ENHANCED EFFICIENCY OF RNA-GUIDED CAS12a VERSUS CAS9 TRANSGENE KNOCK-IN AND ACTIVITY AT A *SCHISTOSOMA MANSONI* GENOME SAFE HARBOR

Max F. Moescheid[1], Prapakorn Wisitphongpun[2,3], Victoria H. Mann[2], Thomas Quack[1], Christoph Grunau[4], Christoph G. Grevelding[1,*], Wannaporn Ittiprasert[2,*], Paul J. Brindley[2,*]

[1] Institute of Parasitology, Biomedical Research Center Seltersberg (BFS), Justus Liebig University Giessen, Giessen 35392 Germany

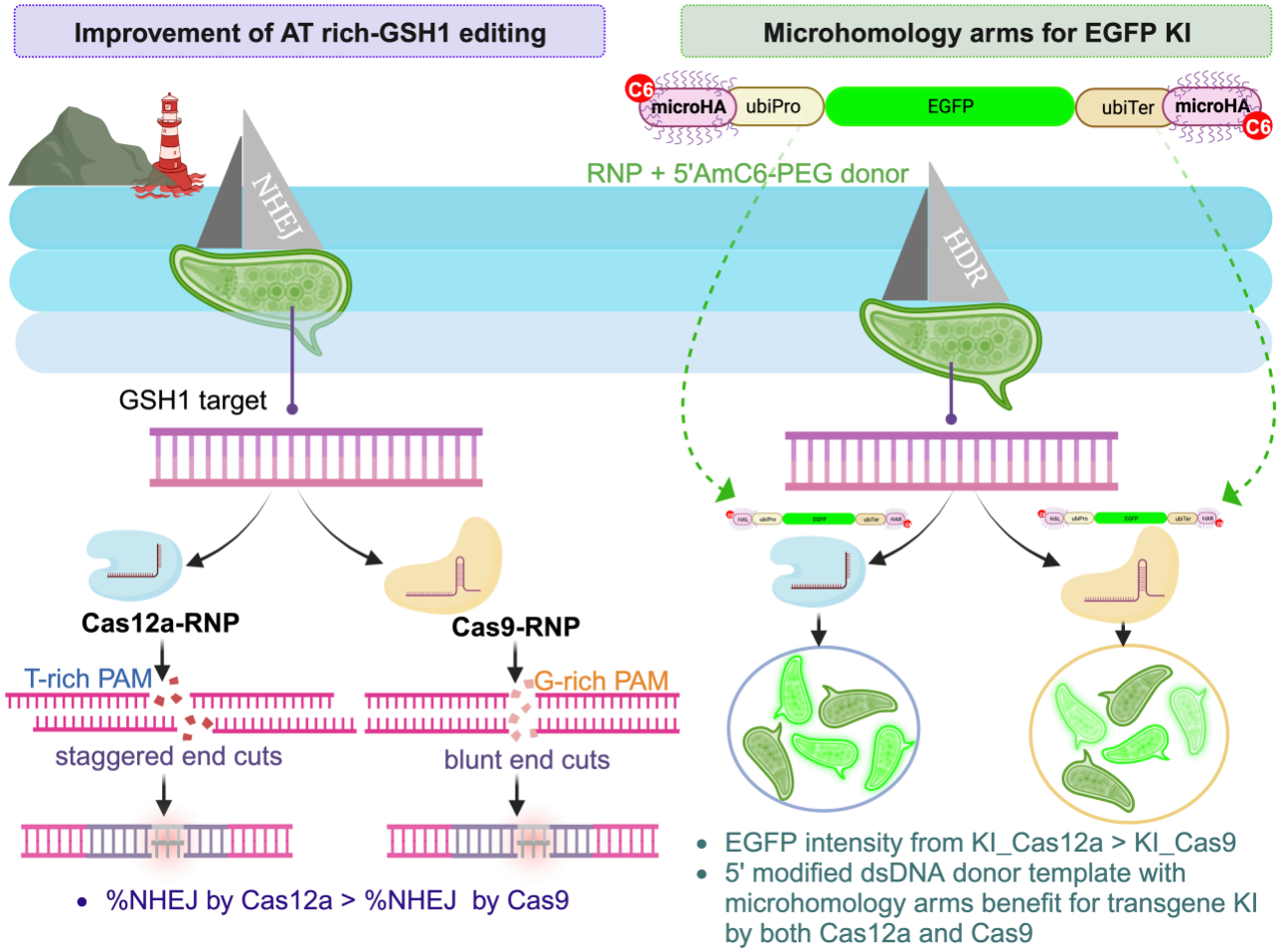
[2] Department of Microbiology, Immunology & Tropical Medicine, George Washington University, Washington, D.C. 20037 USA

[3] Faculty of Medical Technology, Rangsit University, Pathum Thani, 12000 Thailand

[4] IHPE Interactions Hôtes-Pathogènes-Environnements, Université de Perpignan via Domitia, Perpignan, 66860 France

* To whom correspondence should be addressed. Paul J. Brindley. Tel: +1-202-994-7499; Email: pbrindley@gwu.edu. Correspondence may also be addressed to Christoph G. Grevelding. Tel: +49-641-9938466; Email Christoph.Grevelding@vetmed.uni-giessen.de and Wannaporn Ittiprasert. Tel: +1-202-994-3122; Email: wannaporni@gwu.edu

GRAPHICAL ABSTRACT



ABSTRACT

Recently, we reported programmed Cas9 mediated insertion of a reporter gene into a gene safe harbor site, GSH1, of *Schistosoma mansoni* via homology-directed repair (HDR) using overlapping guide RNAs. Here, we report efficient and precise CRISPR/Cas12a-mediated homology-directed insertion of a 5′ C6-PEG10-modified double-stranded transgene bearing microhomology arms, 50 nt length at GSH1. At the outset, we undertook bioinformatic and computational analysis following by experimental verification of the regulatory activity of schistosome ubiquitin (SmUbi) promoter and terminator, aiming to drive strong reporter gene expression. Green fluorescent protein activity driven by this endogenous promoter followed electroporation-mediated transfection of schistosome eggs. HDR induced by CRISPR/Cas12a delivered more efficient transgene integration compared to CRISPR/Cas9, with precise integration of the transgene at Cas12a cleavage sites, which releases overhanding DNA strands at 18-24 nt during programmed cleavage. The transfection design for CRISPR/Cas-mediated genome editing facilitated precise knock-in, specifically chromosomal integration of the SmUbi-EGFP-SmUbi reporter-gene with microhomology arms into GSH1. In this non-model system, the 5′-C6-PEG10 modified-DNA template enhanced knockin efficiency of Cas9 and Cas12a. This approach advances schistosome transgenesis and may be functional also in related parasitic and non-parasitic helminths, which hitherto lack functional genomics and transfection methods.

INTRODUCTION

Schistosoma mansoni, a blood-dwelling parasite of humans and animals, is the cause of the infectious disease schistosomiasis. As with all trematodes, *S. mansoni* and related species exhibit complex life cycles including an invertebrate snail host and a vertebrate final host. Schistosomiasis treatment relies on a single drug, praziquantel (PZQ) (1). Due to the fear of upcoming PZQ resistance (2), there is the urgent need for the development of alternative treatment strategies. For this, and to understand parasite biology as such, it is inevitable to characterize the function of target genes and their potential role(s) in, for example, parasite development and homeostasis. Substantial effort has been exerted to decipher the *S. mansoni* genome (3-5). In addition, ongoing efforts have aimed at deciphering the *S. mansoni* transcriptome of the adult as well further developmental stages (6-16). Now, in the rising post-genomic/transcriptomic era of *S. mansoni* research, new tools are needed for the functional characterization of genes of interest (GOI). Until now, RNA interference (RNAi) has been proven as the most suitable method for functional gene characterization. However, RNAi efficiency varies, and it can lead to ectopic effects (17-20). To study GOI, knock-out (KO) models are common for various model organisms, but not yet established for trematodes or other helminths. Towards this end, a first breakthrough was recently achieved showing a way to precisely integrate a reporter-gene construct by CRISPR/Cas9 into a GSH site within the *S. mansoni* genome (21).

CRISPR/Cas-mediated genome editing has revolutionized, or soon will, research in biomedicine and the life sciences generally (22,23). An increasing repertoire of Cas homologues, in addition to Cas9, possessing divergent programmed nuclease activity including alternative PAMs, selectivity for divergent nucleic acid substrates, and sensitivity to metal ions, to temperature, and other characteristics have been reported (24,25). Cas12a was shown to be enzymatically active in eggs of *S. mansoni* where a Cas12a orthologue from *Acidaminococcus* (AsCas12a) more efficiently edited the *omega-1* locus compared to Cas9 (26). CRISPR/Cas9 RNPs consist of two components: guide RNA (gRNA) and Cas9 nuclease (27-29). This nuclease induces blunt end-DSBs (double-strand breaks) within a targeted locus, which can be repaired by genome

maintenance mechanisms including non-homologous end joining (NHEJ) and homology-directed repair (HDR). In contrast, CRISPR/Cas12a components consist of a crRNA and a Cas12a nuclease that induce sticky-end DSBs (30). CRISPR/Cas9 recognizes the NGG protospacer-adjacent motif (PAM) to catalyze the DSB upstream of PAM, while CRISPR/Cas12a recognizes a T-rich PAM, TTTV, with the DSB generated downstream of its PAM (31). Both nucleases have utility for gene editing in discrete genomic contexts; Cas9 may have advantages for editing GC-rich regions whereas Cas12a may be better suited for editing AT-rich regions (32-34).

Transgene insertion in the genome can negatively impact the host cell function or the cell may silence the transgene due to chromatin structure. To surmount these potential problems, the gene therapy field developed the concept of genome safe harbors (GSHs) (35). An ideal GSH has been defined as a region that does not overlap functional genomic elements and that lacks heterochromatic marks that could impede transcription. We have predicted several GSH in *S. mansoni* using criteria that included a location in euchromatin to avoid transgene silencing, a unique genome-target sequence to minimize off-target events, avoidance of long noncoding RNA-encoding genes, and the presence of histone marks for open chromatin structure and, conversely, the absence of epigenetic marks indicating heterochromatin (21). We reported the Cas9-mediated integration of an exogenous DNA sequence into GSH1 (genome safe harbor number 1) of *S. mansoni*. Programmed cleavage was enhanced using three overlapping guide RNAs complexed with the Cas9 nuclease from *Streptococcus pyogenes* (21). As biological targets, we used liver-stage *S. mansoni* eggs (LE), which represent a heterogenous population at different stages of development. LE comprise newly laid eggs, intermediate, and mature eggs (from stages 1 through 8 of the staging system of Jurberg and coworkers (36)), and they were transfected with Cas9 or Cas12 ribonuclear protein complexes (RNPs). Gene knock-in was achieved in the presence of a donor repair template encoding enhanced green fluorescent protein (EGFP) bearing 5'-phosphorothioate-modified terminus HA of 300-600 bp in length. The complete integration construct was 4,551 bp in length. The use of triple, overlapping gRNAs induced higher levels of precise transgene knock-in (KI) in comparison to single and double gRNAs. Nonetheless, further optimization aiming for even simpler, less expensive, and enhanced levels of efficacy likely can be achieved.

To this end, we evaluated the editing efficiencies of Cas9 versus Cas12a. In addition, we deployed a single gRNA in tandem with short, 5´-modified homology arms of the repair template, and finally investigated whether also newly laid eggs (NLE), which contain only germ cell(s), can be used as biological targets for transformation. This latter manipulation should ensure biologically standardized editing approaches compared to LE. We observed that both programmable nucleases generated insertions and deletions (InDels) at the target GSH, and that the editing efficiency of Cas12a significantly exceeded that of Cas9. Furthermore, with 5´-C6amine-PEG10-modified donor template and micro length HA of 50 bp flanking the donor transgene, we observed successful editing of GSH1 at the genome sequence level as well as by the expression of the genome-integrated EGFP reporter-gene. These results demonstrated a significantly advanced genome-editing approach for *S. mansoni*, with hitherto unprecedented efficiency of gene editing in schistosomes.

MATERIAL AND METHODS

Maintenance and isolation of *S. mansoni* developmental stages

Mice (female, Swiss Webster) exposed to about 180 cercariae of *S. mansoni* were provided by the Schistosomiasis Resource Center, Biomedical Research Institute, Rockville, MD. The mice were housed at the Animal Research Facility of George Washington University (GWU), which is accredited by the American Association for Accreditation of Laboratory Animal Care (AAALAC no. 000347) and has the Animal Welfare Assurance on file with the National Institutes of Health, Office of Laboratory Animal Welfare, OLAW assurance number A3205. All procedures employed were consistent with the Guide for the Care and Use of Laboratory Animals. The Institutional Animal Care and Use Committee at GWU approved the protocol used for mice in this study. Mice were euthanized at about day 49 following infection. Schistosomes were recovered by hepatic portal vein perfusion with 150 mM sodium chloride, 15 mM sodium citrate, pH 7. Thereafter, the worms were washed with 1× phosphate buffer saline (PBS), pH 7.4, 2% penicillin/streptomycin/fungizone (PSF) and maintained thereafter in DMEM, 10% heat inactivated fetal bovine serum (FBS) (Corning Life Sciences, Corning, NY) supplemented with 2% antibiotic/antimycotic (Corning) in the atmosphere of 5% CO_2, 37°C incubator (Panasonic,

Newark, NJ) (37). Adult schistosomes were processed within 24 h following recovery from the mice. To prepare *S. mansoni* eggs, mouse livers were rinsed in 70% ethanol, then twice washed with 1× PBS, 2% PSF, and finally homogenized in 1× PBS, 2% PSF in a mechanical blender. The homogenate was digested with clostridial collagenase (0.7 mg/ml) at 37°C for 18 h, after which cellular debris was removed using sieves of 250 μm and 125 μm mesh (Gilson, Middleton, WI). Sequentially, followed by sucrose/Percoll gradient centrifugation, eggs were recovered (38). The eggs, termed "liver eggs" (LE), were maintained at 37°C in 20% FBS/ 2% PSF/ DMEM, and electroporated with CRISPR/Cas materials with or without donor (below) within 24 h**.**

In silico **analyses of the ubiquitin regulatory promoter elements**

Enhanced green fluorescent protein, EGFP (39) under the control of the promotor and terminator sequences of Smp_335990 (40), a high-abundantly transcribed ubiquitin gene of *S. mansoni* (SmUbi) (Supplementary Table S1) was used as a reporter for transfection/transformation analyses. Nucleotide sequences of upstream and downstream untranslated regions (UTRs) of Smp_335990 were retrieved from WormBase ParaSite (40,41), version 7 of the draft genome of *S. mansoni* (3,5). To confirm this annotation, we used the protein homology/analogy recognition engine Phyre2 with the protein structure modelling mode, http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index (42). To predict the promoter region of SmUbi, adjacent sequences up to the three kb upstream of the start codon were scanned. Promoter elements were identified by the Berkeley *Drosophila* Genome Project Neural Network Promoter Prediction Tool, using a minimum promoter score of 0.8 for eukaryotic promoters (43), https://www.fruitfly.org/seq_tools/promoter.html. In addition, we used the AliBaba2.1 TRANSFAC 4.0 transcription factor binding site prediction tool, http://gene-regulation.com/pub/programs/alibaba2/ (44) to identify potential transcription factor binding sites. We analyzed potential promoter sequences also by AliBaba2.1 TRANSFAC (45), applying the following parameters: the Pairsim function was set to 50, the matrix width set to 10 bp, a minimum number of overlapping predicted transcription factor binding sites was limited to a total number of four for each site with a minimum sequence conversation of 75% coverage, and the factor class was set to 4.

In addition, the sequence was analyzed for canonical core promotor elements by the YAPP eukaryotic core promoter predictor using a cut-off score of 0.8 (46,47). To conduct the promoter analysis, the sequence including the predicted core promoter fused with the sequence of EGFP (717 bp) was used to analyze potential transcription start sites (TSS). Subsequently, TSS prediction was performed by YAPP (46,47). In order to identify the schistosomal Kozak consensus sequence, we performed a comparative analysis applying the AGAT software-package (48,49) to determine also potential Kozak sequences of other genes in the *S. mansoni* genome version 9 (50). To this end, ubiquitously and abundantly transcribed genes were screened for the presence of potential Kozak sequences and patterns similar to SmUbi (51). Specifically, the sequences of Smp_009580, Smp_106930, Smp_042160, Smp_056970, Smp_054160, Smp_182890, Smp_099870, Smp_179300, Smp_111340, Smp_090120, Smp_072330, Smp_003770, Smp_017430, Smp_040130, Smp_155060 and Smp_335990 were compared (52).

In addition, the 3´UTR of SmUbi defined by WormBase ParaSite (40,41), version 7 of the draft genome of *S. mansoni* (3,5), was analyzed. To identify the polyadenylation signal, 600 bp downstream from the stop codon were analyzed with the DNAFSMiner-DNA Functional Site Miner polyadenylation side prediction tool (53). Subsequently, the RNA secondary structure of the predicted 3´UTR was calculated, including the stop codon and predicted rank 1 (highest) polyadenylation site within *in silico* added poly-adenosine tail (10 adenosines). The investigation of potential RNA secondary structures characteristic for translation terminator-loops was done using the RNAfold Webserver Vienna tool, which calculates secondary structures by applying the minimum free energy (MFE) model and activated partition function, while the model avoids RNA folding with isolated base pairs, http://rna.tbi.univie.ac.at (53,54).

EGFP reporter gene plasmid

Based on the pJC53.2 vector backbone (add gene, Plasmid #26536) (55), we cloned the reporter gene construct consisting of EGFP and the promoter and terminator sequences of SmUbi (Figure 1; Supplementary Table S1). To this end, 1 µg pJC53.2 plasmid DNA was cleaved using *Ahd* I (New England Biolabs (NEB), Ipswich, MA) in 50 µl 1× CutSmart buffer (NEB) at 37°C for 2 h. The two processed DNA fragments were size separated by agarose gel electrophoresis. The plasmid

fragment of 3,240 bp was recovered from the gel (Monarch Gel Extraction Kit, NEB, Ipswich, MA). For cloning, essential DNA fragments; full-length EGFP sequence, putative promoter (5´Ubi) and terminator (3´Ubi) of schistosome ubiquitin were amplified by PCR by the primers as listed in Supplementary Table S4. Primers for the amplification of the 5´Ubi and 3´Ubi were designed according to data base information of WormBase ParaSite V7 for Smp_335990 (40,41). PCR amplicon of each fragment was a synthesized in 20 µl reaction volume with either 100 ng of genomic DNA (gDNA) or 10 ng pET-30a_SmDHFR-EGFP (the EGFP-DNA fragment kindly gifted by Jim Collins) as template for the reporter gene EGFP (717 bp). Expected sizes of the specific fragments for 5´Ubi and 3´Ubi were 2,056 bp and 580 bp, respectively. PCR reactions were set up by Q5 High-Fidelity Polymerase kit (NEB) according to the manufacturer's recommendation with thermal cycles; denatured at 98°C for 3 min, followed by 35 cycles of 95°C for 30s, 55°C for 30s and elongation at 72°C for 75s. Products were analyzed by agarose gel electrophoresis and recovered as described above. This step was followed by an additional PCR to add homology arms flanking the reporter transgene, to enable Gibson assembly (56). Primers were designed with homology arms of 20 to 26 bp to both the promoter and the terminator sequences of the backbone pJC53.2 flanking the EGFP fragment. Amplification of the promoter sequence was carried out using primers for Gibson assembly (Supplementary Table S4). A two-step PCR protocol was employed, in which the template DNA was first denatured at 98°C for 5 min, followed by 10 cycles of 95°C for 30s, 54°C for 30s, 72°C for 75s, before 25 cycles followed with 62°C for 30s, and 5 min at 72°C. The entire construct pJC53.2_SmUbi-EGFP-SmUbi was assembled using the NEBuilder Hifi DNA Assembly Cloning Kit (NEB). Competent *Escherichia coli* DH5α (NEB) were transformed with the ligation products, plated on LB agar supplemented with kanamycin and ampicillin (Sigma Aldrich, Darmstadt, Germany) for 18 h at 37°C. Presence of the correct insert in plasmids was confirmed by Sanger sequencing (Microsynth SeqLab, Göttingen, Germany).

Worm transfection by particle bombardment

Particle bombardment (PB) was carried out using the stationary PDS 1000/He system (Bio-Rad, Hercules, CA) (57,58) to test the activity of the SmUbi-EGFP-SmUbi transgene construct. Plasmid

DNA was isolated by the PureLink Expi Endotoxin-Free Maxi Plasmid Purification Kit (Thermo Fisher Scientific, Waltham, MA), and precipitated onto gold particles (0.6 µm) using $CaCl_2$, spermidine, and ethanol (58); 5 µg DNA/600 ng gold. Most of the culture medium was removed, and male worms were positioned centrally in a petri dish; 10 male schistosomes were included in each biological replicate (n=5). The ballistic parameters applied were helium gas pressure at 1,550 psi, 3 cm target distance (stage 1), 15 inHg (381 mmHg) atmosphere, and 23°C. After PB, fresh culture medium was added to the worms, which soon recovered. The worms were maintained for 48 h in culture, before further analysis. To prepare microscopic analyses, the nuclei of the worms were counterstained with Hoechst 33342 (Invitrogen, Waltham, MA) for 120 min, as recommended by manufacturer. Subsequently, the worms were examined by fluorescence microscopy (IX71 Inverted Fluorescence Microscope Pred IX73, Olympus) and confocal laser scanning microscopy (CLSM; TCS SP5 vis confocal laser scanning microscope, Leica Microsystems GmbH, Wetzlar, Germany). The obtained images were analyzed with the cellSens software (Olympus) and LAS X software (Leica), respectively. Worms showing EGFP-induced fluorescence were washed in 1× PBS, transferred to 50 µl DNA/RNA protection buffer (Monarch Total RNA Miniprep Kit, NEB), and frozen in liquid N_2. Subsequently, total RNA was isolated from these worms and its concentration and integrity assessed by electropherogram analyses using Agilent RNA 6000 Nano Chips in the Bioanalyzer 2100 System (Agilent Technologies, Santa Clara, CA). Reverse transcription of 100 ng total RNA from each sample to cDNA was performed with the QuantiTect Reverse Transcription Kit (Qiagen) at 42°C, 30 min, and the reaction terminated at 95°C, 3 min. PCRs using these cDNAs were performed in 25 µl reaction volumes with 200 µM dNTPs, 0.5 µM each primer, 1 µl cDNA, and FIREPol® DNA Polymerase (Solis BioDyne, Tartu, Estonia). The EGFP transcript of 717 nt was amplified using the primers EGFP_fw and EGFP_rev (Supplementary Table S4), and the following PCR conditions: 98°C, 30s, 35 cycles of 98°C for 30s, 55°C for 30s, and 72°C for 60s; 10 ng pJC53.2_ SmUbi-EGFP-SmUbi plasmid-DNA served as a positive PCR control.

Ribonucleoprotein complexes

To prepare CRISPR/Cas-mediated editing of the AT-rich GSH1 (65.82% AT), target protospacer adjacent motifs (PAM) and adjacent sequences for both Cas9 and Cas12a were predicted using CHOPCHOP (48,59,60) applying the default function against the *S. mansoni* genome. To compare the CRISPR/Cas efficiency between these two programmable endonucleases without cleavage site-specific bias, we identified guide RNA (gRNA) sites for Cas9 and Cas12a predicted to exhibit similar CRISPR efficiency (~50%), absence of off-targets, and notably, programmed cleavages at practically identical positions on the targeted region. Both, Cas9 and Cas12a CRISPR/Cas target sites without any predicted-off-targets on the *S. mansoni* genome were used in this study (Figure 2A). We purchased the customer-defined single CRISPR RNA (crRNA) of Cas9; sgRNA containing both crRNA and tracrRNA and crRNA of Cas12a containing chemical modifications to protect the RNA from degradation by cellular RNases (Alt-R™ system) from Integrated DNA Technology (IDT, Coralville, IA). The customer-defined crRNA sequences for Cas9 and Cas12a were 5′-GAGAUCAAUUAUCUGACAAU-3′ and 5′-GAGAUCAAUUAUCUGACAAUGGAA-3′, respectively (49,50)

The Cas9-sgRNA and Cas12a-crRNA were each labeled with CX-rhodamine (Label IT® Nucleic Acid Labeling Reagents Rhodamine Kit, Mirus Bio, Madison, WI) (54,61), after which RNPs of Cas12a were assembled by mixing of 1:1 (w/w, gRNA:enzyme) Alt-R AsCas12a V3 with purified, labeled and purified crRNA; and Alt-R™ SpCas9 Nuclease V3 with labeled, purified sgRNA, as described (54). Briefly, 10 µg Cas9 and 10 µg sgRNA were mixed by pipetting in a final volume of 100 µl Opti-MEM (Gibco, Thermo Fisher, Waltham, MA) and incubated 15 min, at 23°C in the dark. Prior to RNP transfection, ~10,000 eggs were exposed to 0.25% trypsin in EDTA at 37°C for 5 min, aiming to remove the protein lining the pores of the egg (62,63), then washed twice in 1x PBS, and resuspended in 100 µl Opti-MEM in an electroporation cuvette (4 mm) (Bio-Rad, Hercules, CA, USA) and 100 µl labeled RNPs were added. Trypsinized eggs were subjected to 125 volt, 1 pulse for 20 ms with square-wave electroporation by Gene Pulser Xcell Electroporation System (Bio-Rad) or the BTX, ECM 830 Electro Square Porator (Holliston, MA) (26,64). Subsequently, eggs were cultured in complete medium as described above for 120 min

before counterstaining with Hoechst 333420 (H3570, Invitrogen, Waltham, MA, USA) for 30 min. RNP entry into the LE were monitored under using fluorescence microscopy. CX rhodamine-specific fluorescence signals were localized by CLSM at 576 nm excitation (TCS SP5 vis confocal laser scanning microscope, Leica) (54), and images processed using LAS X software (Leica). Control groups included eggs not subjected to electroporation, eggs that were electroporated solely in Opti-MEM without CRISPR materials (mock treatment), and eggs transfected with an irrelevant (scramble) guide RNA/RNP complex (64).

Donor template

The donor DNA template consisting of 3,351 bp of SmUbi promoter (2,056 bp), EGFP (717 bp) and SmUbi terminator (578 bp), and flanked by HAs of 50 bp each (Figure 3A), were prepared by PCR using 5' C6-PEG10-modified primers with 50 bp GSH1 (Supplementary Table S4). Yu et al. proved that DNA donor containing microhomology arms of 50 bp with 5′ modifications including an amine group, a C6 linker (AmC6) and polyethylene glycol (PEG10) treatment improve the efficient gene knock-in of a EGFP reporter gene using the CRISPR/Cas9 system (65). Here, we prepared a DNA donor template by PCR using primers specific to SmUbi promoter and terminator including 50 nt of HA (Supplementary Table S4), which 5´ends were chemically PEG10-AmC6-modified (65,66). Briefly, 10 µM of 5′amine C6 primers (IDT, Coralville, IA) were incubated with 1 mM Bis-PEG10-NHS ester (BroadPharm, San Diego, CA) in borate buffer (Thermo Fisher, Waltham, MA) for 18h at 23°C. Primers were desalted and dissolved in Tris-buffer using Bio-Spin 30 column (Bio-Rad) for use in PCR. These PEG10-5´AmC6-modified primers were used for the amplification of the donor template using GoTaq® Green Master Mix DNA polymerase (Promega). The SmUbi-EGFP-SmUbi construct with the 600 bp HA sequence (21) targeting GSH1 served as template DNA, with the following PCR conditions: 98°C, 5 min, 6 cycles of 95°C, 20 s, 56°C 30 s, and 72°C 3:35 min, which was followed by 25 cycles of 95°C, 20s, 68°C 3:35 min, and 72°C 3:35 min. The presence of an amplicon of the expected size, 3,451 bp, was confirmed following agarose gel electrophoresis, and isolated as described above.

CRSIPR editing at GSH1

Genomic DNA from LE were prepared as previously described (21,26). Briefly, LE were homogenized in DNAzol® (Molecular Research Center, Inc., Cincinnati, OH) using a motorized pestle (Kimble Chase, Thermo Fisher or PowerMasher II Nippi, Funakoshi, Tokyo, Japan). Nanodrop ND-1000 quantified the DNA sample. The PCR fragment flanking the expected CRISPR/Cas cut site on target GSH1 locus (~808 bp) was amplified by GSH1 specific primers (Supplementary Table S4) with thermal cycling of 98°C for 3 min, 35 cycles of 95°C for 20s, 60°C for 30s and 72°C for 60s. Agarose gel electrophoresis and staining with ethidium bromide enabled sizing the amplicons, after which bands of interest were isolated using the NucleoSpin Gel and PCR Clean-up kit (Takara Bio Inc., San Jose, CA). Sanger Direct Sequencing was performed at Azenta Life Sciences (South Plainfield, NJ) or Microsynth Seqlab (Göttingen, Germany). The CRISPR efficiency catalyzed by the Cas9 and Cas12a nucleases was estimated by Tracking of InDels by Decomposition (TIDE) (67) comparing with the control groups (mock and/or wild type).

To reveal the positions and frequencies of insertion-deletions (InDel) around the target GSH1 site, a fragment-harboring target GSH1 CRISPR/Cas cleavage site was amplified by specific primers containing partial Illumina adaptor sequences (Supplementary Table S4). Generated amplicons were suitable for the AmEZ service for next generation sequencing (NGS) (Azenta Life Sciences). PCR was carried out by following thermal cycling program: 98°C 3 min, 10 cycles of 95°C 20s, 55°C 30s, 72°C for 35s, and 25 cycles of 95°C 20s, 62°C 30s, and 72°C 35s. The evidence of InDels around target GSH1 cut site were revealed from >100,000 NGS paired reads from GSH1 edited DNA by CRISPREssoV2 and RGEN Tool; Cas-Analyzer (68,69). The CRISPResso2 analysis was evaluated using NGS reads as paired-end reads, using the optimal parameters: minimum homology for alignment on amplicon, 60%, base editing target base C and result base T, pegRNA extension quantification window size to 5. The quantification window was set to -3, and the plot window size to 30. The TruSeq2PE adapter-trimming setting was used. With CRISPR RGEN, both sense and antisense reads of the sequence reads were used. In addition, the nuclease detection mode was set to single nuclease. All data were tested against

the wild type reference, and the comparison range was set to 70. The wild-type marker was set to 5, and minimum frequency was set to 1.

Precise transgene integration

After electroporation, LE of experimental or control groups were harvested at 5- and 10-days post transfection. Genomic DNA of each sample was analyzed by PCR for 5′- and 3′ transgene integration into the GSH1. For checking the 5´terminus, we used a primer pair spanning 799 nt of the upstream region of GSH1 and part of the ubiquitin promoter (Supplementary Table S4). For the 3´terminus, primers spanning 1,290 nt of the targeted mutation, specifically a chimeric primer spanning the EGFP coding sequence and the terminator paired with a primer for the GSH1(Supplementary Table S4), were used. For KI investigation, we designed specific primers amplifying the GSH1 located in a significant distance away from HAs of the donor DNA template combined. These primers were each combined with primers specific for the donor DNA to investigate the success of transgene integration (Figure. 3B, additional information also find in Supplementary Table S4). Both PCRs were performed by using the GoTaq® Master Mix DNA polymerase (Promega, Madison, WI), 10 ng of gDNA and in a 10 µl reaction volume, as follows. PCR conditions were 98°C 5 min, 38 cycles of 95°C 30 s, 60°C 30s, and 72°C 80s. Ten biological replicates were examined. The internal control PCR was performed by amplifying a 1,466 bp fragment by using the GSH1 specific primers (Figure 2B, Supplementary Table S4). The PCR conditions did not allow amplification of the 4,732 bp fragment representing the complete reporter gene integrated into GSH1.

Quantitative PCR (qPCR) was performed by analyzing 10 ng gDNA (in triplicates) in total volume of 10 µL SYBR green PCR master mix (Bio-Rad, Hercules, CA) applying the Bio-Rad Read time system using 400 nM of EGFP-specific primers as described above (Supplementary Table S4). Thereby, pJC53.2_SmUbi-EGFP-SmUbi plasmid DNA served as a template in the qPCR to create an external and logarithmic standard curve based on the EGFP copy numbers, which ranged from 2.7 to 1.17×10^{7} copies. Subsequently, the copy numbers of integrated EGFP sequences were determined by performing qPCR from100 ng gDNA of KI LE by analyzing the standard curve (26,70).

EGFP reporter-gene integration analysis

EGFP expression was assessed using CLSM (Zeiss Examiner.Z1 AX10; Camera, Carl Zeiss AxioCam MRc) with an inverse objective at 20× magnification (Zeiss W Plan-Apochromat 20x/1.0 DIC (UV) VIS-IR 421402-9800), an approach that enabled imaging of the eggs in culture (20% FBS in DMEM) at 509 nm excitation (21). Image analysis was accomplished using ZEN 2011 (Zeiss) at days 5 and 10 following transfection using ~3,000 LE in each replicate. Subsequently, egg hatching was induced, or eggs were lysed for recovery of DNA and RNA (above). Transcripts of the transgene were amplified by RT-PCR. At day 10 following transfection, 3,000 LEs of each biological replicate were transferred to 50 µl RNAzol RT (Molecular Research Center, Inc., Cincinnati, OH), homogenized using a motorized pestle, and RNA isolated and quantified. Synthesis of cDNA from 30 µg total RNA per sample was carried out using the Maxima First Strand cDNA Synthesis Kit with DNase I (Thermo Fisher) and employed as template to confirm transcripts of the transgene using EGFP-specific primers. SmGAPDH served as the normalization control (Supplementary Table S4) and 10 ng donor dsDNA as a positive control. The PCR reaction volumes were 25 µl, with 3 µl cDNA, 1 µM each primer and GoTaq polymerase reagent (Promega, Madison, WI, USA), with thermal cycling of 98°C 3 min, 40 cycles of 95°C 30s, 55°C 30s, 72°C 60s. Products were sized-controlled by agarose gel electrophoresis and the amplicon density quantified (ChemiDoc Imaging System, Bio-Rad). Values were normalized using those of the reference gene to determine the relative EGFP transcript levels.

RESULTS

Prediction of promoter elements and episomal EGFP expression driven by *S. mansoni* ubiquitin promoter

To establish chromosomally modified schistosomes, it is necessary to integrate a selectable marker gene under the control of regulatory elements, which are active in the desired way. Because we intended to have strong and stage-independent reporter-gene activity, we searched for appropriate candidate genes, the regulatory elements of which will ensure constitutive and robust transgene expression in *S. mansoni*. Based on the results of previous transcriptomics approaches (6), we became attentive of a gene that appears to be strongly, ubiquitously (8), and

stage-independently expressed, Smp_335990 (Supplementary Figure S1). Annotation in WormBase ParaSite (41) indicated its identity as ubiquitin-like domain-containing protein (A0A5K4F9I1; UniProt). Protein structure prediction by Phyre2 (71) supported the annotation and identified Smp_335990 as an ortholog of UBB, the human ubiquitin gene (72,73) (Supplementary Figure S2). Next, the 2,056 bp region upstream of the coding sequence of SmUbi was analyzed for major promoter elements (Figure 1A). Typical promoter and enhancer elements were identified (Supplementary Table S1). Promoter analysis applying the Neural Network Promoter Prediction tool (74) predicted two core promoters at position 1,981 to 2,030 bp ; promoter rank 1 (Supplementary Tables S1 and S2) and 156 to 205 bp ; promoter rank 2 (Supplementary Table S3). Subsequently, these regions were analyzed more detailed applying AliBaba2.1 (42), TRANSFAC 4.0 (43), YAPP eukaryotic core promoter predictor (44-46,75), and by comparison with known consensus promoter elements (47,76,77) (Supplementary Tables S2 and S3). In the core promoter rank 1, typical sequence elements needed for transcription initiation were present including TATA- and CAAT-boxes, the B recognition element (BRE, the transcription initiator element (INR), and the downstream promoter element (DPE). Consensus sequences of enhancer elements including Fushi tarazu (FTz and CAAT/enhancer-binding sites (C/EBPalp,) were present as were binding sites for the serum response factor (SRF), hepatocyte nuclear factor 1 binding element (HNF-1), heat shock transcription factor 1 (HSF), and octamer-binding transcription factor 1 (Oct-1). Furthermore, the start-codon containing Kozak sequence, required for ribosome binding and translation initiation (78), was identified *in silico* by comparing the potential Kozak sequences of *S. mansoni* genes known to be abundantly transcribed and sharing comparable transcription strengths (51,79) and expression patterns (6,51) to that of SmUbi. For these genes, we predicted the consensus sequence 5´-NNT/CA/GNA/G/TATGNC/AN-3´ 6 bp downstream and 5 bp upstream of the translation start (Supplementary Table S2).

We also surveyed the potential terminator sequences including the 3´UTR (3,5) of SmUbi (Figure 1A; Supplementary Table S3). Applying the DNAFSMiner prediction tool for the analysis of functional sites in DNA sequences (53), a polyadenylation signal within the SmUbi 3´UTR was apparent. Moreover, RNA secondary structure prediction by RNAfold (80) revealed several stemloop structures that are typical for transcription termination (81). In addition, to test

promoter activity of the cloned SmUbi-EGFP-SmUbi reporter gene, adult male schistosomes were transfected with a plasmid construct by particle bombardment (PB) (Figures 1B, 1C and 1D), a ballistic gene-transfer method suitable for the transient transformation of adult and larval schistosomes (57,82). Green fluorescing regions were seen by 48 h later (Figure 1B, C, D), and the EGFP transcript expression was confirmed by RT-PCR (DNase I-treated RNA) in SmUbi-EGFP-SmUbi plasmid transfected worms (Figure 1E).

Programmed editing at GSH1 in the schistosome egg

The *S. mansoni* GSH1 is an AT rich (65.82%) sequence, which limits the use of the GC-content dependent Cas9 for genome editing, whereby Cas12a is best suited for editing AT-rich regions. Furthermore, our previous study proved that overlapping of 3 gRNAs combined with Cas9 cleaved the DNA into staggered ends (3,21,26). This led to a notable increase of donor DNA integration by HDR by up to 70% into LE. Here, we hypothesized that DNA cleavage by Cas12a combined with a single crRNA might result in staggered DNA ends, which could increase the efficiency of gene editing and lead to higher InDel frequencies mediated by NHEJ and HDR compared to Cas9 combined with a single gRNA. To prove this hypothesis, we designed gRNAs of bothCas12a and Cas9 with the same expected cut site (Figure 2A). Furthermore, gRNAs were selected sharing the same predicted CRISPR/Cas-efficiency, and which showed no predicted off-targets.

To establish genome editing in schistosomes by chromosomally integrated reporter genes, a suitable life stage(s) of the parasite should be chosen, a compatible method to introduce gRNAs, replacement constructs in case of KI experiments, and the optimal enzyme to efficiently edit the target region of choice. To this end, the egg stage was used as biological target, given its suitability for electroporation-mediated transformation. First, the accessibility of LEs by RNPs of both enzymes was determined using an RNP-tracking approach, where fluorescence-labelled RNPs were introduced into LEs by electroporation. For this, each guide RNA (sgRNA of Cas9 or crRNA of Cas12a) was labeled with a CX-rhodamine fluorescent tag prior to assembly of RNP complexes with either *Streptococcus pyogenes* Cas9 (SpCas9) or *Acidaminococcus* sp. Cas12a (AsCas12a) nucleases (26,65,66). Subsequently, LE were immediately exposed to these RNPs by

square-wave electroporation. To confirm the entry of RNPs into LEs, we employed confocal laser scanning microscopy (CLSM). Within 24 h post electroporation, we observed RNP introduction into the larval tissue and colocalization with Hoechst 33342-stained nuclei for both Cas-enzymes (Figure 2B). Here, we proved the successful delivery of Cas12a-RNPs into LEs by electroporation. Labeled RNPs were able to penetrate the eggshell and were detected in larval cells. For Cas9-RNP, we observed that the majority of Cas9-RNP passed through the eggshell, whereby comparably lower amounts of labeled RNPs were found in larval cells. From these results, we concluded that both RNPs were successfully delivered into larval cells inside the eggs. Moreover, the majority of Cas9 RNPs appeared to accumulate in the eggshell, whereas a wider distribution of signals was observed in Cas12a RNP-treated LEs.

After demonstrating the accessibility of LEs to RNPs of both enzymes, LEs were transfected with unlabeled, freshly prepared RNPs of either Cas9 or Cas12a to demonstrate the accessibility of GSH1 and the cleavage capability of RNPs containing previously designed crRNAs. Ten days after electroporation, genomic DNA was extracted and processed for downstream analysis by PCR for the execution of target InDel analysis by Sanger Direct sequencing followed by TIDE analysis as previously described (26,67). From nine independent biological replicates (sample no. 1-9) applying TIDE-analysis (67), we obtained editing efficiencies of 2.2±0.9% and 3.0±0.9% resulting from Cas9-RNP and Cas12aRNP, respectively. The wild type (untreated), the electroporation control (mock), and LEs treated with RNPs formed by a crRNA with a scrambled sequence; GCACTACCAGAGCTAACTCA (64) were included for this study to investigate background mutations and CRISPR/Cas off-target activity on GSH1. After normalizing the obtained sequences to the background noise, which occurred as an artefact of the sequencing procedure, we subtracted the editing efficiency of the experimental groups (RNP treatment) with the values obtained for wild type and mock controls (~0.1%), to adjust the editing efficiencies. The mutation rate in control group LEs treated with RNPs containing crRNA with a scrambled crRNA sequence was 1.4±0.7%, significantly lower than Cas9 and Cas12a programmed to cleave GSH1 (Figure 2C). Furthermore, we combined gDNAs (after TIDE analysis above) from sample no. 1-3, 4-6 and 8-9 for target amplicon NGS sequencing. NGS paired-reads (~100,000 reads of each sample) from three independent NGS libraries were analyzed by both CRISPRessV2 and

Rgen Tools confirmed the precise GSH1 editing by both enzymes (Figure 2D and 2E) (59,68,69). In this study, we revealed nucleotide deletion on expected cleavage site showing deletion sizes of 1, 7 and 13 bp by Cas12a, and 1, 11 and 13 bp by Cas9 (Figure 2D, 2E). Obtained GSH1 deletion patterns were similar from three NGS libraries.

Additionally, comparable data were shown for the editing efficiency of the GSH1 locus of treated adult *S. mansoni* couples using Cas12a. Here, a significantly higher editing efficiency was shown for Cas12a-treated adult worms compared to Cas9-treated worms (Supplementary Figure S3A). Subsequently, the editing efficiency Cas12a was 100-fold greater compared to editing by Cas9. The size of deleted DNA fragments of GSH1 ranged from 9-18 nt in the Cas12a-treated group, with 1 nt deleted by Cas9 (Supplement Figure S3B and S3C). These findings highlighted the capability of successful genome editing for both Cas nucleases in *S. mansoni* eggs and adults, but in addition our results demonstrated a significantly higher editing efficiency of Cas12a over Cas9.

Precise integration of transgene flanked by microhomology arms

We further investigated whether both enzymes equally enable transgene integration into GSH1 resulting from CRISPR mediated-HDR. Here, we focused to use the 5′ modifications of a double-strand DNA donor as a template for transgene KI into GSH1 with amine C6 with PEG10 treatment at 5' termini of DNA donor containing microhomology arms of 50 nt. In independent studies, 5′C6amine-PEG10(C6-PEG10)-modified donors (with ~50 nt HA) showed highest HDR efficiencies to edit a target gene of HEK293 cells with single gRNAs (65,66). Here, we investigated the utility of PCR amplicons as the donor template for knock-in of the SmUbi-EGFP-SmUbi transgene into GSH1 catalyzed by Cas9 and Cas12a. For both nucleases, transgene integration was achieved at ratios of 1:1:1 (w:w:w) of enzyme: Cas9 sgRNA or Cas12a crRNA: donor EGFP in each case. The 5′ HA of the donor template contained the Cas12a PAM recognition site (TTTC) and a partial proportion of the Cas12a target sequence, one base of the Cas12a PAM TTTC was modified to T<u>C</u>TC, to eliminate the Cas12a PAM sequence. This was done to obviate donor-DNA cleavage by Cas12a RNPs. In addition, the design of the gRNAs and selection of the corresponding PAM recognition sites of both nucleases accomplished transgene

integration by using the same HA-sequences of the donor template targeting GSH1. This was achieved by the selection of predicted and almost identical cleavage sites of both Cas-enzymes at GSH1 (Figure 2 A). For transfection of LE with SmUbi-EGFP-SmUbi, we co-electroporated RNPs and donor template. Mock electroporated wildtype LEs (mock) served as a transfection control, which were transfected with the donor template in absence of RNP. Ten days post electroporation, the eggs were collected, and genomic DNA was isolated and analyzed for transgene integration. For this, we designed two primer pairs amplifying (i) the 5´- integration site including sequences within the GSH1 as well in the SmUbi-promoter region of the transgene (5′-KI, 799 bp, Figure 3B), and (ii) amplifying the 3´-integration site including the 3´sequence of the EGFP and GSH1 (3′-KI, 1,290 bp, Figure 3B). With both prime pairs, we amplified 5′-KI and 3′-KI integration fragments and thereby demonstrated GSH1 editing and KI by both nucleases (Figure 3C). Compared to amplicons obtained by Cas12a-treated LEs, the band intensity of 5′-KI amplicons in the Cas9 group was less intense (Figure 3C). Nonetheless, Sanger sequencing confirmed the integration of the transgene at the programmed cleavage site. Moreover, the precise integration of SmUbi-EGFP-SmUbi into GSH1 resulting from HDR was demonstrated by alignment analysis of these sequences, which were obtained from the purified 5′-KI and 3′-KI amplicons (Figure 3D). As expected, no 5′-KI and 3′-KI fragments were amplified with DNA of the control groups. A 1,466 bp fragment of the unedited GSH, which was amplified in all experimental groups, served as a PCR control.

Next, we quantified the copy number of SmUbi-EGFP-SmUbi transgenes integrated into the genome in edited LEs compared to LEs, which were transfected only with the donor template in the absence of RNPs (Figure 3E). For this, the external standard curve design was employed to estimate integrated-EGFP in genomic DNA of the schistosome eggs (n = 3) as described (26). The standard curve is based on linear regression ($y = 3.3857x+35.093$, $R^2 = 0.9979$) established by a logarithm(20)-transformed initial DNA input, the pJCR53.2-SmUbi-EGFP-SmUbi plasmid was used as the dependent and the Ct value from the qPCRs as the independent variation (not shown). We estimated the EGFP transgene copy number in GSH1 by converting the measured Ct values to the logarithmic copy numbers using the equations obtained from the standard curve. There were 6.594±0.196 and 7.235±0.458 EGFP copies in ~10 ng DNA resulting from Cas9 and

Cas12a, respectively. The control group (donor template only) showed 0.516±0.082 copies (10 ng of gDNA), which probably reflected residual non-integrated SmUbi-EGFP-SmUbi donor DNA. Together with this finding, EGFP transcripts were not detected by RT-PCR of the RNA/cDNA of the control groups, and only found in KI in DNase I-treated RNA samples (Figure 3F). In summary, the number of integrated SmUbi-EGFP-SmUbi copies was significantly higher in Cas12a compared with the Cas9 groups of LEs.

EGFP expression in the miracidium

To evaluate EGFP-reporter gene expression in eggs/miracidia obtained from KI experiments, LEs were analyzed after 5 and 10 days after transfection by CLSM. Differences in the emission spectra were detected and analyzed as described (21). This analysis revealed EGFP expression in numerous cells of the transfected eggs including miracidia, which had developed in some of these eggs. Fluorescence intensity at an emission wavelength of 509 nm was recorded and quantified in eggs from the editing experiments using Cas9 and Cas12a. With respect to fluorescence intensity, eggs of the Cas12a group significantly outnumbered eGFP-positive eggs of the Cas9 Group at days 5 and 10 post electroporation (Figure 4A). For this analysis, only eggs containing a miracidium with >50% EGFP expression in the investigated section plane (two dimensions) were counted as EGFP-positive. No differences in terms of growth and survival were noted between successfully transfected eggs and eggs from the control groups. Furthermore, the viability of treated eggs was demonstrated by successful hatching of miracidia (not shown). From six biological replicates, the percentages of EGFP-positive eggs averaged at 43.6 ±19.9%, 26.3±7.2%, in the Cas12a group at days 5 and 10, and at 20.2±10.1%, and 21±16.2% in the Cas9 group at the same time points (Figure 4B). A weak EGFP signal at 509 nm was detected in LEs transfected with the donor only (no RNP) (Figure 4B). Subsequent RT-PCR analyses with egg RNA/cDNA of all groups confirmed EGFP transcripts in the treatment groups but not in the control groups, except a weak EGFP background signal in LEs transfected with the donor template in the absence of RNPs (Figure 3C).

The EGFP intensity at 509 nm, which represents the EGFP-emission peak as shown in Figure 4C, of miracidia inside LEs were determined 10 days post electroporation, whereby autofluorescence

signals were subtracted. One hundred eggs of each experimental group were investigated for EGFP signal-intensity from live larvae inside the eggs, as described previously (21). The EGFP-specific signal intensities from larvae inside the eggs, which resulted from SmUbi-EGFP-SmUbi integration, - were not significantly different between the CRISPR-Cas9 and -Cas12 groups with average intensities of 6,848.71±1,169.82 and 6,897.25±1,163.45, respectively. No EGFP signals were detected in the control groups, donor only, and mock transfections (Figure 4D).

DISCUSSION

Neglected tropical diseases affect more than one billion people, frequently residents of in some of the world's most impoverished communities and regions. Despite the devastating health consequences, these diseases receive too little attention from global funding agencies, pharmaceutical companies and biotech research and development. They can be difficult to study given many are not laboratory-accessible in that they can have multiple developmental hosts in the life cycle, they may hypoxic or other non-standard physicochemical culture, and they possess complex diploid genomes (83-86). These challenges notwithstanding, to continue to advance functional genomics of helminths (87-89), here we report progress in genome editing at the genome safe harbor site GSH1 in *S. mansoni* (21) by comparing the editing capacities of Cas9 and Cas12a, and optimized RNP delivery conditions and donor DNA template structure. We obtained efficient integration of a reporter gene into GSH1 using 5´-C6-PEG10 modified 50 bp microhomology arms by programmed CRISPR-Cas9/Cas12a-induced HDR repair, as confirmed by amplicon sequencing as well as fluorescence reporter gene activity.

Knockout models are widely used for functional studies of gene(s) of interest. However, in the *omics* era of schistosome and other helminth research, facile protocols for stable transformation are not readily available. CRISPR/Cas gene editing is now the state of the art for genome manipulation (90), and there are recent reports involving CRISPR gene editing in trematodes (26,61,87,91). Genome editing was limited to eggs, and genome manipulation probably accessed the outer cell layers during embryogenesis (36,92,93). Yet, to establish stably transformed parasite lines, approaches need optimize germline transgenesis. We had established a protocol for RNA-programmed, Cas9-catalysed editing at GSH1 in *S. mansoni* that

used multiple, overlapping guide RNAs, after computation prediction of the genome coordinates for this GSH based on genome annotation and chromosome structure followed by experimental validation (21). Here, we boosted both this method and enhanced gene-editing efficiency using a single gRNA-programmed Cas12a in the presence of a donor repair template with chemically modified, microhomology arms that, in unison, enhanced programmed transgene insertion.

Both LEs and adult schistosomes were transfected by electroporation of Cas-containing RNPs. Treatment of adult schistosomes with RNPs transferred by electroporation failed to achieve a penetration and distribution RNPs in all tissues (61). Indeed, exposure of adult stage schistosomes to CRISPR by electroporation can be expected to result in genetic mosaicism. Our findings conformed with earlier reports of deletion mutations at CRISPR targets in LE, schistosomula, and adult schistosomes (26,61,87,94,95). Nonetheless, the outcome can be informative for proof-of-principle studies of genome editing or as an alternative for RNAi of genes transcribed in the (sub)tegument.

To overcome these problems and to establish a genetically homogeneously modified multicellular parasite, it is necessary to target life stages with a limited number of cells, in the best with many stem/germinal cells, such as schistosomula (9), or at a very early developmental stage such as the egg (36). For this purpose, we aimed to introduce RNPs into LEs, which were isolated from livers of infected final hosts, and which consisted of egg subpopulations representing different developmental stages. Importantly, the LE population included also non-embryonated eggs, which consisted of a zygote and a few vitelline cells (36). Confocal microscopy of LEs transfected with RNPs, which incorporated rhodamine-labelled Cas9-gRNAs (61) or Cas12a-gRNAs, demonstrated the suitability of this approach for permeation of the eggshell, whereby Cas12a RNPs showed a higher efficiency in accessing embryonic cells. Moreover, transfection of the zygotes of eggs at an early developmental stage was visualized by RNP-tracking experiments (Supplemental Figure S5). The higher efficiency might reasonably be explained by the size difference of RNPs given that the Cas12a-RNP complex (Cas12a: 150.9 kDa) was smaller than Cas9-RNP complex (Cas9: 161.3 kDa) (30). This likely resulted in a

comparatively better uptake of the Cas12a-RNP complex through the pores in the egg (96). Moreover, Cas12a gRNA (41-44 nt) is smaller than Cas9 gRNA (100 nt) (30). Furthermore, after delivery we observed RNPs of both Cas-enzymes in numerous cells of the eggs. A limitation of this approach was that not every cell of egg appeared positive for the RNPs. To optimize this, alternative transfection methods such as lipid nanoparticles (97) and other methods (82,98-102) could be tested. Alternatively, one could envisage to use larval stages such as sporocysts as editing templates to enhance cell-transfection efficiency (61).

In addition to the observed differences in the efficiency of eggshell penetration between the two Cas enzymes, analysis of the deletion frequencies of GSH1 revealed a significantly higher editing efficiency of Cas12a over Cas9 in eggs and adult schistosomes. The editing efficiency of Cas12a with a single gRNA was comparable to the efficiency of Cas9 with two to three gRNAs (21). The differences might result from i) the efficiency bias of Cas12a in penetrating the eggshell, ii) differences in the cleavage profiles of the nucleases (Cas9, blunt-ended; Cas12a, staggered ends) (30), iii) differences in the efficiency of NHEJ, which relies on the gRNA sequence and current cell-cycle state of the cells during treatment (103-105), and/or iv) combinations of these factors.

Intriguingly, the deletion patterns of Cas12a and Cas9 in eggs substantially differed from those found in adult schistosomes. DNA repair preferences may differ between eggs and adults. The schistosome egg in utero is largely composed of vitelline cells, S4 cells (106), and a fertilized oocyte in the early phase, or the miracidium with surrounding acellular and cellular layers, Reynolds' layer and von Lichtenberg's envelope, respectively (62), at a late stage (36,107). We speculate that these cell types which perform discrete functions differ in DNA repair processes. Findings in other metazoans support the conjecture (108-111). Aim to buttress support for this hypothesis, we performed a first i*n silico* approach to identify orthologs of genes that are part of the double-strand DNA break repair machineries (NHEJ and HDR) in *S. mansoni* based on data available for the nematodes *Caenorhabditis elegans* and *Pristionchus pacificus* (Supplementary Table S5 and Supplementary Figure S6) (112-117). Furthermore, no schistosome proteins presumably involved in double stranded break (DSB) repair showed high similarities to mammalian DSB-repair pathways (105,118). Analysis of the expression patterns of *S. mansoni*

protein-coding genes, predicted to be involved in DSB-repair (7,51,79), showed clear differences in the tissue-preferential transcript profiles of genes attributable to NHEJ (*e.g.*, SmCKU70, SmCKU80) or HDR (*e.g.*, SmHIM6, SmRAD51, SmRPA). NHEJ-associated genes were transcribed throughout the tissue whereas HDR-associated proteins were preferentially transcribed in gonads (7,111,119). Genes predicted to be involved in HDR, such as SmRAD51, SmRPA and SmBRC1 were not transcribed in S4 vitellocytes (51). These variations may explain the differences in observed InDel-rates and -patterns between the egg and adult stages of this schistosome. In addition, the main repair mechanism in proliferating cells, such as embryonic cells, is expected to be HDR (105), where deletions are not expected to accompany perfect DSB repair (109). Correspondingly, increased transcript levels of HDR-associated genes during the parasitic life cycle was noted for the larval stages, including the sporocyst, which exhibit high rates of cell proliferation (Table S6) (79,120). By contrast, cells and tissue that are terminally differentiated without further need for proliferation, such as S4 vitelline cells, are likely to use NHEJ (105,118) for DNA repair (103,108).

We also demonstrated the utility of microhomology arms for CRISPR-Cas-induced HDR repair, as demonstrated by constitutive reporter gene expression in embryonating miracidia. To this end, Cas9 and Cas12a were each applied together with a 5'C6-PEG10-modified donor template encoding an EGFP reporter-gene under control of the strong native ubiquitin promoter of *S. mansoni* (21). We analyzed this promoter in more detail and identified several core promoter elements (76,77,121-123), including a schistosomal Kozak sequence (78). The promoter elements identified may contribute to the constitutive transcriptional activity of the affiliated ubiquitin gene throughout the schistosome life cycle (79). Recent studies have determined the limitation of transgene integration by 4 kb by HDR (21,124,125), whereas the use of arms of only 50 bp in length, i.e., "microhomology" arms deployed here, can facilitate increase the transgene cargo length up to 9 kb (66,126-128).

For reporter gene integration, we used a 5´-C6-PEG10-modified donor DNA template to prevent DNA degeneration by endonucleases and to enhance HDR effects as shown in human cell lines before (HCT116, HEK293T, H1, and WTC G3) (66). For human cell-line genome-editing by Cas9,

superior transgene integration- was observed for the donor template when it was combined with 5'C6-PEG10 modification and short HAs (65,66). Transgene insertion into GSH1 using this adapted approach for schistosomes represents a significant optimization over our original protocol (21) as both HAs were significantly shorter and dsDNA could be used instead of a single-stranded DNA donor. Likewise, by using RNPs of Cas12a and just a single crRNA, we achieved comparable and higher integration rates with the modified donor template compared with Cas9 and up to three gRNAs (21). Cas12a was preferable to Cas9 for the integration of a transgene. Furthermore, Cas12a is particularly well-suited for *S. mansoni* because of its preference for AT-rich genomes through the specific PAM 5´-TTTN-3´ (3,30). In overview, Cas12a is more flexible in terms of applicability compared to Cas9, which is more restricted by its PAM 5´-NGG-3´ in selecting suitable gRNA binding sites in a target gene.

Optimized genome editing for schistosomes will be a useful addition to post-genomic research. Our previous report defining a panel of genome safe harbors in *S. mansoni* provided a salient innovation for CRISPR-based functional genomics and transgenesis in *S. mansoni*. Now, we have also overcome other earlier limitations through deployment of Cas12a as the RNA-guided nuclease and the modification of the donor transgene with chemically protected, diminutive length (microhomology) arms. Along with loss-of-function mutation, these improvements will allow gain-of-function approaches by e.g., integrating selectable marker genes into a GSH as an alternative for, or in addition to, fluorescence reporters. This would enable maintenance of transgenic lines with drug selection pressure. Indeed, oxamniquine showed selectivity for sporocysts and adult schistosomes (61). Beyond that, this model could be used for testing target genes/proteins for novel of helminth parasites as potential targets with the goal of developing new drugs and vaccines to treat and prevent schistosomiasis and other similar diseases. This method also opens the way for studying genes inaccessible to RNA interference (17,19,129). We respectfully posit that this new workflow can provide a blueprint for the design of more advanced, tissue-stage investigation and contribute to deciphering the complex biology of schistosomes, which in turn, provides a role model for parasitic platyhelminths generally.

DATA AVAILABILITY

The nucleotide sequence reads are available at the NIH Sequence Read Archive, BioProject PRJNA950942, accessions SRR24035153 to 24035171 and SRR24034375 to 24035409. All other data are available in the main text and the supporting information.

AUTHOR CONTRIBUTIONS

Max F. Moescheid: Methodology, Formal Analysis, Validation, Writing-original draft, Writing-review & editing. Prapakorn Wisitpongpun: Methodology, Formal Analysis. Victoria H. Mann: Methodology. Thomas Quack: Methodology. Christoph Grunau:GSH1 identification, Writing-review & editing. Christoph G. Grevelding: Conceptualization, Supervision, Writing-review & editing. Wannaporn Ittiprasert: Conceptualization, Supervision, Formal Analysis, Validation, Writing-original draft, Writing-review & editing. Paul J. Brindley: Conceptualization, Supervision, Writing-review & editing

ACKNOWLEDGEMENTS

Schistosoma mansoni-infected mice were provided by the Schistosomiasis Resource Center of Biomedical Research Institute, Rockville, MD through NIH-NIAID contract HHSN272201700014I for distribution through BEI Resources. The schistosome lifecycle at the Institute of Parasitology in Giessen was maintained by Christina Scheld and Georgette Stovall (Institute of Parasitology, Biomedical Research Center Seltersberg (BFS), Justus Liebig University Giessen). We thank Oliver Puckelwaldt (Institute of Parasitology, Justus Liebig University Giessen, Germany) for bioinformatical support analyzing the schistosomal Kozak-sequences. We thank Dr. James J. Collins for kindly supplying the EGFP reporter gene fragment. The scholarship of Max Möscheid was covered by in part by the Wellcome Trust FUGI grant 107475/Z/15/Z, the International Giessen Graduate Centre for the Life Sciences (GGL) of the Justus Liebig University Giessen and the German Society of Parasitology. For open access, the author applied a CC by public copyright license to any author-accepted manuscript version arising from this submission. This study also is set within the framework of the Laboratoire d'Excellence (LabEx) TULIP (ANR-10-LABX-41) with the support of LabEx CeMEB, an ANR Investments d'avenir program (ANR-10-

LABX-04-01) and the Environmental Epigenomics Core Service at IHPE (Interactions Hôtes Pathogènes Environnement). Award number RR025565 (PI, Anastas Popratiloff) from the NIH Office of Research Infrastructure's S10 program enabled the purchase of the Zeiss 710 confocal microscope.

FUNDING

This work was funded in part by the Wellcome Trust grant 107475/Z/15/Z (Flatworm Functional Genomics Initiative, PI, Karl F. Hoffmann). For open access, the author applied a CC by public copyright license to any author-accepted manuscript version arising from this submission.

CONFLICT OF INTEREST

All other authors declare they have no competing interests.

FIGURE LEGENDS

Figure 1. Bioinformatic prediction and experimental verification of the regulatory activity of *S. mansoni* ubiquitin (SmUbi) promoter and terminator. **(A)** Overview of SmUbi (Smp_335990) promotor and terminator elements (additional details can be found in Supplementary Tables S1and S2). The elements were identified using Berkeley *Drosophila* Genome Project Neural Network PromoterPrediction Tool, AliBaba2.1 TRANSFAC 4.0 (43), and YAPP eukaryotic core promoter predictor (44-46,75), respectively. Two core promoters were predicted (brown box and arrow) containing a CCAAT enhancer-binding site (light blue arrow), hepatocyte nuclear factor 1 binding element (NHF-1), B recognition element (BRE), TATA-Box-containing region (TBP), transcription initiator element (INR), serum response factor (SRF), and downstream promoter element (DPE). The presumptive terminator region contained three terminator stem-loop structures and the polyadenylation signal, which were predicted by the RNAfold Webserver Vienna tools (80) and the DNAFSMiner-DNA Functional Site Miner polyadenylation side prediction tool (53), respectively. SmUbi promoter-induced EGFP reporter-gene expression was monitored in male *S. mansoni*, which had been transiently transfected by particle bombardment (PB) (101) with the Ubi-EGFP-Ubi plasmid construct. Following cloning of the SmUbi-EGFP-SmUbi reporter-gene construct (using the rank 1 core promoter sequence) into the plasmid vector pJC53.2 (55), adult male schistosomes were transfected by PB. After 48 h, transfected as well as untreated control worms were counterstained with Hoechst33342 (blue) and analyzed by CLSM **(Panels (B) and (C))**. Green fluorescing regions were observed by fluorescence microscopy, while no EGFP signal were found in control worms. As expected, brightfield microscopy also confirmed the presence of clusters of gold particles in the tissue (white arrows, **Panel (D)**). EGFP signals matched with the presence of gold particles (green arrows, **Panel (D)**). **(E)** RT-PCR confirmed the presence of EGFP transcripts only in transfected worms. No EGFP transcripts occurred in wild type (WT) worms (control) or in a control without template. +ve = positive RT-PCR control; H_2O, control RT-PCR without template. *Abbreviations: CLSM, confocal laser scanning microscopy; EGFP, enhanced green fluorescent protein.*

Figure 2. Editing GSH1 in *S. mansoni* showed an efficiency bias for Cas12a vs. Cas9 and prominent deletion patterns at this target locus. Synopsis of the comparative approach using Cas9 and Cas12a for programmed gene editing with the *S. mansoni* egg as biological target for transformation by electroporation. **(A)** Shown is a schematic overview of the GSH1 (SM_V7_Chr3; 13380432-13381848; 1,416 bp) target sites for Cas9 (blue arrow) and Cas12a (red arrow). DNA cleavage by Cas9 was expected to result in a blunt-ended cut, whereas Cas12a digested released overhanging DNA strands. The non-target DNA strand at GSH1was expected to be cleaved at the same position as illustrated. GSH1 regions highlighted in magenta represent the short homology arms, 50 nt used for the donor template. **(B)** LEs were subjected to electroporation in the presence of CX Rhodamine-labelled RNPs (red), after which they were counterstained with Hoechst33342 (blue) and examined by CLSM 24 h later. Introduction of RNPs was demonstrated (arrows) with both nucleases. The majority of Cas9 RNPs accumulated around the eggshell, a dissimilar pattern to that observed in the Cas12a RNP-treated LEs, where signals were also found inside the eggs. **(C)** Results of the determination of editing efficiencies by InDel abundance generated by Cas9 and Cas12a RNPs in GSH1, as analyzed by TIDE analysis (67). The editing efficiencies of GSH1-targeting RNPs normalized to the wild-type sequence (KO-Cas9, dark blue bar; KO-Cas12a, red bar) and RNPs non-specific to the *S. mansoni* genome (scramble-Cas9, purple bar; scramble-Cas12a, orange bar), respectively are shown. The editing efficiency of LEs treated with RNPs were significantly higher compared to LEs from the scramble control (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$; t test, $n = 9$). Compared with Cas9-RNPs, treatment with Cas12a-RNPs (red) showed a statistically significant higher editing efficiency. **(D)**, Representative examples of NGS reads or Cas12a and Cas9, respectively, to show alleles with frequently seen profiles of deletions at the programmed cleavage positions e (red arrows) ($n = 3$). Abbreviations: CLSM, confocal laser scanning microscopy; KO, knock-out, ns, not statistic significant; wt, wild type.

Figure 3. Successful RNP-mediated chromosomal integration of the SmUbi-EGFP-SmUbi reporter-gene with microhomology arms into GSH1. **(A)** Illustration of the HDR- mediated integration of the SmUbi-EGFP-SmUbi reporter-gene into GSH1. Shown are the *S. mansoni* ubiquitin promoter (SmUbi; 2,056 kb) and terminator (SmUbi; 302 bp) in orange, flanking the

EGFP reporter (green; 717 bp). The microhomology arms (50 nt) are indicated in pink. To avoid Cas12a-mediated cleavage of the SmUbi-EGFP-SmUbi donor template, the 5′-TTTV-3′ Cas12a PAM inside the microhomology arms was changed to 5′-TCTV′3′ (red letters). Furthermore, the donor template was 5´C6-amine and PEG(10)-modified (65). **(B)** Shown is a schematic overview of the GSH1 locus before and after successful KI of the SmUbi-EGFP-SmUbi (orange and green) reporter-gene construct. The illustration displays the PAM site (red) and primer binding sites for KI analyses by PCR. **(C)** The success of the chromosomal integration of the SmUbi-EGFP-SmUbi donor template was demonstrated by PCR, amplifying the 3'- and 5'-integration sites. The WT locus of GSH1 served as control, whereas the complete edited regions were not amplified due to time limitation of the PCR. DNA fragments specific for the 5´and 3´ integrations of the donor template were obtained by PCR and sequenced by Sanger direct sequencing. **(D)** For both Cas enzymes, alignment analyses of these fragments confirmed the successful integration of the donor template (yellow and green) into GSH1 (grey), as indicated by red bars. **(E)** Estimation of EGFP transgene copy-numbers in genomic DNA (day 10 post electroporation) based on applying the external standard curve plotted from plasmid DNA-containing DNA, as described (26). The EGFP copy numbers in the Cas9 and Cas12a groups were significantly higher compared to copy number in LEs transfected with the control donor template only (no RNP). KI analysis of the Cas9 group showed a very wide range of EGFP copy numbers compared to the Cas12a group (2Way ANOVA, ***P = 0.0001, **** P< 0.0001). Significantly lower copy numbers of residual (episomal) donor DNA were detected in the control donor-only control; grey violin plot). **(F)** RNA/cDNA of eggs of all experimental groups were analyzed by RT-PCR to monitor EGFP transcription. RT-PCR showed EGFP transcripts (236 bp) in LEs only the from KI experiments using Cas12a and Cas9. DNase I treated-RNA was used for RT-PCR process to eliminate the contaminated DNA and/or remaining donor. Lanes b1 and b2 indicate the outcome from two independent biological replicates. No EGFP transcripts occurred in the control groups (top panel). The housekeeping gene GAPDH (Smp_0569701, 285 bp, bottom panel) served as control for successful cDNA synthesis and RT-PCR in each case.

Figure 4. HDR induced by CRISPR/Cas12a resulted higher efficiency of transgene integration into the GSH1 locus of schistosome cells compared to CRISPR/Cas9. **(A)** Inside the eggs, green-

fluorescent cells and miracidia were observed by CLSM at days 5 and 10 days after treatment with Cas12a or Cas9 RNPs targeting GSH1. Autofluorescence is shown in orange. Scale bars: 50 µm. **(B)** The proportion of EGFP-positive eggs was examined by CLSM analysis, which revealed the proportion of eggs with EGFP-positive larvae at 5 and 10 days after treatment with RNP plus donor DNA. More than 10,000 eggs of each biological replicates (n = 6) were studied. Comparing Cas9 and Cas12a (red) displayed higher EGFP-positive rates in LEs after 5 day. Cas12a- and SpCas9 (blue)-treated eggs contained significantly more fluorescing larvae as the donor-only control (grey). No fluorescence signals occurred in untreated eggs. *P<0.05, **P<0.01, ***P<0.001 by t-test. **(Panels (C) and (D))** Verification of EGFP signals in LEs transformed with Cas12a RNPs after 5 days. Using the Zeiss Zen blue 3.4 software, different positions of the fluorescence spectrum were sampled to distinguish them from the background fluorescence (orange). The typical EGFP spectrum (left part; blue and green lines) was evident in the larvae (right part; green and blue + signs) and was distinct from the autofluorescence spectrum (right part, red + sign) of the egg, which mainly occurred in the eggshell (red + sign). The non-HDR EGFP-positive miracidium from donor only and mock showed fluorescent intensity at ~1333.84±235.39 a.u. and 1331.12±1169.31 a.u., respectively. The EGFP-KI mediated by Cas12a showed higher signal intensity than Cas9 at 6897.25±1169.31 a.u. and 6548.71±941.11 a.u., respectively. The EGFP signals from KI_Cas12a eggs were statistically significantly different from signals of KI_Cas9 eggs by ordinary one-way ANOVA ****P< 0.0001, **P = 0.0079 (n = 100). The spectral intensity of the autofluorescence in donor only (no CRISPR materials) and mock electroporated-eggs were not statistically significant. *Abbreviations: auto fl., auto fluorescence; CLSM, confocal laser scanning microscopy; CTRL, control; HA, homology arms; KI, knock-in; LE, liver eggs; primer fw, forward primer; primer rev, reverse primer.*

SUPPLEMENTARY INFORMATION

Supplementary Figure S1. Protein structure modelling and sequence analysis confirmed Smp_335990 (SmUbi) as a human tetra-ubiquitin orthologue. **(A)** The protein structure of SmUbi was calculated by Phyre2 (https://doi.org/10.1016/j.retram.2022.103333) based on previously known structures of orthologous ubiquitin proteins (72,73). **(B)** The amino acid sequence of SmUbi was compared to characterized tetra-ubiquitin orthologs from model organisms, such as *Homo sapiens*, using Phyre2. Using alignment coverage, SmUbi was verified as an ortholog of human tetra- ubiquitin.

Supplementary Figure S2. Existing transcript data for Smp_335990 (SmUbi) revealed its ubiquitous transcript pattern in both sexes and all cell types of adults and in the different life stages of *S. mansoni*. Compilation of expression data for Smp_335990 (SmUbi) based on different RNAseq data sets that comprised **(Panel (A))** transcript levels in adult schistosomes and their gonads (6). **(B)** The transcript levels across all life stages of (paired (top) and unpaired (bottom) *S. mansoni*, and the cell type-independent transcript profile of Smp_335990 as disclosed in the single-cell atlas of *S. mansoni* (8). *Abbreviations: bM, bisex male (pairing experienced male); sM, single-sex males (pairing unexperienced male), bT, testes of bM; sT, testes of sM; bF, bisex female (pairing experienced female); sF, single-sex (pairing unexperienced female); bO, ovary of bF; sO, ovary of sF.*

Supplementary Figure S3. Insertions and deletions (InDels) after programmed editing of *S. mansoni* GSH1. **(A)** CRISPR/Cas efficiencies of Cas9- and Cas12a-mediated KO as determined by TIDE-analysis of adult schistosome DNA after normalization with sequencing results from the treated and control (untreated) worms. Shown are the gene editing efficiencies of the LEs, KO-Cas9 (dark blue bar), scramble-Cas9 (purple bar), scramble-Cas12a (orange bar), and KO-Cas12a (red bar) groups. The programmed CRISPR mutation efficiency was significantly higher in the GSH-targeted groups than in the scramble groups (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$; t test, $n = 8$). Finally, Cas12a RNPs (red) showed significantly higher CRISPR efficiency than Cas9 RNPs. **(Panels (B) and (C))** Summary of the major InDels found at the programmed cleavage sites (red

arrows) of Cas12a and Cas9, respectively, based on deep sequencing reads from three biological replicates. *Abbreviations: CTRL, control; KO, knockout; ns, not significant.*

Supplementary Figure S4. Cas9/12-mediated mutation-rates of the GSH1 upon KI-experiments in *S. mansoni* eggs. **(A)** The CRISPR/Cas-mediated deletion rate of the programmed cleavage site in GSH1 of treated LEs during trans-gene integration experiments was less than 1% in the KI-Cas9 (blue) and the KI-Cas12a (red) groups after 5 and 10 days post-electroporation with RNPs together with the SmUbi-EGFP-SmUbi donor template. The genomic deletion sites were created by NHEJ during transgene integration. The resulting deletions had sizes of up to 94 nt in the KI-Cas12a group **(Panel (B))** and up to 18 nt in the KI-Cas9 group **(Panel (C))**.

Supplementary Figure S5. Successful transfection of eggs at an early developmental stage using labelled CRISPR/Cas12a RNPs. **(A)** Confocal microscopy of early schistosome eggs (0-24 h, NLEs), 2 h after transfection with RNPs, which included rhodamine-labelled Cas12a-gRNAs. Control NLEs were electroporated in the absence of RNPs. In control NLEs, only eggshell autofluorescence was observed, whereas fluorescent signals were detected on the eggshell and in the cells of a portion of NLEs treated with rhodamine-labeled RNPs. **(B)** Detailed image of NLEs transfected with rhodamine-labelled Cas12a-RNPs at 60 min after electroporation. Successful delivery of CRISPR/Cas-material (arrows) into numerous cells of the eggs was observed. Fluorescence spectra were determined using a Leica TCS SP5 vis confocal laser scanning microscope with Leica LasX software. Scale bars: 50 µm. *Abbreviations: auto fl., auto fluorescence; CLSM, confocal laser scanning microscopy; Ctr, control; Zy, Zygote.*

Supplementary Figure S6. Schematic overview of predicted DSB-repair pathways in *S. mansoni*. Shown is the bioinformatic identification of the DNA repair machinery, and involved genes including *S. mansoni* orthologs (Smp numbers; reference Howe et al. ….), as described for nematodes. Orthologs to the key actors of the double-stranded break (DSB)-repair pathway, as described for *Caenorhabditis elegans* (112,113) were found in *S. mansoni*. Consequently, orthologs essential for both NHEJ and HDR were identified. For this purpose, the protein sequences of previously characterized DSB repair proteins from the nematodes *Caenorhabditis elegans* and *Pristionchus pacificus* were analyzed using BLAST against the proteome of *S.*

mansoni (parasite.wormbase.org, genome V10). Furthermore, both the sequence coverage and domain structures of the analyzed proteins were compared using the SMART Domain analyzer (smart.embl-heidelberg. de, (130,131)), and orthologs were annotated as described previously(113). Additional annotation details can be found in Supplementary Table S6.

REFERENCES

1. Vale, N., Gouveia, M.J., Rinaldi, G., Brindley, P.J., Gartner, F. and Correia da Costa, J.M. (2017) Praziquantel for Schistosomiasis: Single-Drug Metabolism Revisited, Mode of Action, and Resistance. *Antimicrob Agents Chemother*, **61**.
2. Wang, W., Wang, L. and Liang, Y.S. (2012) Susceptibility or resistance of praziquantel in human schistosomiasis: a review. *Parasitol Res*, **111**, 1871-1877.
3. Berriman, M., Haas, B.J., LoVerde, P.T., Wilson, R.A., Dillon, G.P., Cerqueira, G.C., Mashiyama, S.T., Al-Lazikani, B., Andrade, L.F., Ashton, P.D. *et al.* (2009) The genome of the blood fluke Schistosoma mansoni. *Nature*, **460**, 352-358.
4. Crellen, T., Allan, F., David, S., Durrant, C., Huckvale, T., Holroyd, N., Emery, A.M., Rollinson, D., Aanensen, D.M., Berriman, M. *et al.* (2016) Whole genome resequencing of the human parasite *Schistosoma mansoni* reveals population history and effects of selection. *Sci Rep*, **6**, 20954.
5. Protasio, A.V., Tsai, I.J., Babbage, A., Nichol, S., Hunt, M., Aslett, M.A., De Silva, N., Velarde, G.S., Anderson, T.J., Clark, R.C. *et al.* (2012) A systematically improved high quality genome and transcriptome of the human blood fluke Schistosoma mansoni. *PLoS Negl Trop Dis*, **6**, e1455.
6. Lu, Z., Sessler, F., Holroyd, N., Hahnel, S., Quack, T., Berriman, M. and Grevelding, C.G. (2016) Schistosome sex matters: a deep view into gonad-specific and pairing-dependent transcriptomes reveals a complex gender interplay. *Sci Rep*, **6**, 31150.
7. Lu, Z., Sessler, F., Holroyd, N., Hahnel, S., Quack, T., Berriman, M. and Grevelding, C.G. (2017) A gene expression atlas of adult *Schistosoma mansoni* and their gonads. *Sci Data*, **4**, 170118.
8. Wendt, G.R., Reese, M.L. and Collins, J.J., 3rd. (2021) SchistoCyte Atlas: A Single-Cell Transcriptome Resource for Adult Schistosomes. *Trends Parasitol*, **37**, 585-587.
9. Diaz Soria, C.L., Lee, J., Chong, T., Coghlan, A., Tracey, A., Young, M.D., Andrews, T., Hall, C., Ng, B.L., Rawlinson, K. *et al.* (2020) Single-cell atlas of the first intra-mammalian developmental stage of the human parasite *Schistosoma mansoni*. *Nat Commun*, **11**, 6411.
10. Lu, Z., Sankaranarayanan, G., Rawlinson, K.A., Offord, V., Brindley, P.J., Berriman, M. and Rinaldi, G. (2021) The Transcriptome of *Schistosoma mansoni* Developing Eggs Reveals Key Mediators in Pathogenesis and Life Cycle Propagation. *Front Trop Dis*, **2**, 713123.
11. Wangwiwatsin, A., Protasio, A.V., Wilson, S., Owusu, C., Holroyd, N.E., Sanders, M.J., Keane, J., Doenhoff, M.J., Rinaldi, G. and Berriman, M. (2020) Transcriptome of the parasitic flatworm *Schistosoma mansoni* during intra-mammalian development. *PLoS Negl Trop Dis*, **14**, e0007743.
12. Rawlinson, K.A., Reid, A.J., Lu, Z., Driguez, P., Wawer, A., Coghlan, A., Sankaranarayanan, G., Buddenborg, S.K., Soria, C.D., McCarthy, C. *et al.* (2021) Daily rhythms in gene expression of the human parasite *Schistosoma mansoni*. *BMC Biol*, **19**, 255.
13. Vasconcelos, E.J.R., daSilva, L.F., Pires, D.S., Lavezzo, G.M., Pereira, A.S.A., Amaral, M.S. and Verjovski-Almeida, S. (2017) The *Schistosoma mansoni* genome encodes thousands of long non-coding RNAs predicted to be functional at different parasite life-cycle stages. *Sci Rep*, **7**, 10508.

14. Vasconcelos, E.J.R., Mesel, V.C., daSilva, L.F., Pires, D.S., Lavezzo, G.M., Pereira, A.S.A., Amaral, M.S. and Verjovski-Almeida, S. (2018) Atlas of *Schistosoma mansoni* long non-coding RNAs and their expression correlation to protein-coding genes. *Database (Oxford)*, **2018**.
15. Anderson, L., Amaral, M.S., Beckedorff, F., Silva, L.F., Dazzani, B., Oliveira, K.C., Almeida, G.T., Gomes, M.R., Pires, D.S., Setubal, J.C. *et al.* (2015) *Schistosoma mansoni* Egg, Adult Male and Female Comparative Gene Expression Analysis and Identification of Novel Genes by RNA-Seq. *PLoS Negl Trop Dis*, **9**, e0004334.
16. Berger, D.J., Crellen, T., Lamberton, P.H.L., Allan, F., Tracey, A., Noonan, J.D., Kabatereine, N.B., Tukahebwa, E.M., Adriko, M., Holroyd, N. *et al.* (2021) Whole-genome sequencing of *Schistosoma mansoni* reveals extensive diversity with limited selection despite mass drug administration. *Nat Commun*, **12**, 4776.
17. Da'dara, A.A. and Skelly, P.J. (2015) Gene suppression in schistosomes using RNAi. *Methods Mol Biol*, **1201**, 143-164.
18. Gava, S.G., Tavares, N.C., Salim, A.C.M., Araujo, F.M.G., Oliveira, G. and Mourao, M.M. (2017) *Schistosoma mansoni*: Off-target analyses using nonspecific double-stranded RNAs as control for RNAi experiments in schistosomula. *Exp Parasitol*, **177**, 98-103.
19. Stefanic, S., Dvorak, J., Horn, M., Braschi, S., Sojka, D., Ruelas, D.S., Suzuki, B., Lim, K.C., Hopkins, S.D., McKerrow, J.H. *et al.* (2010) RNA interference in *Schistosoma mansoni* schistosomula: selectivity, sensitivity and operation for larger-scale screening. *PLoS Negl Trop Dis*, **4**, e850.
20. Moescheid, M.F., Puckelwaldt, O., Beutler, M., Haeberlein, S. and Grevelding, C.G. (2023) Defining an optimal control for RNAi experiments with adult *Schistosoma mansoni*. *Sci Rep*, **13**, 9766.
21. Ittiprasert, W., Moescheid, M.F., Chaparro, C., Mann, V.H., Quack, T., Rodpai, R., Miller, A., Wisitpongpun, P., Buakaew, W., Mentink-Kane, M. *et al.* (2023) Targeted insertion and reporter transgene activity at a gene safe harbor of the human blood fluke, *Schistosoma mansoni*. *Cell Rep Methods*, **3**, 100535.
22. Zhang, B. (2021) CRISPR/Cas gene therapy. *J Cell Physiol*, **236**, 2459-2481.
23. Zhang, D., Zhang, Z., Unver, T. and Zhang, B. (2021) CRISPR/Cas: A powerful tool for gene function study and crop improvement. *J Adv Res*, **29**, 207-221.
24. Makarova, K.S., Wolf, Y.I., Iranzo, J., Shmakov, S.A., Alkhnbashi, O.S., Brouns, S.J.J., Charpentier, E., Cheng, D., Haft, D.H., Horvath, P. *et al.* (2020) Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat Rev Microbiol*, **18**, 67-83.
25. Pickar-Oliver, A. and Gersbach, C.A. (2019) The next generation of CRISPR-Cas technologies and applications. *Nat Rev Mol Cell Biol*, **20**, 490-507.
26. Ittiprasert, W., Chatupheeraphat, C., Mann, V.H., Li, W., Miller, A., Ogunbayo, T., Tran, K., Alrefaei, Y.N., Mentink-Kane, M. and Brindley, P.J. (2022) RNA-Guided AsCas12a- and SpCas9-Catalyzed Knockout and Homology Directed Repair of the Omega-1 Locus of the Human Blood Fluke, Schistosoma mansoni. *Int J Mol Sci*, **23**.
27. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819-823.

28. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816-821.
29. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823-826.
30. Paul, B. and Montoya, G. (2020) CRISPR-Cas12a: Functional overview and applications. *Biomed J*, **43**, 8-17.
31. Swarts, D.C. and Jinek, M. (2018) Cas9 versus Cas12a/Cpf1: Structure-function comparisons and implications for genome editing. *Wiley Interdiscip Rev RNA*, **9**, e1481.
32. Kim, H., Kim, S.T., Ryu, J., Kang, B.C., Kim, J.S. and Kim, S.G. (2017) CRISPR/Cpf1-mediated DNA-free plant genome editing. *Nat Commun*, **8**, 14406.
33. Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D. and Kamoun, S. (2013) Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nat Biotechnol*, **31**, 691-693.
34. Svitashev, S., Young, J.K., Schwartz, C., Gao, H., Falco, S.C. and Cigan, A.M. (2015) Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. *Plant Physiol*, **169**, 931-945.
35. Papapetrou, E.P. and Schambach, A. (2016) Gene Insertion Into Genomic Safe Harbors for Human Gene Therapy. *Mol Ther*, **24**, 678-684.
36. Jurberg, A.D., Goncalves, T., Costa, T.A., de Mattos, A.C., Pascarelli, B.M., de Manso, P.P., Ribeiro-Alves, M., Pelajo-Machado, M., Peralta, J.M., Coelho, P.M. *et al.* (2009) The embryonic development of Schistosoma mansoni eggs: proposal for a new staging system. *Dev Genes Evol*, **219**, 219-234.
37. Mann, V.H., Morales, M.E., Rinaldi, G. and Brindley, P.J. (2010) Culture for genetic manipulation of developmental stages of Schistosoma mansoni. *Parasitology*, **137**, 451-462.
38. Dalton, J.P., Clough, K.A., Jones, M.K. and Brindley, P.J. (1997) The cysteine proteinases of Schistosoma mansoni cercariae. *Parasitology*, **114 (Pt 2)**, 105-112.
39. Cormack, B.P., Valdivia, R.H. and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, **173**, 33-38.
40. Howe, K.L., Bolt, B.J., Cain, S., Chan, J., Chen, W.J., Davis, P., Done, J., Down, T., Gao, S., Grove, C. *et al.* (2016) WormBase 2016: expanding to enable helminth genomic research. *Nucleic Acids Res*, **44**, D774-780.
41. Howe, K.L., Bolt, B.J., Shafie, M., Kersey, P. and Berriman, M. (2017) WormBase ParaSite - a comprehensive resource for helminth genomics. *Mol Biochem Parasitol*, **215**, 2-10.
42. Roberts, S.B., Segil, N. and Heintz, N. (1991) Differential phosphorylation of the transcription factor Oct1 during the cell cycle. *Science*, **253**, 1022-1026.
43. Wingender, E. (2004) TRANSFAC, TRANSPATH and CYTOMER as starting points for an ontology of regulatory networks. *In Silico Biol*, **4**, 55-61.
44. Gershenzon, N.I. and Ioshikhes, I.P. (2005) Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. *Bioinformatics*, **21**, 1295-1300.
45. Jin, V.X., Singer, G.A., Agosto-Perez, F.J., Liyanarachchi, S. and Davuluri, R.V. (2006) Genome-wide analysis of core promoter elements from conserved human and mouse orthologous pairs. *BMC Bioinformatics*, **7**, 114.

46. Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M. and Werner, T. (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*, **21**, 2933-2942.
47. Kovacs, K.A., Steinmann, M., Magistretti, P.J., Halfon, O. and Cardinaux, J.R. (2003) CCAAT/enhancer-binding protein family members recruit the coactivator CREB-binding protein and trigger its phosphorylation. *J Biol Chem*, **278**, 36959-36965.
48. Labun, K., Montague, T.G., Krause, M., Torres Cleuren, Y.N., Tjeldnes, H. and Valen, E. (2019) CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res*, **47**, W171-W174.
49. Jeon, Y., Choi, Y.H., Jang, Y., Yu, J., Goo, J., Lee, G., Jeong, Y.K., Lee, S.H., Kim, I.S., Kim, J.S. *et al.* (2018) Direct observation of DNA target searching and cleavage by CRISPR-Cas12a. *Nat Commun*, **9**, 2777.
50. Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A. *et al.* (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, **163**, 759-771.
51. Wendt, G., Zhao, L., Chen, R., Liu, C., O'Donoghue, A.J., Caffrey, C.R., Reese, M.L. and Collins, J.J., 3rd. (2020) A single-cell RNA-seq atlas of Schistosoma mansoni identifies a key regulator of blood feeding. *Science*, **369**, 1644-1649.
52. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res*, **14**, 1188-1190.
53. Liu, H., Han, H., Li, J. and Wong, L. (2005) DNAFSMiner: a web-based software toolbox to recognize two types of functional sites in DNA sequences. *Bioinformatics*, **21**, 671-673.
54. Nasri, M., Mir, P., Dannenmann, B., Amend, D., Skroblyn, T., Xu, Y., Schulze-Osthoff, K., Klimiankou, M., Welte, K. and Skokowa, J. (2019) Fluorescent labeling of CRISPR/Cas9 RNP for gene knockout in HSPCs and iPSCs reveals an essential role for GADD45b in stress response. *Blood Adv*, **3**, 63-71.
55. Collins, J.J., 3rd, Hou, X., Romanova, E.V., Lambrus, B.G., Miller, C.M., Saberi, A., Sweedler, J.V. and Newmark, P.A. (2010) Genome-wide analyses reveal a role for peptide hormones in planarian germline development. *PLoS Biol*, **8**, e1000509.
56. Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*, **6**, 343-345.
57. Wippersteg, V., Kapp, K., Kunz, W., Jackstadt, W.P., Zahner, H. and Grevelding, C.G. (2002) HSP70-controlled GFP expression in transiently transformed schistosomes. *Mol Biochem Parasitol*, **120**, 141-150.
58. Beckmann, S., Wippersteg, V., El-Bahay, A., Hirzmann, J., Oliveira, G. and Grevelding, C.G. (2007) Schistosoma mansoni: germ-line transformation approaches and actin-promoter analysis. *Exp Parasitol*, **117**, 292-303.
59. Labun, K., Montague, T.G., Gagnon, J.A., Thyme, S.B. and Valen, E. (2016) CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res*, **44**, W272-276.
60. Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M. and Valen, E. (2014) CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res*, **42**, W401-407.

61. Sankaranarayanan, G., Coghlan, A., Driguez, P., Lotkowska, M.E., Sanders, M., Holroyd, N., Tracey, A., Berriman, M. and Rinaldi, G. (2020) Large CRISPR-Cas-induced deletions in the oxamniquine resistance locus of the human parasite Schistosoma mansoni. *Wellcome Open Res*, **5**, 178.
62. Neill, P.J., Smith, J.H., Doughty, B.L. and Kemp, M. (1988) The ultrastructure of the Schistosoma mansoni egg. *Am J Trop Med Hyg*, **39**, 52-65.
63. Candido, R.R.F., Morassutti, A.L., Graeff-Teixeira, C., St Pierre, T.G. and Jones, M.K. (2018) Exploring Structural and Physical Properties of Schistosome Eggs: Potential Pathways for Novel Diagnostics? *Adv Parasitol*, **100**, 209-237.
64. Hulme, B.J., Geyer, K.K., Forde-Thomas, J.E., Padalino, G., Phillips, D.W., Ittiprasert, W., Karinshak, S.E., Mann, V.H., Chalmers, I.W., Brindley, P.J. *et al.* (2022) Schistosoma mansoni alpha-N-acetylgalactosaminidase (SmNAGAL) regulates coordinated parasite movement and egg production. *PLoS Pathog*, **18**, e1009828.
65. Yu, Y., Guo, Y., Tian, Q., Lan, Y., Yeh, H., Zhang, M., Tasan, I., Jain, S. and Zhao, H. (2020) Publisher Correction: An efficient gene knock-in strategy using 5'-modified double-stranded DNA donors with short homology arms. *Nat Chem Biol*, **16**, 479.
66. Yu, Y., Guo, Y., Tian, Q., Lan, Y., Yeh, H., Zhang, M., Tasan, I., Jain, S. and Zhao, H. (2020) An efficient gene knock-in strategy using 5'-modified double-stranded DNA donors with short homology arms. *Nat Chem Biol*, **16**, 387-390.
67. Brinkman, E.K. and van Steensel, B. (2019) Rapid Quantitative Evaluation of CRISPR Genome Editing by TIDE and TIDER. *Methods Mol Biol*, **1961**, 29-44.
68. Clement, K., Rees, H., Canver, M.C., Gehrke, J.M., Farouni, R., Hsu, J.Y., Cole, M.A., Liu, D.R., Joung, J.K., Bauer, D.E. *et al.* (2019) CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol*, **37**, 224-226.
69. Park, J., Lim, K., Kim, J.S. and Bae, S. (2017) Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics*, **33**, 286-288.
70. Joshi, M., Keith Pittman, H., Haisch, C. and Verbanac, K. (2008) Real-time PCR to determine transgene copy number and to quantitate the biolocalization of adoptively transferred cells from EGFP-transgenic mice. *Biotechniques*, **45**, 247-258.
71. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*, **10**, 845-858.
72. Gao, S., Pan, M., Zheng, Y., Huang, Y., Zheng, Q., Sun, D., Lu, L., Tan, X., Tan, X., Lan, H. *et al.* (2016) Monomer/Oligomer Quasi-Racemic Protein Crystallography. *J Am Chem Soc*, **138**, 14497-14502.
73. Silver, C.E., Cusumano, R.J., Fell, S.C. and Strauch, B. (1989) Replacement of upper esophagus: results with myocutaneous flap and with gastric transposition. *Laryngoscope*, **99**, 819-821.
74. Reese, M.G. (2001) Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. *Comput Chem*, **26**, 51-56.
75. Chalkley, G.E. and Verrijzer, C.P. (1999) DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. *EMBO J*, **18**, 4835-4845.
76. Butler, J.E. and Kadonaga, J.T. (2002) The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev*, **16**, 2583-2592.
77. Godbey, W.T. (2021) *Biotechnology and its Applications*. 2nd ed. Elsevier.

78. Kozak, M. (1987) At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J Mol Biol*, **196**, 947-950.
79. Lu, Z., Zhang, Y. and Berriman, M. (2018) A web portal for gene expression across all life stages of Schistosoma mansoni.
80. Gruber, A.R., Lorenz, R., Bernhart, S.H., Neubock, R. and Hofacker, I.L. (2008) The Vienna RNA websuite. *Nucleic Acids Res*, **36**, W70-74.
81. Cheng, S.W., Lynch, E.C., Leason, K.R., Court, D.L., Shapiro, B.A. and Friedman, D.I. (1991) Functional importance of sequence in the stem-loop of a transcription terminator. *Science*, **254**, 1205-1207.
82. Wippersteg, V., Ribeiro, F., Liedtke, S., Kusel, J.R. and Grevelding, C.G. (2003) The uptake of Texas Red-BSA in the excretory system of schistosomes and its colocalisation with ER60 promoter-induced GFP in transiently transformed adult males. *Int J Parasitol*, **33**, 1139-1143.
83. Hotez, P.J., Aksoy, S., Brindley, P.J. and Kamhawi, S. (2020) What constitutes a neglected tropical disease? *PLoS Negl Trop Dis*, **14**, e0008001.
84. Mitra, A.K. and Mawson, A.R. (2017) Neglected Tropical Diseases: Epidemiology and Global Burden. *Trop Med Infect Dis*, **2**.
85. Tidman, R., Kanankege, K.S.T., Bangert, M. and Abela-Ridder, B. (2023) Global prevalence of 4 neglected foodborne trematodes targeted for control by WHO: A scoping review to highlight the gaps. *PLoS Negl Trop Dis*, **17**, e0011073.
86. Aya Pastrana, N., Beran, D., Somerville, C., Heller, O., Correia, J.C. and Suggs, L.S. (2020) The process of building the priority of neglected tropical diseases: A global policy analysis. *PLoS Negl Trop Dis*, **14**, e0008498.
87. Ittiprasert, W., Mann, V.H., Karinshak, S.E., Coghlan, A., Rinaldi, G., Sankaranarayanan, G., Chaidee, A., Tanno, T., Kumkhaek, C., Prangtaworn, P. *et al.* (2019) Programmed genome editing of the omega-1 ribonuclease of the blood fluke, Schistosoma mansoni. *Elife*, **8:e41337**.
88. Stitz, M., Chaparro, C., Lu, Z., Olzog, V.J., Weinberg, C.E., Blom, J., Goesmann, A., Grunau, C. and Grevelding, C.G. (2021) Satellite-Like W-Elements: Repetitive, Transcribed, and Putative Mobile Genetic Factors with Potential Roles for Biology and Evolution of Schistosoma mansoni. *Genome Biol Evol*, **13**.
89. Roquis, D., Taudt, A., Geyer, K.K., Padalino, G., Hoffmann, K.F., Holroyd, N., Berriman, M., Aliaga, B., Chaparro, C., Grunau, C. *et al.* (2018) Histone methylation changes are required for life cycle progression in the human parasite Schistosoma mansoni. *PLoS Pathog*, **14**, e1007066.
90. Wang, X., Tang, Y., Lu, J., Shao, Y., Qin, X., Li, Y., Wang, L., Li, D. and Liu, M. (2016) Characterization of novel cytochrome P450 2E1 knockout rat model generated by CRISPR/Cas9. *Biochem Pharmacol*, **105**, 80-90.
91. Chaiyadet, S., Tangkawattana, S., Smout, M.J., Ittiprasert, W., Mann, V.H., Deenonpoe, R., Arunsan, P., Loukas, A., Brindley, P.J. and Laha, T. (2022) Knockout of liver fluke granulin, Ov-grn-1, impedes malignant transformation during chronic infection with *Opisthorchis viverrini*. *PLoS Pathog*, **18**, e1010839.
92. Khampoosa, P., Jones, M.K., Lovas, E.M., Srisawangwong, T., Laha, T., Piratae, S., Thammasiri, C., Suwannatrai, A., Sripanidkulchai, B., Eursitthichai, V. *et al.* (2012) Light and

electron microscopy observations of embryogenesis and egg development in the human liver fluke, *Opisthorchis viverrini* (Platyhelminthes, Digenea). *Parasitol Res*, **110**, 799-808.

93. Born-Torrijos, A., Holzer, A.S., Raga, J.A., van Beest, G.S. and Yoneva, A. (2017) Description of embryonic development and ultrastructure in miracidia of *Cardiocephaloides longicollis* (Digenea, Strigeidae) in relation to active host finding strategy in a marine environment. *J Morphol*, **278**, 1137-1148.
94. You, H., Mayer, J.U., Johnston, R.L., Sivakumaran, H., Ranasinghe, S., Rivera, V., Kondrashova, O., Koufariotis, L.T., Du, X., Driguez, P. *et al.* (2021) CRISPR/Cas9-mediated genome editing of *Schistosoma mansoni* acetylcholinesterase. *FASEB J*, **35**, e21205.
95. Du, X., McManus, D.P., French, J.D., Collinson, N., Sivakumaran, H., MacGregor, S.R., Fogarty, C.E., Jones, M.K. and You, H. (2022) CRISPR interference for sequence-specific regulation of fibroblast growth factor receptor A in Schistosoma mansoni. *Front Immunol*, **13**, 1105719.
96. Race, G.J., Michaels, R.M., Martin, J.H., Larsh, J.E., Jr. and Matthews, J.L. (1969) *Schistosoma mansoni* eggs: an electron microscopic study of shell pores and microbarbs. *Proc Soc Exp Biol Med*, **130**, 990-992.
97. Kazemian, P., Yu, S.Y., Thomson, S.B., Birkenshaw, A., Leavitt, B.R. and Ross, C.J.D. (2022) Lipid-Nanoparticle-Based Delivery of CRISPR/Cas9 Genome-Editing Components. *Mol Pharm*, **19**, 1669-1686.
98. Kuppers, D.A., Linton, J., Ortiz Espinosa, S., McKenna, K.M., Rongvaux, A. and Paddison, P.J. (2023) Gene knock-outs in human CD34+ hematopoietic stem and progenitor cells and in the human immune system of mice. *PLoS One*, **18**, e0287052.
99. Mazurov, D., Ramadan, L. and Kruglova, N. (2023) Packaging and Uncoating of CRISPR/Cas Ribonucleoproteins for Efficient Gene Editing with Viral and Non-Viral Extracellular Nanoparticles. *Viruses*, **15**.
100. Bhandawat, A., Sharma, V., Rishi, V. and J, K.R. (2020) Biolistic Delivery of Programmable Nuclease (CRISPR/Cas9) in Bread Wheat. *Methods Mol Biol*, **2124**, 309-329.
101. Heyers, O., Walduck, A.K., Brindley, P.J., Bleiss, W., Lucius, R., Dorbic, T., Wittig, B. and Kalinna, B.H. (2003) Schistosoma mansoni miracidia transformed by particle bombardment infect Biomphalaria glabrata snails and develop into transgenic sporocysts. *Exp Parasitol*, **105**, 174-178.
102. Wippersteg, V., Kapp, K., Kunz, W. and Grevelding, C.G. (2002) Characterisation of the cysteine protease ER60 in transgenic *Schistosoma mansoni* larvae. *Int J Parasitol*, **32**, 1219-1224.
103. Khan, S. and Sallard, E. (2023) Current and Prospective Applications of CRISPR-Cas12a in Pluricellular Organisms. *Mol Biotechnol*, **65**, 196-205.
104. Moreno-Mateos, M.A., Fernandez, J.P., Rouet, R., Vejnar, C.E., Lane, M.A., Mis, E., Khokha, M.K., Doudna, J.A. and Giraldez, A.J. (2017) CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. *Nat Commun*, **8**, 2024.
105. Mao, Z., Bozzella, M., Seluanov, A. and Gorbunova, V. (2008) DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle*, **7**, 2902-2906.
106. Lu, Z., Quack, T., Hahnel, S., Gelmedin, V., Pouokam, E., Diener, M., Hardt, M., Michel, G., Baal, N., Hackstein, H. *et al.* (2015) Isolation, enrichment and primary characterisation of

vitelline cells from *Schistosoma mansoni* obtained by the organ isolation method. *Int J Parasitol*, **45**, 663-672.
107. Ashton, P.D., Harrop, R., Shah, B. and Wilson, R.A. (2001) The schistosome egg: development and secretions. *Parasitology*, **122**, 329-338.
108. Fattah, F., Lee, E.H., Weisensel, N., Wang, Y., Lichter, N. and Hendrickson, E.A. (2010) Ku regulates the non-homologous end joining pathway choice of DNA double-strand break repair in human somatic cells. *PLoS Genet*, **6**, e1000855.
109. Sato, M., Takabayashi, S., Akasaka, E. and Nakamura, S. (2020) Recent Advances and Future Perspectives of In Vivo Targeted Delivery of Genome-Editing Reagents to Germ Cells, Embryos, and Fetuses in Mice. *Cells*, **9**.
110. Hoch, N.C. (2023) Tissue Specificity of DNA Damage and Repair. *Physiology (Bethesda)*, **38**, 0.
111. Iyama, T. and Wilson, D.M., 3rd. (2013) DNA repair mechanisms in dividing and non-dividing cells. *DNA Repair (Amst)*, **12**, 620-636.
112. Bae, W., Hong, S., Park, M.S., Jeong, H.K., Lee, M.H. and Koo, H.S. (2019) Single-strand annealing mediates the conservative repair of double-strand DNA breaks in homologous recombination-defective germ cells of *Caenorhabditis elegans*. *DNA Repair (Amst)*, **75**, 18-28.
113. Lemmens, B.B. and Tijsterman, M. (2011) DNA double-strand break repair in *Caenorhabditis elegans*. *Chromosoma*, **120**, 1-21.
114. Dieterich, C., Clifton, S.W., Schuster, L.N., Chinwalla, A., Delehaunty, K., Dinkelacker, I., Fulton, L., Fulton, R., Godfrey, J., Minx, P. *et al.* (2008) The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. *Nat Genet*, **40**, 1193-1198.
115. Rillo-Bohn, R., Adilardi, R., Mitros, T., Avsaroglu, B., Stevens, L., Kohler, S., Bayes, J., Wang, C., Lin, S., Baskevitch, K.A. *et al.* (2021) Analysis of meiosis in *Pristionchus pacificus* reveals plasticity in homolog pairing and synapsis in the nematode lineage. *Elife*, **10**.
116. Consortium, C.e.S. (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*, **282**, 2012-2018.
117. Martin, J.S., Winkelmann, N., Petalcorin, M.I., McIlwraith, M.J. and Boulton, S.J. (2005) RAD-51-dependent and -independent roles of a *Caenorhabditis elegans* BRCA2-related protein during DNA double-strand break repair. *Mol Cell Biol*, **25**, 3127-3139.
118. Lieber, M.R. (2008) The mechanism of human nonhomologous DNA end joining. *J Biol Chem*, **283**, 1-5.
119. Ozturk, S. and Demir, N. (2011) DNA repair mechanisms in mammalian germ cells. *Histol Histopathol*, **26**, 505-517.
120. You, H., Jones, M.K., Whitworth, D.J. and McManus, D.P. (2021) Innovations and Advances in Schistosome Stem Cell Research. *Front Immunol*, **12**, 599014.
121. Ohler, U., Liao, G.C., Niemann, H. and Rubin, G.M. (2002) Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol*, **3**, RESEARCH0087.
122. Juven-Gershon, T. and Kadonaga, J.T. (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev Biol*, **339**, 225-229.
123. Busek, S.U., Fantappie, M., Malaquias, L.C., Wilson, R.A., Correa-Oliveira, R. and Oliveira, G.C. (2002) Cis-acting elements, CArG-, E-, CCAAT- and TATA-boxes may be involved in

sexually regulated gene transcription in *Schistosoma mansoni*. *Mem Inst Oswaldo Cruz*, **97 Suppl 1**, 85-90.

124. Byrne, S.M., Ortiz, L., Mali, P., Aach, J. and Church, G.M. (2015) Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. *Nucleic Acids Res*, **43**, e21.
125. Zhang, J.P., Li, X.L., Li, G.H., Chen, W., Arakaki, C., Botimer, G.D., Baylink, D., Zhang, L., Wen, W., Fu, Y.W. *et al.* (2017) Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol*, **18**, 35.
126. Sakuma, T., Takenaga, M., Kawabe, Y., Nakamura, T., Kamihira, M. and Yamamoto, T. (2015) Homologous Recombination-Independent Large Gene Cassette Knock-in in CHO Cells Using TALEN and MMEJ-Directed Donor Plasmids. *Int J Mol Sci*, **16**, 23849-23866.
127. Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K.T. and Yamamoto, T. (2016) MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat Protoc*, **11**, 118-133.
128. Li, G., Zhang, X., Wang, H., Mo, J., Zhong, C., Shi, J., Zhou, R., Li, Z., Yang, H., Wu, Z. *et al.* (2020) CRISPR/Cas9-Mediated Integration of Large Transgene into Pig CEP112 Locus. *G3 (Bethesda)*, **10**, 467-473.
129. Geldhof, P., Visser, A., Clark, D., Saunders, G., Britton, C., Gilleard, J., Berriman, M. and Knox, D. (2007) RNA interference in parasitic helminths: current situation, potential pitfalls and future prospects. *Parasitology*, **134**, 609-619.
130. Letunic, I. and Bork, P. (2018) 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res*, **46**, D493-D496.
131. Letunic, I., Khedkar, S. and Bork, P. (2021) SMART: recent updates, new developments and status in 2020. *Nucleic Acids Res*, **49**, D458-D460.

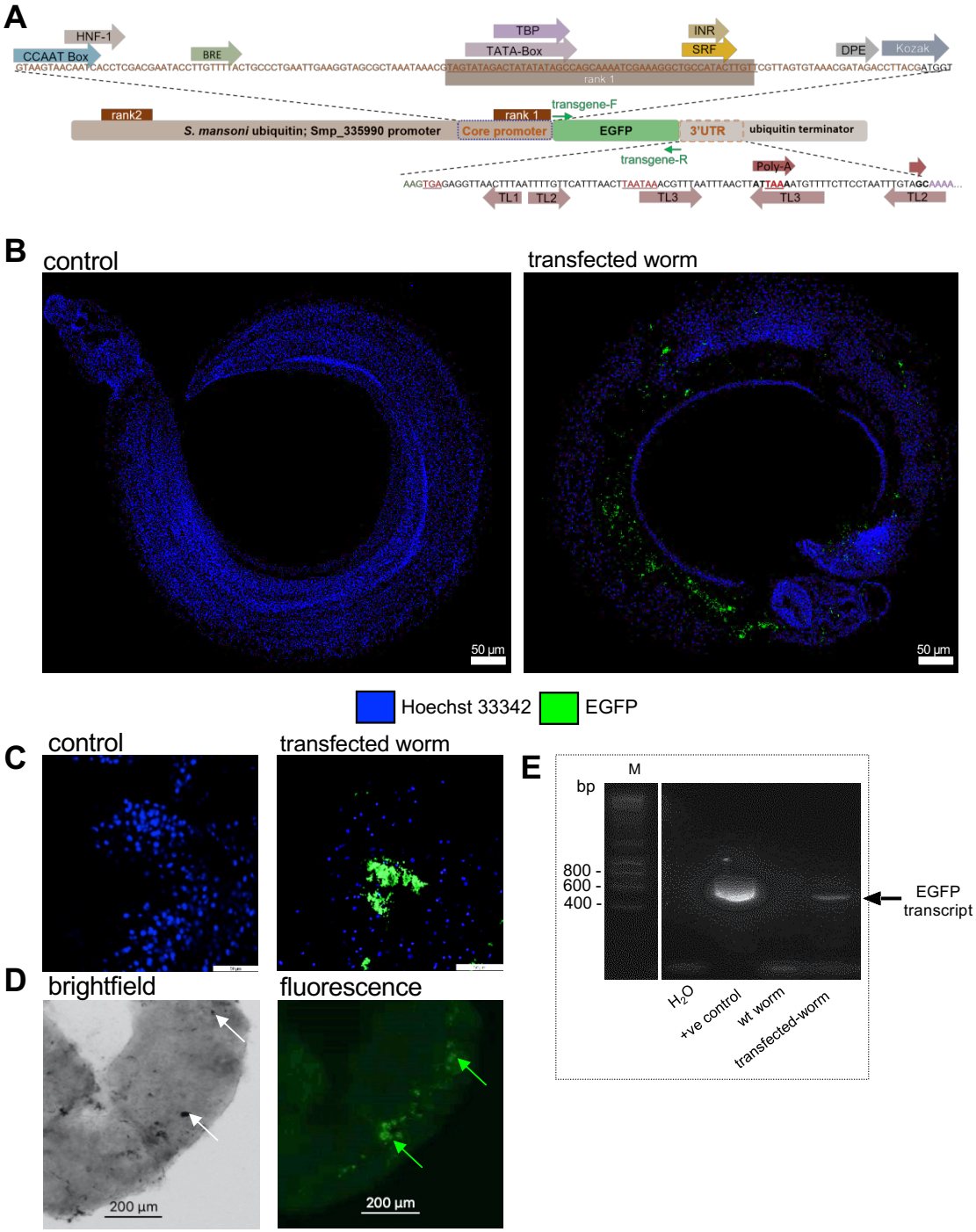

A
HNF-1
CCAAT Box
BRE
TBP
TATA-Box
INR
SRF
DPE
Kozak
rank 1
rank2
rank 1
transgene-F
S. mansoni ubiquitin; Smp_335990 promoter
Core promoter
EGFP
3'UTR
ubiquitin terminator
transgene-R
Poly-A
TL1
TL2
TL3
TL3
TL2
B
control
transfected worm
50 µm
50 µm
Hoechst 33342
EGFP
C
control
transfected worm
D
brightfield
fluorescence
200 µm
200 µm
E
M
bp
800 -
600 -
400 -
EGFP transcript
H2O
+ve control
wt worm
transfected-worm

Figure 1

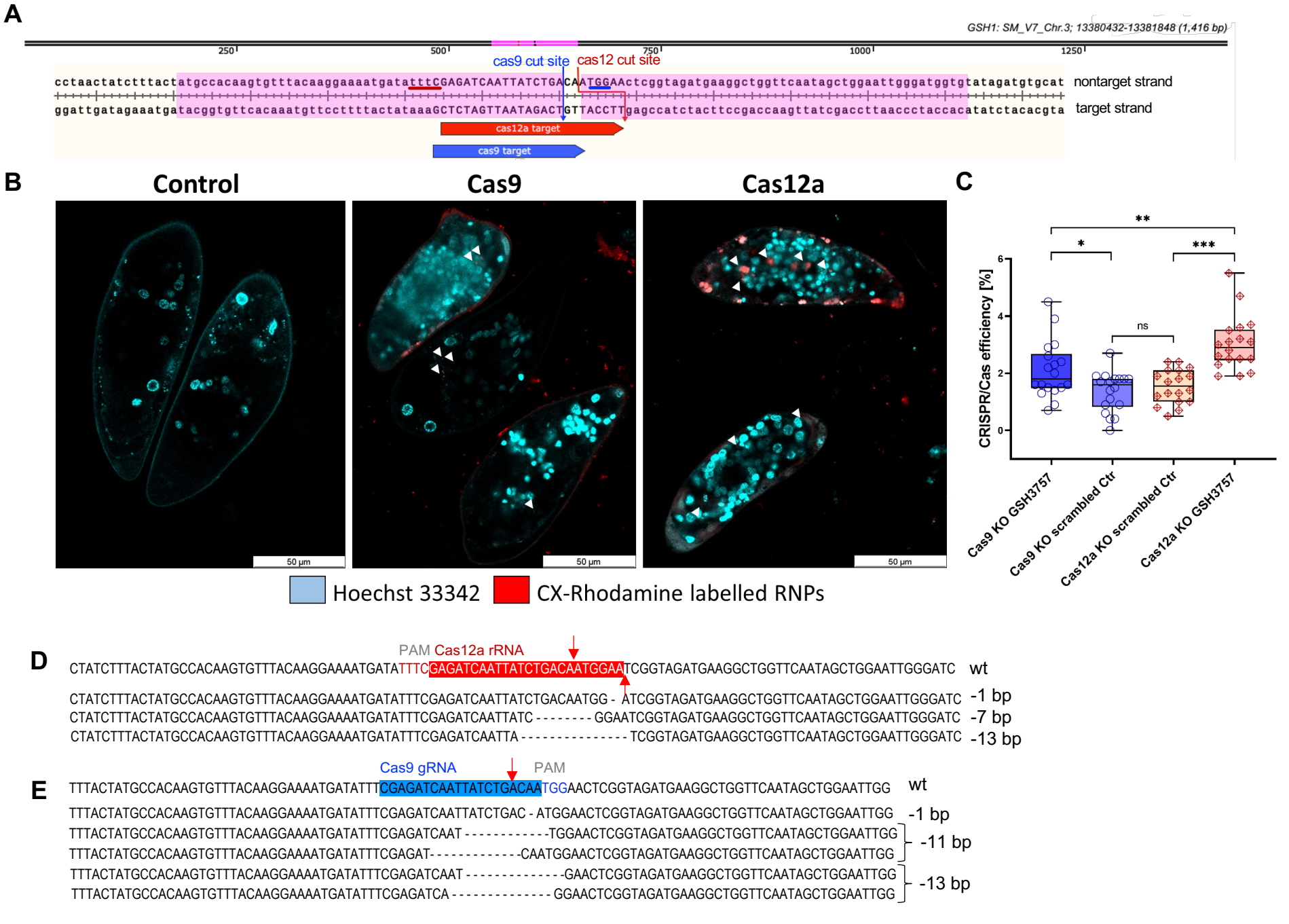

A
GSH1: SM_V7_Chr.3; 13380432-13381848 (1,416 bp)
250
500
750
1000
1250
cas9 cut site
cas12 cut site
cctaactatctttactatgccacaagtgtttacaaggaaaatgatatttCGAGATCAATTATCTGACAATGGAActcggtagatgaaggctggttcaatagctggaattgggatggtgtatagatgtgcat
nontarget strand
ggattgatagaaatgatacggtgttcacaaatgttccttttactataaaGCTCTAGTTAATAGACTGTTACCTTgagccatctacttccgaccaagttatcgaccttaaccctaccacatatctacacgta
target strand
cas12a target
cas9 target
B
Control
Cas9
Cas12a
50 µm
50 µm
50 µm
Hoechst 33342
CX-Rhodamine labelled RNPs
C
CRISPR/Cas efficiency [%]
0
2
4
6
*
**
***
ns
Cas9 KO GSH3757
Cas9 KO scrambled Ctr
Cas12a KO scrambled Ctr
Cas12a KO GSH3757
D
PAM Cas12a rRNA
CTATCTTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAATTATCTGACAATGGAATCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGGGATC
wt
CTATCTTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAATTATCTGACAATGG - ATCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGGGATC
-1 bp
CTATCTTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAATTATC - - - - - - - - GGAATCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGGGATC
-7 bp
CTATCTTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAATTA - - - - - - - - - - - - - - - TCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGGGATC
-13 bp
E
Cas9 gRNA
PAM
TTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAATTATCTGACAATGGAACTCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGG
wt
TTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAATTATCTGAC - ATGGAACTCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGG
-1 bp
TTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAAT - - - - - - - - - - - - TGGAACTCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGG
TTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGAT- - - - - - - - - - - - - CAATGGAACTCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGG
-11 bp
TTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAAT - - - - - - - - - - - - - - GAACTCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGG
TTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCA - - - - - - - - - - - - - - GGAACTCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGG
-13 bp

Figure 2

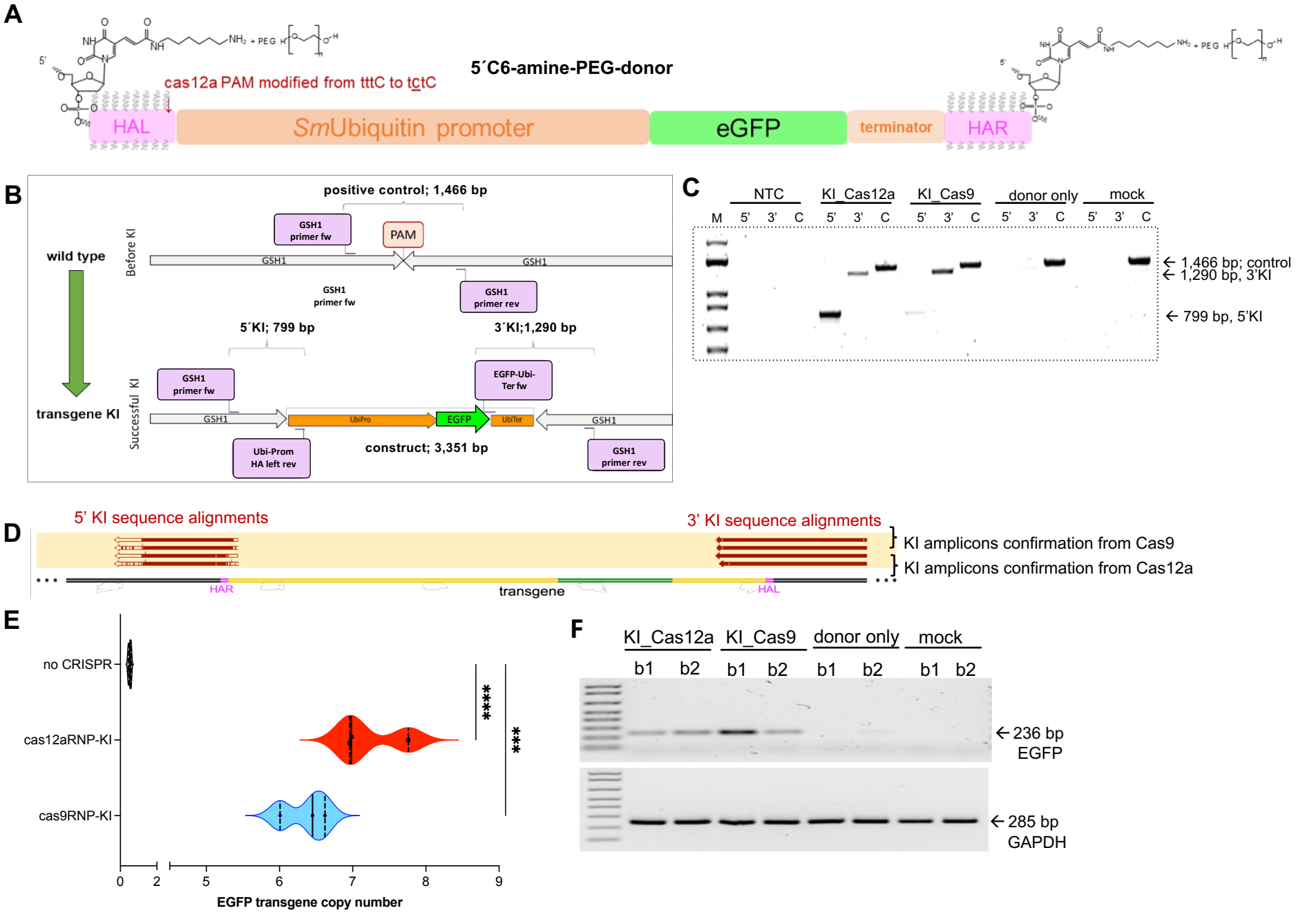

A
5´C6-amine-PEG-donor
cas12a PAM modified from tttC to tctC
HAL
SmUbiquitin promoter
eGFP
terminator
HAR
B
positive control; 1,466 bp
GSH1 primer fw
PAM
wild type
Before KI
GSH1
GSH1 primer fw
GSH1 primer rev
5´KI; 799 bp
3´KI;1,290 bp
transgene KI
Successful KI
EGFP-Ubi-Ter fw
UbiPro
EGFP
UbiTer
Ubi-Prom HA left rev
construct; 3,351 bp
C
NTC
KI_Cas12a
KI_Cas9
donor only
mock
M 5' 3' C 5' 3' C 5' 3' C 5' 3' C 5' 3' C
1,466 bp; control
1,290 bp, 3'KI
799 bp, 5'KI
D
5' KI sequence alignments
3' KI sequence alignments
KI amplicons confirmation from Cas9
KI amplicons confirmation from Cas12a
HAR
transgene
HAL
E
no CRISPR
cas12aRNP-KI
cas9RNP-KI
****
***
0 2 5 6 7 8 9
EGFP transgene copy number
F
KI_Cas12a KI_Cas9 donor only mock
b1 b2 b1 b2 b1 b2 b1 b2
236 bp EGFP
285 bp GAPDH

Figure 3

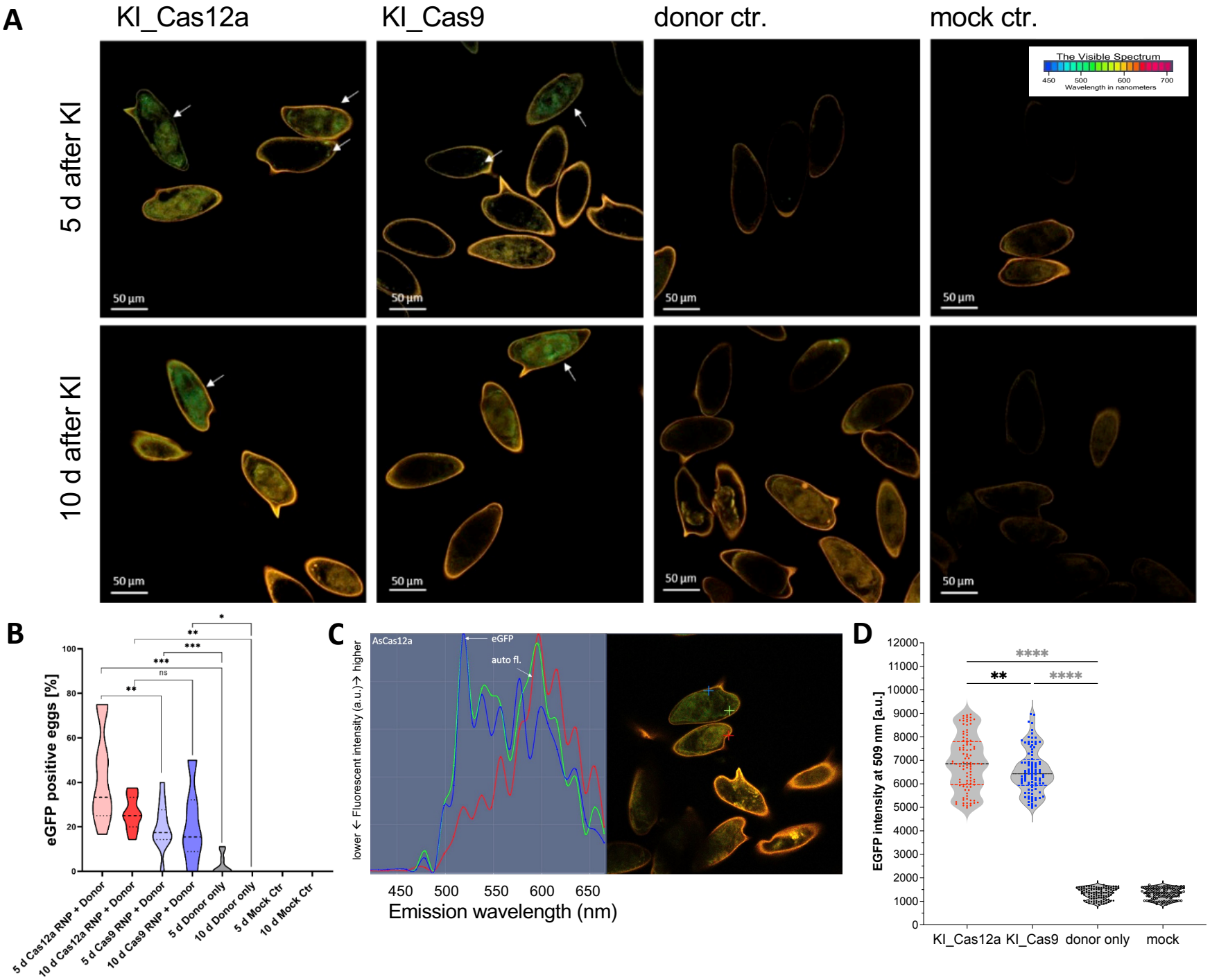
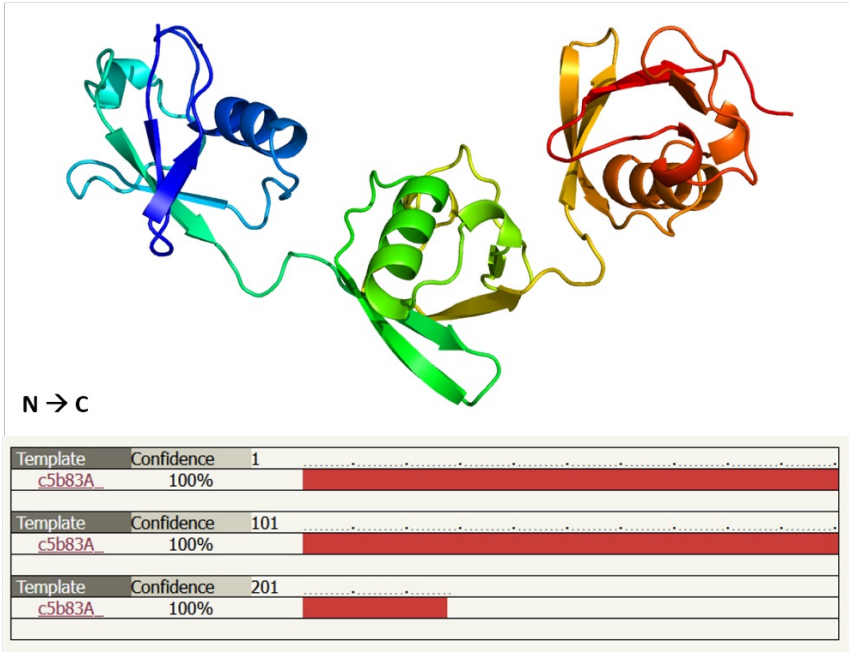

A
KI_Cas12a
KI_Cas9
donor ctr.
mock ctr.
5 d after KI
10 d after KI
The Visible Spectrum
450 500 600 700
Wavelength in nanometers
50 µm
B
eGFP positive eggs [%]
5 d Cas12a RNP + Donor
10 d Cas12a RNP + Donor
5 d Cas9 RNP + Donor
10 d Cas9 RNP + Donor
5 d Donor only
10 d Donor only
5 d Mock Ctr
10 d Mock Ctr
ns
C
AsCas12a
eGFP
auto fl.
lower ← Fluorescent intensity (a.u.) → higher
Emission wavelength (nm)
D
EGFP intensity at 509 nm [a.u.]
KI_Cas12a
KI_Cas9
donor only
mock

Figure 4

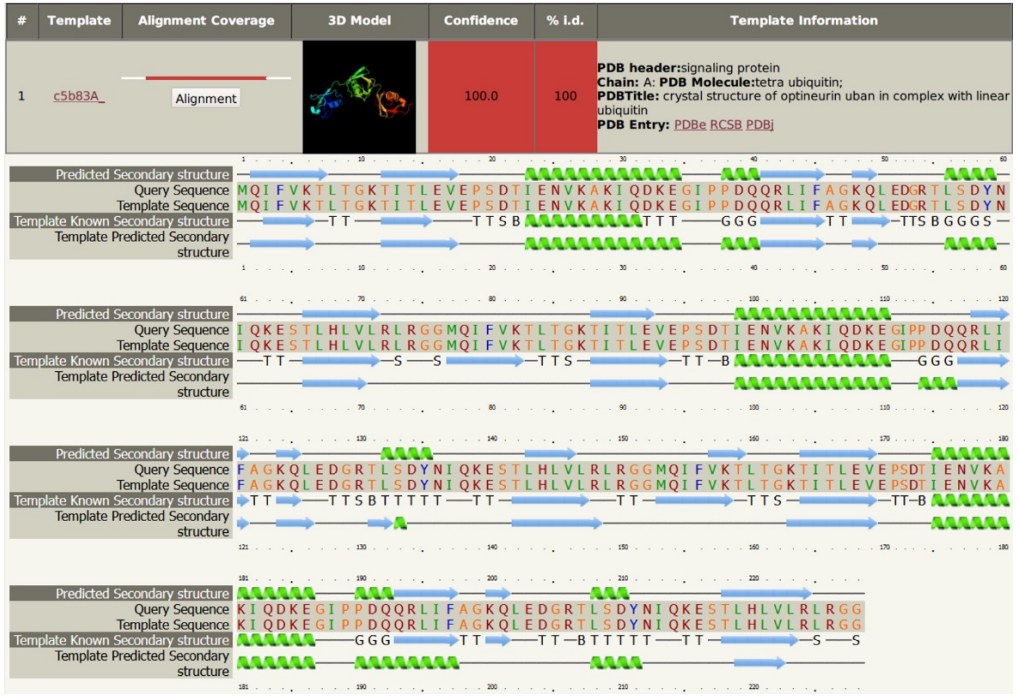
A

Predicted protein structure of Smp_335990



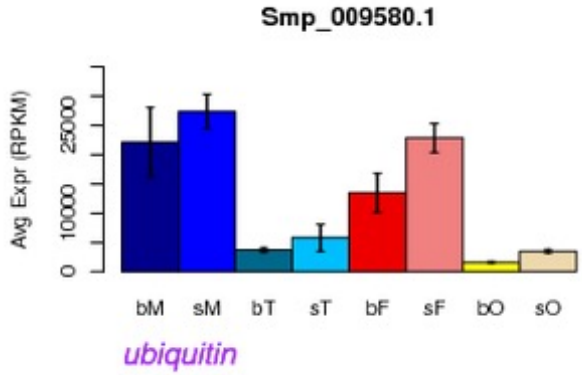
B

Alignment coverage with reference c5b83A (tetra ubiquitin, *Homo sapiens*)

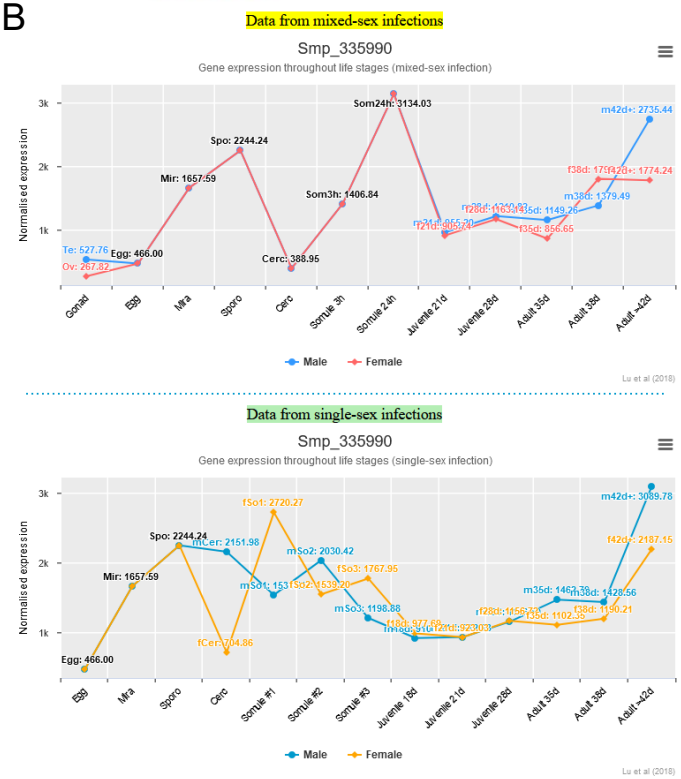


Supplementary Figure S1

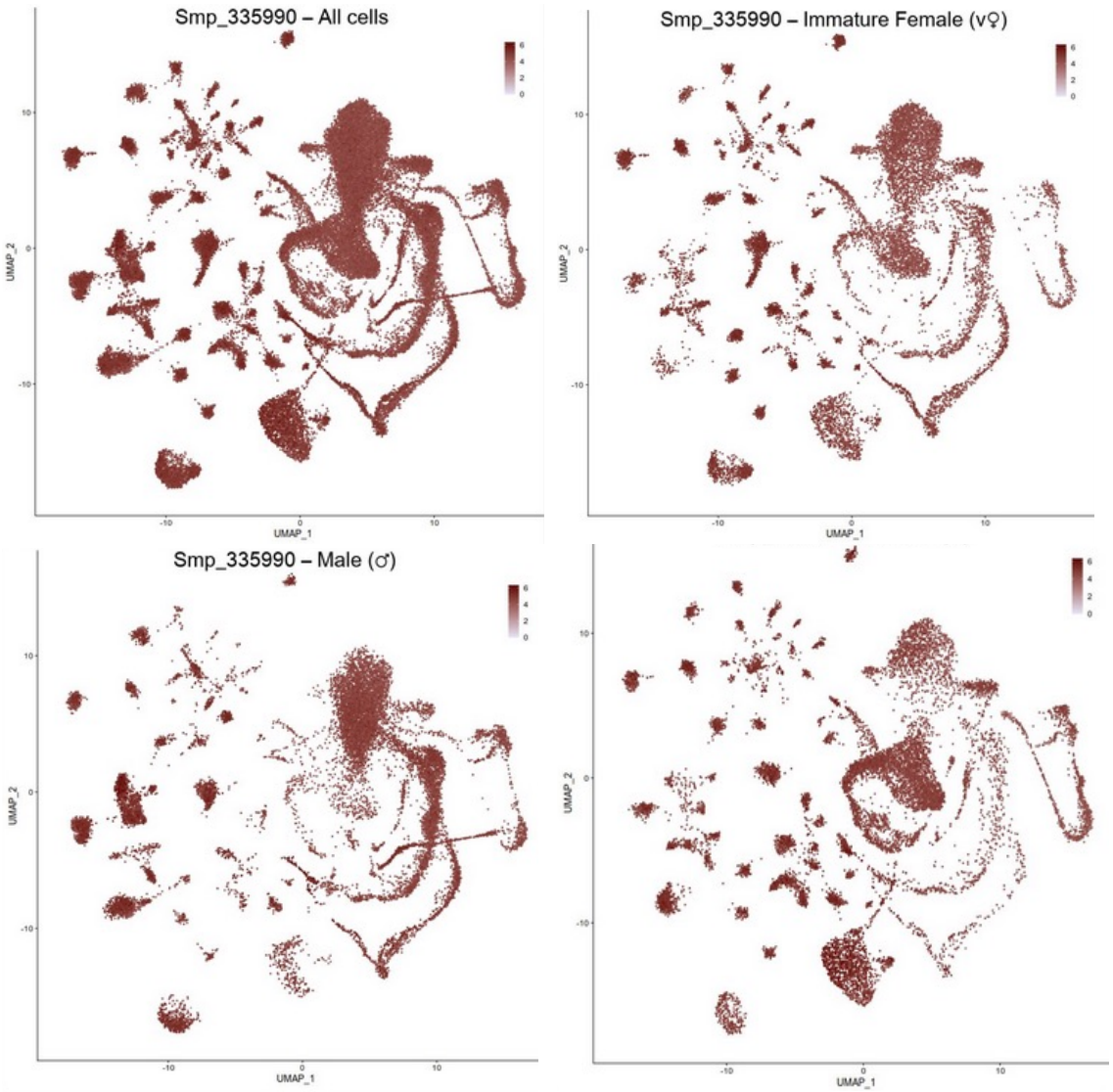
A



B

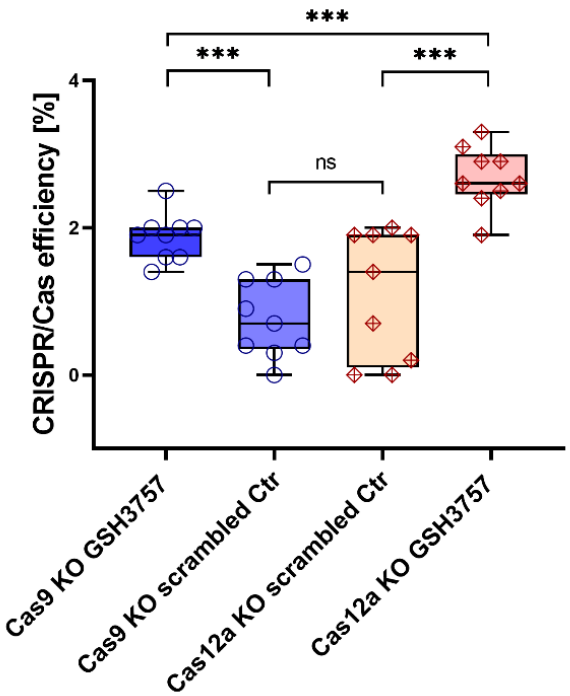


C

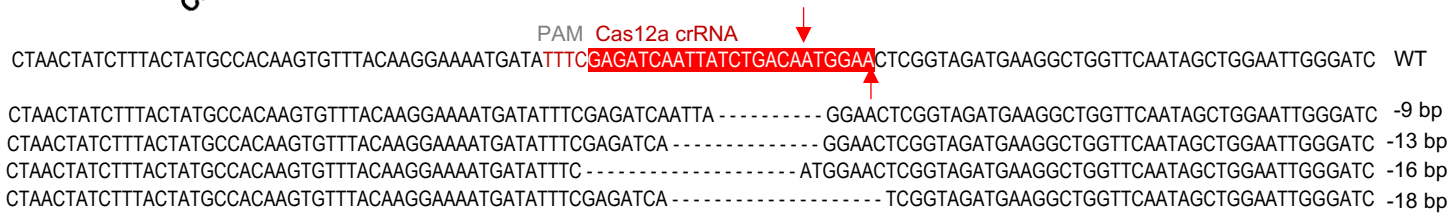


Supplementary Figure S2

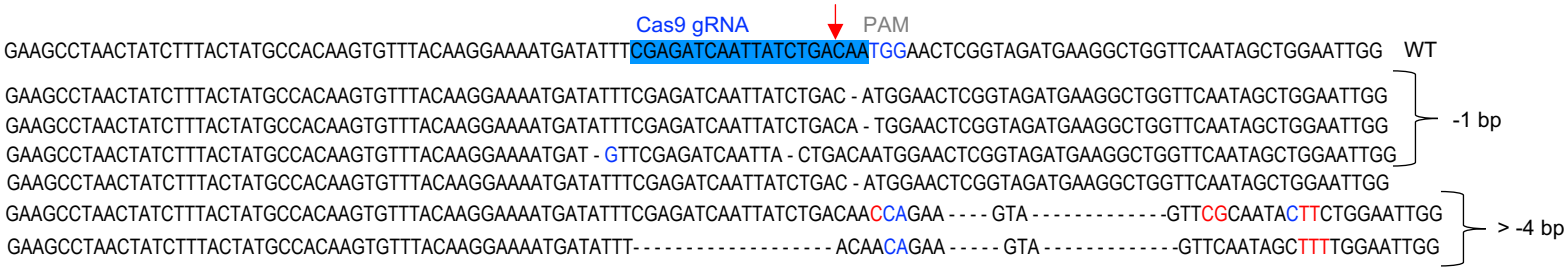
A



B

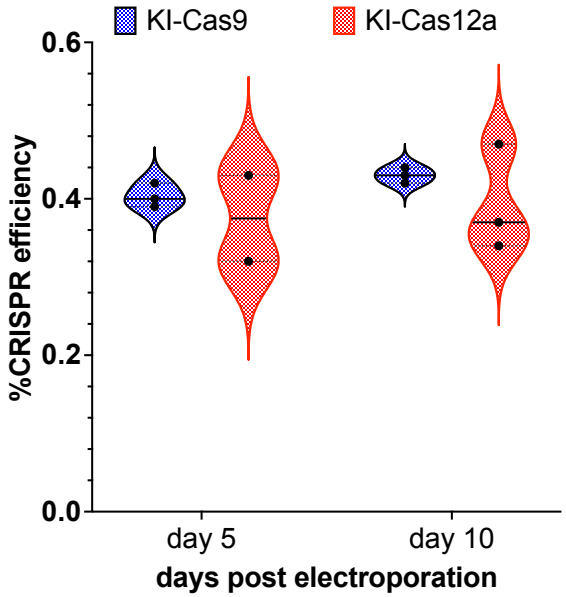


C



Supplementary Figure S3

A



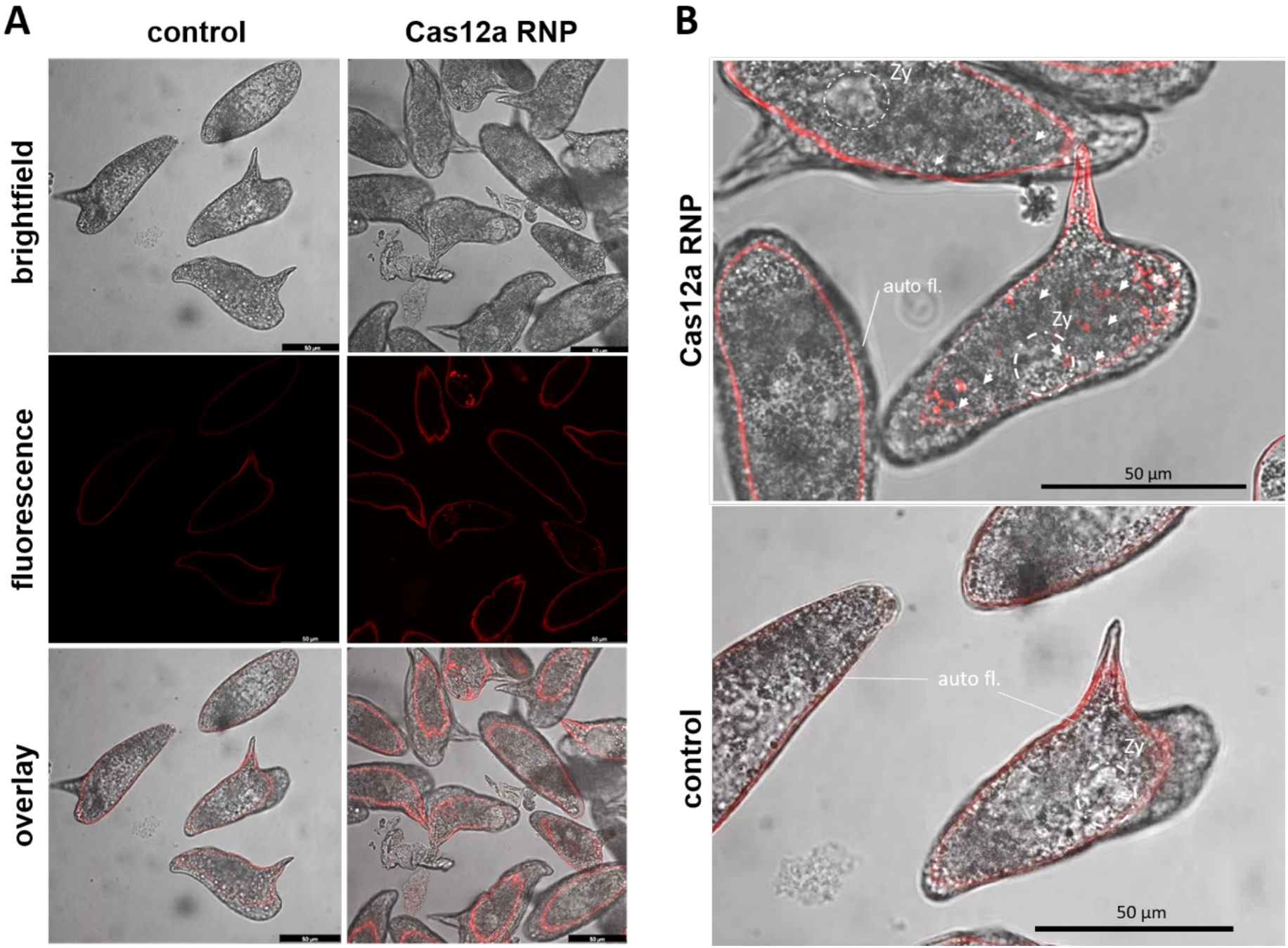
B

```
                                                PAM  Cas12a crRNA
CTAACTATCTTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAATTATCTGACAATGGAACTCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGGGGATCGTGTATAGATGTGCATTAGTCACAGA   WT
CTAACTATCTTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAATTATCTGAC - ATGGAACTCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGGGATCGTGTATAGATGTGCATTAGTCACAGA   -1 bp
CTAACTATCTTTACTATGCC ----------------------------------------------------------------------------------------------------------GTGTATAGATGTGCATTAGTCACAGA   -94 bp
```

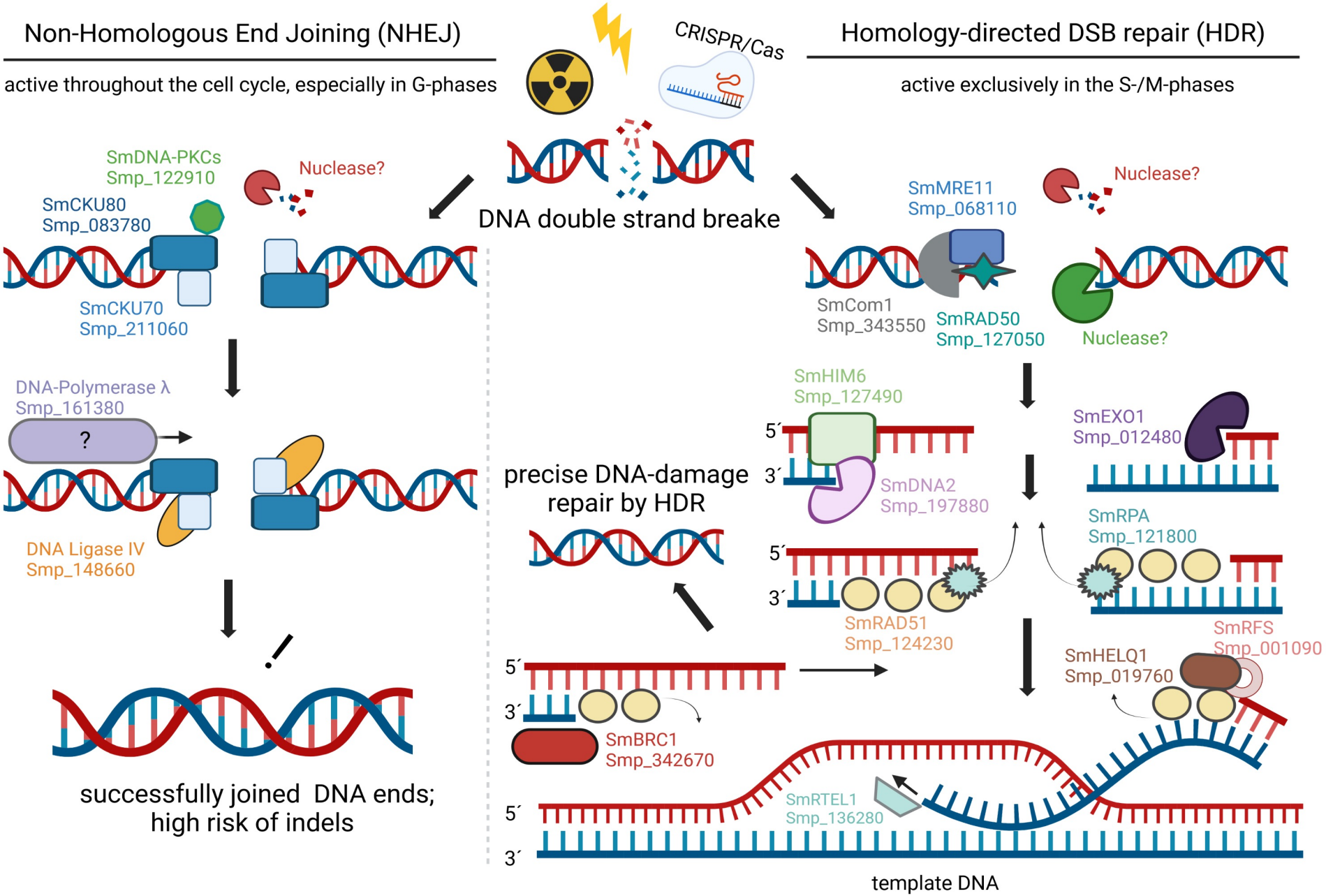
C

```
                                   Cas9 gRNA          PAM
TTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAATTATCTGACAATGGAACTCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGG   WT
TTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAATTATCTGAC - ATGGAACTCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGG   -1 bp
TTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTTCGAGATCAATTA ----------------CGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGG   -15 bp
TTTACTATGCCACAAGTGTTTACAAGGAAAATGATATTT ------------------CAATGGAACTCGGTAGATGAAGGCTGGTTCAATAGCTGGAATTGG   -18 bp
```

Supplementary Figure S4



Supplementary Figure S5

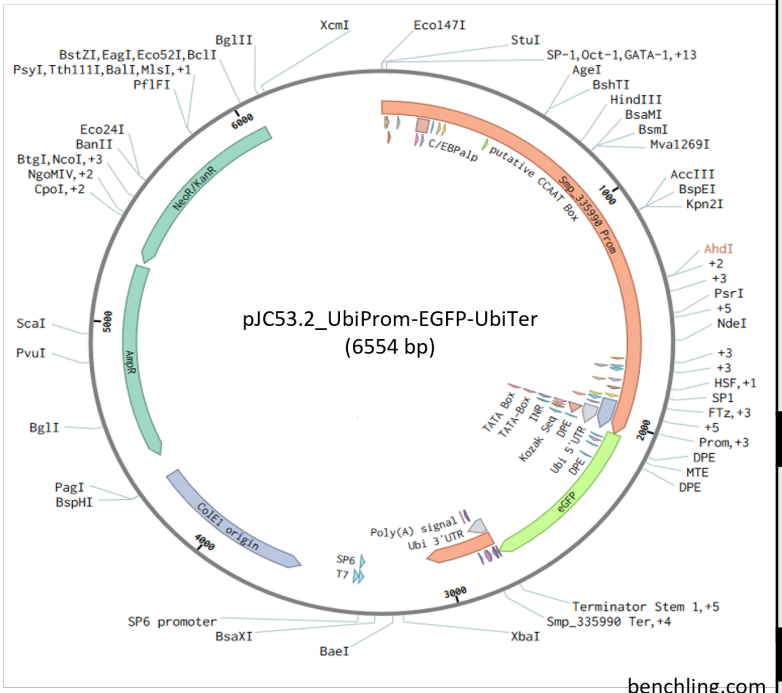


Supplementary Figure S6

Supplementary Table 1

Plasmid card of the reporter-gene construct

pJC53.2_UbiProm-EGFP-UbiTer (6554 bp; AmpR; KanR)



benchling.com

Sequences of the reporter-gene construct

Ubiquitin (Smp_335990) promoter 5´-3´

CCTTTCCAGATTGTCTCCGATTGATTTCACTTATCTCAATCACCACCAGATTGTGTCTTCTCACTTGGACACGATCTGAACGGTCCTCAATTGTGTGCGAACACTTCA
CTTTCCGATATTGGCTGCCATTCTGGTCATTGTGAAGGTCTACGTTTTTGTGGCTCATAAAAAGTCCCAGCTTTCGGCACAAGTAATTGTAAGACTGCCAAACTACC
AAGGAAGAGGTTAAATGACTTATCTGGTAAGTGTTTTTCGGAAATATTTTGCTGTATGGAGTCACTAACCACACACGACCTGTTTTCTCTACTCGAGCTTGAGAAAG
TCACTAACCCAAGAAAATTGAAGCCTGTAATGTTCGTCATCCCTTCTCTTAGCGTTTATCAAGCCCACGAATACGTCCAGCCAAAAGAGAATTATTGCTGCCGGAAT
AGCGGAGGTGAGTGCTTCTAAGTCACTGAGGTTGTCAACATGGTCAATGTAAGGAAACTTCTGACTGTTCCTTCTCTCAGAATCGAGGTAGATAATATGTGGACCA
AGAAATGTGTTTTGTAGACAGCGTATTTTCTGGTGTGTTAGCCTCGTACATAAGCTAACATCAGAGATTCATGGGAAACCGGTGTTGATGTTATCACTAGTGTGAA
TACTAGGGAATGATCTGGGGTTTGGAAAGTATTAGTTGGTCCCATCATTATTCTTCACACTCGGAGCATCAGGGCTTGATGACTTGAAAACGCCATCTTGTGTCTCC
GTGTAAGTGGAAATATAACATAAAATAAGCTTTTTACGTTCGAGACACATATCGACCAGCTTAGTTGGGCAATGAATGCAATCTTCTCTTTGGGATTCAGTCTAGG
AATTTCCCATTTTTAAAGGAGAAAGTTCCTTTAAAATCTTTTTAAGAGTCTAATTTCGAGAAACTACATTAAACAGTGTGCGATTTTATCCTCATGTGTTCGATGAAG
GAGCTAGAAAAACATTACTCGGCAGTCATCCTTCCGCAAAGAAACCATAGTGAGAGGTTGTTGGCGAAAACGTTTTTGCCAGTAAAGATATAATATTAGCAGACCT
GTGACTGGATGCTGATCCCACATTGCAGAATAACTTGACTTCTCCGGACTTGAATTGTCCCTATTAATGAACTCATCTCCTCCACTATATGTGTTCTGAAAAATTAAC
CTAGAAAATATTCAAATATCTTCAGGCTATAATATGAAGTTTCATAGTCTAAGAAAAGAGATTATTCCATCCAAAACCAGATTTTCACGCTCAATTTCAGAACCGTA
GTTCTTGTCAACATTTCTGATGAGTGCTAGTTAGTTTCTGACAAATTGAAGTGATAAGAAAGTAAGAAGTGATGTCGGAAAGAAGATGAGGGTGGTGACAAGCAG
TCGAAATCCAGGAGTCACGGCTAATCCAATGGTCGCACCACTTTGTAAGTCGGAATTCTTATAACCTTCGTGGACAGTCGTGTTATTTTCACGCTCAATTGTTGTGA
ACTGGGTCTTCCAATTTAGCGTCTGCAGAAGTATCTGAACGTCTCTTACTCGCTTAAGTCAGAGACAAACTAGTGAGACATTTTCAAGGTCTTCACAATTGTTTTCT
GTTGATTTAGATGTGAAAAATCATTCATATGATGCACTTGGAGATGTTTAAGTTCCGATCTAACGAAGCATATCTCCAACAGTTAACTGCTCTAATAAATTGATGAC
TTCACGTGAACTAAAACCCATAAAGATAAAATGGCAAACTTTATTTTACATTATTACCTAAACTTTTGTCATCTTGAAACTCGAAAAATGTGGCACTTGAATGAGAA
GTTGCTGACTAACAAAAACGTTTCATAATGAATTAGGAGGTTCAGACTGTAACAAGTTATTTTCTGAAAGTAAATGTGACTAGAATGTGGTTTGAGTAAGTAACAA
TCACCTCGACGAATACCTTGTTTTACTGCCCTGAATTGAAGGTAGCGCTAAATAAACGTAGTATAGACTATATATAGCCAGCAAAATCGAAAGGCTGCCATACTTG
TTCGTTAGTGTAAACGATAGACCTTACG

EGFP-reporter gene 5´-3´

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCG
AGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGC
AGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACG
GCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGG
GCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGG
ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGA
GCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTG

Ubiquitin (Smp_335990) terminator 5´-3´

GAGGTTAACTTTAATTTTGTTCATTTAACTTAATAAACGTTTAATTTAACTTATTAAAATGTTTTCTTCCTAATTTGTAGCGTATCTTAGTGTAACTGTGCTTCCTATTA
GATTGAAATTCTATCGTAATCTGCAATTAGGCTCATGGGGATATGTTATGCCAATTTTCCGAAAATTGTTAGCACTTAGGTTTCGGTGTTTTTTTACTAGGATGTCAT
GTTTTATTCATATTTGCGTCCTTGCTCTTGTGTGTGGTATGCGTCGCTATGGTCTCAGCTTTGGGCTGCATTGTTAAGTTTGAG

Supplementary Table 2

Predicted promoter element	Abbreviation	5‘ to 3’ sequence	Position	Sources
Promoter Rank 1	**Prom rank 1**	TAGTATAGACTATATATAG CCAGCAAAATCGAAAGGC TGCCATACTTGTT	1981 - 2030 bp	Berkeley Drosophila Genome Project: Neural Network Promoter Prediction
Core Promoter	**Core Prom**	GTAAGTAACAATCACCTC GACGAATACCTTGTTTTAC TGCCCTGAATTGAAGGTA GCGCTAAATAAACGTAGT ATAGACTATATATAGCCA GCAAAATCGAAAGGCTGC CATACTTGTTCGTTAGTGT AAACGATAGACC	1912 – 2051 bp	Butler and Kadonaga 2002
Enhancer Box	**E-Box**	CACGTG	1712 – 1717 bp	Chaudhary and Skinner 1999
CCAAT/enhancer-binding protein binding site	**C/EBPalp**	ACTTTATTTT	1747 – 1756 bp	Kovács *et al.*, 2003; AliBaba2.1 TRANSFAC 4.0
Octamer-Binding Transcription Factor 1	**Oct-1**	TATTTTACAT	1751 – 1760 bp	Roberts *et al.,* 1991; AliBaba2.1 TRANSFAC 4.0
CCAAT/enhancer-binding binding site	**C/EBPalp**	TATTACCTAAA	1761 – 1771 bp	Kovács *et al.,* 2003; AliBaba2.1 TRANSFAC 4.0
Heat Shock Transcription Factor 1	**HSF**	GAGAAGTTGCTG	1812 – 1823 bp	The Human Gene Database genecards.org; AliBaba2.1 TRANSFAC 4.0
Activator protein 1	**AP-1**	TGCTGACTAA	1819 – 1828 bp	Hess *et al.,* 2004; AliBaba2.1 TRANSFAC 4.0
Transcription Factor SP-1	**SP1**	TAGGAGGTTCA	1850 -1860 bp	The Human Gene Database genecards.org; AliBaba2.1 TRANSFAC 4.0
Fushi Tarazu element	**FTz**	AAATGTGACTA	1888 – 1898 bp	Ohtsuki *et al.,* 1998; AliBaba2.1 TRANSFAC 4.0
GATA-binding factor 1	**GATA-1**	ACTAGAATGT	1895 – 1904 bp	AliBaba2.1 TRANSFAC 4.0
CAAT-Box like sequence	**CAAT-Box**	GTAAGTAACAATC	1912 – 1924 bp	Busek *et al.,* 2002; Juven-Gershon *et al.,* 2010; Godbey 2021
Hepatocyte nuclear factor 1 binding element	**HNF-1**	AACAATCACCT	1918 – 1928 bp	Au *et al.,* 2008; AliBaba2.1 TRANSFAC 4.0
B recognition element	**BRE**	TTGTTTT	1940 – 1946 bp	Ohler *et al.,* 2002; Juven-Gershon *et al.,* 2010; Godbey 2021; YAPP eukaryotic core promoter predictor
TATA-Box containing region	**TATA-Box**	TATAGACTATATATAGCCA	1984 – 2002 bp	Godbey 2021; YAPP eukaryotic core promoter predictor
TATA-Box binding protein	**TBP**	GACTATATATAG	1988 – 1999 bp	AliBaba2.1 TRANSFAC 4.0
CCAAT/enhancer-binding protein binding site	**C/EBPalp**	CAAAATCGAA	2004 – 2003 bp	AliBaba2.1 TRANSFAC 4.0
Serum response factor	**SRF**	GCCATACTTG	2019 – 2028 bp	AliBaba2.1 TRANSFAC 4.0
Transcription Initiator Element	**INR**	CCATACT	2020 – 2026 bp	Godbey 2021; YAPP eukaryotic core promoter predictor
downstream promoter element	**DPE**	AGACC	2047 – 2051 bp	YAPP eukaryotic core promoter predictor
Kozak-Sequence	**Kozak**	CTTACGATG GT Consensus: [N N (T/C) (A/G) N (A/G/T) ATG N (C/A/T) N)]	2051 – 2061 bp	Kozak 1987, Seeber *et al.,* 1997; WEBLOGO tool

Supplementary Table 3

Predicted promoter element	Abbreviation	*S. mansoni* Sequence 5´- 3	Position in construct	Source
Promoter rank 2	**Prom rank 2**	TTGTGGCTCATAAAAAGTCCCAGCTTTCGGCA CAAGTAATTGTAAGACTG	156 – 205 bp	Berkeley Drosophila Genome Project: Neural Network Promoter Prediction
Transcription Factor SP-1	**SP-1**	GTCTCCGAT	13 – 21 bp	AliBaba2.1 TRANSFAC 4.0
Octamer-Binding Transcription Factor 1	**Oct-1**	TGATTTCACT	22 -31 bp	AliBaba2.1 TRANSFAC 4.0
GATA-binding factor 1	**GATA-1**	TTATCTCAAT	31-40 bp	AliBaba2.1 TRANSFAC 4.0
CCAAT/enhancer-binding site	**C/EBPalp**	CACGATCTGA	70 – 79 bp	AliBaba2.1 TRANSFAC 4.0
putative TATA-Box containing region	**TATA-Box**	CTCATAAAAAGT	162 – 173 bp	YAPP eukaryotic core promoter predictor
CCAAT/enhancer-binding site	**C/EBPalp**	AAGTAATTGT	189 – 198 bp	AliBaba2.1 TRANSFAC 4.0
downstream promoter element	**DPE**	GGTT	224 – 227 bp	YAPP eukaryotic core promoter predictor
CCAAT/enhancer-binding site	**C/EBPgam**	TTTCGGAAAT	251 – 260 bp	AliBaba2.1 TRANSFAC 4.0
Activator protein 1	**AP-1**	AGTCACTAAC	275 – 284 bp	Hess *et al.,* 2004; AliBaba2.1 TRANSFAC 4.0
CAAT-Box like sequence	**CAAT-Box**	GGTCAATGT	471 – 479 bp	Godbey 2021; AliBaba2.1 TRANSFAC 4.0

Predicted terminator element	Abbreviation	*S. mansoni* Sequence 5´- 3	Position in construct	Source
mRNA 3´Untranslated Region	**3´UTR**	GAGGTTAACTTTAATTTTGTTCATTTAACTTAA TAAACGTTTAATTTAACTTATTAAAATGTTTTCT TCCTAATTTGTAGC	2777 – 2857 bp	Berriman *et al.,* 2009; Protasio *et al.,* 2012
Polyadenylation signal	**Poly(A)**	TGA (52 bp) ATTAAA (21 bp) GC	2774 – 2857 bp	DNAFSMiner-DNA Functional Site Miner polyadenylation side prediction too
Terminator Stem-loop 1	**TL1**	AGAG	2776 – 2779 bp	RNAfold Webserver: Vienna tools
Terminator Stem-loop 1	**TL1**	CUUU	2785 – 2788 bp	RNAfold Webserver: Vienna tools
Terminator Stem-loop 2	**TL2**	UUUUGUUCA	2791 – 2799 bp	RNAfold Webserver: Vienna tools
Terminator Stem-loop 3	**TL3**	AAACGUUUAAU	2811 – 2821 bp	RNAfold Webserver: Vienna tools
Terminator Stem-loop 3	**TL3**	AUU	2829 – 1831 bp	RNAfold Webserver: Vienna tools
Terminator Stem-loop 3	**TL3**	AAAUGUUU	2833 – 2840 bp	RNAfold Webserver: Vienna tools
Terminator Stem-loop 2	**TL2**	UGUAGCAAAA	2852 – 2857 bp	RNAfold Webserver: Vienna tools

Supplemental Table S4. Compilation of primers used for cloning procedure, PCR, RT-PCR analysis amd crRNAs

Methods Chapter in which the primers were used	Primer name	Primer sequence	Primer purpose
Construction of the EGFP reporter gene	SmUbi_promoter_fw	5´-CCT TTC CAG ATT GTC TCC GA-3´	cloning pJC53.2_SmUbi-EGFP-SmUbi, initial amplification of Sm*ubi* promoter squence
Construction of the EGFP reporter gene	SmUbi _promoter_rev	5´-CGT AAG GTC TAT CGT TTA CAC T-3´	cloning pJC53.2_SmUbi-EGFP-SmUbi, initial amplification of Smubi promoter squence
Construction of the EGFP reporter gene	SmUbi _terminator_fw	5´- TGG AGG TTA ACT TTA ATT TTG-3´	cloning pJC53.2_SmUbi-EGFP-SmUbi, initial amplification of Smubi terminator squence
Construction of the EGFP reporter gene	SmUbi _terminator_rev	5´- CCA AAG GTA TTG CGA ACT -3´	cloning pJC53.2_SmUbi-EGFP-SmUbi, initial amplification of Smubi terminator squence
Construction of the EGFP reporter gene, Particle bombardment	EGFP_fw	5´- ATG GTG AGC AAG GGC GAG -3´	cloning pJC53.2_SmUbi-EGFP-SmUbi, Gibson assembly, validation of EGFP transcription after particle bombardment
Construction of the EGFP reporter gene, Particle bombardment	EGFP_rev	5´ CTT GTA CAG CTC GTC CAT-3´	cloning pJC53.2_SmUbi-EGFP-SmUbi, Gibson assembly, validation of EGFP transcription after particle bombardment
Construction of the EGFP reporter gene	SmUbi_promoter_gibson_fw	5´-GGA GAA TTA ACC CTC ACT AAA GGG AGA CCT TCC TTT CCA GAT TGT CTC CGA-3´	cloning pJC53.2_SmUbi-EGFP-SmUbi, Gibson assembly
Construction of the EGFP reporter gene	SmUbi_promoter_gibson_rev	5´- GGA GAA TTA ACC CTC ACT AAA GGG AGA CCT TCC TTT CCA GAT TGT CTC CGA-3	cloning pJC53.2_SmUbi-EGFP-SmUbi, Gibson assembly
Construction of the EGFP reporter gene	SmUbi_terminator_gibson_fw	5´-GCA TGG ACG AGC TGT ACA AGT GGA GGT TAA CTT TAA TTT TG 3´	cloning pJC53.2_SmUbi-EGFP-SmUbi, Gibson assembly
Construction of the EGFP reporter gene	SmUbi_terminator_gibson_rev	5´-GGG AGA TTT AGG TGA CAC TAT AGA AGT GAC CTT CCA AAG GTA TTG CGA ACT AC-3´	cloning pJC53.2_SmUbi-EGFP-SmUbi, Gibson assembly
Donor template synthesis	50HA_GSH1_Donor_fw	5'Amine-ATG CCA CAA GTG TTT ACA AGG AAA ATG ATA CCT CGA GAT CAA TTA TCT Gac ctt tcc aga ttg tct ccg-3'	Synthesis of SmUbi-EGFP-SmUbi dsDNA donor template
Donor template synthesis	50HA_GSH1_Donor_rev	5'Amine-ACA CCA TCC CAA TTC CAG CTA TTGA ACC AGC CTT CAT CTA CCG AGT TCC Acc aaa ggt att gcg aac tac-3'	Synthesis of SmUbi-EGFP-SmUbi dsDNA donor template
Editing of GSH1 and determination of indel frequencies	GSH1-Tide_fw	5'-GAT GAG TGC ATC GAT CGA TTA CTA AGG AAT TG-3'	Library prediction for determination of editing efficiency by TIDE-analysis
Editing of GSH1 and determination of indel frequencies	GSH1-Tide_rev	5'- GAC TCC TGA TTG AGG GTT GCA GGT G -3'	Library prediction for determination of editing efficiency by TIDE-analysis
Editing of GSH1 and determination of indel frequencies	GSH1-NGS_fw	5'- ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT CCT TTA GTT ATG GCT AAT TTT TAA CAA CCG-3'	NGS library for pricise analysis of indels after Cas induced editing of GSH1, including Illumina adapters
Editing of GSH1 and determination of indel frequencies	GSH1-NGS_rev	5'- GAC TGG AGT TCA GAC GTG TGC TCT TCC GAT CTT GTA TTT AAT TTT GAG GTA GTA GGT CAA AC	NGS library for pricise analysis of indels after Cas induced editing of GSH1, including Illumina adapters
Precise transgene integration	5´-KI_GSH1 primer fw	5´- GTGTATTTCCTATGCTATGAAAATACCT-3´	amplification of 5´-integration site of SmUbi-EGFP-SmUbi at GSH1, amplification wildetyp GSH1
Precise transgene integration	5´-KI_Ubi-Prom HA left rev	5´- att gag gac cgt tca gat cgt gtc-3	amplification of 5´-integration site of SmUbi-EGFP-SmUbi at GSH1
Precise transgene integration	3´-KI_EGFP-Ubi-Ter fw	5´- CAT GGA CGA GCT GTA CAA GTG GAG GT-3	amplification of 3´-integration site of SmUbi-EGFP-SmUbi at GSH1
Precise transgene integration	3´-KI_GSH1 primer rev	5´- GGT TCC ATA CTA TGC AGT TTC CTT AAA C-3´	amplification of 3´-integration site of SmUbi-EGFP-SmUbi at GSH1, amplification of wildetype GSH1
Determination of transgene activity	RT-PCR_EGFP_fw	5'-GAC TGG GTG CTC AGG TAG TG-3'	RT-PCR, validation of transgene activity after auccessful integration
Determination of transgene activity	RT-PCR_EGFP_rev	5'-CAA GGA CGA CGG CAA ATA CA-3'	RT-PCR, validation of transgene activity after auccessful integration
Determination of transgene activity	RT-PCR_SmGAPDH_fw	5'-ATG GGA CAT TTC CAG GCG AG-3'	amplification of RT-PCR control transcript
Determination of transgene activity	RT-PCR_SmGAPDH_rev	5'-CCA ACA ACG AAC ATG GGT-3'	amplification of RT-PCR control transcript

Enzyme	Name	Target locus	Sequence
***Sp* Cas9**	crRNA_Cas9_GSH1	GSH1	5'- GAG AUC AAU UAU CUG ACA AU-3
***As* Cas12a**	crRNA_Cas12a_GSH1	GSH1	5' GAG AUC AAU UAU CUG ACA AUG GAA 3'