68. Schipler A, Iliakis G. DNA double-strand-break complexity levels and their possible contributions to the probability for error-prone processing and repair pathway choice. Nucleic Acids Res. 2013 Sep;41(16):7589-605.
69. Soni A, Siemann M, Grabos M, Murmann T, Pantelias GE, Iliakis G. Requirement for parp-1 and DNA ligases 1 or 3 but not of Xrcc1 in chromosomal translocation formation by backup end joining. Nucleic Acids Res. 2014 Apr 19.
70. Iliakis G. Backup pathways of NHEJ in cells of higher eukaryotes: Cell cycle dependence. Radiother Oncol. 2009 Sep;92(3):310-5.
71. Mansour WY, Schumacher S, Rosskopf R, Rhein T, Schmidt-Petersen F, Gatzemeier F, et al. Hierarchy of nonhomologous end-joining, single-strand annealing and gene conversion at site-directed DNA double-strand breaks. Nucleic Acids Res. 2008 Jul;36(12):4088-98.
72. Bennardo N, Cheng A, Huang N, Stark JM. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. PLoS Genet. 2008 Jun 27;4(6):e1000110.
73. George CM, Alani E. Multiple cellular mechanisms prevent chromosomal rearrangements involving repetitive DNA. Crit Rev Biochem Mol Biol. 2012 May-Jun;47(3):297-313.
74. Certo MT, Ryu BY, Annis JE, Garibov M, Jarjour J, Rawlings DJ, et al. Tracking genome engineering outcome at individual DNA breakpoints. Nat Methods. 2011 Jul 10;8(8):671-6.
75. Doyon Y, Choi VM, Xia DF, Vo TD, Gregory PD, Holmes MC. Transient cold shock enhances zinc-finger nuclease-mediated gene disruption. Nat Methods. 2010 Jun;7(6):459-60.
76. Adimoolam S, Sirisawad M, Chen J, Thiemann P, Ford JM, Buggy JJ. HDAC inhibitor PCI-24781 decreases RAD51 expression and inhibits homologous recombination. Proc Natl Acad Sci U S A. 2007 Dec 4;104(49):19482-7.
77. Mazur DJ, Perrino FW. Excision of 3' termini by the Trex1 and TREX2 3'-->5' exonucleases. characterization of the recombinant proteins. J Biol Chem. 2001 May 18;276(20):17022-9.
78. Mazur DJ, Perrino FW. Identification and expression of the TREX1 and TREX2 cDNA sequences encoding mammalian 3'-->5' exonucleases. J Biol Chem. 1999 Jul 9;274(28):19655-60.
79. Chen MJ, Ma SM, Dumitrache LC, Hasty P. Biochemical and cellular characteristics of the 3' -> 5' exonuclease TREX2. Nucleic Acids Res. 2007;35(8):2682-94.
80. Perrino FW, de Silva U, Harvey S, Pryor EE,Jr, Cole DW, Hollis T. Cooperative DNA binding and communication across the dimer interface in the TREX2 3' --> 5'-exonuclease. J Biol Chem. 2008 Aug 1;283(31):21441-52.
81. Perrino FW, Harvey S, McMillin S, Hollis T. The human TREX2 3' -> 5'-exonuclease structure suggests a mechanism for efficient nonprocessive DNA catalysis. J Biol Chem. 2005 Apr 15;280(15):15212-8.
82. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, et al. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol. 2011 Feb;29(2):143-8.
83. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, et al. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell. 2008 Jul 25;31(2):294-301.

84. Woodbine L, Brunton H, Goodarzi AA, Shibata A, Jeggo PA. Endogenously induced DNA double strand breaks arise in heterochromatic DNA regions and require ataxia telangiectasia mutated and artemis for their repair. Nucleic Acids Res. 2011 Sep 1;39(16):6986-97.
85. Kanikarla-Marie P, Ronald S, De Benedetti A. Nucleosome resection at a double-strand break during non-homologous ends joining in mammalian cells - implications from repressive chromatin organization and the role of ARTEMIS. BMC Res Notes. 2011 Jan 21;4:13,0500-4-13.
86. de Silva U, Perrino FW, Hollis T. DNA binding induces active site conformational change in the human TREX2 3'-exonuclease. Nucleic Acids Res. 2009 Apr;37(7):2411-7.
87. Shevelev IV, Ramadan K, Hubscher U. The TREX2 3'-->5' exonuclease physically interacts with DNA polymerase delta and increases its accuracy. ScientificWorldJournal. 2002 Feb 1;2:275-81.
88. de Villartay JP, Shimazaki N, Charbonnier JB, Fischer A, Mornon JP, Lieber MR, et al. A histidine in the beta-CASP domain of artemis is critical for its full in vitro and in vivo functions. DNA Repair (Amst). 2009 Feb 1;8(2):202-8.
89. Kurosawa A, Adachi N. Functions and regulation of artemis: A goddess in the maintenance of genome integrity. J Radiat Res. 2010;51(5):503-9.
90. Malu S, De Ioannes P, Kozlov M, Greene M, Francis D, Hanna M, et al. Artemis C-terminal region facilitates V(D)J recombination through its interactions with DNA ligase IV and DNA-PKcs. J Exp Med. 2012 May 7;209(5):955-63.
91. Soubeyrand S, Pope L, De Chasseval R, Gosselin D, Dong F, de Villartay JP, et al. Artemis phosphorylated by DNA-dependent protein kinase associates preferentially with discrete regions of chromatin. J Mol Biol. 2006 May 19;358(5):1200-11.
92. Segal DJ, Carroll D. Endonuclease-induced, targeted homologous extrachromosomal recombination in xenopus oocytes. Proc Natl Acad Sci U S A. 1995 Jan 31;92(3):806-10.
93. Puchta H, Dujon B, Hohn B. Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. Nucleic Acids Res. 1993 Nov 11;21(22):5034-40.
94. Beumer K, Bhattacharyya G, Bibikova M, Trautman JK, Carroll D. Efficient gene targeting in drosophila with zinc-finger nucleases. Genetics. 2006 Apr;172(4):2391-403.
95. Goncalves MA, van Nierop GP, Holkers M, de Vries AA. Concerted nicking of donor and chromosomal acceptor DNA promotes homology-directed gene targeting in human cells. Nucleic Acids Res. 2012 Apr;40(8):3443-55.
96. Cristea S, Freyvert Y, Santiago Y, Holmes MC, Urnov FD, Gregory PD, et al. In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. Biotechnol Bioeng. 2013 Mar;110(3):871-80.
97. Zhang W, Wang D, Liu S, Zheng X, Ji H, Xia H, et al. Multiple copies of a linear donor fragment released in situ from a vector improve the efficiency of zinc-finger nuclease-mediated genome editing. Gene Ther. 2014 Mar;21(3):282-8.
98. Chen F, Pruett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J, et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. Nat Methods. 2011 Jul 17;8(9):753-5.

99. Rosas LE, Grieves JL, Zaraspe K, La Perle KM, Fu H, McCarty DM. Patterns of scAAV vector insertion associated with oncogenic events in a mouse model for genotoxicity. Mol Ther. 2012 Nov;20(11):2098-110.
100. Daley JM, Wilson TE. Rejoining of DNA double-strand breaks as a function of overhang length. Mol Cell Biol. 2005 Feb;25(3):896-906.
101. Dellaire G, Kepkay R, Bazett-Jones DP. High resolution imaging of changes in the structure and spatial organization of chromatin, gamma-H2A.X and the MRN complex within etoposide-induced DNA repair foci. Cell Cycle. 2009 Nov 15;8(22):3750-69.
102. Oza P, Jaspersen SL, Miele A, Dekker J, Peterson CL. Mechanisms that regulate localization of a DNA double-strand break to the nuclear periphery. Genes Dev. 2009 Apr 15;23(8):912-27.
103. van der Heijden T, Modesti M, Hage S, Kanaar R, Wyman C, Dekker C. Homologous recombination in real time: DNA strand exchange by RecA. Mol Cell. 2008 May 23;30(4):530-8.
104. Ristic D, Kanaar R, Wyman C. Visualizing RAD51-mediated joint molecules: Implications for recombination mechanism and the effect of sequence heterology. Nucleic Acids Research. 2011 January 01;39(1):155-67.
105. Beumer KJ, Trautman JK, Mukherjee K, Carroll D. Donor DNA utilization during gene targeting with zinc-finger nucleases. G3 (Bethesda). 2013 Mar 22.
106. Lin FL, Sperle K, Sternberg N. Model for homologous recombination during transfer of DNA into mouse L cells: Role for DNA ends in the recombination process. Mol Cell Biol. 1984 Jun;4(6):1020-34.
107. Delacote F, Perez C, Guyot V, Duhamel M, Rochon C, Ollivier N, et al. High frequency targeted mutagenesis using engineered endonucleases and DNA-end processing enzymes. PLoS One. 2013;8(1):e53217.
108. Choi VW, Samulski RJ, McCarty DM. Effects of adeno-associated virus DNA hairpin structure on recombination. J Virol. 2005 Jun;79(11):6801-7.
109. Cataldi MP, McCarty DM. Hairpin-end conformation of adeno-associated virus genome determines interactions with DNA-repair pathways. Gene Ther. 2013 Jun;20(6):686-93.
110. Chen ZY, He CY, Ehrhardt A, Kay MA. Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. Mol Ther. 2003 Sep;8(3):495-500.
111. Schumacher AJ, Mohni KN, Kan Y, Hendrickson EA, Stark JM, Weller SK. The HSV-1 exonuclease, UL12, stimulates recombination by a single strand annealing mechanism. PLoS Pathog. 2012;8(8):e1002862.
112. Balasubramanian N, Bai P, Buchek G, Korza G, Weller SK. Physical interaction between the herpes simplex virus type 1 exonuclease, UL12, and the DNA double-strand break-sensing MRN complex. J Virol. 2010 Dec;84(24):12504-14.