


High-efficiency homology-directed insertion into the genome using the engineered homing endonuclease ARCUS

Laura Christian Resly, Alan L. Tubbs, Alexander J. Vogel, Jo Ann Hux, Ian A. MacDonald, Jason Harris, Adam Mischler, Ginger H. Tomberlin, Kathryn S. Evans, Rhese Thompson, Jeffrey Sunman, Janel Lape, J. Jeff Smith, Aaron J. Martin  *

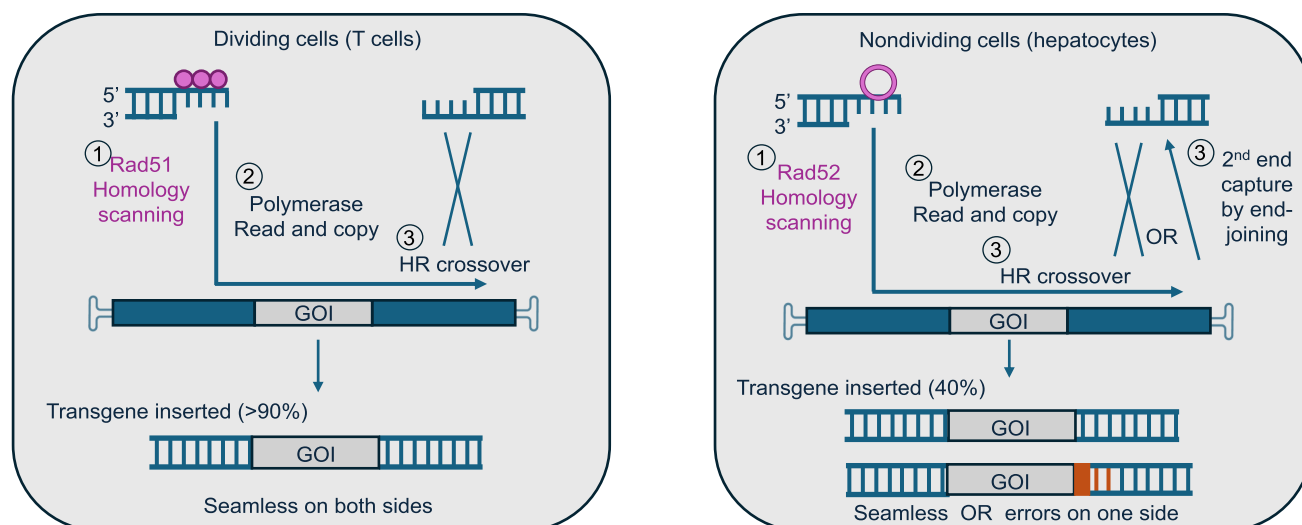
Precision BioSciences, Inc. Durham, NC, United States

*To whom correspondence should be addressed. Email: aaron.martin@precisionbiosciences.com

Abstract

Several gene editing tools have entered the clinic, representing varied options for eliminating or correcting mutations. Although gene editing by homologous recombination (HR) can potentially accomplish any type of gene edit (insertions, deletions, and replacements), as the outcome is defined by a recombinant repair template, gene editing enzymes that support efficient HR are rare. ARCUS nucleases, engineered from the homing endonuclease I-Crel, have programmable sequence specificity and support precise, high-frequency transgene insertion. In this study, we demonstrate that the 3' overhangs that ARCUS nucleases generate when cutting DNA are key to triggering high rates of HR. We show that a single editor can be used to accomplish the full range of currently understood DNA editing approaches, allowing all combinations of single base changes, introducing small, specific deletions, small and large insertions, and the ability to replace large segments of genomic DNA with efficiencies ranging from 60% to 90% in lymphocytes. ARCUS also supports precise, efficient insertion (30%–40%) in noncycling hepatocytes via nonclassical HR pathways. Collectively, this work characterizes a flexible and efficient gene insertion system for potential therapeutic use.

Graphical abstract



Introduction

Insertion of recombinant DNA into precise locations in the genome can permanently alleviate diseases caused by mutations and can serve therapeutic areas that gene deletion cannot, such as loss-of-function or autosomal dominant disorders

[1–3]. Ideally, a therapeutic nuclease is delivered, which introduces a double-stranded break (DSB) in the genomic DNA and activates a DNA damage response (DDR). A recombinant DNA template can be delivered simultaneously to work in concert with cellular DDR factors to repair the DSB [4]. This

Received: November 13, 2024. Revised: August 21, 2025. Accepted: September 5, 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

results in insertion of the synthetic template into the genomic site, thereby disrupting the pathogenic mutant gene and allowing expression of the healthy copy from the synthetic template. In order to avoid random insertion and other undesired mutations, genomic integration should be mediated by a precise repair pathway, such as homologous recombination (HR), rather than by more error-prone end-joining mechanisms [4, 5].

Development of DNA editing enzymes has yielded several options for deleting or correcting mutations (reviewed in [6]). In addressing the challenge of some specific recessive and autosomal dominant disorders, the preferred approaches involve replacement or re-coding of the diseased allele [7]. The CRISPR–Cas endonuclease system has been a popular way of generating DSBs and allowing an insertion of synthetic DNA templates; however, rates of precise insertion using this technology are typically low [8–11]. Despite the precision of HR, it is slower and more complex than the more error-prone end-joining methods, which cells can deploy extremely quickly and efficiently [12]. In order to achieve appreciable homology based insertion, studies suggest it is necessary to use high quantities of synthetic template (e.g. 10^5 or 10^6 AAV viral genomes per cell), alter the biology of the target cell (e.g. inhibit end-joining repair, manipulate cell-cycle progression, etc.), or both [13–25]. Consensus therefore is that HR-mediated repair of mutations using CRISPR–Cas, zinc finger, or TALE endonucleases and recombinogenic DNA templates is inefficient in rapidly dividing cells and very inefficient, on the verge of nonexistent, in slow- or nondividing cell types [1, 8, 10, 26].

Other editing technologies are also advancing, such as base editors, prime editors, and PASTE, which utilize the targeting capability of CRISPR nickases in combination with additional editing domains [1, 27]. Base editors repair defects by targeting point mutations with a Cas9 nickase fused to either a cytidine or adenine deaminase. Consequently, only adenine or cytidine bases in the genome are amenable to changing to guanine or thymine, respectively, thereby constraining the scope of targets and mutations that can be addressed [28]. Prime editors, in contrast, can introduce small prime editing guide RNA (pegRNA) templated changes in the genome to accomplish all possible base changes as well as small insertions and deletions [29]. Prime editing approaches involve a guided Cas9 enzyme fused to reverse transcriptases, which are more error-prone than other DNA polymerases [30]. Furthermore, the size of an edit that can be made by prime editors is limited by the amount of template in their pegRNA (50–200 bp) which precludes whole gene insertion applications [29]. PASTE seeks to overcome the limited size of insertions for prime editors by inserting an integrase recognition site into the genome using a pair of prime editors and then inserting a larger DNA fragment at the newly created site with either an independent or fused integrase or recombinase [31]. PASTE has been shown to successfully insert 1–10 kb DNA fragments but is challenged by delivery constraints inherent to large multi-subunit protein complexes. A small gene editor that stimulates precise, efficient insertion could potentially address multiple types of edits without these various platform limitations.

I-CreI is a particularly small protein belonging to a class of homing endonucleases with defined DNA sequence specificity [32]. These endonucleases, which are encoded in mobile introns in the genomes of many species, evolved to introduce a DSB in unoccupied cognate alleles, and to thereby facilitate

the insertion of a copy of their coding sequence [33, 34]. I-CreI is able to achieve this result, in part, because it generates a 4 nucleotide (nt) staggered cut, leaving behind a 3' single-stranded overhang that serves as a catalyst for the cell's HR pathway [33]. Our group, using iterative directed evolution on the I-CreI active site to confer high specificity for a DNA sequence of interest, has generated enzymes termed ARCUS nucleases for a variety of agricultural and human health applications [35–38]. We previously reported efficient integration of a chimeric antigen receptor (CAR) gene into the T-cell receptor α chain constant region (*TRAC*) locus of T lymphocytes using nucleofected ARCUS messenger RNA (mRNA) and AAV6-delivered repair template encoding the CAR sequence [39].

Here, we investigate the extent and manner by which ARCUS achieves transgene integration. First, we show that our optimized ARCUS nuclease targeting *TRAC* enables insertion of an AAV6-delivered DNA template with high frequency (>80% of primary human cells) at low multiplicities of infection (MOI; as low as 5000 v.g./cell). We then demonstrate that, like its evolutionary ancestor I-CreI, ARCUS nucleases enable cells to use HR with high efficiency, rather than nonhomologous end-joining (NHEJ) or microhomology-mediated end-joining (MMEJ), with contributions being made by the 3' 4 nt overhang, the vector homology, and DDR machinery. As a result, this platform enables a wide variety of genetic changes, including all base-pair changes and large-sequence insertions, either proximal to the cut-site or kilobases distal, because all of these changes can be achieved through HR-mediated integration of defined recombinant DNA templates. Finally, we demonstrate that this phenomenon is not limited to actively-dividing T lymphocytes, but high-frequency targeted insertion (30%–40%) can also be seen in nondividing primary human hepatocytes (PHHs) likely via synthesis-dependent strand annealing (SDSA), a nonclassical HR pathway [40–42]. As in lymphocytes, the staggered DSB, vector homology, and DDR machinery were shown to be critical determinants of high insertion efficiency. Taken together, these data show that the characteristic enzymology of this class of endonucleases can unlock the potential for ideal repair promised by HR for human health and disease.

Materials and methods

Cell culture and addition of DNA repair inhibitors

Apheresis products were purchased from StemExpress (Folsom, CA) and human T lymphocytes were enriched using CD3 Positive Selection reagents from Stem Cell Technologies (Vancouver, Canada). T cells were cultured in Xuri medium (GE) supplemented with 10% fetal bovine serum (GeminiBio, Sacramento, CA) and 10 ng/ml interleukin-2 (IL-2) (Gibco, Waltham, MA). To prepare for editing, T cells were stimulated with ImmunoCult anti-CD2/3/28 according to manufacturer's instructions (StemCell Technologies, 10991). All T-cell culture media were supplemented with 1% antibiotic–antimycotic (Gibco, 15240012). PHHs were sourced from BioIVT (Westbury, NY) and were cultured on Corning BioCoat collagen plates (Cat# 326408, Corning, NY) INVITROGRO CP medium (BioIVT, Z99209) supplemented with TORPEDO antibiotics (BioIVT, Z99000) or Cellartis Power Hepatocyte medium (Takara Bio, San Jose, CA, Cat# Y20020). For inhibitor experiments, select DNA repair inhibitors were added to the primary cell cultures

concurrent with editing. The following inhibitors and concentration ranges were used: trifluorothymidine, TFT (0.01–45 μ M, Fisher Scientific, Waltham, MA, Cat# AC452912500), B02 (0.01–50 μ M, SelleckChem, Houston, TX, Cat# S8434), NU7026 (10 nM–90 μ M, SelleckChem, Cat# S2893), Rucaparib camsylate (0.1–40 μ M, Sigma, Allentown, PA, Cat# PZ0036), 6-hydroxy dopamine (1–200 μ M, MedChemExpress Cat#HY-110286 Monmouth Junction, NJ), TDRL-551 (1–1000 μ M, MedChemExpress, Cat# HY-114842), and zelpolib (1–100 μ M, Aobious Inc, Gloucester, MA, Cat#AOB13830). Reported results were obtained using the maximum tolerated dose. All inhibitors were resuspended and aliquoted per manufacturer's instructions.

RNA preparation

The ARCUS and ARCUS nuclease fusion proteins were cloned downstream of a T7 promoter. Linearized plasmids were used as template for *in vitro* transcription (IVT) using a HiScribe T7 kit (New England Biolabs, E2040S). A co-transcriptional capping RNA synthesis procedure was performed according to the manufacturer's protocol with the following exceptions. Uridine was substituted completely with N1-methylpseudouridine (m1 Ψ , Hongene Biotech Union City, CA, Cat# R5-207). CleanCap Reagent AG (Trilink, San Diego, CA) was included in the synthesis. After IVT, the reactions were treated with TURBO DNase (ThermoFisher, AM2238). Post DNase treatment, the RNA was column purified with the SV Total RNA Isolation System (Promega, Madison, WI, Cat# Z3105) normalized to 1.0 μ g/ μ l, aliquoted, and stored at –80°C until ready for use. Cas9 RNA was purchased from TriLink (Cat# L-7206) and guide RNA was purchased from Integrated DNA Technologies.

Encapsulation of RNA into lipid nanoparticles

Lipid nanoparticles (LNPs) were produced using a Benchtop NanoAssemble (Precision Nanosystems, Vancouver, Canada) with Benchtop cartridges. A prewarmed ethanolic lipid mixture consisting of 40:48.5:10:1.5 parts of D-Lin-MC3-DMA (MedKoo Biosciences, Durham, NC), cholesterol (Spectrum Chemical, New Brunswick, NJ), DSPC (NOF America Corporation, Boston MA), and DMG-PEG2k (Avanti Polar Lipids, Alabaster, AL) were mixed with mRNA in citrate buffer (50 mM, pH 4) at an N:P ratio of 8, flow rate ratio of 3:1 (mRNA:lipid mixture), and total flow rate of 12 ml/min. LNPs were dialyzed overnight in PBS (1 \times , pH 7.2) at room temperature before being concentrated using an Amicon spin filter (Ultra 0.5 ml, Ultracel 30K) to >200 ng/ μ l at 4°C. Size and dispersity were determined by diluting 1:500 in phosphate-buffered saline (PBS) (1 \times , pH 7.2) and characterized by dynamic light scattering using a Zetasizer Nano ZSP (Malvern Instruments, Malvern UK). Encapsulation efficiency and concentration was determined using the Quant-IT Ribogreen RNA Assay Kit (Invitrogen, Eugene, OR).

AAV generation

Reporter constructs expressing green fluorescent protein (GFP) under control of the JeT promoter [39] were cloned between left and right arms sharing homology with genomic regions in the *TRAC*, *TGFBR2*, or *HAO1* loci adjacent to nuclease cleavage sites or where indicated. Construct for *HAO1* insertion analysis in PHH was cloned to contain 550 bp of noncoding DNA flanked by 300 bp left and right arms shar-

ing homology with the genomic regions in the *HAO1* locus. Seamless insertion construct targeting *HAO1* used for insertion analysis in PHH was clone to contain an in frame P2A-furin GFP and BGH polyA tail sequence flanked by 500 bp left homology arm (LHA) and right homology arm (RHA). Lambda DNA stuffer was added outside of the RHA to ensure proper AAV packaging. Addition reporter constructs developed for homology arm (HA) variant testing replaced the HA sequence with stuffer DNA to maintain AAV size. Reporter constructs were cloned into a vector containing AAV inverted terminal repeats (ITRs) and the plasmids were transfected along with AAV6 RepCap and Helper plasmids into HEK293-F cells using Jet-PEI. After 3 days, nuclear fractions were collected and AAV6 particles were purified by CsCl centrifugation and column chromatography. Titration was performed via digital polymerase chain reaction (PCR) after a digest of nonencapsulated DNA.

Gene editing and viral transduction

T cells were edited as described previously [39]. Samples of 1×10^6 cells received 500 ng of RNA encoding a *TRAC* or *TGFBR2*-specific ARCUS nuclease, or 5000 ng of RNA encoding *Streptococcus pyogenes* Cas9 and a *TRAC*-specific gRNA at a 1:4.2 ratio. Immediately upon plating in serum-free medium containing any inhibitors, cells were transduced with AAV6 preparations at the indicated MOI. Following an overnight incubation, cultures were fed with complete medium. ARCUS nucleases were generated as previously described [36, 38, 39, 43]. A table detailing the amino acid sequences of ARCUS nucleases versus I-CreI appears in [Supplementary Table S3](#) while each enzyme's 22 bp recognition sequence appears in [Supplementary Table S4](#). PHHs were edited using LNP doses containing 25–500 ng of RNA (concentrations selected following a titration to optimize efficacy and tolerability).

Flow cytometry

At 7 days post electroporation, single-cell suspensions were labeled with antibodies against human CD3 (BioLegend, San Diego, CA, Cat# 344818), TGFBR2 (R&D Systems, Minneapolis, MN, Cat# AF-241-NA), or B2M (Becton-Dickinson, Franklin Lakes, NJ, Cat# 551337). Anti-TGFBR2 was biotinylated in-house using EZ-Link Sulfo NHS Biotin (ThermoFisher, 21335), and streptavidin-PE was used as a secondary staining reagent (BD Biosciences, 551337). Data were acquired on a Beckman-Coulter CytoFLEX-S and analyzed using CytExpert software (Beckman-Coulter, Brea, CA) or FlowJo software (FlowJo, Ashland, OR). For KI-67 staining, primary human T cells stimulated for 3 days with anti-CD3/CD28/CD2 (ImmunoCult, Stemcell technologies, #10971), AML14 myeloid leukemia cells, and PHHs were isolated from culture and spun in a 96-well round bottom plate for 8 min at 400 \times g. Supernatant was decanted and cell pellets were washed with PBS and respun and decanted. Cell pellets were fixed and permeabilized using the FOXP3/Transcription Factor Staining Buffer Set (eBioscience, #00-5523-00), and stained using PE anti-human KI67 (BioLegend, #350504) according to manufacturer recommended protocols. KI67 expression was measured by flow cytometry using a CytoFlex S and further visualized using FlowJo compared to unstained controls or noncycling controls.

Droplet digital polymerase chain reaction

PHH monolayers were dissociated from the 24-well collagen treated plate by incubation with TrypLE Express solution (Gibco, 12604021) and spun in a microfuge for 7 min at $300 \times g$. Cell pellets were washed with PBS pH 7.4 (Gibco, 10010-023) and respun for 7 min at $300 \times g$. Genomic DNA was isolated from pelleted PHHs after RNase cocktail (ThermoFisher Scientific, AM2286) and proteinase K treatment according to the Macherey-Nagel NucleoSpin Blood Quick-Pure (Macherey-Nagel, Duren, Germany, 740569.250) manufacturer protocol. DNA concentration was measured using a Qubit (ThermoFisher Scientific, Waltham, MA). Quantification of total gene editing (INDELs) and insertion events at the *HAO1* locus in PHHs were measured by duplex digital droplet PCR. For INDEL analysis, the assay amplifies across the *HAO1* ARCUS cut site and indicates the presence of wild-type un-edited DNA. For insertion analysis, the assay amplifies from inside the inserted repair template into the genomic DNA which will only generate a product if an insertion event is present. A second reference assay amplifying a region >50 kb away from the *HAO1* ARCUS cut site is also included in the reactions for normalization purposes. PCR amplifications were duplexed to contain the target and reference primers and probes sets. Twenty-four microliters of reactions contained $1 \times$ ddPCR Supermix no dUTP (Bio-Rad, Philadelphia, PA, Item 1863024), 250 nM probes, 900 nM primers, 20 U/ μ l HinDIII-HF (New England Biolabs, R3104L), and 200 ng genomic DNA. Droplets were generated using the Bio-Rad Automated Droplet Generator and PCR amplification performed using the Bio-Rad C1000 Touch Thermal Cycler. PCR cycling conditions for INDELs were as follows: a 10-min incubation at 95°C , 45 cycles of 95°C for 30 s, 62°C for 30 s, 72°C for 2 min, and one cycle at 98°C for 10 min concluding with a 4°C hold. PCR cycling conditions for insertions were as follows: a 10-min incubation at 95°C , 45 cycles of 95°C for 30 s, 59°C for 30 s, 72°C for 3 min, and one cycle at 98°C for 10 min concluding with a 4°C hold. Samples were run on the Bio-Rad QX 200 Droplet Reader and analyzed using the Bio-Rad QX manager software. Total gene editing events (INDELs) are measured as (reference copies/ μ l–INDELs copies/ μ l/reference copies/ μ l) while insertion events are measured as (insertion copies/ μ l/reference copies/ μ l).

Long-amplicon sequencing

Genomic DNA (gDNA) was isolated from T cells using the Nucleospin Blood Quick Spin Kit (Macherey-Nagel) according to the manufacturer's instructions. The genomic DNA was quantified using the Nanodrop One (Thermo Scientific). PCR was conducted using primers amplifying the region outside of the ITRs. PCR amplifications were multiplexed in a 50 μ l reaction containing $2 \times$ repliQa HiFi Toughmix (QuantaBio, 95200-500), 0.3 μ M of each primer, and ~ 100 ng of gDNA. PCR was performed on a 2720 Thermal Cycler (Applied Biosystems, Waltham, MA). Cycling conditions were as follows: 1 cycle of 98°C for 2 min, 35 cycles of 98°C for 10 s, and 68°C for 10 s, 4°C hold.

For PHH insertion analysis by nanopore sequencing, DNA was isolated as described above. PCR was conducted using genomic primers amplifying a 2659 bp region across the *HAO1* cut site. PCR amplifications were carried out in a 50 μ l reaction containing $2 \times$ repliQa HiFi Toughmix (QuantaBio, 95200-500), 0.3 μ M of each primer, and ~ 100 ng of gDNA.

PCR was performed on a 2720 Thermal Cycler (Applied Biosystems, Waltham, MA). Cycling conditions were as follows: 1 cycle of 98°C for 2 min, 35 cycles of 98°C for 10 s, 61°C for 5 s, and 68°C for 30 s, 4°C hold.

Resultant amplicons were prepared for sequencing using the Native Barcoding Kit 24 V14 (SQK-NBD114.24) from Oxford Nanopore Technologies per manufacturer's instructions. Samples were sequenced using R10.4.1 flow cells on a MinION or PromethION device with MinKNOW v23.04.6, v23.11.7, or v24.11.8 and super-accuracy base-calling mode selected.

Long-amplicon nanopore data analysis

Guppy v6.5.7 or Dorado v7.2.13 or v7.6.7 was used to extract the bases from pod5 or fast5 data and turn them into standard FASTQ files. Resultant FASTQ files were analyzed using a custom Nextflow pipeline. Reads were trimmed with porechop (v0.2.4) to remove sequencing adapters and filtered to keep full-length reads containing both the forward and reverse primer sequences. Fused reads were identified and split using pyclop (v2.7.10). All reads were aligned with ngmlr (v0.2.7) to the full amplicon sequence with desired insert or replacement. Reads with the desired insert or replacement, were checked for ITR sequence. Reads that did not contain the insert sequence were subsequently mapped to the genomic amplicon sequence.

Sam files for both "insert" (reads with desired insert or replacement) and "no insert" (reads without insert, mapped to genomic amplicon) were loaded into R statistical software for further analysis. Alignments were analyzed by parsing CIGAR strings using GenomicAlignments (v1.38.0) to determine the match length, location and size of indels (insertions/deletions), and number of alignments per read. Nanopore LSK114 chemistry has a background indel rate of 3% for Q15 reads (97% accuracy). A high number of small indels (1–3 bp) were observed randomly throughout the amplicon, not focused just at the cut-site, and were presumed to be background noise of Nanopore sequencing and not as a result of ARCUS editing. To avoid inflating false NHEJ edits due to the overlap of these artificial 1–3 bp deletions near the cut-site, only indels within 10 bp of the cut-sites with a size \geq the 99th-percentile of the background indel sizes (≥ 4 bp) were included in further analysis (smaller indels were ignored and considered a sequence match).

Short-amplicon sequencing

Genomic DNA was isolated from T cells using the Nucleospin Blood Quick Spin Kit (Macherey-Nagel) according to the manufacturer's instructions. The genomic DNA was quantified using the Nanodrop One (Thermo Scientific). A two-step PCR reaction was conducted: the first using a forward primer amplifying a region in the genomic DNA, and reverse primer amplifying a region in the insert to eliminate amplification of free-floating AAV. After gel purification on a 1.5% agarose gel, DNA was extracted using the Monarch PCR & DNA Cleanup Kit (NEB, Ipswich, MA, T1030L), following the manufacturer's instructions. A second PCR was conducted using nested primers which amplify a ~ 200 bp sequence amenable for short-range sequencing. PCR was performed on a 2720 Thermal Cycler (Applied Biosystems). Cycling conditions were as follows: 1 cycle of 98°C for 2 min, 35 cycles of 98°C for 10 s, 68°C for 10 s, 4°C hold.

Resultant amplicons were prepared for sequencing using the NEBNext Ultra II Library Prep Kit for Illumina (NEB) per manufacturer's instructions. Samples were sequenced using a 2 × 150 bp paired-end run on the NextSeq 2000 (Illumina).

Short-amplicon data analysis

To analyze sequencing data, paired-end reads were first merged with flash (v1.2.11) with a max overlap of 195 bp. Merged reads were aligned to the target amplicon sequences using bwa mem (v0.7.18) and reads per amplicon were counted.

Data analysis and statistics

Data were plotted and hypothesis tests were performed using GraphPad Prism software.

Results

DSB made with ARCUS nucleases support high rates of targeted integration via homology directed repair in dividing cells

Primary human T cells can repair DNA damage using a variety of mechanisms, and are therefore a useful model to explore targeted insertion strategies (reviewed in [44]). Our group previously showed that electroporating T cells with an optimized amount of mRNA encoding a highly specific TRAC-targeting ARCUS nuclease led to efficient disruption of the TRAC gene and insertion of cargo [39] and show representative data in [Supplementary Fig. S1](#). There are several factors which likely contribute to high insertion efficiencies, including: the presence of HAs to the regions flanking the cut-site, the nature of the DSB generated by ARCUS, and/or the relative contribution of cellular DNA repair pathways. We therefore set about to examine each of these factors individually.

We first evaluated the role of HAs by building AAV vectors configured as pictured in [Fig. 1A](#). We compared insertion of AAV6 repair templates that comprised: 1100 bp HAs on both sides of the insert ([Fig. 1B](#)), no HA on either side ([Fig. 1C](#)), a single 1100 bp LHA ([Fig. 1D](#)), a single 1100 bp RHA ([Fig. 1E](#)), or microhomology arms corresponding to the 13 bp halves of the ARCUS-binding site on either side of the insert ([Fig. 1F](#)). We selected 1100 bp as our HA length due to precedent established elsewhere and because this arm length brings the size of the transgene to 3.2 kb, which is in the optimal range for AAV6 packaging [8, 16]. To avoid size discrepancies among the HA variants, noncoding lambda phage DNA was included in place of an HA where indicated. As expected for a recombinogenic DNA repair template, insertion was highly dependent on both HAs being present; loss of both arms reduced integration by 90%, and loss of either arm individually reduced integration by ~50% ([Figs. 1A–D](#)). Loss of HAs was accompanied by a concomitant loss in GFP median fluorescence intensity (MFI), reflecting reduced transgene expression, though the construct with a single RHA displayed significantly higher MFI than the LHA or no-HA vectors ([Fig. 1G](#)). Finally, because the 3' overhangs generated by ARCUS enzymes are short, we also tested whether insertion required full length HAs, or could be achieved using microhomology arms corresponding to the 13 bp halves of the ARCUS-binding site. The repair template containing microhomology arms was indistinguishable from one without any HAs in both their low integration efficiency and GFP MFI ([Fig. 1C and F](#)). Taken together,

one HA appears to be sufficient for some amount of insertion, whereas repair vectors that are nonhomologous or only contain regions of microhomology do not integrate efficiently ([Fig. 1H](#)). These observations suggest that HR mediates template insertion when a sufficient HA is present, while nonhomologous end joining (NHEJ), or spontaneous ITR capture [45] mediates integration otherwise. Sequencing the insert-bearing alleles largely confirmed these observations, indicating that scarless integration was occurring in the presence of two HAs (i.e. only rare detection of extraneous indels at the cut site, integrated ITR sequences detected, or HA duplications), while no significant insertion of an armless vector was detected ([Fig. 1I](#)). Depictions of the six most common outcomes are shown in [Supplementary Fig. S2](#). Vectors with only one HA were found to integrate at significantly lower rates and via HDR on the side with a HA, but repair occurred via endjoining on the nonhomologous side. The higher MFI observed for the RHA only versus the LHA only could potentially be due to error-free repair on the HA side preserving the end of the GFP gene and poly-adenine signal which are more important for gene function and mRNA stability versus potential truncations into the promoter from NHEJ DNA capture which would likely still be partially functional. Importantly, the absence of extraneous indels in the presence of two HAs is suggestive of repair by homology directed repair (HDR), the prototypical HR pathway in dividing cells.

We next investigated the contributions made by the DNA DSB generated by ARCUS relative to other endonucleases. As described previously [27], ARCUS nucleases create a staggered DSB with a 4 nt, 3' overhang ([Fig. 2A](#)).

Unlike blunt-ended cuts or those with 5' overhangs, we reasoned that a 3' single strand overhang could more directly initiate the process of Rad51-mediated strand invasion and culminate in insertion into the target location by primed DNA extension with polymerase. We hypothesized that removing the 3' overhangs and creating a blunt-ended ARCUS DSB would disrupt this potential mechanism of insertion and alter the likelihood that HR is leveraged to integrate the GFP transgene into the DSB, as described in other systems [46]. To achieve this, we fused the ARCUS nuclease to a TREX1 3'-to-5' exonuclease (ARCUS-TREX) domain using a Gly-Ser linker. TREX is reported to reduce HDR efficiency in *Drosophila*, mice, and humans [47]. NGS analysis on human T lymphocytes electroporated with an optimized dose of ARCUS-TREX mRNA supported this reasoning, as most indels were small deletions with a median size of 4 bp ([Supplementary Fig. S3A](#)). The most commonly detected sequences lacked the 4 overhanging nucleotides normally generated by ARCUS, suggesting the ARCUS-TREX-generated DSB was being directly end-joined by NHEJ. Cas9 endonucleases used in CRISPR systems also leave behind blunt-ended DSBs. To add additional support to the findings, we also used a TRAC-targeting single-guide RNA (sgRNA) and an SpCas9 nuclease to target and cleave an identical site in the TRAC locus ([Fig. 2C](#)). Therefore, the genetic context of the repair template and the genomic locus would not complicate comparison because they were the same for all three cuts, despite the fact that the gRNA-Cas9 pair was selected due to position and not due to efficiency or specificity of nuclease activity at this site.

We tested the effect of these three different DSBs on integration of our fully recombinogenic (i.e. containing both 1100 bp left and right HAs) GFP repair template. We titrated the amount of AAV added into each nuclease system (ARCUS,

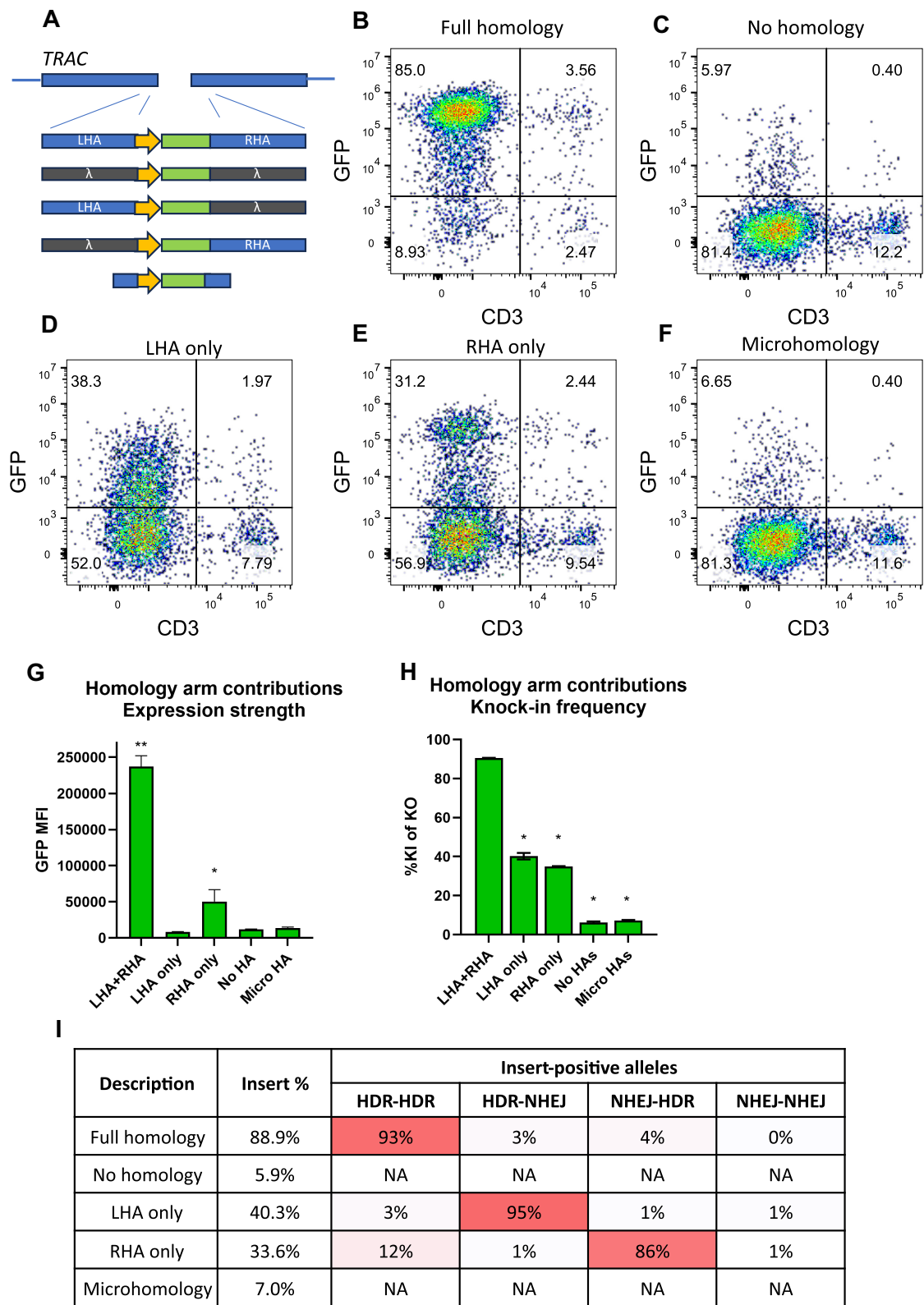


Figure 1. HAs on both sides of the DNA insert are required for both optimal transgene expression and efficient gene insertion. *TRAC* loci were edited in stimulated human T cells and repaired using templates with varying HAs. **(A)** Diagram of repair templates used. Cells were edited and transduced with either a standard GFP repair template containing **(B)** 1100 bp HAs on both sides of the DNA insert, **(C)** constructs with no HAs but λ phage DNA on either side of GFP insert, **(D)** only the LHA plus λ phage DNA, **(E)** only the RHA plus λ phage DNA, or **(F)** short (13 bp) HAs that encompass half of the ARCUS-binding site (microhomology). Bar graphs of this data are displayed as **(G)** intensity of GFP expression and **(H)** transgene insertion, as measured by % GFP KI of total CD3 KO. **(I)** Repair mechanisms resulting in GFP cassette insertion. Experiments are representative of three biological replicates each containing three technical replicates using two different human donors. * $P < 0.05$, ** $P < 0.01$ as determined using ANOVA and a Tukey-Kramer multiple comparisons correction, $n = 3$. Cells received 500 ng/1e⁶ of ARCUS mRNA and AAV at an MOI of 4.0e⁴.

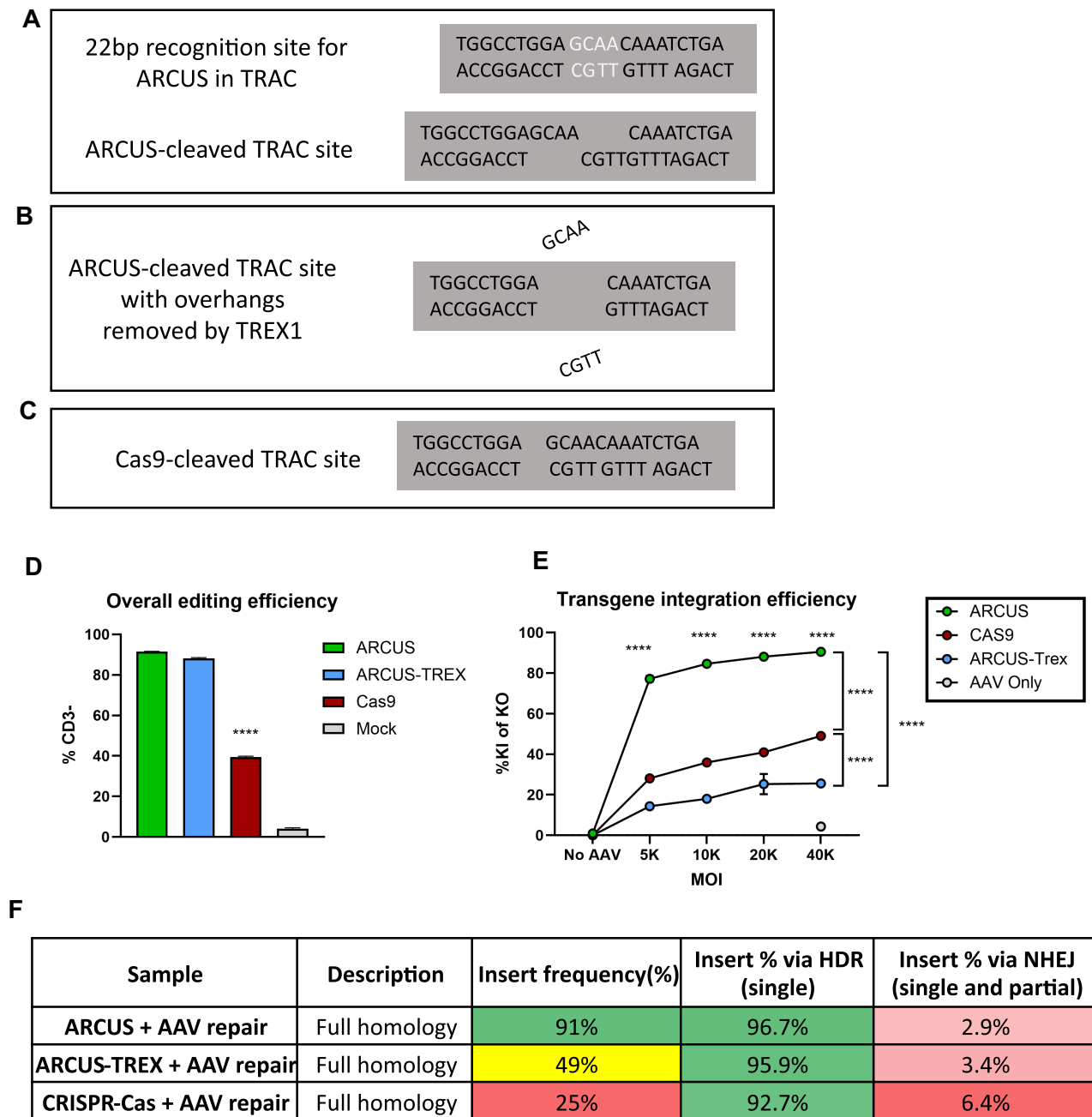


Figure 2. Double-stranded breaks made with ARCUS support high rates of insertion via HDR. In activated human T cells, TRAC loci were edited with various endonucleases and repaired with a common HDR template. The enzyme recognition sites and resulting DSBs are shown in panels (A–C). Cells were electroporated with either ARCUS mRNA (500 ng/1e6 cells), ARCUS–TREX fusion mRNA (500 ng/1e6 cells), or Cas9 mRNA (1 µg/1e6 cells) plus sgRNA for TRAC (5 µg/1e6 cells) and were transduced with the HDR template across a range of MOIs. Overall editing rates appear in panel (D) while insertion rates are plotted in relation to MOI in panel (E) as the percentage of knockouts (KO) that are also knockins (KI). Nanopore long-read analysis was performed across the inserts and percentages of reads detecting HDR or some combination of HDR/NHEJ appear in panel (F). Experiments are representative of three biological replicates using cells from different human donors. Significance was determined using ANOVA and Tukey–Kramer multiple comparisons correction; **** $P < 0.0001$, $n = 3$.

ARCUS–TREX, and Cas9) in order to capture their respective sensitivity to repair template. The CRISPR sgRNA was chosen for comparability of location and repair over optimal editing efficiency, and Cas9 drove editing at roughly half the efficiency of either ARCUS or ARCUS–TREX (Fig. 3D). By normalizing insert rates to overall editing efficiency, we account for differences in nuclease activity at the target site, so we can focus on the contribution that the different types of DSB make towards insertion efficiency. Of those cells that were successfully edited

in all three systems (i.e. CD3 KO), those with an ARCUS-generated DSB most efficiently integrated the AAV6-delivered repair template, with the fraction of GFP⁺ (i.e. knock-in or KI) cells ranging from 75% of edited cells at an MOI of 5000 vg/cell to 90% at an MOI of 40 000 vg/cell (Fig. 2E). In contrast, while the ARCUS–TREX fusion protein yielded a similar frequency of CD3 KO (Fig. 2D), the loss of the 4 nt, 3' overhangs greatly reduced insertion efficiency, with only 15%–25% of edited cells expressing GFP over the same range of

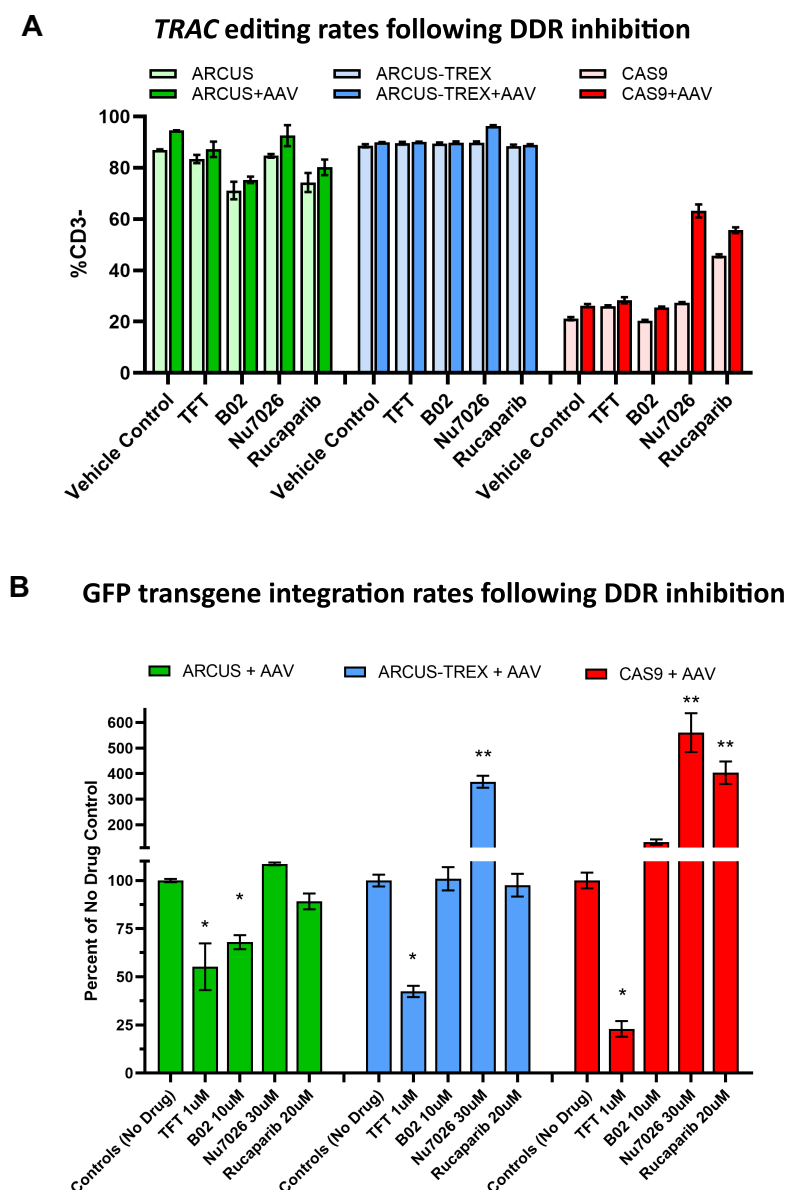


Figure 3. ARCUS lesions are repaired primarily by HDR. The effects of small molecule inhibitors of DNA damage response (DDR) pathways on editing of *TRAC* alleles (**A**) and insertion of GFP into DSBs (**B**) in primary human T cells are displayed. Cells were electroporated with mRNA encoding ARCUS, ARCUS fused to TREX1, or Cas9 plus a sgRNA followed by transduction with a GFP encoding AAV repair template at an MOI of 5000. Cells were treated with TFT (1 μ M), B02 (10 μ M), NU7026 (30 μ M), or Rucaparib (20 μ M) for 24 h. Seven days after editing, cells were evaluated for CD3 and GFP frequencies by flow cytometry. (A) Quantification of editing frequency as measured by CD3 loss in the presence or absence of AAV repair template. (B) Quantification of insertion as measured by flow cytometry and plotted as knock-in (GFP+) normalized to the no drug control (i.e. <100% indicates inhibition of integration and >100% indicates enhancement relative to the no drug control normalized to 100%). Significance was determined using ANOVA and a Dunnett's multiple comparisons correction, $n = 3$, * $P < 0.02$, ** $P < 0.001$. Data are representative of three independent experiments using cells from two human donors

MOI (Fig. 2E). Similarly, our *TRAC* sgRNA–Cas9 system also generated a lower overall frequency of GFP inserted, edited cells, ranging from 25% to 50% over the same range of MOI (Fig. 2D and E). A recent report demonstrated that TREX exonuclease activity could be directed against exogenous single stranded DNA (ssDNA) repair templates, including those delivered by AAV6 [48]. We assessed AAV6 copy number using ddPCR and found that at 24 h following editing and transduction, AAV copy number in ARCUS–TREX samples was not reduced (Supplementary Fig. S3B), suggesting that HDR inhibition by TREX is most likely due to alterations made to the genomic DSB.

We conclude that the 3' overhangs generated by ARCUS are critical to high-frequency transgene integration and that they catalyze HDR more efficiently than a blunt-ended DSB such as generated by Cas9. However, because T lymphocytes are efficient at repairing DNA damage via HR as well as other mechanisms, we sought to confirm that the GFP+ integrants after ARCUS, ARCUS–TREX, and Cas9 editing were derived from HR. Nanopore sequencing analysis of insert-bearing alleles confirmed that all three of these editing methods can drive DSB repair without extraneous indels (Fig. 2F). Sequencing analysis of GFP+ cells showed that nontemplated insertions or deletions at the junctions between the repair tem-

plate and genomic sequences occurred rarely in GFP⁺ alleles (2.9%–6.4%), consistent with homology-directed repair. In summary, repair of staggered DSBs resulted in transgene insertion more frequently than blunt DSBs (Fig. 2E). When insertion occurred, however, it appears to take place via HDR irrespective of the features of the DSB (Fig. 2F).

Finally, we assessed the relative importance of different DNA damage repair pathways using pharmacological inhibitors. Specifically, cells were exposed to inhibitors of HDR (TFT or B02), NHEJ (NU7026), or MMEJ (Rucaparib) simultaneous to introduction of nucleases and repair templates. All drugs were titrated to identify the maximum tolerated dose (>75% cell recovery) and these concentrations are indicated in Fig. 3, in which the frequency of overall TRAC editing (panel A) and the frequency of GFP insertion (Fig. 3B) are also shown. Insertion data in panel B are expressed as the percent of edited alleles carrying an insert (CD3⁺GFP⁺/total CD3⁺) and then normalized to the vehicle only control. Percentages below 100 indicate reduced integration frequency relative to control while percentages >100 indicate an enhancement of integration relative to control. DNA repair inhibitors had no consistent impact on either ARCUS or ARCUS–TREX overall editing rates between replicates, while inhibition of NHEJ or MMEJ nearly doubled the Cas9 insertion rate (Fig. 3A). Notably, NHEJ inhibition increased Cas9 overall editing rates only in the presence of a repair template. Integration was reduced in the presence of the thymidine analogue TFT, irrespective of the nuclease used to introduce the DSB, supporting the conclusion that transgene integration is occurring primarily by HDR under all editing conditions (Fig. 3B). In contrast, treatment with B02, a Rad51 inhibitor, significantly reduced the frequency of integration into ARCUS DSBs, but had no effect on integration into blunt DSBs. NU7026 measurably increased the integration rate in ARCUS–TREX and Cas9 DSBs but in contrast, had no effect on integration into ARCUS DSBs. This is consistent with the notion that blunt-ended DSBs are preferentially repaired via end-joining mechanisms; when this pathway is inhibited, the cells may permit an HDR-mediated repair process more frequently [5, 44, 49]. The lack of effect of an NHEJ-inhibitor on integration efficiency supports HDR as the primary method of repair for ARCUS DSBs. Finally, Rucaparib treatment resulted in a slight but measurable decrease in integration into ARCUS DSBs, but increased the efficiency of integration into Cas9 DSBs, indicating microhomology-driven mechanisms may contribute to integration into ARCUS DSBs but confound integration into blunt-ended DSBs (Fig. 3B). Representative dot plots are shown in Supplementary Fig. S4.

Perfect homology between the template and genome is not required for high-frequency insertion

Previous reports and our own unpublished observations have demonstrated that HA length contributes to integration efficiency, suggesting that a certain extent of base pairing is required during the process of strand invasion or resolution [15, 50]. We postulated that HAs of sufficient length may be able to tolerate mismatches while still supporting efficient integration. To test this hypothesis, we designed GFP reporter inserts at the TRAC locus with 1, 3, 5, or 10 bp contiguous mismatches in either the RHA or LHA, positioned 50 bp away from the DSB

site to provide sufficient homology to still support HDR [51] (Fig. 4A).

Using flow cytometry, we observed no reduction in integration frequency when mismatches were engineered into the repair vector HAs (Fig. 4B). We performed Illumina sequencing and analysis on these samples to determine the extent of mismatch incorporation into the genomic DNA. In general, increasing the number of mismatches decreased the likelihood of their integration into the genome, with 1 bp mismatches being integrated in 65%–80% of insert-positive alleles and 10 bp mismatches incorporated in roughly 30% of insert-positive alleles, regardless of the HA (left or right) that carried the mismatches (Fig. 4C). Sequencing analysis also revealed that insert-positive alleles aligned perfectly to either the genomic sequence or to the mismatched HA sequence; either the whole set of mismatches were incorporated, or none of them were. These data highlight that perfect homology between the template and genome is not required for effective transgene integration and provides a potential route to modify bases further afield from the recognition site without having to produce a large-scale insertion.

Interestingly, these data argue that HDR via imperfectly homologous templates can achieve gene editing outcomes similar to base editors. To demonstrate this, we designed an ARCUS nuclease with specificity to the *TGFBR2* locus at a site in the 5' UTR of exon 1, upstream of the translation initiation site (Fig. 5A).

We incorporated a mismatch in the RHA so that integration would result in an A to G substitution in the genomic sequence, thereby changing the ATG start codon to GTG and ablating initiation of protein synthesis. Treatment of cells with *TGFBR2* ARCUS mRNA alone slightly reduced TGFBR2 expression on 25% of cells (Fig. 5B, second panel) presumably due to indels produced in the 5'UTR, which has been shown to govern the translation levels and production of alternative isoforms [52]. ARCUS mRNA plus ATG repair vector AAV had minimal effect on TGFBR2 surface display with only 11% of events falling into the TGFBR2⁺ region (Fig. 5B, third panel). There was no appreciable knockout of TGFBR2 when GTG repair was delivered without ARCUS (2%). In contrast, ARCUS mRNA plus GTG repair produced integration of the mismatched GTG template into the *TGFBR2* DSB resulting in loss of receptor expression in 65% of cells, suggesting a high rate of incorporation of the mismatched G into the start codon 214 bp downstream of the cut.

Since modifications are being produced by HDR off a template, this suggested it could also be possible to make small insertions and indels similar to prime editors. To test this, we cloned vectors with a RHA featuring a 3 bp deletion (removal of the ATG start codon) or a 3 bp insertion (insertion of the TGA stop codon immediately following the ATG). Compared to vector alone controls, we observed insertion into the *TGFBR2* locus and ablation of receptor display on 50% of cells for 3 bp insertion and 52% of cells for 3 bp deletion (Fig. 5B). These data highlight the ability to leverage staggered ARCUS DSBs to introduce small, specific changes, deletions, and insertions via HDR at a considerable distance from the DSB.

ARCUS enables high-frequency base editing

Presently, adenine and cytosine base editors are limited in the bases that can be targeted for substitution (A and C), as well as the bases to which they can be changed. With the large num-

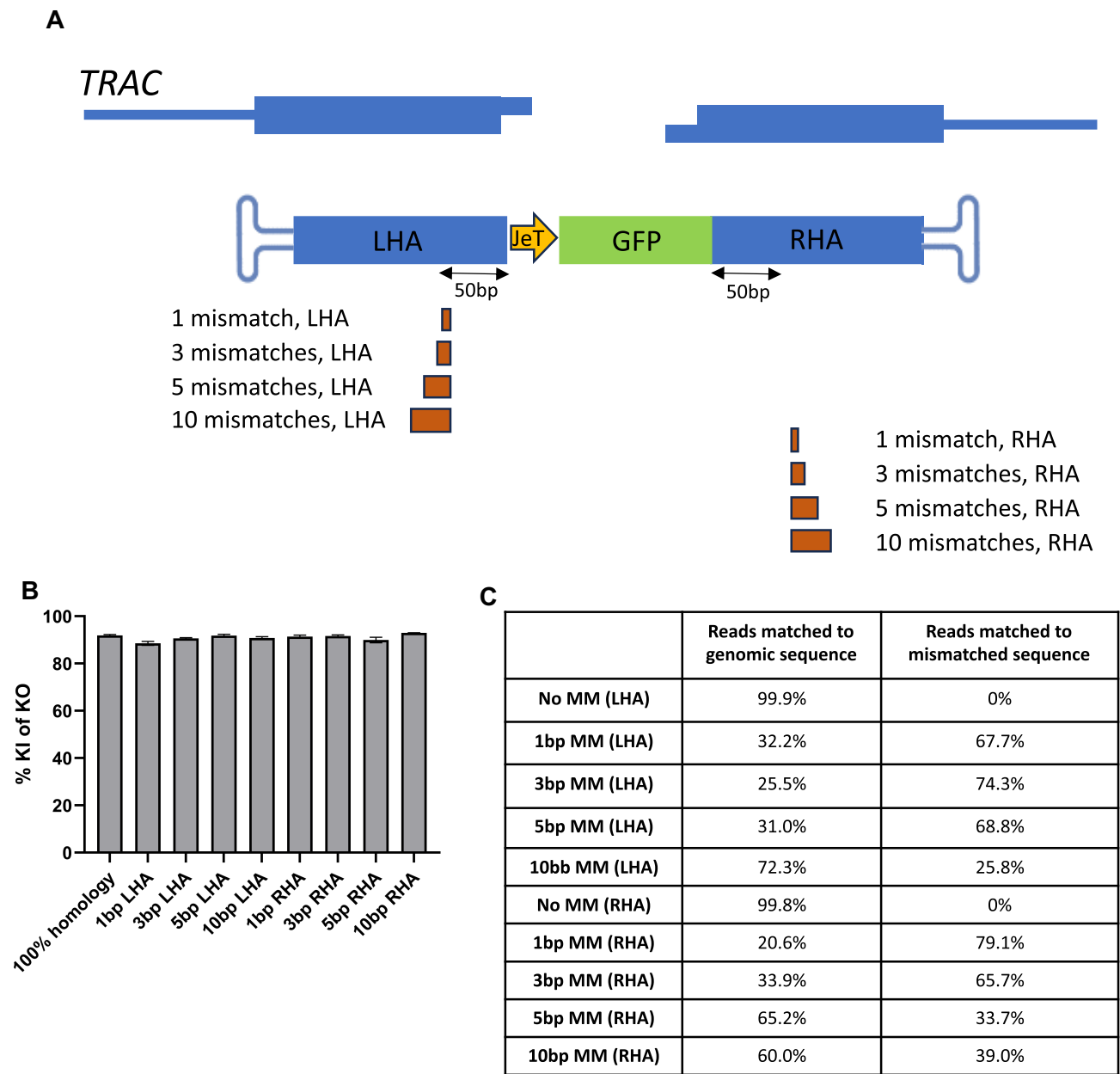


Figure 4. Perfect homology between the template and genome is not required for integration. Activated human T cells were electroporated with mRNA (500 ng/1e⁶ cells) encoding a *TRAC*-specific ARCUS nuclease and were transduced with AAVs encoding 1100 bp HAs flanking the ARCUS recognition site at an MOI of 4.0e⁴. **(A)** Schematic detailing the location of consecutive mismatches (MM) introduced into the *TRAC* genomic locus, in either the LHA or RHA. **(B)** Flow cytometry data showing the overall knock in of knockout percentage. **(C)** Table of next generation sequencing results showing the percentage of insert⁺ alleles in which alignment to mismatched sequences was detected.

ber of genetic disorders caused by single point mutations—not all of which are able to be treated with A and C base substitutions—there is a need for additional technologies that can fill this gap. Because ARCUS can support high rates of HDR-mediated repair and insertion without requiring perfect homology between the template and the genome, we hypothesized that this tool could be used to introduce all combinations of base changes into the genome. Therefore, we designed a library of four AAV6 vectors with imperfect homology at the *TGFBR2* locus by incorporating degeneracy at the ATG site as well as to an adjacent C 213 bp downstream of the ARCUS cut (Fig. 6A). We evaluated the frequency of all four bases at the indicated positions present in the plas-

mid library, the AAV6 product, and genomic DNA of cells treated with the *TGFBR2* ARCUS nuclease and the degenerate AAVs. We found that all four bases were present in the plasmid and AAV at roughly equal frequencies, indicating that cloning and packaging was not a source of meaningful amounts of bias (Supplementary Fig. S5A and B). Illumina sequencing and analysis from recipient cell genomes indicated that bases were changed in 39%–47% of sequenced alleles, while wild-type sequences were recovered in 53%–61% of reads (Supplementary Fig. S5C). The percentage of edited alleles is an underestimate since the degenerate repair template contained the wild-type base so edits that resulted in restoration of wild-type are indistinguishable from unedited wild-

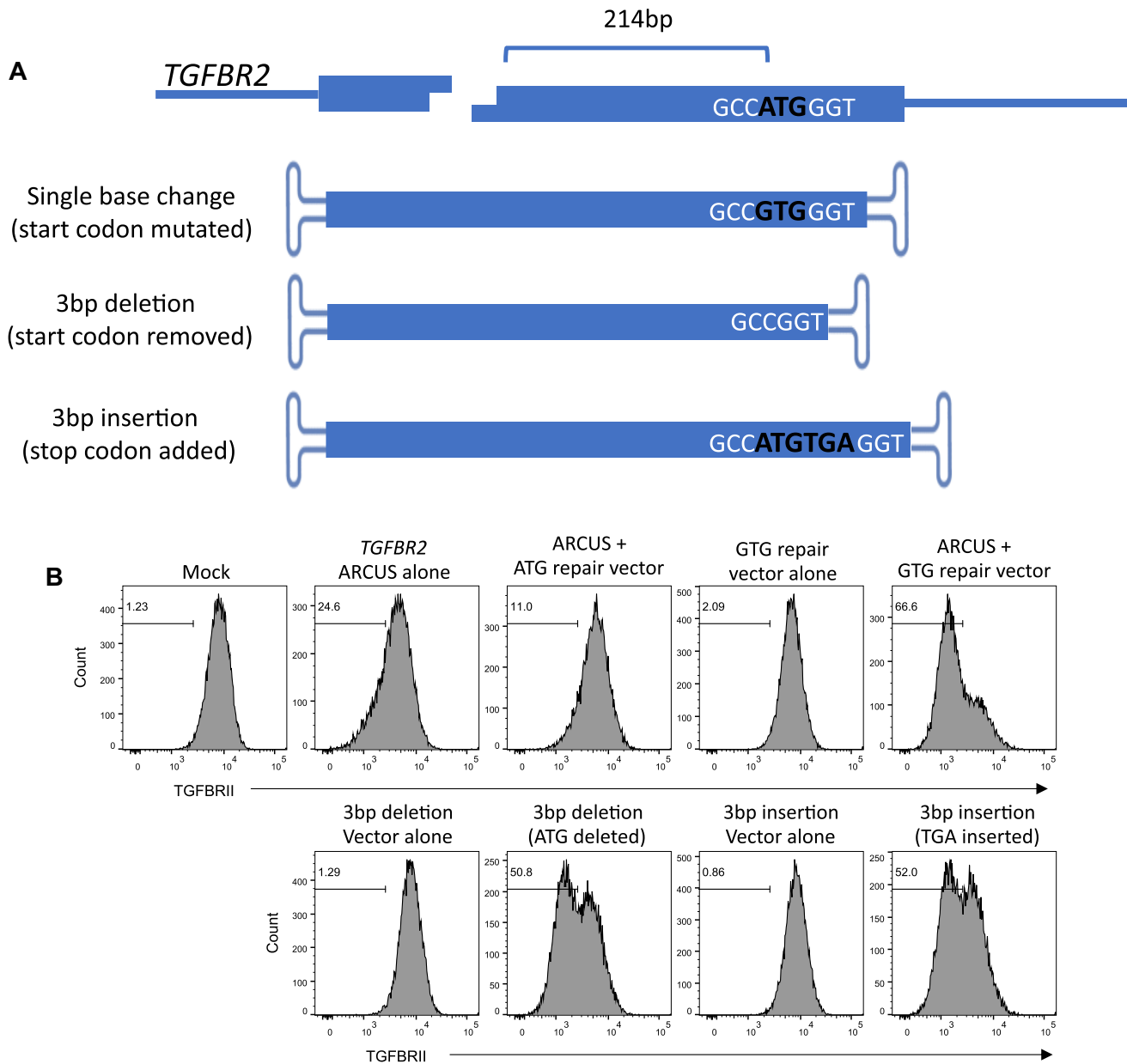


Figure 5. Highly efficient and specific small insertions, deletions, and substitutions are enabled using ARCUS and AAV. T cells were electroporated with mRNA encoding a *TGFB2*-specific ARCUS nuclease and transduced with one of four AAV6 templates displaying 500 bp HAs flanking the cut site. Each AAV6 vector in this study contained mismatches in the RHA as denoted in (A). Representative histograms shown in (B) show loss of TGFB2II display on insert⁺ cells, compared to cells treated with repair vector or nuclease alone. Data are representative of three independent experiments using T cells from two human donors.

type. We found that when a base was changed, the frequencies of the other three bases appearing in the allele were approximately equal in each respective library (Fig. 6B). Although changes to C bases appeared to be slightly less frequent, it should be noted that C variants were the least represented in each degenerate library (Supplementary Fig. S5). We conclude that HDR-mediated base change using HAs with 1 bp mismatches can accomplish all 12 base changes with comparable efficiency to generate any potential edit.

Homology arms targeting DSB-distal regions can replace genomic sequence

Until now, we have designed repair constructs with HAs directly flanking the ARCUS cut. However, HAs that target

DSB-distal genomic regions can enable changes to be made to stretches of genomic DNA in excess of the carrying capacity of an AAV. This, in turn, could enable replacement of multiple exons within the center of a large gene. To test whether ARCUS DSBs could enable this phenomenon, we designed an AAV6 vector targeting the *TRAC* locus using one DSB-flanking HA, and a second HA targeting a region 2.8 kb downstream of the DSB site (Fig. 7A). The underlying strategy for this construct was to insert a new artificial exon and an intron containing a β -2 microglobulin (B2M)-targeting microRNA as a reporter of insertion, which, through HDR crossover, would replace the endogenous sequence spanning the 3' end of exon 1 through a 5' portion of exon 3. In this insert and replace strategy, only correct insertion and removal will repair func-

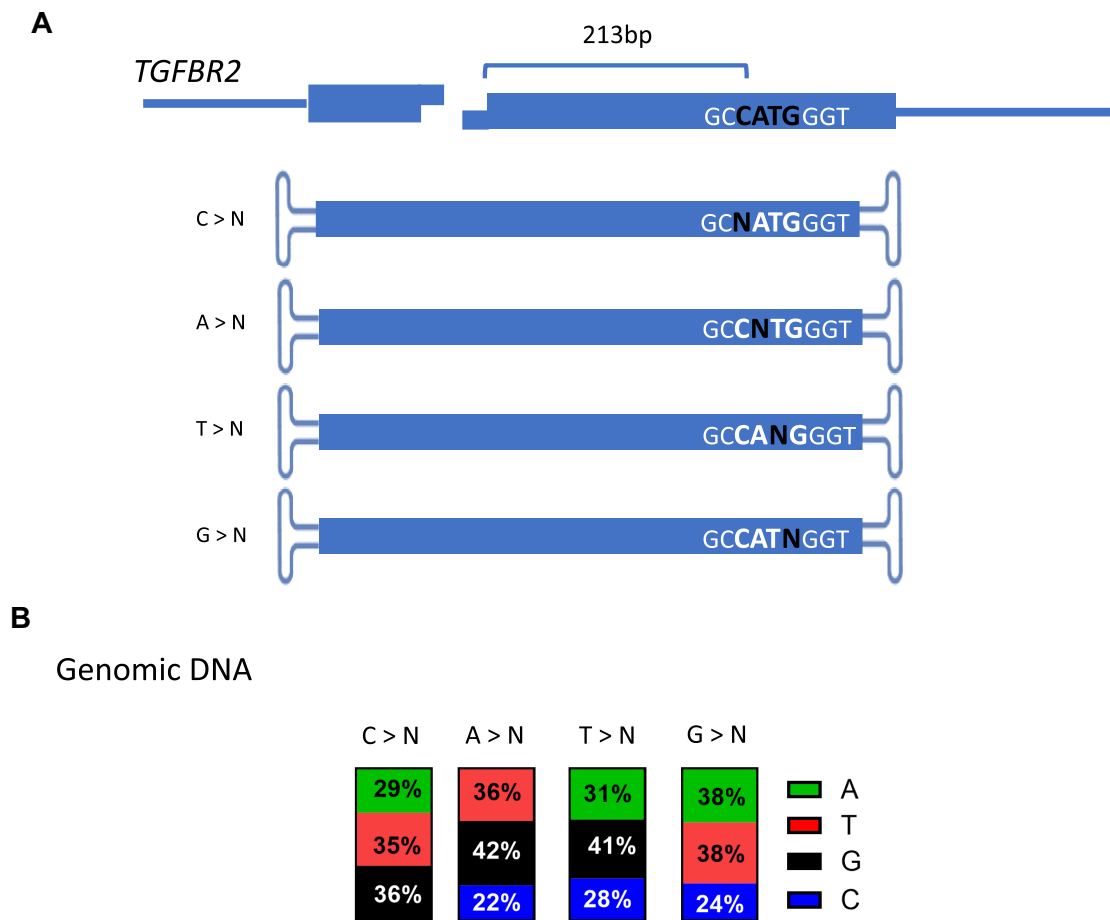


Figure 6. ARCUS can enable all base changes through HDR. **(A)** Activated human T cells were electroporated with mRNA encoding a *TGFBR2*-specific ARCUS nuclease (500 ng/10⁶ cells) and transduced with one of four AAV6 vectors displaying 500 bp HAs flanking the cut site (MOI: 4.0e⁴). Each vector contained a degenerate sequence at the indicated position (N). **(B)** Seven days after editing and transduction, Illumina sequencing of T-cell genomic DNA measured the frequencies of changes to the other three bases at each position. Data are representative of three biological replicates.

tion of CD3 while expression of the microRNA from the intron will suppress B2M expression.

To avoid premature crossover between the exons being replaced, the repair vector contained alternative (wobbled) coding sequences for the exons that were excised by this skip. Indels formed following ARCUS editing would ablate CD3 display (CD3⁻) while cells capturing a repair template would lose B2M expression (B2M⁻). CD3⁻B2M⁻ cells have likely captured the AAV template via NHEJ and disrupted *TRAC* expression. Alternatively, this CD3⁻B2M⁻ population may have an ARCUS-driven indel at the transcriptionally active TCR allele and carry an edit and insert at the second TCR allele. Only cells that correctly inserted the repair template into the ARCUS DSBs would exhibit B2M loss but retain CD3 expression (CD3⁺B2M⁻). These possibilities are diagrammed in Fig. 7B. An example allele carrying the correct insert is detailed in Fig. 7A, with native exon 1 sequence in dark blue, some transgenic, re-coded DNA in light blue, the artificial intron in purple, and transgenic, re-coded exon 2 fused directly to native exon 3 in light blue again.

A majority of T cells electroporated with only the TRC1-2 nuclease were edited but as expected contained no insert (70% CD3⁻B2M⁺). Cells treated with the nuclease and repair template achieved integration ~60% of the time (total

B2M⁻), with 14.6% of events expressing CD3 as intended (Fig. 7B). We reasoned that a longer RHA may increase likelihood of correct crossover. We found that lengthening the RHA from 500 bp to 2.3 kb resulted in a substantial increase in the frequency of correct integration events (14.6% versus 62%) that supported proper expression of TCR and display of CD3 subunits on the cell surface and suppression of B2M (CD3⁺B2M⁻).

We then altered the position of the DSB-distal HA and targeted regions that were ~5, 7.5, or 10 kb downstream of the DSB, or 2.5 kb upstream of the DSB (Fig. 7C). We found that efficiency of transgene integration was inversely proportional to the distance between the distal HA and the DSB (Fig. 7D), with the insertion efficiency decreasing as larger regions of endogenous sequence were crossed out of the genome. Additionally, the vector designed to perform an insertion and replacement upstream of the DSB functioned with ~25% efficiency, while a vector achieving a downstream skip of similar distance functioned at >40% efficiency. On the surface, this result indicates that, at least in *TRAC*, a genomic insertion and replacement downstream of the DSB is more productive than upstream of the DSB.

To evaluate the mechanism of DNA repair used in these replacement strategies, we fluorescence-sorted the CD3⁺B2M⁻

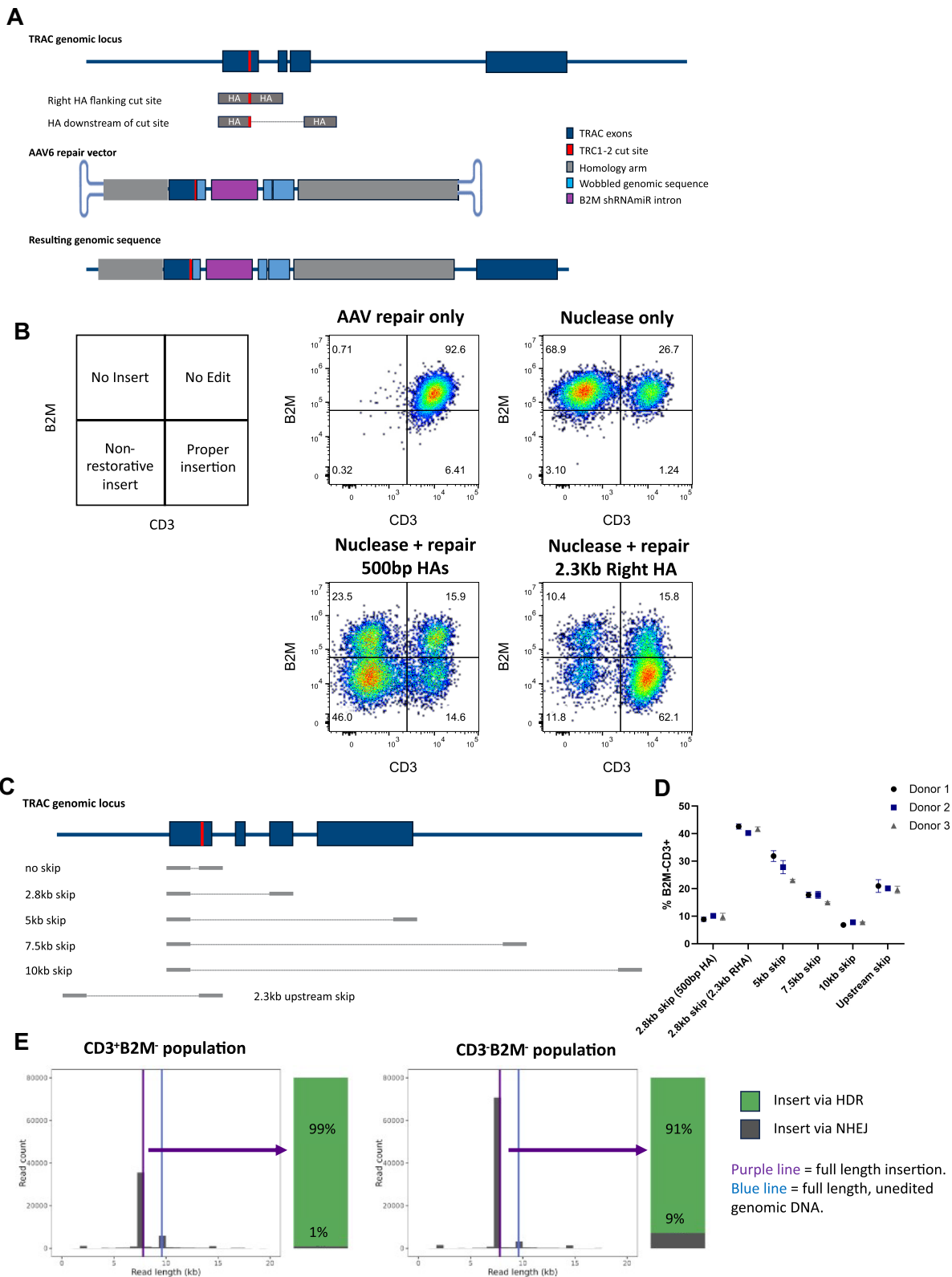


Figure 7. HAs distal to the ARCUS DSB mediate efficient replacement of genomic sequence. Activated human T cells were electroporated with mRNA (500 ng/1e⁶ cells) encoding a *TRAC*-specific ARCUS nuclease and transduced with AAV6 repair templates at an MOI of 4e4 v.g./cell. **(A)** Schematic detailing the *TRAC* genomic locus, the AAV6 repair template with HAs targeting genomic regions nonadjacent to the cut site, and the resulting genomic sequence. **(B)** Representative flow cytometry analysis of editing and insertion 7 days after electroporation, as measured by CD3 and B2M expression. **(C)** HAs were placed at different lengths both upstream and downstream of the genomic cut site, detailed in the schematic. **(D)** Data compiled from three individual T-cell donors via flow cytometry shows that the percentage of inserted cells is inversely proportional to the genomic distance skipped. **(E)** T cells were electroporated with *TRAC*-specific ARCUS nuclease mRNA as described above, flow sorted to isolate CD3⁺B2M⁻ expressing insert cells and PCR amplified across the *TRAC*-specific ARCUS cut site, producing a 9.7 kb amplicon. Nanopore sequencing shows that 92% of insert-containing cells repaired by seamless HDR. Data are representative of three technical replicates, three independent experiments, and two human donors.

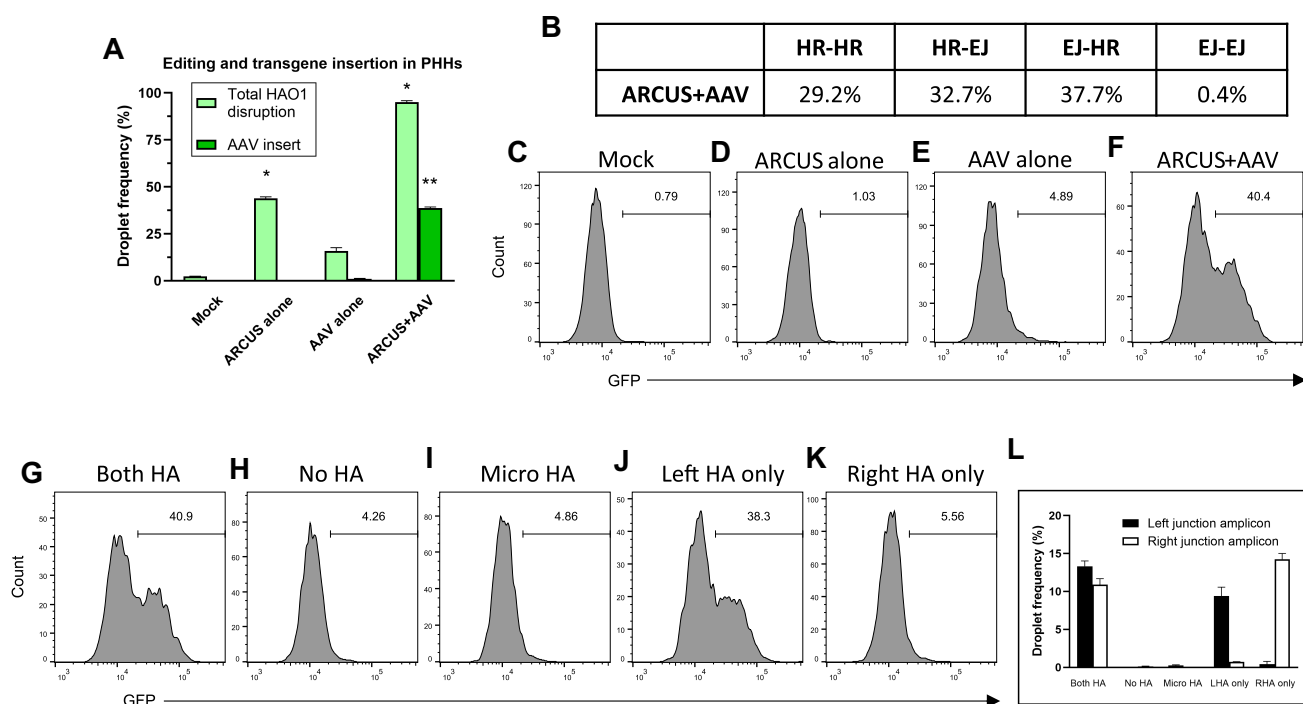


Figure 8. Efficient transgene insertion in PHHs requires homology. PHHs were thawed, plated, and cultured overnight prior to being treated with LNPs containing 500 ng/1e6 cells of mRNA encoding a *HAO1*-specific ARCUS nuclease and $1e^5$ v.g./cell of an AAV6 vector encoding a droplet digital PCR amplicon. Seven days after editing, DNA was harvested and *HAO1* disruption and transgene insertion were measured by ddPCR. The frequency of disrupted *HAO1* alleles (sum of small indels plus AAV gene insertions in light green, and AAV insertions in dark green) is plotted in panel (A). Repair outcomes were determined by Nanopore long read sequencing and the results appear in panel (B). Alternatively, PHHs were edited and transduced with an AAV6 vector encoding a P2A-furin controlled GFP cassette. Cells were analyzed by flow cytometry (panels C–F), to quantify AAV insertions. Significance was determined using ANOVA and a Tukey–Kramer multiple comparisons correction. Data are representative of three biological replicates using hepatocytes from two different donors; * $P < 0.05$, ** $P < 0.01$, $n = 3$. Edited PHHs were also transduced with AAV6 vectors encoding HA variants (both HAs, no homology, microhomology, left arm only, right arm only) and insertion was measured by flow cytometry (G–K) or ddPCR (L).

population from cells treated with the nuclease and repair template with the 2.3 kb right HA (Fig. 7B, bottom right plot). Genomic DNA was amplified across the TRC1-2 site, resulting in a 9.7 kb amplicon and long-range Nanopore sequencing was performed. The majority of the sorted $CD3^+B2M^-$ cells were full length sequences containing the insert; 99% of these integrants repaired perfectly by HDR (Fig. 7E). We also sorted the $CD3^-B2M^-$ population for long-range Nanopore sequencing. Surprisingly, a large majority (91%) of the sequenced integrants in this quadrant repaired by perfect HDR insertion into the *TRAC* locus. In T-cells, a single *TRAC* allele is actively expressed and we assume that $CD3^-B2M^-$ events are due to insertions into the inactive allele and an indel at the active allele. These experiments highlight the unique capabilities of ARCUS-mediated long-distance genomic skipping and replacement of functional, template cargo.

High-efficiency gene insertion in nondividing primary human hepatocytes using ARCUS

Targeted integration via HDR occurs efficiently in activated T lymphocytes, as shown here (Figs 1–4), but HDR is reported to be inefficient in nondividing cells such as PHHs [53, 54]. We aimed to test our capability for gene insertion in stationary PHH using ARCUS and AAV6. First, we engineered an ARCUS nuclease with specificity to the hydroxy-acid oxidase (*HAO1*) gene, which is expressed in liver and may alleviate high systemic oxalate levels in the rare disorder primary hyperoxaluria type 1 (PH1) if disrupted [55,

56]. We plated PHHs and administered mRNA encoding the *HAO1*-specific ARCUS nuclease encapsulated in LNPs along with a repair template carrying a noncoding insert flanked by *HAO1* HAs in an AAV6 vector (Supplementary Fig. S6A). Seven days following editing and transduction, we harvested DNA and measured intact and disrupted *HAO1* alleles and the frequency of integration-positive alleles by ddPCR (diagrammed in Figure S6). As shown in Fig. 8A, we observed 47% *HAO1* disruption (small insertions/deletions) following ARCUS–LNP administration alone and 95% disruption (small insertions/deletions plus AAV transgene inserts) following ARCUS–LNP plus an AAV6 HR template. We also observed integration of the ddPCR target in 39% of *HAO1* alleles, indicating that the combination of a 3' staggered DSB and a homologous template can drive high targeted integration rates in noncycling primary human cells.

To ascertain which repair mechanisms are contributing to gene insertion in noncycling PHHs, we characterized repaired alleles at 7d post-editing via Nanopore long-read sequencing. In contrast to our results in T cells (Fig. 2), sequence analysis of insert-bearing alleles found that only a minority of repairs (29%) were error-free on both ends, while the remaining 70% of repaired alleles featured error-free repair at one end or the other, but contained insertions or deletions at the opposite end (Fig. 8B). The most commonly observed imperfections were duplications in HA sequences or capture via ITRs, suggestive of an end-joining event. Errors were detected on either side of the DSB with equal frequency (EJ–HR versus HR–EJ, Fig. 8B), indicating no directional bias in the process.

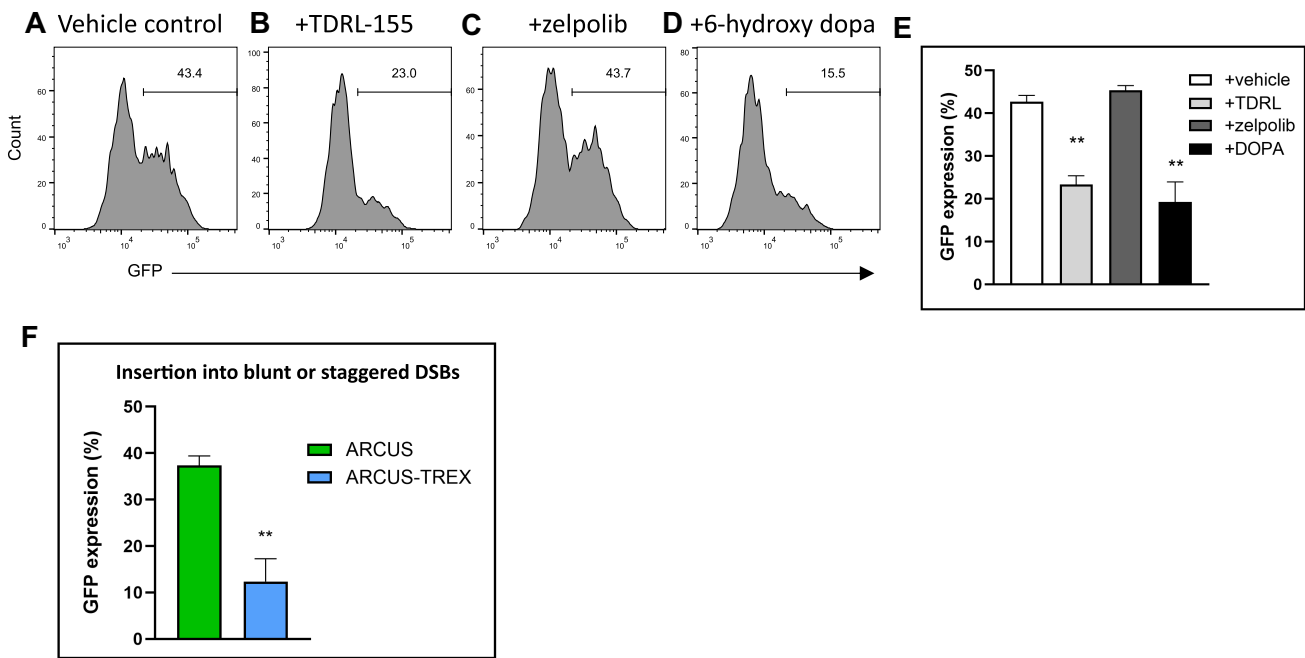


Figure 9. SDSA and 3' overhangs drive transgene insertion in PHHs. PHHs were thawed, plated, and cultured overnight prior to being treated with LNPs containing mRNA encoding a *HAO1*-specific ARCUS nuclease and an AAV6 vector encoding a P2A-furin controlled GFP cassette. Editing was carried out in the presence of vehicle alone, 500 μ M TDRL-551, 90 μ M zelpolib, or 200 μ M 6-hydroxy dopamine. Seven days following editing, cells were analyzed for GFP frequency via flow cytometry. Representative GFP histograms are shown in panels (A–D) and the means of three technical replicates appear in panel (E). Significance was determined using ANOVA and a Tukey–Kramer multiple comparisons correction. Data are representative of three biological replicates using hepatocytes from two different donors; ** $P < 0.01$, $n = 3$. PHHs were also treated with LNPs carrying either mRNA encoding a *HAO1*-specific ARCUS nuclease or the same nuclease linked to a TREX1 domain. Seven days later GFP frequencies were measured by flow cytometry and the results are plotted in panel (F). Significance was assessed using a two-tailed Student's *t*-test; ** $P < 0.013$, $n = 3$.

To confirm these results using a more functional readout, we designed a promoterless GFP insert vector controlled by a P2A-furin element, which requires seamless insertion and maintenance of the *HAO1* reading frame for reporter detection (Supplementary Fig. S6B) [57]. Seven days following editing and transduction of PHHs, we analyzed GFP frequency by flow cytometry. Relative to mock, LNP only, and AAV only controls, we observed GFP signal in just over 40% of PHHs (Fig. 8C–F), corroborating the results obtained from ddPCR measurements and sequencing analysis. Sequence observations in Fig. 8 are consistent with the analysis of TRAC alleles repaired with single-arm repair vectors (Fig. 3) and suggest that homology may only be triggering insertion from one end of the template at the *HAO1* locus in PHHs in the majority of events. To explore this, we measured insert frequency (by GFP analysis and ddPCR) of repair vectors bearing two HA, no HAs, microhomology arms, or only the left or right HA targeting the *HAO1* locus in PHHs. Compared to a fully homologous repair template (40.9%), GFP signals from non-homologous templates or those with micro-HAs were virtually nonexistent (Fig. 8G–K). Surprisingly, the vector bearing only the left HA supported GFP expression at nearly the same frequency as vectors with both HAs while homology on the right side supported very little GFP expression. We confirmed insertion via ddPCR on amplicons spanning the left and right junctions and observed comparable frequencies on both sides for the fully homologous vector (Fig. 8L). However, amplification was only observed on the left side for alleles repaired with LHA only vectors and only on the right side for RHA only vectors. Importantly, despite the absence of GFP signal from RHA only vectors, ddPCR detected insertion at levels

that were roughly equivalent to LHA only or fully homologous vectors. From these results we conclude that ARCUS and AAV6 can support efficient transgene insertion in PHHs, but HDR is unlikely to be the mechanism by which this occurs, as errors on one end or the other are common. We also conclude that while classical HDR is not likely occurring, homology between the repair template and the DSB is required on at least one end of the vector, and this may initiate repair while either an end joining event (more commonly) or homology-mediated crossover (less commonly) repairs the opposite side of the lesion.

Aside from HDR, HR can also occur via a pair of non-classical pathways called SDSA and break-induced replication (BIR) [41]. To determine if these mechanisms contribute to repair of ARCUS DSBs using AAV6, we administered ARCUS RNA encapsulated in LNPs and a GFP cassette controlled by P2A-furin (as above) packaged into AAV6 capsids in the presence of a panel of small molecule inhibitors to perturb nonclassical HR pathways. We targeted replication protein A (RepA) using TDRL-551 to disrupt both BIR and SDSA [58]. Polymerase δ was targeted using zelpolib to disrupt BIR [59], and Rad52 was targeted using 6-hydroxy-dopamine to inhibit SDSA [40, 60]. Treatment with TDRL-551 resulted in significantly reduced GFP frequency relative to vehicle control (Fig. 9A–E), suggesting a role for nonclassical HR pathways in repair of ARCUS DSBs in noncycling cells. Zelpolib treatment did not affect GFP frequency, suggesting that BIR mediated by Pol δ is not likely to be required. 6-Hydroxy-dopamine treatment resulted in a significant reduction in GFP frequency similar in magnitude to TDRL-551 treatment and suggests that Rad52-mediated SDSA or Rad51-independent

BIR may be key contributors to ARCUS DSB repair in PHHs.

Finally, we confirmed in noncycling cells that the 4nt overhangs on the 3' ends of ARCUS DSBs are crucial to triggering the high rates of insertion (Fig. 9F). As shown in lymphocytes in Fig. 2, an *HAO1*-specific ARCUS supports ~40% GFP insertion in PHHs while the same enzyme fused to a TREX1 exonuclease domain allows significantly lower rates of GFP insertion. Collectively, these results in noncycling PHHs suggest that transgene insertion into staggered DSBs is dependent on homology for initiation, but that the second end can be repaired via a mix of mechanisms based on sequencing outcomes, the use of single-arm repair templates, and small-molecule inhibition. In contrast to what is observed in actively cycling cells, gene insertion appears to be accomplished in PHHs primarily via the nonclassical HR pathway SDSA.

Discussion

Cells utilize different DDR pathways based on cell cycle and relative abundance of repair factors; all cells have NHEJ machinery at the ready and end-joining appears to be the preferred first option, as the response time is reported to be short [12, 26]. HDR, on the other hand, takes longer to carry out and attempts to leverage this pathway using a variety of endonuclease systems have yielded low efficiencies, especially in noncycling cells. To solve for this, the field encourages the low-efficiency reaction by supplying ample substrate (i.e. high AAV MOI) [13, 14, 24, 54], altering the cell cycle [54], inhibiting NHEJ [5, 19, 21, 22, 26, 61], or attaching other functional domains to the nuclease [62, 63]. These approaches lead to unwieldy therapeutic scenarios for gene therapy, as simultaneous delivery of guide RNAs, nucleases, associated enzyme complexes, repair templates and drugs are a considerable challenge and may incur compounding toxicity risks.

In this work, we show that the ARCUS nuclease/AAV6 system supports rates of transgene insertion exceeding 85% in T lymphocytes, and that this occurs by HDR. The template, the ends of the DSB, and the cell machinery all make key contributions to this process, which we confirmed by repair template variations, small molecule inhibitors, and NGS to be error-free HDR. Importantly, this was achieved using a relatively low abundance of repair template and was observed in rapidly dividing cells with MOIs as low as 5000 vg/cell. Insertion into ARCUS DSBs via HDR is dependent on HAs for efficient insertion (Fig. 1). We observed a nearly 50% decrease in insertion rates when either the LHA or RHA was removed and a 90% reduction when both arms were removed. Second, undesired insertions, deletions, duplications, or ITRs were only rarely found by long-read sequence analysis of repaired alleles, supporting HDR as the primary repair mechanism in T cells. Finally, both the thymidine analog TFT as well as the Rad51 inhibitor B02 significantly reduced HDR rates in T cells (Fig. 3). In contrast, the NHEJ inhibitor, NU7026, had no effect on ARCUS-mediated insertion (Fig. 3) suggesting that it is not the predominant repair mechanism for inserting AAV cargo into staggered DSBs. The reported MMEJ inhibitor Rucaparib had no effects on editing or insertion when ARCUS was used to introduce DSBs, but slightly increased overall editing and insertion using Cas9. This is consistent with results obtained with other end-joining inhibitors in this study (Fig. 3).

For HR to occur following a double stranded DNA break, at least the following are required: a resected end display-

ing a 3' overhang and a repair template displaying homology for DSB-adjacent sequences (reviewed in [5]). In dividing cells, the homologous repair template is typically a sister chromatid. Its DNA strands are invaded by the resected 3' ends of the DSB before DNA polymerases read through the template and repair the damaged DNA via 5'-to-3' DNA synthesis. We posit that the DDR pathway decisions made after a gene editing event depend on the substrates made available to cells by the editing enzymes. Blunt-ended DSBs are good NHEJ substrates; however, they are unfavorable HR substrates because they first require end-resection, and consequently, are preferentially repaired by the rapid end joining response. Our result, showing that HDR at blunt-ended DSBs was inefficient unless NHEJ was inhibited (Fig. 3), as well as similar observations published elsewhere, supports this concept [11, 19, 21, 22, 25]. In contrast, ARCUS DSBs are potent HDR substrates (Fig. 2) by virtue of their characteristic 3' overhangs, which mimic resection by 5'-to-3' exonucleases. This may explain the higher efficiency of integration of AAV6 templates into ARCUS DSBs, and why HDR inhibitors were the only molecules that perturbed integration into staggered-ended sites (Figs 2 and 3). The key differentiator of ARCUS appears to lie in the ability of these nucleases to generate a 3' overhang without the activity of additional cell cycle-dependent end resection machinery. The potential importance of the 3' overhang cut is exemplified by the significant loss of insertion efficiency (90% to 25%) when TREX is fused to ARCUS to remove the 3' overhangs or by the comparison to the blunt cut at the same site by CRISPR-Cas9 (Fig. 2). TREX1 has also been shown to interface with DDR pathways by way of PARP1 agonism [64] which may offer an alternative explanation for the reduced insert rates in ARCUS-TREX treated genomes despite similar overall editing frequencies. This may also partly explain why PARP1 inhibition (Rucaparib) enhanced insertion into Cas9 DSBs but not ARCUS-TREX DSBs, even though both lack overhangs (Fig. 3). Aside from ARCUS nucleases, Fanzor proteins, Cas12i2, Tef-based or fused enzymes, and paired sets of nickases generate 3' overhangs [33, 65–68], and while we found the 3' overhangs to be critical to promote HDR via ARCUS (Fig. 2), other studies have reported no benefit to a 3' overhang [66]. This could simply be attributed to the differences in the proteins since the scaffold for ARCUS, I-CreI, and homing endonucleases within the same family of proteins evolved to trigger HDR efficiently and facilitate propagation of the I-CreI coding sequence, whereas Cas enzymes evolved as DNA ablative bacterial defense systems [11, 33].

The ability to use HDR for gene editing allows a wide variety of different editing options. The most obvious option is to insert a full gene into the genome. We have previously demonstrated the highly efficient insertion of a large CAR construct into T cells for a clinically tested CAR T product. In this work, we demonstrate insertion of a GFP expression cassette at the TRAC target site of our ARCUS nuclease for simplicity of data collection. ARCUS editing by HDR can also be used to make smaller substitutions and edits away from the cut site (Figs 4 and 5). Despite the importance of HAs for integration, we found that mismatches could be tolerated with no decrease in GFP integration frequency and that the incorporation of mismatched bases was surprisingly an all-or-nothing phenomenon. This suggests that the crossover stage of HDR can take place at a variety of sites in the HA. Failure of mismatches to become incorporated into the genome could be due to the crossover event taking place in the 50 bp between the

HA mismatches and the GFP coding sequence. This is likely as 40 bp [51] is frequently mentioned as a minimum requirement to support HDR, and we designed the mismatches into the vector following 50 bp of homology. Indeed, we have observed efficient integration using vectors with 50 bp HAs, and the frequency was only slightly lower than was observed with longer arms (Supplementary Table S1). This suggests that the efficiency for substitutions away from the cut site could be improved by decreasing the distance between the cut and the substitution. Importantly, mismatch tolerance in the HA can be leveraged to change DNA sequences beyond just the single point of contact between the DNA and the editing tool. We extend previous findings of base changing capability via HDR, reporting here 60%–80% efficiency in primary human cells using ARCUS and AAV6 (Fig. 4C) while previous works report mid-single digit efficiencies [69]. As in previous reports, base changing via recombinant repair templates demonstrated complete freedom to make single base substitutions, as well as small insertions and deletions at high efficiency in primary human cells in excess of 200 bp from the cut site [70, 71]. Based on the correlation between distance and substitution efficiency (Fig. 7D), we would predict that the efficiency would increase further as we decreased this 214 bp separation. Because the base changes are dictated by a homology repair template, we demonstrated that all 12 possible base changes can be accomplished with similar efficiency (Fig. 5), which improves upon the capabilities of base editors. Furthermore, since the base changes are templated off the repair vector, this method of base change should not be capable of generating bystander base changes. We also show that the ARCUS-AAV6 pairing can achieve larger scale insert and replace operations with considerable efficiency at regions both upstream and downstream of the genomic cut site. This strategy allowed us to simultaneously insert sequence and remove sequence from the genome for a net outcome that replaces the center of a native gene with a new gene sequence. We demonstrated that the efficiency of this insert and replace is dependent on the length of the crossover distal HA and the distance in the genome that is being skipped (Fig. 7B–D). This is consistent with previous reports using a CRISPR–Cas9 system in which 20%–30% of alleles were corrected by a 1.5 kb replacement [15]. Results shown in Fig. 7 suggest that the higher HDR rates achievable at staggered DSBs from an ARCUS cut can enable a doubling of insert-and-replace efficiency (62%) while allowing larger distances to be skipped. In alignment with the other results in this paper, a large majority (99%) of the inserted outcomes in the insert and replace operation were produced by HDR [15]. Many *in vivo* gene editing applications are dependent on AAV to deliver the insertion template into the nucleus of cells [7]. As a result of this, the size of a gene that can be delivered in gene therapy and gene editing applications is severely limited by the cargo capacity of AAV at ≤ 4.7 kb. However, the insert and replace strategy has the potential to replace the center of a large gene via HDR by removing introns and fusing exons, hence expanding the types of mutations that can be addressed with this editing strategy beyond the smaller, whole gene insertion approaches.

Addressing *in vivo* gene editing especially in adults will often require a strategy that can work in quiescent cells as many pertinent target cells in adults are not rapidly dividing. Classical HDR in quiescent cells is frequently viewed as having too low of an efficiency for *in vivo* gene editing in these cells without additional manipulation. A recent publication demon-

strated integration into PHH with CRISPR–Cas9; however, this study encouraged PHH cell division using growth factors and reported 20%–30% integration frequency in a proliferating subset of hepatocytes [54] whereas the studies reported in Fig. 8 used stationary PHH cultures (Supplementary Fig. S7). We have demonstrated that ARCUS is able to generate high rates of gene insertion (30%–40%) in nondividing PHH (Fig. 8), and while insertion is clearly dependent on homology, long read sequencing detected extraneous indels on one side of the DSB or the other in 70% of insert-bearing alleles, suggesting that only one HA in the template was triggering repair. This was supported by results in Fig. 8G–K where it was found that one HA was sufficient to achieve integration in 30%–40% of target sites.

Nonclassical HR mechanisms appear to be prominent contributors to repair in PHHs as evidenced by this partial involvement of homology and the loss of insert frequency following RPA and Rad52 inhibition (Figs 8 and 9). These results are highly suggestive that SDSA is important for inserting AAV templates into ARCUS DSBs, but we cannot rule out other nonclassical HR mechanisms (e.g. BIR) for a few reasons. It's surprising that no effect was observed upon inhibiting polymerases generally with TFT (not shown), nor with inhibition of Pol δ specifically with zepolib, given the involvement of polymerases in the SDSA process [41]. TFT, which requires phosphorylation by thymidine kinases to exhibit repression, likely lacked activity in the PHH model due to low expression of such kinases. *TK1* and *TK2* are reported by several public datasets to be 8- to 16-fold less abundant at the transcript level in hepatocytes compared to activated lymphocytes, or to have low protein detection scores in liver samples [72, 73]. Specific targeting of Pol δ was also ineffective, perhaps due to the involvement of other polymerases, such as Pol η or other such trans-lesion polymerases in the insert/repair process [42, 74]. It is also possible that the doses of zepolib, a highly competitive but reversible inhibitor, were not sufficiently durable to inhibit repair activity over the course of our assay [59]. It is also surprising that polymerases would generate errors in SDSA as frequently as observed in Fig. 8. This may be because a relatively short AAV rather than a long chromosomal sequence is functioning as the template in this process and one of the HAs must serve as a substrate for second end capture. The HAs used in this study may be of insufficient length or composition to permit capture by homology. In this scenario, the polymerases may read through to the terminus of the AAV template before dissociating and leaving a 3' end that is captured by the other side of the DSB and repaired via end joining. It could also be that the choice between second end capture via RAD52-dependent annealing versus end-joining is inherently stochastic or is directed toward end-joining outcomes via the recombinogenic nature of AAV [42, 45]. This may also be ascribed to the abundance or scarcity of key DDR pathway determinants in PHHs such as helicases or ssDNA-binding proteins [75].

We believe this work relates efficient transgene integration via HR to the nature of DDR substrates provided by the editing operation, and that the explanation for why HR is considered inefficient is because most editing processes do not provide cells with appropriate HR substrates. Accounting for this when devising a correction strategy for mutations could increase the probability of success in clinical gene insertion efforts. To this end, we have demonstrated that the engineered homing endonuclease ARCUS is well suited to produce a DSB

that supports a high rate of insertion via HR mechanisms in dividing and stationary cells. Unlike insertion strategies where end-joining dominates, the involvement of at least one HA ensures that the cargo is inserted into the genome in the proper orientation and will support appropriate gene expression. Additionally, understanding the HR–NHEJ nature of insert in quiescent cells informs the strategy to place the NHEJ portion in an intron or UTR where it will not have impact on expression. This work also demonstrates the flexibility of the ARCUS platform, which through efficient HR, can support the breadth of currently understood gene editing outcomes (knockout, insertions, base changes, and replacements). Accomplishing this using one small protein (~1 kb of coding space) eases delivery restrictions that are a recognized challenge in gene therapy, theoretically enabling the packaging of its coding sequence along with a repair template in the same AAV capsid. Based on the findings described here, the small size and broad capabilities are advantages to this platform in diverse applications, and this is being validated in an ongoing human trial addressing the monogenic inherited disorder ornithine transcarbamylase deficiency by targeted gene insertion.

Acknowledgements

The authors wish to thank Jiminique Currie and Carolyn Reid for AAV6 packaging, Cassandra Gorsuch for thoughtful review, Rebecca van de Beek and Jamie Juett for technical assistance, and Chris Godsey for data management support.

Author contributions: Conceptualization was performed by J.J.S., J.S., J.H., and A.J.M. Investigation was carried out by L.C.R., A.L.T., A.J.V., A.M., K.S.E., I.A.M., and J.L. Resources were provided by A.M., J.H., G.H.T., and R.T. The original manuscript was written by A.J.M., A.L.T., K.S.E., and L.C.R., while J.J.S. and A.J.M. reviewed and edited the work.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

All authors are employees of Precision BioSciences, Inc, a publicly-traded company, and may have financial stake in the company. This work was supported by the company's R&D budget.

Funding

Precision BioSciences is a publicly held company and publication costs are paid from the corporate budget. Funding to pay the Open Access publication charges for this article was provided by the corporate budget.

Data availability

Raw sequence data have been posted to NCBI and can be found using the BioProject accession number PRJNA1183863. The list of software used for analysis and download locations appear in [Supplementary Table S2](#). Other data can be made available upon reasonable request.

References

- Chen X, Du J, Yun S *et al*. Recent advances in CRISPR–Cas9-based genome insertion technologies. *Mol Ther Nucleic Acids* 2024;35:102138. <https://doi.org/10.1016/j.omtn.2024.102138>
- Anguela XM, High KA. Entering the modern era of gene therapy. *Annu Rev Med* 2019;70:273–88. <https://doi.org/10.1146/annurev-med-012017-043332>
- Ay C, Reinisch A. Gene therapy: principles, challenges and use in clinical practice. *Wien Klin Wochenschr* 2025;137:261–71. <https://doi.org/10.1007/s00508-024-02368-8>
- Jasin M, Rothstein R. Repair of strand breaks by homologous recombination. *Cold Spring Harb Perspect Biol* 2013;5:a012740. <https://doi.org/10.1101/cshperspect.a012740>
- Scully R, Panday A, Elango R *et al*. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat Rev Mol Cell Biol* 2019;20:698–714. <https://doi.org/10.1038/s41580-019-0152-0>
- van der Oost J, Patinios C. The genome editing revolution. *Trends Biotechnol* 2023;41:396–409. <https://doi.org/10.1016/j.tibtech.2022.12.022>
- Bijlani S, Pang KM, Sivanandam V *et al*. The role of recombinant AAV in precise genome editing. *Front Genome Ed* 2021;3:799722. <https://doi.org/10.3389/fgeed.2021.799722>
- Suzuki K, Tsunekawa Y, Hernandez-Benitez R *et al*. In vivo genome editing via CRISPR–Cas9 mediated homology-independent targeted integration. *Nature* 2016;540:144–9. <https://doi.org/10.1038/nature20565>
- Fu B, Liao J, Chen S *et al*. CRISPR–Cas9-mediated gene editing of the BCL11A enhancer for pediatric beta(0)/beta(0) transfusion-dependent beta-thalassemia. *Nat Med* 2022;28:1573–80. <https://doi.org/10.1038/s41591-022-01906-z>
- Stephenson AA, Nicolau S, Vetter TA *et al*. CRISPR–Cas9 homology-independent targeted integration of exons 1–19 restores full-length dystrophin in mice. *Mol Ther Meth Clin Develop* 2023;30:486–99. <https://doi.org/10.1016/j.omtm.2023.08.009>
- Leal AF, Herreno-Pachon AM, Benincore-Florez E *et al*. Current strategies for increasing knock-in efficiency in CRISPR–Cas9-based approaches. *Int J Mol Sci* 2024;25. <https://doi.org/10.3390/ijms25052456>
- Mao Z, Bozzella M, Seluanov A *et al*. Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair* 2008;7:1765–71. <https://doi.org/10.1016/j.dnarep.2008.06.018>
- Degagne E, Donohoue PD, Roy S *et al*. High-specificity CRISPR-mediated genome engineering in anti-BCMA allogeneic CAR T cells suppresses allograft rejection in preclinical models. *Cancer Immunol Res* 2024;12:462–77. <https://doi.org/10.1158/2326-6066.CIR-23-0679>
- Naeimi Kararoudi M, Likhite S, Elmas E *et al*. Optimization and validation of CAR transduction into human primary NK cells using CRISPR and AAV. *Cell Rep Meth* 2022;2:100236. <https://doi.org/10.1016/j.crmeth.2022.100236>
- Allen D, Knop O, Itkowitz B *et al*. CRISPR–Cas9 engineering of the RAG2 locus via complete coding sequence replacement for therapeutic applications. *Nat Commun* 2023;14:6771. <https://doi.org/10.1038/s41467-023-42036-5>
- Cromer MK, Camarena J, Martin RM *et al*. Gene replacement of alpha-globin with beta-globin restores hemoglobin balance in beta-thalassemia-derived hematopoietic stem and progenitor cells. *Nat Med* 2021;27:677–87. <https://doi.org/10.1038/s41591-021-01284-y>
- Allen D, Kalter N, Rosenberg M *et al*. Homology-directed-repair-based genome editing in HSPCs for the treatment of inborn errors of immunity and blood disorders. *Pharmaceutics* 2023;15. <https://doi.org/10.3390/pharmaceutics15051329>
- Schubert MS, Thommandru B, Woodley J *et al*. Optimized design parameters for CRISPR–Cas9 and Cas12a homology-directed

- repair. *Sci Rep* 2021;11:19482. <https://doi.org/10.1038/s41598-021-98965-y>
19. Hu Z, Shi Z, Guo X *et al.* Ligase IV inhibitor SCR7 enhances gene editing directed by CRISPR–Cas9 and ssODN in human cancer cells. *Cell Biosci* 2018;8:12. <https://doi.org/10.1186/s13578-018-0200-z>
 20. Yu C, Liu Y, Ma T *et al.* Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* 2015;16:142–7. <https://doi.org/10.1016/j.stem.2015.01.003>
 21. Molugu K, Khajanchi N, Lazzarotto CR *et al.* Trichostatin A for efficient CRISPR–Cas9 gene editing of human pluripotent stem cells. *CRISPR J* 2023;6:473–85. <https://doi.org/10.1089/crispr.2023.0033>
 22. Ray U, Vartak SV, Raghavan SC. NHEJ inhibitor SCR7 and its different forms: promising CRISPR tools for genome engineering. *Gene* 2020;763:144997. <https://doi.org/10.1016/j.gene.2020.144997>
 23. Lamas-Toranzo I, Martinez-Moro A, OC E *et al.* RS-1 enhances CRISPR-mediated targeted knock-in in bovine embryos. *Mol Reprod Dev* 2020;87:542–9. <https://doi.org/10.1002/mrd.23341>
 24. Eyquem J, Mansilla-Soto J, Giavridis T *et al.* Targeting a CAR to the TRAC locus with CRISPR–Cas9 enhances tumour rejection. *Nature* 2017;543:113–7. <https://doi.org/10.1038/nature21405>
 25. Fu YW, Dai XY, Wang WT *et al.* Dynamics and competition of CRISPR–Cas9 ribonucleoproteins and AAV donor-mediated NHEJ, MMEJ and HDR editing. *Nucleic Acids Res* 2021;49:969–85. <https://doi.org/10.1093/nar/gkaa1251>
 26. Xue C, Greene EC. DNA repair pathway choices in CRISPR–Cas9-mediated genome editing. *Trends Genet* 2021;37:639–56. <https://doi.org/10.1016/j.tig.2021.02.008>
 27. Porto EM, Komor AC. In the business of base editors: evolution from bench to bedside. *PLoS Biol* 2023;21:e3002071. <https://doi.org/10.1371/journal.pbio.3002071>
 28. Dadush A, Merdler-Rabinowicz R, Gorelik D *et al.* DNA and RNA base editors can correct the majority of pathogenic single nucleotide variants. *Npj Genom Med* 2024;9:16. <https://doi.org/10.1038/s41525-024-00397-w>
 29. Anzalone AV, Randolph PB, Davis JR *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 2019;576:149–57. <https://doi.org/10.1038/s41586-019-1711-4>
 30. Roberts JD, Bebenek K, Kunkel TA. The accuracy of reverse transcriptase from HIV-1. *Science* 1988;242:1171–3. <https://doi.org/10.1126/science.2460925>
 31. Yarnall MTN, Ioannidi EI, Schmitt-Ulms C *et al.* Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases. *Nat Biotechnol* 2023;41:500–12. <https://doi.org/10.1038/s41587-022-01527-4>
 32. Chevalier BS, Monnat RJ, Stoddard BL. The homing endonuclease I-CreI uses three metals, one of which is shared between the two active sites. *Nat Struct Biol* 2001;8:312–6. <https://doi.org/10.1038/86181>
 33. Chevalier BS, Stoddard BL. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res* 2001;29:3757–74. <https://doi.org/10.1093/nar/29.18.3757>
 34. Jurica MS, Stoddard BL. Homing endonucleases: structure, function and evolution. *Cell Mol Life Sci* 1999;55:1304–26. <https://doi.org/10.1007/s000180050372>
 35. Antunes MS, Smith JJ, Jantz D *et al.* Targeted DNA excision in Arabidopsis by a re-engineered homing endonuclease. *BMC Biotechnol* 2012;12:86. <https://doi.org/10.1186/1472-6750-12-86>
 36. Gorsuch CL, Nemec P, Yu M *et al.* Targeting the hepatitis B cccDNA with a sequence-specific ARCUS nuclease to eliminate hepatitis B virus *in vivo*. *Mol Ther* 2022;30:2909–22. <https://doi.org/10.1016/j.ymthe.2022.05.013>
 37. Shoop WK, Lape J, Trum M *et al.* Efficient elimination of MELAS-associated m.3243G mutant mitochondrial DNA by an engineered mitoARCUS nuclease. *Nat Metab* 2023;5:2169–83. <https://doi.org/10.1038/s42255-023-00932-6>
 38. Djukanovic V, Smith J, Lowe K *et al.* Male-sterile maize plants produced by targeted mutagenesis of the cytochrome P450-like gene (MS26) using a re-designed I-CreI homing endonuclease. *Plant J* 2013;76:888–99. <https://doi.org/10.1111/tpj.12335>
 39. MacLeod DT, Antony J, Martin AJ *et al.* Integration of a CD19 CAR into the TCR alpha chain locus streamlines production of allogeneic gene-edited CAR T cells. *Mol Ther* 2017;25:949–61. <https://doi.org/10.1016/j.ymthe.2017.02.005>
 40. Liang CC, Greenhough LA, Masino L *et al.* Mechanism of single-stranded DNA annealing by RAD52–RPA complex. *Nature* 2024;629:697–703. <https://doi.org/10.1038/s41586-024-07347-7>
 41. Liu L, Malkova A. Break-induced replication: unraveling each step. *Trends Genet* 2022;38:752–65. <https://doi.org/10.1016/j.tig.2022.03.011>
 42. McIlwraith MJ, West SC. DNA repair synthesis facilitates RAD52-mediated second-end capture during DSB repair. *Mol Cell* 2008;29:510–6. <https://doi.org/10.1016/j.molcel.2007.11.037>
 43. Wang L, Smith J, Breton C *et al.* Meganuclease targeting of PCSK9 in macaque liver leads to stable reduction in serum cholesterol. *Nat Biotechnol* 2018;36:717–25. <https://doi.org/10.1038/nbt.4182>
 44. Rezaiofi A, Fritz L, Forster R *et al.* Challenges of CRISPR-based gene editing in primary T cells. *Int J Mol Sci* 2022;23. <https://doi.org/10.3390/ijms23031689>
 45. Hanlon KS, Kleinstiver BP, Garcia SP *et al.* High levels of AAV vector integration into CRISPR-induced DNA breaks. *Nat Commun* 2019;10:4439. <https://doi.org/10.1038/s41467-019-12449-2>
 46. Certo MT, Gwiazda KS, Kuhar R *et al.* Coupling endonucleases with DNA end-processing enzymes to drive gene disruption. *Nat Methods* 2012;9:973–5. <https://doi.org/10.1038/nmeth.2177>
 47. Chauvin SD, Ando S, Holley JA *et al.* Inherited C-terminal TREX1 variants disrupt homology-directed repair to cause senescence and DNA damage phenotypes in Drosophila, mice, and humans. *Nat Commun* 2024;15:4696. <https://doi.org/10.1038/s41467-024-49066-7>
 48. Karasu ME, Toufektchan E, Chen Y *et al.* Removal of TREX1 activity enhances CRISPR–Cas9-mediated homologous recombination. *Nat Biotechnol* 2024;43:1168–1176.
 49. Mimori T, Hardin JA. Mechanism of interaction between Ku protein and DNA. *J Biol Chem* 1986;261:10375–9. [https://doi.org/10.1016/S0021-9258\(18\)67534-9](https://doi.org/10.1016/S0021-9258(18)67534-9)
 50. Hermantara R, Richmond L, Taqi AF *et al.* Improving CRISPR–Cas9 directed faithful transgene integration outcomes by reducing unwanted random DNA integration. *J Biomed Sci* 2024;31:32. <https://doi.org/10.1186/s12929-024-01020-x>
 51. Chen F, Pruett-Miller SM, Huang Y *et al.* High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat Methods* 2011;8:753–5. <https://doi.org/10.1038/nmeth.1653>
 52. Ning B, Huang J, Xu H *et al.* Genomic organization, intragenic tandem duplication, and expression analysis of chicken TGFBR2 gene. *Poult Sci* 2022;101:102169. <https://doi.org/10.1016/j.psj.2022.102169>
 53. VanLith CJ, Guthman RM, Nicolas CT *et al.* Ex vivo hepatocyte reprogramming promotes homology-directed DNA repair to correct metabolic disease in mice after transplantation. *Hepatology* 2019;3:558–73. <https://doi.org/10.1002/hep4.1315>
 54. Zhang K, Wan P, Wang L *et al.* Efficient expansion and CRISPR–Cas9-mediated gene correction of patient-derived hepatocytes for treatment of inherited liver diseases. *Cell Stem Cell* 2024;31:1187–1202.e8. <https://doi.org/10.1016/j.stem.2024.04.022>
 55. Owens GL, Whitney, Jordan-Steele M *et al.* American Society for Cell and Gene Therapy Abstract and Poster Presentation. Cell Press, Washington, DC, 2022.
 56. Zabaleta N, Barberia M, Martin-Higueras C *et al.* CRISPR–Cas9-mediated glycolate oxidase disruption is an efficacious and safe treatment for primary hyperoxaluria type I.

- Nat Commun* 2018;9:5454. <https://doi.org/10.1038/s41467-018-07827-1>
57. Fang J, Yi S, Simmons A *et al.* An antibody delivery system for regulated expression of therapeutic levels of monoclonal antibodies *in vivo*. *Mol Ther* 2007;15:1153–9. <https://doi.org/10.1038/sj.mt.6300142>
 58. Mishra AK, Dormi SS, Turchi AM *et al.* Chemical inhibitor targeting the replication protein A–DNA interaction increases the efficacy of Pt-based chemotherapy in lung and ovarian cancer. *Biochem Pharmacol* 2015;93:25–33. <https://doi.org/10.1016/j.bcp.2014.10.013>
 59. Mishra B, Zhang S, Zhao H *et al.* Discovery of a novel DNA polymerase inhibitor and characterization of its antiproliferative properties. *Cancer Biol Ther* 2019;20:474–86. <https://doi.org/10.1080/15384047.2018.1529126>
 60. Chandramouly G, McDevitt S, Sullivan K *et al.* Small-molecule disruption of RAD52 rings as a mechanism for precision medicine in BRCA-deficient cancers. *Chem Biol* 2015;22:1491–504. <https://doi.org/10.1016/j.chembiol.2015.10.003>
 61. Li G, Zhang X, Wang H *et al.* Increasing CRISPR/Cas9-mediated homology-directed DNA repair by histone deacetylase inhibitors. *Int J Biochem Cell Biol* 2020;125:105790. <https://doi.org/10.1016/j.biocel.2020.105790>
 62. Pallares-Masmitja M, Ivancic D, Mir-Pedrol J *et al.* Find and cut-and-transfer (FiCAT) mammalian genome engineering. *Nat Commun* 2021;12:7071. <https://doi.org/10.1038/s41467-021-27183-x>
 63. Matsumoto D, Tamamura H, Nomura W. A cell cycle-dependent CRISPR–Cas9 activation system based on an anti-CRISPR protein shows improved genome editing accuracy. *Commun Biol* 2020;3:601. <https://doi.org/10.1038/s42003-020-01340-2>
 64. Miyazaki T, Kim YS, Yoon J *et al.* 3rd. The 3′–5′ DNA exonuclease TREX1 directly interacts with poly(ADP-ribose) polymerase-1 (PARP1) during the DNA damage response. *J Biol Chem* 2014;289:32548–58. <https://doi.org/10.1074/jbc.M114.547331>
 65. Huang X, Sun W, Cheng Z *et al.* Structural basis for two metal-ion catalysis of DNA cleavage by Cas12i2. *Nat Commun* 2020;11:5241. <https://doi.org/10.1038/s41467-020-19072-6>
 66. Ran FA, Hsu PD, Lin CY *et al.* Double nicking by RNA-guided CRISPR–Cas9 for enhanced genome editing specificity. *Cell* 2013;154:1380–9. <https://doi.org/10.1016/j.cell.2013.08.021>
 67. Han W, Li Z, Guo Y *et al.* Efficient precise integration of large DNA sequences with 3′-overhang dsDNA donors using CRISPR–Cas9. *Proc Natl Acad Sci USA* 2023;120:e2221127120. <https://doi.org/10.1073/pnas.2221127120>
 68. Saito M, Xu P, Faure G *et al.* Fanzor is a eukaryotic programmable RNA-guided endonuclease. *Nature* 2023;620:660–8. <https://doi.org/10.1038/s41586-023-06356-2>
 69. Kan Y, Ruis B, Takasugi T *et al.* Mechanisms of precise genome editing using oligonucleotide donors. *Genome Res* 2017;27:1099–111. <https://doi.org/10.1101/gr.214775.116>
 70. Riesenberger S, Kanis P, Macak D *et al.* Efficient high-precision homology-directed repair-dependent genome editing by HDRobust. *Nat Methods* 2023;20:1388–99. <https://doi.org/10.1038/s41592-023-01949-1>
 71. O'Brien AR, Wilson LOW, Burgio G *et al.* Unlocking HDR-mediated nucleotide editing by identifying high-efficiency target sites using machine learning. *Sci Rep* 2019;9:2788. <https://doi.org/10.1038/s41598-019-39142-0>
 72. Uhlen M, Fagerberg L, Hallstrom BM *et al.* Proteomics. Tissue-based map of the human proteome. *Science* 2015;347:1260419. <https://doi.org/10.1126/science.1260419>
 73. Karlsson M, Zhang C, Mear L *et al.* A single-cell type transcriptomics map of human tissues. *Sci Adv* 2021;7:154243.
 74. Guo X, Jinks-Robertson S. Roles of exonucleases and translesion synthesis DNA polymerases during mitotic gap repair in yeast. *DNA Repair (Amst)* 2013;12:1024–30. <https://doi.org/10.1016/j.dnarep.2013.10.001>
 75. Elbakry A, Lobrich M. Homologous recombination subpathways: a tangle to resolve. *Front Genet* 2021;12:723847. <https://doi.org/10.3389/fgene.2021.723847>