Check for updates

RESEARCH ARTICLE

REVISED

Large CRISPR-Cas-induced deletions in the oxamniquine resistance locus of the human parasite *Schistosoma mansoni* [version 2; peer review: 4 approved]

Geetha Sankaranarayanan*, Avril Coghlan (ID)*, Patrick Driguez (ID), Magda E. Lotkowska, Mandy Sanders, Nancy Holroyd, Alan Tracey (ID), Matthew Berriman (ID), Gabriel Rinaldi (ID)

Wellcome Sanger Institute, Hinxton, CB10 1SA, UK

* Equal contributors

Abstract

Background. At least 250 million people worldwide suffer from schistosomiasis, caused by *Schistosoma* worms. Genome sequences for several *Schistosoma* species are available, including a high-quality annotated reference for *Schistosoma mansoni*. There is a pressing need to develop a reliable functional toolkit to translate these data into new biological insights and targets for intervention. CRISPR-Cas9 was recently demonstrated for the first time in *S. mansoni*, to produce somatic mutations in the *omega-1* (*ω1*) gene.

Methods. We employed CRISPR-Cas9 to introduce somatic mutations in a second gene, *SULT-OR*, a sulfotransferase expressed in the parasitic stages of *S. mansoni*, in which mutations confer resistance to the drug oxamniquine. A 262-bp PCR product spanning the region targeted by the gRNA against *SULT-OR* was amplified, and mutations identified in it by high-throughput sequencing.

Results. We found that 0.3-2.0% of aligned reads from CRISPR-Cas9-treated adult worms showed deletions spanning the predicted Cas9 cut site, compared to 0.1-0.2% for sporocysts, while deletions were extremely rare in eggs. The most common deletion observed in adults and sporocysts was a 34 bp-deletion directly upstream of the predicted cut site, but rarer deletions reaching as far as 102 bp upstream of the cut site were also detected. The CRISPR-Cas9-induced deletions, if homozygous, are predicted to cause resistance to oxamniquine by producing frameshifts, ablating *SULT-OR* transcription, or leading to mRNA degradation *via* the nonsense-mediated mRNA decay pathway. However, no *SULT-OR* knock down at the mRNA level was observed, presumably because the cells in which CRISPR-Cas9 did induce mutations represented a small fraction of all

Open Peer Review

Reviewer Status ✓ ✓ ✓ ✓

	Invited Reviewers			
	1	**2**	**3**	**4**
version 2 (revision) 20 Jan 2021				
version 1 23 Jul 2020	✓ report	✓ report	✓ report	✓ report

1. **Phillip LoVerde** (ID), University of Texas Health Science Center at San Antonio, San Antonio, USA

2. **Arnon Jurberg**, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

3. **Mattie Pawlowic** (ID), University of Dundee, Dundee, UK

4. **Patrick Skelly** (ID), Tufts University, North Grafton, USA

Any reports and responses or comments on the article can be found at the end of the article.

cells expressing *SULT-OR*.

Conclusions. Further optimisation of CRISPR-Cas protocols for different developmental stages and particular cell types, including germline cells, will contribute to the generation of a homozygous knock-out in any gene of interest, and in particular the *SULT-OR* gene to derive an oxamniquine-resistant stable transgenic line.

Keywords

Schistosoma mansoni, Sulfotransferase, Transfection, Transgenesis, Genome editing, CRISPR-Cas9, Amplicon sequencing, CRISPResso

This article is included in the Wellcome Sanger Institute gateway.

Corresponding author: Gabriel Rinaldi (gr10@sanger.ac.uk)

Author roles: Sankaranarayanan G: Formal Analysis, Investigation, Methodology, Validation, Writing – Review & Editing; **Coghlan A**: Data Curation, Formal Analysis, Investigation, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Driguez P**: Formal Analysis, Investigation, Validation, Writing – Review & Editing; **Lotkowska ME**: Investigation, Writing – Review & Editing; **Sanders M**: Project Administration, Writing – Review & Editing; **Holroyd N**: Project Administration, Writing – Review & Editing; **Tracey A**: Data Curation, Software, Writing – Review & Editing; **Berriman M**: Conceptualization, Funding Acquisition, Resources, Supervision, Writing – Review & Editing; **Rinaldi G**: Conceptualization, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

> **REVISED** **Amendments from Version 1**
>
> The questions raised by the four reviewers have been fully addressed, and accordingly, the text has been modified in the sections Methods, Results and Discussion as it has been indicated in both the V2 version of the manuscript, and in the Responses to the reviewers' comments. In addition, the Extended Data word file has been edited as the Supplementary Figure S1 legend was updated.
>
> **Any further responses from the reviewers can be found at the end of the article**

Introduction

Schistosomiasis is a major neglected tropical disease (NTD) affecting more than 250 million people worldwide[1]. *Schistosoma mansoni* and *S. japonicum* are the agents of hepato-intestinal schistosomiasis manifested by abdominal pain, liver inflammation and fibrosis that leads to portal hypertension. Infection with *S. haematobium,* agent of urogenital schistosomiasis, is associated with infertility, haematuria, kidney pathology and squamous cell carcinoma of the bladder. In addition, all forms of schistosomiasis are associated with systemic morbidities that include malnutrition, anaemia, physical and/or cognitive impairment and stunted development in children[2]. Currently, praziquantel is the single effective drug to treat the infection, and is employed in mass drug administration programmes across endemic areas, which could eventually lead to drug resistance emerging[3]. Therefore, there is an urgent need for the development of novel drugs and vaccines[4]. Understanding the basic biology of schistosomes at the cellular and molecular levels is critical to identify exploitable vulnerabilities of the parasite. High-throughput datasets, including high quality reference genomes for the three main species of schistosomes[5–7], have been generated. More recently, a thorough transcriptome analysis during the parasite's intra-mammalian development[8], and the identification of different cell types by single-cell RNA sequencing of various life cycle stages[9,10] represent significant steps towards deciphering cell fate and pathways involved in parasite development and host-parasite interactions.

In parallel to the generation of large-scale datasets, a functional genomics toolkit is needed to experimentally investigate hypotheses that emerge from these data, to confirm biological insights and validate targets for intervention. Recently, a large RNAi-based gene silencing screen, encompassing almost one third of *S. mansoni* protein-coding genes, revealed genes associated with parasite viability and potential targets for drug development[11]. However, not every gene is susceptible to RNAi, the effect is transient and highly variable depending on the expression level, tissue localisation and half-life of the target mRNA and protein. In addition, off-target effects are common, in particular when long dsRNA molecules are used, and the gene silencing is typically not heritable unless an RNAi-based construct is employed as a transgene expressed in the germ line[12]. Therefore, to truly examine gene function across the life-cycle, transgenesis-based approaches already available for model organisms[13,14] need to be developed for *S. mansoni,* including protocols to create genetically-modified parasite strains with homozygous gene knock-outs, and site-specific gene mutations.

Promising progress with transgenesis and genome editing has been achieved. Retrovirus transduction of schistosome developmental stages, including eggs, has proved effective, and will likely be a key delivery system in the generation of stable transgenic lines[15–18]. Site-specific integration of transgenes and highly precise site-specific genome editing using CRISPR-Cas technology will be a key step[19]. CRISPR-Cas9 has recently been used to create a heritable gene knock-out line in the parasitic nematode *Strongyloides stercoralis*[20]. In *S. mansoni*, the technology has been used to produce mutations in the *omega-1* (*ω1*) gene in somatic cells of the egg[21], and in a related parasitic flatworm, the liver fluke *Opisthorchis viverrini*, CRISPR-Cas9 mutations have been introduced into the granulin gene in somatic cells of adult worms[22]. Somatic mutations in these two flatworms were associated with dramatic reductions in *ω1* and granulin mRNA levels, respectively, and produced *in vitro* and *in vivo* phenotypic effects shedding new light on their functional roles and contributions to pathogenesis[21,22].

Whether different CRISPR-Cas protocols are needed to deliver site-specific mutations in *S. mansoni* genes expressed in other tissues or developmental stages remains to be determined. Likewise, the types of mutations to be expected and the degree of mRNA knock-down in different genes is not yet known. In the current study we have used CRISPR-Cas9 to introduce site-specific mutations in a second *S. mansoni* gene, to better understand how the CRISPR-Cas system works when applied to *S. mansoni.* We compared the efficiency of the approach in different developmental stages of the parasite: eggs, mother sporocysts (the first intramolluscan stage) and adult worms. The mutations produced by CRISPR-Cas9, including their sizes and locations, were characterised. We chose the *SULT-OR* sulfotransferase (*Smp_089320*) gene as a target because recessive mutations in this gene, both induced in laboratory conditions and detected in field samples, confer resistance to the drug oxamniquine (OXA)[23,24]. In addition, it is mostly expressed in the intra-mammalian stages of the life cycle (schistosomula and adults)[25] that would likely be the target of any new intervention strategy. Our findings provide insights that will help pave the way towards using CRISPR-Cas to achieve the generation of stable genetically-engineered schistosomes.

Methods
Ethics statement
The complete life cycle of *Schistosoma mansoni* NMRI (Puerto Rican) strain is maintained at the Wellcome Sanger Institute (WSI) by breeding and infecting susceptible *Biomphalaria glabrata* snails and mice. The mouse experimental infections and rest of regulated procedures were conducted under the Home Office Project Licence No. P77E8A062 held by GR. All protocols were revised and approved by the Animal Welfare and Ethical Review Body (AWERB) of the WSI. The AWERB is constituted as required by the UK Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012.

Animal procedures

To obtain the parasite material described below, susceptible *Biomphalaria glabrata* snails and mice are routinely infected. In brief, snails exposed to 30 *S. mansoni* miracidia are maintained in aerated tanks in water and moved into dark cupboards at 28°C when they start shedding cercariae. For the mouse infections, cercariae are collected by placing ~50 infected snails in a 200ml glass beaker containing water and exposed to bright light for an hour. To quantify the cercariae, 12 5µL aliquots of cercarial water are sampled, mixed with Lugol (cat.# 62650, Sigma Aldrich), and the larvae counted under a dissecting microscope. Thereafter, the cercarial solution is diluted to a final concentration of 500 cercariae/ml, and immediately used for percutaneous infection. Eight to 12 weeks old outbred HsdOla:TO female mice are infected with 250 *S. mansoni* cercariae for 40 minutes by percutaneous infection through the tail. Briefly, tubes containing 5.5 ml of conditioned water[26] are prefilled and placed onto a bespoke anaesthesia rig. The mice are anaesthetised in an induction box using 4% isoflurane (Vetflurane®); 1 l/min oxygen, and eye ointment used to prevent corneal damage. Under anaesthesia, the mice are carefully transferred onto individual holders on the rigs and their tails inserted into the test tubes. Nose cones are adjusted for each animal, and anaesthesia is maintained at 2% isoflurane;1 l/min oxygen. In each test tube, 500 µL of a stock solution containing 500 cercariae/ml is added (i.e. 250 cercariae per mouse). After 40 minutes, animals are removed from the anaesthesia rigs, placed back into their cage and monitored until full recovery from the anaesthesia.

At 6 weeks post infection the mice are euthanised by intraperitoneal injection of 200 µl of 200 mg/ml pentobarbital (Dolethal®) supplemented with 100 U/ml heparin (cat.# H3393, Sigma Aldrich), adult worms recovered by portal perfusion (the portal vein is sectioned followed by intracardiac perfusion with phenol-red-free DMEM, cat.# 31053-044 ThermoFisher Scientific, containing 10 U/mL heparin), and whole livers collected.

The outbred HsdOla:TO female mice are commercially outsourced (Envigo, UK), housed in GM500 Individually Ventilated Cages or IsoCage N -Biocontainment Systems (Tecniplast) and maintained on individual air handling units at 19 to 23°C and 45–65% humidity. Animals are given access to food and water *ad libitum*, maintained on a 12-hour light/dark cycle, and housed in groups of no more than 5 adults per cage. Welfare assessments are carried out daily, and abnormal signs of behaviour or clinical signs of concern are reported. All personnel at the WSI performing welfare checks on animals are trained and assessed as competent by qualified named individuals.

Parasite material

Developmental stages of *S. mansoni* were collected and maintained as described[27]. In brief, mixed-sex adult worms were collected by portal perfusion of experimentally-infected mice 6 weeks after infection (above), washed with 1x Phosphate-Buffered Saline (PBS, cat.# D8662, Sigma Aldrich), supplemented with 200 U/ml penicillin, 200 µg/ml streptomycin and 500 ng/ml amphotericin B (cat.# 15240062, ThermoFisher Scientific),

and cultured in complete high-glucose DMEM (cat.# 11995065, ThermoFisher Scientific), 10% fetal bovine serum (FBS, cat.# 10500064), 200 U/ml penicillin, 200 µg/ml streptomycin and 500 ng/ml amphotericin B (cat.# 15240062, ThermoFisher Scientific) at 37°C, under 5% CO_2 in air. All media components were purchased from ThermoFisher Scientific. *S. mansoni* eggs were isolated from the livers of experimentally-infected mice removed after the portal perfusion[28]. The livers were finely minced and digested overnight in the presence of 0.5% *Clostridium histolyticum* collagenase (cat.# C5138, Sigma Aldrich), followed by three washes with 1x PBS and filtered through 250 µm and 150 µm sieves. The filtrate was passed through a Percoll-sucrose gradient prepared by mixing 8 ml of Percoll with 32 ml of sterile-filtered 0.25M sucrose (Percoll cat.# P1644, Sucrose cat.# 84097, Sigma Aldrich), and the resulting purified eggs washed in 1x PBS and cultured in complete DMEM medium at 37°C, under 5% CO_2 in air as described[29]. Primary sporocysts were obtained by transferring miracidia hatched from freshly collected eggs into complete sporocyst medium (MEMSE-J, 10% Fetal Bovine Serum, 10mM Hepes, 100 U/ml penicillin, 100 µg/ml streptomycin) and cultured in a hypoxia chamber in a gas mixture of 1% O_2, 3% CO_2 and balance N_2, at 28°C[27].

CRISPR-Cas 9 ribonucleoprotein complex assembly

We explored the activity of a 'two-piece' guide RNA that included a (1) CRISPR RNA (crRNA) molecule of 20 nucleotides target-specific sequence, and (2) the conserved 67 nucleotide trans-activating crRNA (tracrRNA). The crRNA sequence 5'-ACAATCCAAGTTATCTCAGC-3', spanning positions 19-38 from the first codon of exon 1 of *SULT-OR* (*Smp_089320*) and followed by the protospacer adjacent motif (PAM) TGG (Figure 1), was designed using the web-based tool CRISPR RGEN Tools (Computational tools and libraries for RNA-guided endonucleases, RGENs). The crRNA, the fluorescently labelled tracrRNA (Alt-R® CRISPR-Cas9 tracrRNA, ATTO™ 550), and the recombinant *Streptococcus pyogenes* Cas9 nuclease containing a nuclear localization sequence (Alt-R® S.p. Cas9 Nuclease V3) were purchased from IDT. The CRISPR-Cas9 ribonucleoprotein complex (RNP) was assembled *in vitro* following the manufacturer's recommendations slightly modified based on 30 by combining the 'two-piece' gRNA with the Cas9 nuclease (163.7 kDa). Briefly, the 'two-piece' gRNA was generated by mixing equal volumes of 200 µM *SULT-OR* crRNA and 200 µM ATTO™ 550 tracrRNA in IDT buffer. The RNA oligos were annealed by incubating the mixture at 95°C for 5 min followed by a slow cooling to room temperature for at least 10 min. Thereafter, the RNP was assembled by combining 100 pmol Cas9 nuclease (stock concentration, 10 µg/µl = 61 µM) with 150 pmol 'two-piece' gRNA. The RNP was gently mixed avoiding pipetting, incubated at room temperature for 10 min and kept on ice. Immediately before the parasite transfection Opti-MEM media (cat.# 31985070, ThermoFisher Scientific) was added to the RNP to reach a final volume of 100 µl and kept on ice.

CRISPR-Cas9 transfection of schistosome developmental stages

The CRISPR-Cas9 RNP was delivered into *S. mansoni* mixed-sex adult worms, eggs and *in vitro*-transformed mother sporocysts by square-wave electroporation as previously described[32,33]
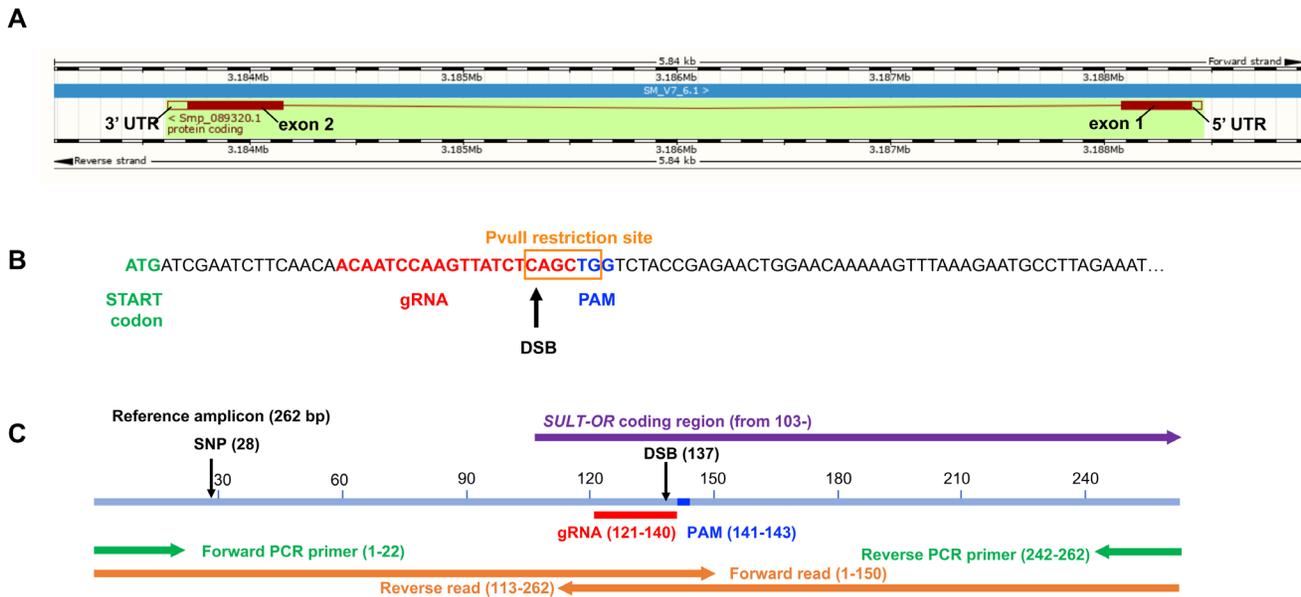
Figure 1. (**A**) Gene model of *SULT-OR* (*Smp_089320*), indicating the position of the two exons, one intron, and UTRs, spanning 4837 bp on the reverse strand of chromosome 6. Boxes filled in dark red represent the protein coding sequence. *Schistosoma mansoni* (PRJEA36577). Assembly: Smansoni_v7 (GCA_000237925.3). Region: Scaffold SM_V7_6:3,183,084-3,188,924. Adapted from WormBase ParaSite 14[31]. (**B**) Nucleotide sequence of the start of exon 1 indicating location and sequence of gRNA target site, predicted double-stranded break (DSB), protospacer adjacent motif (PAM), and the PvuII restriction site. (**C**) Reference PCR amplicon (pale blue), showing the positions of the gRNA, PAM, DSB, forward and reverse PCR primers (green), and forward and reverse sequence reads (orange), as well as a SNP site found in many sequence reads. The diagrams are drawn to scale.

with minor modifications. Briefly, groups of ~16 male and female worms were transferred to a pre-cooled 4-mm electroporation cuvette (BTX), and washed 3 times by gravity with Opti-MEM medium with no FBS and no antibiotic/antimycotic mix. After the last wash, the worms were maintained in 50 µl of Opti-MEM medium and the RNP in 100 µl of Opti-MEM medium (above) was added to the cuvette containing the worms. The eggs isolated from the livers and cultured as described above were collected and washed in Opti-MEM medium (no FBS and no antibiotic/antimycotic mix) 3 times by centrifugation at 400 g for 5 min. After the last wash, the eggs were split into groups of ~10,000, resuspended in 100 µl of Opti-MEM medium containing the RNP (above) plus 50 µl of Opti-MEM to collect all the remaining eggs from the original tube, and transferred to a pre-cooled 4-mm electroporation cuvette (BTX). Three-day old *in vitro*-transformed sporocysts were collected and washed in Opti-MEM medium (no FBS and no antibiotic/antimycotic mix) 3 times by centrifugation at 400 g for 5 min. After the last wash, the sporocysts were split in groups of ~10,000, resuspended in 100 µl of Opti-MEM medium containing the RNP (above) plus 50 µl of Opti-MEM to collect all the remaining sporocysts from the original tube, and transferred to a pre-cooled 4-mm electroporation cuvette (BTX). The final electroporation volume for the schistosome worms, eggs and sporocysts was 150 µl, i.e. the final concentration of the RNP complex in the cuvette was 1.67 µM. The three developmental stages were subjected

to the same electroporation conditions; square-wave, a single pulse of 125V for 20 msec in a BTX Gemini X2 electroporator (BTX). Immediately after electroporation, the schistosome worms and eggs were collected in pre-warmed complete DMEM medium, and the sporocysts in complete sporocysts medium and cultured as described above. Four hours after transfection, three male and female worms and a few thousand eggs and sporocysts were collected for confocal microscopy (below). Four days post-transfection the parasites were collected, washed in 1x PBS and processed for DNA and RNA isolation (below). In addition to the CRISPR-Cas9 experimental condition, i.e. parasites exposed to the CRISPR-Cas9 RNP complex, we included three control groups subjected to the same electroporation protocol: (1) mock-treated group that included parasites exposed to no molecules, (2) parasites exposed to Cas9 nuclease only, and (3) parasites exposed to the 'two-piece' gRNA only. **Extended data Table S1** summarises the experimental conditions and biological replicates performed for each of the three tested developmental stages[34].

DNA isolation and amplicon sequencing libraries

A conventional phenol:chloroform:isoamyl alcohol (25:24:1) protocol was employed to isolate DNA from RNP-transfected parasites and all control groups. Briefly, wet pellets of adult worms, eggs or mother sporocysts stored at -80°C were incubated overnight in the presence of 500 µl genomic DNA lysis buffer (200 mM NaCl, 100 mM Tris-HCl pH 8.5, 50 mM

EDTA pH 8, 0.5 % SDS) and 10 µl of proteinase K (20 mg/ml, cat.# AM2546, Life Technologies) at 56°C with agitation (400 rpm). Thereafter, 5 µl of 4 mg/ml of RNase A (cat.# 7973, Promega) was added to the lysate and incubated at 37°C for 10 min. One volume of phenol-chloroform-isoamyl alcohol (25:24:1) (cat.# p2069, Sigma-Aldrich) was added to the sample, mixed vigorously, incubated at room temperature for 5 min and centrifuged at 14,000 g at room temperature for 15 min. The aqueous top layer was transferred to a new tube, 1 volume (~200 µl) of chloroform:isoamyl alcohol (24:1) (cat.# 327155000, Acros Organic,) was added to the sample, mixed vigorously and centrifuged as above. The aqueous top layer was transferred to a new tube and the DNA precipitated with 0.1 volume of 3 M sodium acetate, 3 volumes of 95%-100% ethanol, and 2 µl of Glycoblue (cat.# AM9516, Thermo Fisher Scientific) overnight at -20°C. The DNA was recovered by centrifugation at 14,000 g at 4°C for 30 min, washed with 500 µl of 70% ethanol, resuspended in pre-warmed nuclease-free water and quantified by Qubit fluorometer. For the indicated samples (**Extended data Table S1**[34]) in order to enrich for *SULT-OR* mutant alleles, 20 ng of DNA was digested with 6 to 12 U of the restriction enzyme PvuII (PvuII-HF, cat.#R3151, NEB) overnight at 37°C.

For the amplicon library preparation, a 2-step PCR protocol was followed. During the first PCR, a 262 bp *SULT-OR*-specific amplicon spanning the predicted double-stranded breaking site (DBS) was generated using 10 ng template DNA (20 µl of 0.5 ng/µl DNA preparation), 300 nM forward and reverse primers (**Extended data Table S2**[34]), and 2x Kapa HiFi Master Mix (cat.# KK2602, Roche) in a 50 µl PCR reaction performed in a Thermocycler (Eppendorf mastercycler X50s). The PCR protocol included an initial denaturation step at 95°C for 3 min, 18 cycles of denaturation step at 98°C for 20 sec, annealing step at 53°C for 15 sec, and extension step at 72°C for 40 sec, followed by a final extension step at 72°C for 5 min. Four PCR reactions per sample were run in parallel and the products were pooled at the end, i.e. a total of 40 ng of each sample DNA preparation was used to generate the amplicon. For sample DNA preparations that were digested with PvuII, two PCR reactions per sample were run in parallel and the products were pooled, i.e. a total of 20 ng of each of two PvuII-digested DNA preparations was used to generate the amplicon. The pooled PCR products for each sample were cleaned up using a column-based kit (cat.# D4014, Zymo DNA Clean and concentrator), eluted in 17 µl of nuclease-free water; 2 µl were used for quantification and the rest entirely used as template in the second PCR for Nextera Indexing (Nextera-XT Index kit FC-131-1001). In a 50 µl reaction the concentrated DNA (15 µl) was mixed with 10 µl of the Nextera index mix (i5 + i7) and 2x Kapa HiFi Master Mix (cat.# KK2602, Roche). The PCR was performed in a Thermocycler (Eppendorf mastercycler X50s) with an initial denaturation step at 95°C for 3 min, 8 cycles of denaturation step at 98°C for 20 sec, annealing step at 55°C for 15 sec, and extension step at 72°C for 40 sec, followed by a final extension step at 72°C for 5 min. The PCR products were purified using a bead-based cleaner kit (cat.# A63880, AMPure XP, Beckman Coulter), eluted in 30 µl of nuclease-free water and quantified using a high sensitivity DNA chip in a Bioanalyzer

(2100 Bioanalyzer Instrument, Agilent Technologies). Equimolar amounts of each library were combined and 20–30% PhiX was added to the mix to introduce complexity into these low-diversity amplicon libraries.

Bioinformatic analysis

Amplicon libraries from the samples summarised in **Extended data Table S1** were sequenced on a MiSeq Illumina sequencing platform spiked with 20-30% PhiX to generate diversity[34]. If a sample had been multiplexed and run on several MiSeq lanes, the fastq files for that particular sample were merged. Trimmomatic version 0.33[35] was used to discard low quality read-pairs where either read had average base quality < 23[36]. To detect CRISPR-induced mutations, the software CRISPResso v1.0.13[37,38] was employed using a window size of 500 bp (-w 500) with the reference amplicon according to *Smp_089320* in the *S. mansoni* V7 assembly from WormBase ParaSite, version 14.0 (August 2019)[31]. In most samples, the majority of reads had a G→A single-nucleotide polymorphism (SNP) at position 28 of the amplicon, presumably due to genetic variation in the population of *S. mansoni* NMRI strain maintained in our laboratory. Thus, although the *S. mansoni* V7 reference assembly has 'G' at this position, we used 'A' at this position in the 'reference amplicon' sequence given to CRISPResso. A window size of 500 bp was used to include the entire amplicon. CRISPResso was run with the -exclude_bp_from_left 30 and -exclude_bp_from_right 30 options in order to disregard the (21-22 bp) primer regions on each end of the amplicon, and the SNP at position 28, when indels and substitutions were being quantified and reads being classified as 'non-homologous end joining' ('NHEJ') or 'unmodified' by CRISPResso.

Gene expression analysis for *SULT-OR* gene

Total RNA was extracted from adult worms, eggs or *in vitro* transformed sporocysts following a phenol:chloroform-based protocol. In brief, four days after transfection, parasites were collected from the culture, washed three times in 1x PBS complemented with antibiotic-antimycotic as described above for each of the three developmental stages, transferred to 1ml of Trizol, incubated at room temperature for ~10 min and stored at -80°C. The parasites in Trizol were mechanically-dissociated using a bead beater homogenizer (Fast Prep-24, MP Biomedicals) using two 20-second pulses at setting four for adult worms and *in vitro* transformed sporocysts, and two 20-second pulses at setting six, after three cycles of freezing-thawing, for eggs. Thereafter, one volume of chloroform was added to the samples, mixed vigorously, centrifuged at 14,000 g at room temperature for 15 min, and the aqueous top layer was carefully transferred to a clean tube. The total RNA was precipitated using an equal volume of 100% molecular biological grade ethanol. Residual DNA was removed by digestion with DNaseI (cat.# E1010, Zymo). RNA was cleaned and concentrated using Zymo RNA clean and concentrator columns, and eluted in 15 µl of nuclease-free water. cDNA was synthesized from 65 -175 ng of total RNA using the iScript cDNA Synthesis Kit (cat.#1708891, Bio-Rad, Hercules, CA). Target-specific primers designed with the assistance of the free web-based PrimerQuest® Tool (IDT) are shown in **Extended data Table S2**, and the amplification efficiencies for each primer

set were estimated to be 90–105% by titration analysis[34,39]. Real time quantitative PCRs (qPCR) were performed in triplicate, in 96-well plates, following an initial denaturation step at 95°C for 3 min followed by 40 cycles of 30 sec at 95°C and 30 sec at 50 °C, and a final melting curve, in a StepOnePlus™ Real-Time PCR System (Applied Biosystems). Reactions run in 10 μl included 300 nM of each target-specific primer, 1 μl of cDNA, and Kapa Sybr FastqPCR Master Mix (cat.# KK4600, Roche). The relative quantification assay[40] was employed using both *S. mansoni* glyceraldehyde-3-phosphate dehydrogenase (*SmGAPDH*, *Smp_056970*) and *S. mansoni* α-tubulin1 (Sm*AT1*, *Smp_090120*) as reference genes. The target gene expression levels were normalised using the control group.

Confocal microscopy

Four hours after electroporation with fluorescently labelled RNP complex, transfected adult worms, eggs or mother sporocysts were collected from the culture, fixed and processed for confocal microscopy imaging. In brief, the parasites were collected and washed three times in 1x PBS complemented with antibiotic-antimycotic solution as described above; adult worms were washed by gravity, and eggs and sporocyst by centrifugation, 400 g for 5 min. After the final wash the parasites were fixed overnight in 4% methanol-free paraformaldehyde (cat.# 28906, Pierce™) diluted in 1x PBS at 4°C, washed three times in 1x PBS, resuspended in mounting media containing 4', 6'-diamidino-2-phenylindole (DAPI) for nuclear staining (cat.# 15596276, Fluoromount-G™ Mounting Medium, with DAPI, Invitrogen), and incubated overnight at 4°C. The parasites were mounted on microscope slides and images taken with a Leica SP8 confocal microscope using appropriate settings to capture DAPI and ATTO 550 fluorochromes. Manipulation of digital images was undertaken with the assistance of the LAS X software (Leica) and was limited to insertion of scale bars, adjustment of brightness and contrast, and cropping. The image enhancement algorithms were applied in a linear fashion across the entire image.

Accession numbers

The sequence data generated in this study are available at the European Nucleotide Archive (ENA) accession number ERP 121238. The accession number for each sample is shown in **Extended data Table S1**, columns P, Q[34].

Statistical analysis

A paired one-sided Wilcoxon test (non-parametric) was used to analyse significant differences between percentages of aligned reads containing deletions (or insertions or substitutions) between CRISPR-Cas9-treated samples and respective matched controls. All Statistical analyses were performed using R, version 3.4.1.

Results

The *SULT-OR* gene belongs to a multi-copy locus on chromosome 6 of *S. mansoni*

The *SULT-OR* gene (*Smp_089320*) belongs to a multi-copy locus containing six other paralogous genes on chromosome 6 of the *S. mansoni* reference genome, version 7 (WormBase ParaSite), (**Extended data Figure S1A, B**[34]). This locus in chromosome 6 has been correctly resolved with no evidence of repetitive regions that frequently appear 'collapsed' within assemblies (**Extended data Figure S1C**[34]). The biological function of SULT-OR remains unknown, except that it converts the pro-drugs OXA and hycanthone to their active forms[23,24]. It displays sulfotransferase activity *in vitro* on exogenous substrates[23], even though the protein shows a low level of sequence similarity to other sulfotransferases, and it is mostly expressed in the intra-mammalian stages of the life cycle (schistosomula and adults, **Extended data Figure S2A**[34])[25]. Intriguingly, *SULT-OR* belongs to a gene family that has expanded in trematodes[41], suggesting it may play an important role in clade-specific biology. However, *ex vivo SULT-OR* RNAi experiments in adult male worms showed no evident phenotypic effects other than becoming resistant to OXA[23]. Single-cell transcriptomic analysis of two-day-old schistosomula[9], adult worms[10], and *in vitro*-transformed mother sporocysts (unpublished) revealed *SULT-OR* mRNA is a marker of parenchymal cell clusters (**Extended data Table S3** and **Extended data Figure S2B**[34]), while its top BLASTP hit in the planarian *Schmidtea mediterranea*, *dd_Smed_v6_9472_0*, is a marker of intestinal cells[42].

A specific gRNA to introduce mutations in exon 1 of *SULT-OR*

The *SULT-OR* gene comprises two coding exons separated by one intron, spans 4837 bp on the reverse strand of chromosome 6, and includes a short 46 bp 5'UTR (Figure 1A). A gRNA was designed to target residues 19 to 38 of the coding region of *SULT-OR* within exon 1, adjacent to a TGG protospacer adjacent motif (PAM) and with the predicted double strand break (DSB; i.e. the predicted Cas9 cut site) 3 bp upstream of the PAM (Figure 1B). Importantly, the sequences homologous to this gRNA's target region are relatively diverged in the paralogous genes on chromosome 6, with many mismatches within the seed region (10-12 bp at its 5' end) (**Extended data Figure S3A**[34]). It has been shown that mismatches in the gRNA 'core' sequence located between 4 to 7 nucleotides upstream of the PAM abolishes off target cleavages[19,43]; hence, our gRNA is expected to be specific to *SULT-OR*.

CRISPR-Cas9 machinery successfully delivered into schistosome developmental stages

To investigate whether the CRISPR-Cas9 machinery, i.e. RNP (ribonucleoprotein) complex containing the Cas9 nuclease and *SULT-OR*-specific gRNA, was successfully delivered into adult worms, sporocysts and eggs, we used fluorescently labelled RNP. Parasites were collected from culture four hours after transfection and fixed for confocal microscopy. The images revealed that the RNP complex entered cells of adult worms, sporocysts and eggs (Figure 2). Even though the parasites were thoroughly washed before fixation, a strong signal outside the tegument was evident, in particular in adult worms, suggesting RNP complex molecules unspecifically bound to the surface of the parasites (Figure 2A and **Movie 1**[44]). However, in addition to the signal in the surface of the parasite, the confocal optical sections revealed fluorescently-labelled
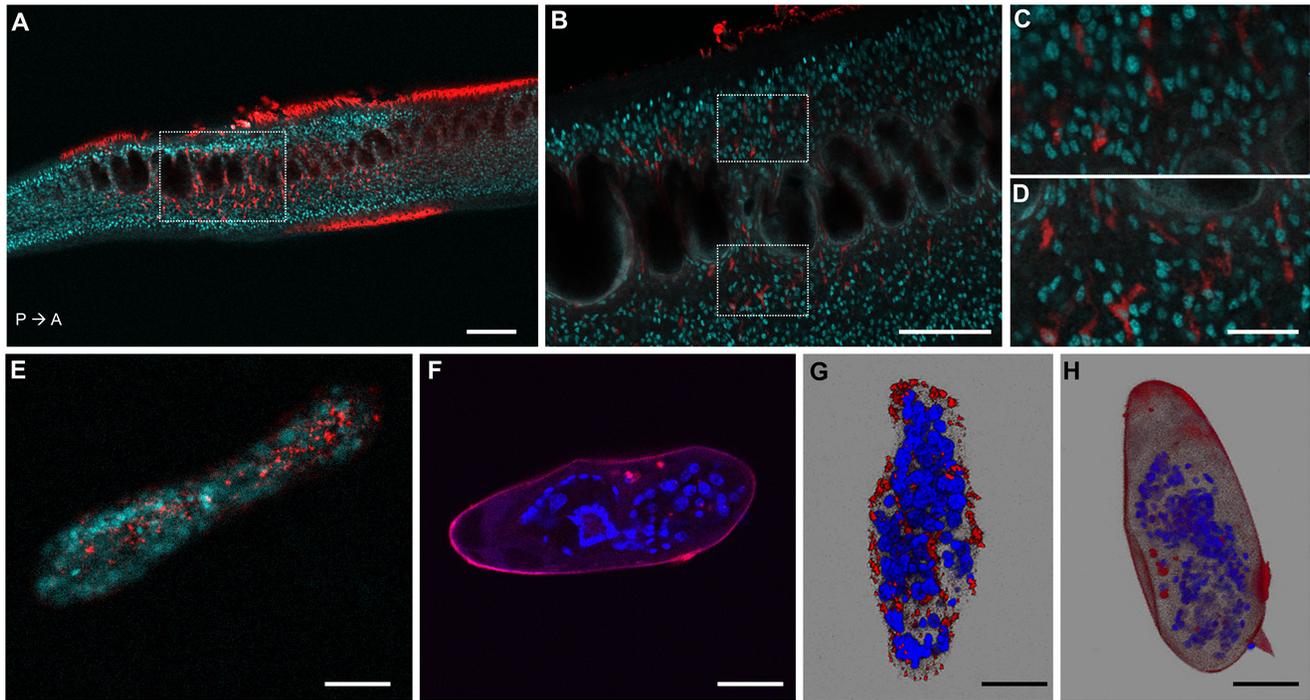
Figure 2. Confocal microscopy images of *S. mansoni* parasites transfected with fluorescently labelled Cas9-gRNA (ATTO™ 550 signal in red), fixed and DAPI-stained (DAPI signal in aqua blue or blue). (**A**) Confocal optical section of a male adult worm. P→ A: indicates the posterior anterior axis. Scale bar: 100 μm. (**B**) Magnified squared-area in (**A**). Scale bar: 50 μm. (**C, D**) Magnified top and bottom squared-areas in (**B**), respectively. Scale bar: 10 μm. (**E, F**) Confocal optical sections of a sporocyst and an egg, respectively. (**G, H**) Maximum intensity projection of z-stack images of a sporocyst and an egg, respectively. Scale bars in **E–H**: 25 μm. The images of worms, sporocysts and eggs, were taken from representative specimens collected from the biological replicate "Experiment 7, tags 50 and 61", "Experiment 11, tag 19 (panel **E**) and "Experiment 2, tag 6 (panel **G**)", and "Experiment 1, tag 4", respectively.

cells within the body of both male and female worms (Figure 2B, **Extended data Figure S4A–C**[34], **Movies 2** and **3**[44]). The staining pattern was similar in all observed specimens, and no evident differences in the staining pattern were evident between sexes. Interestingly, the majority of these successfully transfected cells were located around the intestine (Figures 2B–D and **Movie 1**[44]). The relatively higher concentration of the RNP complex surrounding the adult gut may have resulted from worms swallowing Cas9-gRNA molecules in the suspension before the electroporation step was carried out. The fluorescent signal in sporocysts was evenly distributed within the organism (Figures 2E, G, **Extended data Figure S5A, B**[34], and **Movie 4**[44]), whereas within eggs the signal was mainly localised outside the larvae, but inside the eggshell (Figures 2F, H). Importantly, no autofluorescence signal was seen in control parasites (**Extended data Figures S4D, F**, and **S5C, D**[34]).

Evident CRISPR-Cas9-induced deletions in exon 1 of *SULT-OR*

The CRISPR-Cas9 transfection experiments were performed on adult worms (3 biological replicates), mother sporocysts (2 biological replicates) and eggs (3 biological replicates). All replicates were performed by different experimentalists on

different days as indicated in **Extended data Table S1**[34]. DNA was extracted from the parasites four days after transfection, and a 262-bp PCR product spanning the gRNA region was amplified (Figure 1C) and sequenced on an Illumina MiSeq platform. The *SULT-OR*-specific PCR primers were designed in regions that are divergent between *SULT-OR* and the other paralogous genes (**Extended data Figure S3B, C**[34]). With the assistance of the CRISPResso software[37,38] we searched for mutations in the sequence reads by aligning reads to the reference amplicon. Remarkably, the percent of aligned reads that contained deletions was significantly higher for CRISPR-Cas9-treated samples than for matched controls when we pooled all tested developmental stages (Figure 3A and **Extended data Table S1**[34]; paired one-sided Wilcoxon test: n=8 biological replicates, P=0.04). In contrast, the percent of aligned reads with insertions or substitutions was not consistently higher in CRISPR-Cas9-treated samples than matched controls (P=0.2 for insertions and P=0.9 for substitutions; **Extended data Figure S6B, C**[34]). Importantly, no evident differences were observed among the three types of controls employed in the experiments (**Extended data Table S1**). The apparent substitutions seen in both CRISPR-treated and control samples are likely due to sequencing errors, especially at the ends of the
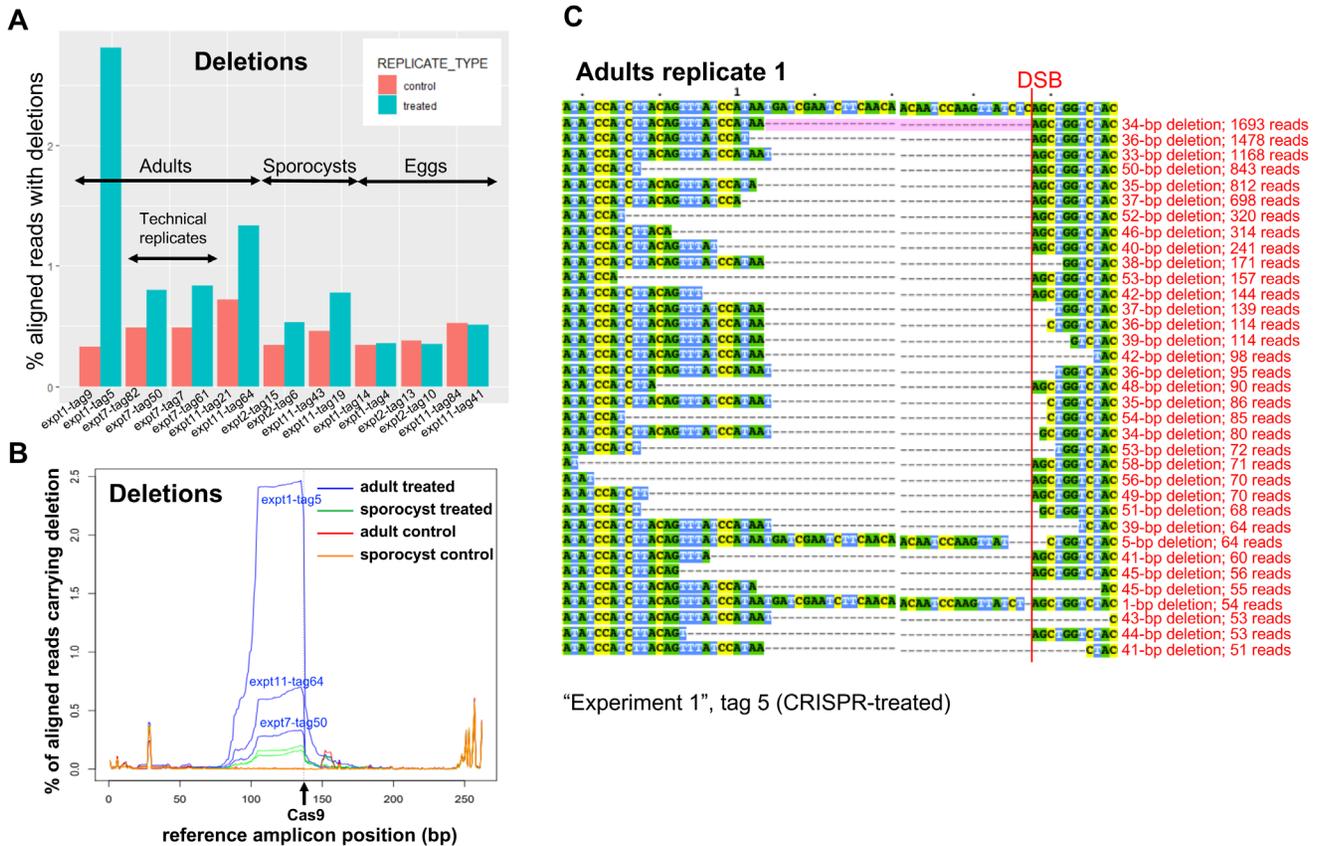
Figure 3. (**A**) Frequency of deletions in NGS sequencing data, identified with the assistance of CRISPResso in three biological replicates from adults, two from sporocysts, and three from eggs, as indicated (sample identifiers at the bottom). The controls include worm-only controls for the adult samples, worm treated with Cas9-only and worm-only controls for the sporocyst samples, eggs treated with Cas9-only and egg-only controls for the egg samples. (**B**) CRISPR-induced deletions in adult worms and sporocysts. The positions of deletions found by CRISPResso in the reference amplicon are indicated, in three biological replicates of CRISPR-Cas9-treated adult samples (blue lines: experiment 1, tag 5; experiment 7, tag 50; experiment 11, tag 64) and matched adult control samples (red lines: experiment 1, tag 9; experiment 7, tag 82; experiment 11, tag 21), and in two biological replicates of CRISPR-Cas9-treated sporocysts (green lines: experiment 2, tag 6; experiment 11, tag 19) and matched sporocyst controls (orange lines: experiment 2, tag 15; experiment 11, tag 43). The black arrow shows the predicted Cas9 cut site. (**C**) Multi-sequence alignment of *SULT-OR* alleles with deletions found in CRISPR-Cas9-treated adult worms that are supported by >=50 reads and span the DSB site indicated with a red line, based on one of the treated adult replicates (experiment 1, tag 5). The common 34-bp deletion is highlighted in pale pink.

amplicon, since the two reads of a read-pair overlap in a 38-bp region in the centre of the amplicon, allowing CRISPResso to infer a higher-quality consensus sequence for that central region (Figure 1C).

Remarkably, a closer examination revealed that all three biological replicates of CRISPR-Cas9-treated adult worms had large deletions absent from control samples, extending from the predicted Cas9 cut site to about 60 bp upstream (Figure 3B, C). Considering the reads that contained a single internal deletion spanning the predicted Cas9 cut site, and no internal insertions, we found that 0.3-2.0% of aligned reads from CRISPR-Cas9-treated adult worms exhibited such deletions, compared to 0.0% of aligned reads from matched controls (**Extended data Table S1**[34]).

Higher CRISPR-induced mutation rate in adults compared to sporocysts and eggs

Interestingly, up to 10 times more reads containing deletions spanning the predicted Cas9 cut site were detected in CRISPR-Cas9-treated adult worms (0.3-2.0% of aligned reads) compared to CRISPR-Cas9-treated sporocysts (0.1-0.2%; **Extended data Table S1**[34]). In contrast, in eggs the rate of such deletions was not any higher than in matched controls (**Extended data Table S1**[34]).

Deletions of the same size, and in the same position, were identified in CRISPR-Cas9-treated sporocysts and adults, being absent from respective matched controls (Figure 3B). In addition, across different biological replicates of CRISPR-Cas9-treated adults and sporocysts, the most frequent deletion

alleles (i.e. those for which we detected the most supporting reads) had roughly the same sizes and positions (Figure 3C and **Extended data Figure S7**[34]). The most common deletion identified in all three adult biological replicates, and in one of the two sporocyst biological replicates, was 34 bp directly upstream of the predicted Cas9 cut site (spanning positions 104-137 in the reference amplicon). We observed rare deletions that were up to three times longer: that is, deletions that extended from the predicted Cas9 cut site to 102 bp upstream (to position 36 in the reference amplicon). Strikingly, none of these deletions were apparent in CRISPR-Cas9-treated eggs.

Almost all the deletions observed extended upstream from the predicted Cas9 cut site; rare deletions extending both upstream and downstream of the cut site were identified but at relatively lower frequency, although often supported by 50 or more reads (**Extended data Figure S7**[34]). In all biological replicates from adults, we did observe extremely low-frequency deletions, supported by few reads (<50 reads, not shown in **Extended data Figure S7**[34]), extending from the predicted Cas9 cut site to 102 bp upstream (position 36 in the reference amplicon), and deletions spanning the cut site that extended as far as 79 bp downstream of the cut site (position 216 in the amplicon).

The percent of aligned reads carrying deletions that spanned the predicted Cas9 cut site did not differ between CRISPR-Cas9-treated eggs and control eggs. This suggested that in eggs either CRISPR-Cas9 did not introduce mutations in *SULT-OR* or they had occurred at an extremely low level. The presence of a recognition site for the restriction enzyme PvuII overlapping the predicted Cas9 cut site (Figure 1B) allowed us to develop a protocol to enrich for mutant alleles. Any CRISPR-Cas9-induced deletions that extended upstream from the Cas9 cut site would remove this PvuII recognition site, so by digesting the DNA from treated parasites with PvuII, we expected to enrich for CRISPR-Cas9-induced deletions. In two out of three biological replicates of CRISPR-Cas9-treated egg samples, after PvuII treatment we were able to detect a slightly higher rate of deletions spanning the predicted Cas9 cut site, compared to in PvuII-treated control egg samples, i.e. an increase of at least 2-fold (**Extended data Table S1**[34]), even though these deletions were still at very low frequency. This finding indicates that CRISPR-Cas was indeed active in eggs, although at very low levels.

Evidence for large deletions
Large CRISPR-induced deletions of >500 bp have been observed in the nematode *Strongyloides stercoralis*[20]. In addition to the most common CRISPR-Cas9-induced deletions observed in *S. mansoni* that extended 34 bp upstream of the predicted Cas9 cut site (Figure 3B), we did observe low-frequency deletions (supported by few reads) extending from the predicted Cas9 cut site to 102 bp upstream (to position 36 in the reference amplicon) (Figure 4 and **Extended data Figure S8**[34]). We simulated reads carrying deletions of every possible length, extending upstream from the predicted Cas9 cut site, that is, a read carrying a deletion of 1-bp upstream of the

Cas9 cut site, a read carrying a deletion of 2-bp upstream of the Cas9 cut site, reads with deletions of 3-bp, 4-bp, 5-bp, and so on. Using our parameter settings, CRISPResso detected the simulated deletions of sizes ranging from 1-bp up to 104-bp upstream of the predicted Cas9 cut site, but could not detect simulated deletions extending 105 bp or further upstream of the Cas9 cut site. This is presumably because, by default, CRISPResso requires that a read aligns with at least 60% identity to the reference amplicon, thus, discarding reads carrying simulated deletions spanning 40% (i.e. 105 bp) or more of our 262-bp reference amplicon. Given that in our real data (Figure 4 and **Extended data Figure S8**[34]), we observed low-frequency deletions extending up to 102 bp upstream of the predicted Cas9 cut site, it is possible that in reality even longer deletions did also occur, but they would not have been detected, either because the reads were discarded by CRISPResso, or one or both PCR primer regions were deleted.

Deletions predicted to cause oxamniquine resistance
Relative to the *SULT-OR* amplicon, the start codon is located at position 103 and the predicted Cas9 cut site at position 137. In adult worms and sporocysts, the CRISPR-Cas9-induced deletions extended upstream from the predicted cut site into the first exon. The most common deletions extended 34 bases upstream, completely removing the first coding exon (barring a single base) and shifting the reading frame of the entire coding region (Figure 1C and **Extended data Figure S7**[34]). This would result in the preferential degradation of the mutant mRNAs, as previously suggested to explain the *ω1* knockdown at the mRNA level observed in CRISPR-Cas9-treated eggs[21]. Moreover, any remaining frame-shifted SULT-OR protein is predicted to have lost its ability to convert OXA to its active drug form. Furthermore, even longer deletions upstream of the predicted Cas9 cut site were observed (**Extended data Figure S7**[34]), which may extend into the *SULT-OR* promoter region given the 5'-UTR spans only 46 bp upstream of the first protein-coding codon, and hence may ablate the transcription of *SULT-OR*. The large CRISPR-Cas9-induced deletions, if they are homozygous, are predicted to cause resistance to OXA, either by producing frameshifts in the *SULT-OR* mRNA or by ablating *SULT-OR* transcription. However, when we analysed the expression levels of *SULT-OR* mRNA by qRTPCR across preparations of whole parasites in CRISPR-Cas9-treated samples versus control adult worms, no differences were evident (not shown).

> **Extended data: Video files for 'Large CRISPR-Cas-induced deletions in the oxamniquine resistance locus of the human parasite Schistosoma mansoni'**
>
> *4 video files*
>
> http://dx.doi.org/10.6084/m9.figshare.12631670.v1

Discussion
Genome editing mediated by CRISPR-Cas has been recently applied to *S. mansoni* to knock out the egg-specific gene *omega-1* (*ω1*)[21]. The CRISPR-Cas9 treatment of eggs by electroporation in the presence of the RNP complex, or egg
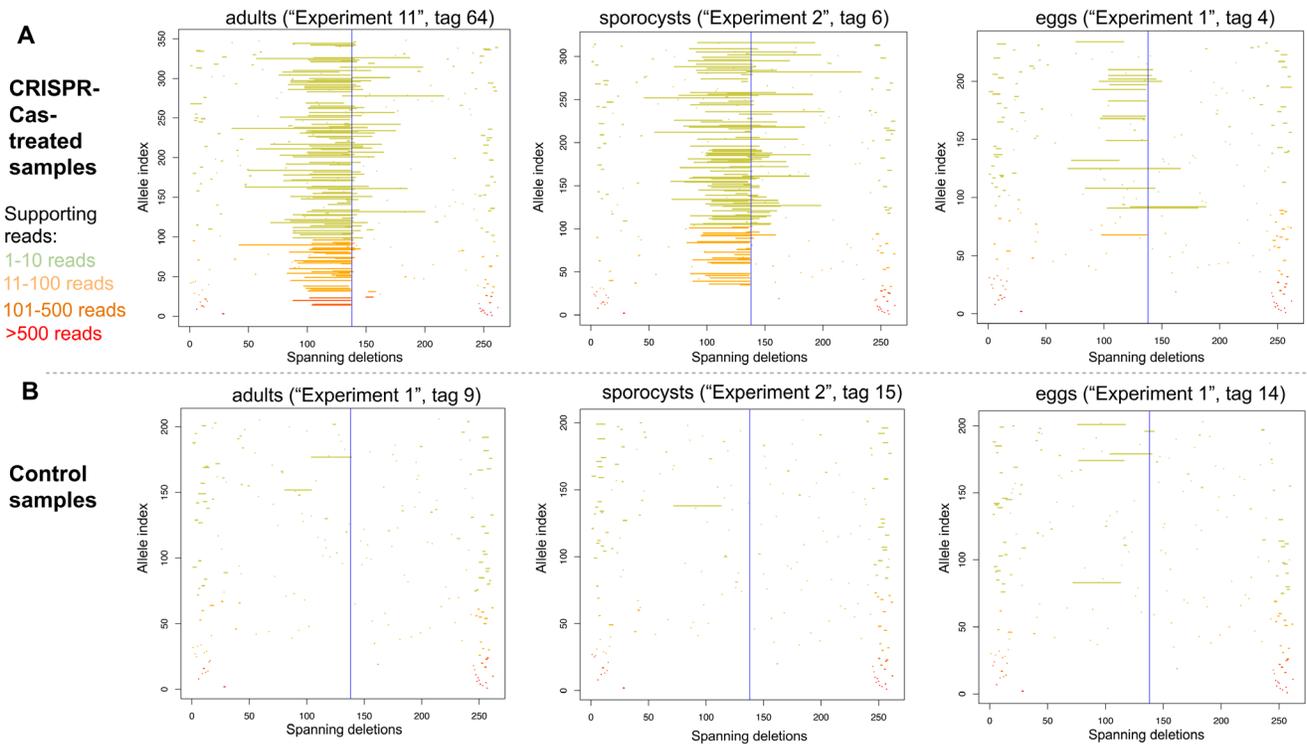
Figure 4. Deletion alleles seen in the *SULT-OR* gene in amplicon sequencing reads from treated (**A**) and control (**B**) adults (left), sporocysts (centre), and eggs (right), showing alleles that contain a single internal deletion and no internal insertions with respect to the reference amplicon. The y-axis shows deletion alleles sorted by the number of reads supporting them, with the alleles supported by the most reads at the bottom. Alleles supported by >500 reads in red, alleles supported by 101-500 reads in dark orange, alleles supported by 11-100 reads in pale orange, and alleles supported by 1-10 reads in pale green. The x-axis shows the position of the deletion along the reference amplicon, with a blue vertical line at the predicted Cas9 cut site.

transduction with lentivirus particles expressing Cas9 and the gRNA, induced a detectable knock-down both at the mRNA and protein levels and a clear phenotype of smaller granulomas in mice exposed to CRISPR-Cas9-treated eggs. In the current study, we decided to employ CRISPR-Cas9 to target the *SULT-OR* sulfotransferase gene in *S. mansoni* with an RNP complex. Strikingly, we detected large deletions of ≥34 bp extending upstream of the predicted Cas9 cut site, whereas deletions extending downstream of the cut site were extremely rare. The tendency for deletions to be upstream of the predicted Cas9 cut site agrees with observations in mouse cell lines[45]. We identified deletions extending up to 102 bp upstream of the predicted Cas9 cut site, reaching the limit detectable with CRISPResso (104 bp upstream, using our own parameter settings), suggesting that even larger deletions may have been missed. Deletions of several hundred base pairs have been described in *Strongyloides*[20], *C. elegans*[46], and mammalian cell-lines[45,47]. In addition, we characterised CRISPR-Cas9-induced mutations across three discrete developmental stages: adult worms, eggs, and *in vitro*-transformed sporocysts. The deletions spanning the predicted Cas9 cut site were most commonly detected in adult worms (0.3-2.0% of aligned reads), followed by sporocysts (0.1-0.2%), and extremely rare in eggs. Interestingly, no evidence for CRISPR-Cas9-induced

insertions or substitutions in the *SULT-OR* gene was observed. In Strongyloides spp, Gang *et al.*[20] showed that in the absence of a repair template, small insertions or deletions (indels) or substitutions were not observed, but instead the authors found deletions of >500 bp at the target site in the unc-22 gene. On the other hand, in *C. elegans*, one study in the absence of a repair template, detected deletions ranging from 7 bp to >2 kb in the dpy-11 and unc-4 genes[46]; while another study found only small insertions and deletions of <20 bp[48]. In mammalian cell lines, the majority of indels are relatively small (1-50 bp), but larger deletions of kilobases in size are also sometimes observed[49]. An interesting question is whether the pattern of CRISPR-Cas-induced mutations varies from gene-to-gene. When targeting the *ω1* gene in the absence of donor molecule, where CRISPR-Cas9-induced mutations presumably occurred by non-homologous end joining (NHEJ)[19], the overall rate of deletions detected in reads from CRISPR-Cas9-treated eggs was not higher than in control eggs[21]. This is consistent with an extremely low rate of NHEJ-mediated deletions in the *ω1* gene in eggs, similar to what we described here for the genome editing of the *SULT-OR* gene in eggs. On the other hand, when the *ω1* gene was targeted using CRISPR-Cas9 in the presence of a ssODN donor molecule, rare larger deletions spanning the predicted Cas9 cut site were detected in

the CRISPR-Cas9-treated eggs compared to controls[21]. Since this effect was only observed in the presence of the donor molecule, it is tempting to speculate that those large deletions in the *ωl* gene may have been related to the homology directed repair (HDR) mechanism involved in the knock-in of the donor molecule rather than driven by NHEJ[50]. The low CRISPR-Cas efficiency in our study may be explained, at least partially, by low efficiency of RNP complex assembly; however, as discussed above, a low CRISPR-Cas efficiency was also described when lentivirus (rather than RNP complex) was used to deliver the CRISPR-Cas9 cargo into the parasite[21]. In addition, differences between protocols employed for *ωl* and *SULT-OR* may have resulted in different spectra of CRISPR-Cas9-induced mutations: the CRISPR-Cas9-treated eggs sequenced in the *ωl* study were transduced with lentivirus encoding a single-gRNA and Cas9 nuclease[21], whereas we induced mutations in *SULT-OR* by electroporating the parasites with an *in vitro*-assembled RNP complex of Cas9 nuclease and a two-piece gRNA. Interestingly, in the liver fluke *Opisthorchis viverrini* electroporated in the presence of a plasmid encoding a single-gRNA and Cas9 nuclease, small deletions of up to ~10 bp and insertions of up to 2 bp were detected near the predicted Cas9 cut site in the granulin gene[22].

Our data suggest that for *SULT-OR* the highest CRISPR efficiency was in adults, followed by sporocysts, and the lowest efficiency in eggs. For the latter we only found such deletions in <0.02% of aligned reads even after using PvuII to enrich for mutant reads. Three possible non-mutually exclusive hypotheses may explain these differences. Firstly, electroporation of the RNP complex may be most efficient in adults and least efficient in eggs, possibly because the surface area:volume ratio of an adult worm is greater than that of a sporocyst or egg, and/or because the egg has a protective coating that makes it hard to penetrate[51]. However, microporous and internal microcanals shown to be scattered across *Schistosoma* eggshells would allow the interchange of macromolecules with the host tissues[52]. Relevant for us, the diameter of the smallest pores in *S. mansoni* eggs is 100 nm, and we have estimated the diameter of the RNP complex, assuming a globular shape, to be ~10 nm[53] indicating that the complex could have entered the egg through the pores. A second possibility is that some key NHEJ repair enzymes required for CRISPR have higher expression in adults than in sporocysts or eggs; according to RNAseq metadata[25] this is the case for *Smp_211060*, previously identified[21] as a homolog of the Ku70/Ku80 genes that play a key role in NHEJ[54]. Finally, CRISPR might be more efficient in inducing mutations in *SULT-OR* in adults than sporocysts or eggs, because *SULT-OR* is expressed more in the former developmental stage, making its chromatin more open and therefore more accessible to the CRISPR machinery[19,55–57].

To create a CRISPR-Cas9-mediated mutant of any schistosome gene in every cell of the animal, the germline cells need to be targeted and mutated by a germline transgenesis approach. So far, only two studies demonstrated germline transmission of exogenous DNA in schistosomes. An early study published in 2007[58] showed that a GFP-expressing plasmid was introduced into the miracidium germ cells by particle bombardment. Subsequently, the transfected miracidia infected snails,

and resulting cercariae were employed to infect hamsters and obtain F1 transformed eggs. However, over a few generations the transformed parasites died and/or (as expected) the plasmid was diluted or lost. A few years later, germline transmission of integrated retroviral transgenes was demonstrated. Murine Leukemia Retrovirus (MLV) transgenes transduced the germ cells of eggs and were propagated through both the intramolluscan asexual developmental stages and intramammalian sexual developmental stages, reaching the F1 eggs[16,17]. However, no germline transgenesis approach has yet been achieved using CRISPR-Cas in schistosomes. Ittiprasert *et al.* in 2019[21] applied the CRISPR technology using a lentivirus expressing Cas9 and the gRNA, plus a donor, to introduce a 24-bp insertion (by HDR) into the *ωl* gene in *S. mansoni* eggs. Intriguingly, although the expression of the *ωl* gene was reduced by 81-83% after CRISPR-Cas9 treatment using the donor molecule, only ~4.5% of reads were identified by amplicon sequencing as mutated by indels (with <1% showing deletions) or substitutions, and only 0.19% of reads contained the 24-bp insertion. Proposed explanations for this discrepancy include a preferential penetration of CRISPR-Cas9 machinery (lentivirus and donor) in the envelope of the egg, where *ωl* may be expressed[51], and/or the presence of large deletions that removed either one or both primer regions and so were not detected by amplicon sequencing (as seen in *Strongyloides*[20]). Similarly, expression of the *Opisthorchis viverrini* granulin gene was reduced by >80% after CRISPR-Cas9 treatment of pooled adults in the absence of a donor, but only 1.3% of amplicon sequencing reads contained indel mutations[22]. This apparent anomaly may be due to the predominant expression of the granulin gene in the *O. viverrini* tegument and gut, where electroporation of the gene editing plasmid may have been most efficient[22]. Furthermore, there may be variation in CRISPR-Cas9 efficiency between individual adults: taking adult *Opisthorchis* worms from hamsters that had been infected 60 days previously with CRISPR-treated newly encysted juveniles (NEJs), individual adults in which there was a greater knock-down at the mRNA level showed a far greater level of mutations upon amplicon sequencing, especially deletions and substitutions[22]. In this species, significant variation was seen in CRISPR-Cas9 efficiency between life stages. Using a plasmid encoding Cas9 and gRNA, a knock-down of >80% of granulin mRNA levels was achieved in adults and NEJs, but of <4% in metacercariae, possibly due to inefficient electroporation of the plasmid through the metacercarial cyst wall[22].

In our study, while amplicon sequencing revealed reads carrying CRISPR-Cas9-induced mutations in *SULT-OR*, a knock-down of *SULT-OR* at the mRNA level was not evident, probably because the deletions occurred in only a small fraction of the adult cells that express *SULT-OR*. Single-cell sequencing data from adults shows *SULT-OR* is a marker of parenchymal cells[10], while our confocal microscopy data suggest the Cas9-gRNA complex penetrated better into the adult tegument and intestine compared to parenchymal tissue. The same phenomenon was previously described in the liver fluke *Fasciola hepatica* when delivering fluorescently labelled molecules by electroporation[59,60], suggesting the flatworm intestine as the main point of entry when this delivery approach is employed.

Furthermore, *SULT-OR* is also expressed at a low level in many other cell types in adults (**Extended data Figure S2B**[34]). Future experiments comparing the delivery of the RNP complex by soaking versus electroporation would inform about the best delivery approach. The absence of nuclear localisation of the RNP complex in the confocal images may indicate that the 4-hour timeframe between the molecule delivery and parasite fixation was too short, and which may explain, at least partially, the low CRISPR-induced deletion efficiency. The large deletions spanning the predicted Cas9 cut site were found in 0.3–2.0% of aligned reads from CRISPR-treated adult worms, so our best estimate of the fraction of adult cells in which CRISPR worked is 0.3–2.0% assuming an even distribution of mutations across the transfected parasites. However, since a pool of five adult worms were transfected with the RNP complex, the efficiency of CRISPR (and the amount of knock-down at the mRNA level) may have varied between worms, as well as between cells of an individual worm.

To conclude, more work is required to optimise CRISPR-Cas protocols to work best at different developmental stages and in particular tissues, and understand whether these differing protocols will result in different spectra of mutations[46] and degrees of mRNA knock-down. To do this, it may be critical to identify the mechanisms underlying CRISPR-Cas-induced mutations in schistosomes in each case (e.g. NHEJ, HDR or other mechanisms such as polymerase theta-mediated end-joining[61]). Addressing these items would help the research community to achieve the holy grail of targeting the germ line and creating a stable knock-out or knock-in strain of any gene of interest.

Data availability

Underlying data

Large CRISPR-Cas induced deletions in the oxamniquine resistance locus of the human parasite *Schistosoma mansoni*, Accession number: ERP 121238 https://identifiers.org/ena.embl:ERP121238

Open Science Framework: Large CRISPR-Cas-induced deletions in the oxamniquine resistance locus of the human parasite Schistosoma mansoni. https://doi.org/10.17605/OSF.IO/B96SR[34]

This project contains the following underlying data:

- qPCR_rawCts_exp2_Sankaranarayanan, Coghlan_etal_WOR.xls (raw Cts values)

- qPCR_rawCts_exp7_Sankaranarayanan, Coghlan_etal_WOR.xls (raw Cts values)

- Fig2A_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure 2, panel A)

- Fig2B-D_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure 2, panels B-D)

- Fig2E_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure 2, panel E)

- Fig2F_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure 2, panel F)

- Fig2G_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure 2, panel G)

- Fig2H_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure 2, panel H)

- FigS4A-C_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure S4, panel A-C)

- FigS4D_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure S4, panel D)

- FigS4E_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure S4, panel E)

- FigS4F_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure S4, panel F)

- FigS4G_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure S4, panel G)

- FigS5A_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure S5, panel A)

- FigS5B_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure S5, panel B)

- FigS5C_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure S5, panel C)

- FigS5D_Sankaranarayanan, Coghlan_etal_WOR.tif (original picture Figure S5, panel D)

Extended data

Open Science Framework: Large CRISPR-Cas-induced deletions in the oxamniquine resistance locus of the human parasite Schistosoma mansoni. https://doi.org/10.17605/OSF.IO/Z45BG[34]

This project contains the following extended data:

- Sankaranarayanan, Coghlan_etal_WOR_extended_data_V2_07Dec2020 (Word document containing legends for extended data)

- TableS1_Sankaranarayanan, Coghlan_etal_WOR.xlsx (Extended Data Table S1)

- TableS2_Sankaranarayanan, Coghlan_etal_WOR.xlsx (Extended Data Table S2)

- TableS3_Sankaranarayanan, Coghlan_etal_WOR.xlsx (Extended Data Table S3)

- FigS1_Sankaranarayanan, Coghlan_etal_WOR.pdf (Extended Data Figure S1)

- FigS2_Sankaranarayanan, Coghlan_etal_WOR.pdf (Extended Data Figure S2)

- FigS3_Sankaranarayanan,   Coghlan_etal_WOR.pdf
  (Extended Data Figure S3)

- FigS4_Sankaranarayanan,   Coghlan_etal_WOR.pdf
  (Extended Data Figure S4)

- FigS5_Sankaranarayanan,   Coghlan_etal_WOR.pdf
  (Extended Data Figure S5)

- FigS6_Sankaranarayanan,   Coghlan_etal_WOR.pdf
  (Extended Data Figure S6)

- FigS7_Sankaranarayanan,   Coghlan_etal_WOR.pdf
  (Extended Data Figure S7)

- FigS8_Sankaranarayanan,   Coghlan_etal_WOR.pdf
  (Extended Data Figure S8)

Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

Figshare: Video files for 'Large CRISPR-Cas-induced deletions in the oxamniquine resistance locus of the human parasite *Schistosoma mansoni*. https://doi.org/10.6084/m9.figshare.12631 670.v1[44]

This project contains the following video files:

- 16031-V1-2-MovieS1_20Apr2020.mov

Movie 1. Serial optical sections of a *S. mansoni* male adult worm transfected with fluorescently labelled Cas9-gRNA (ATTO™ 550 signal in red), fixed and DAPI-stained (DAPI signal in aqua blue). Scale bar: 100 µm.

- 16031-V1-2-MovieS2_20Apr2020.mov

Movie 2. Serial optical sections of a *S. mansoni* male adult worm transfected with fluorescently labelled Cas9-gRNA (ATTO™ 550 signal in red), fixed and DAPI-stained (DAPI

signal in aqua blue). In these series of optical sections, the anterior end of the worm is observed. Scale bar: 100 µm.

- 16031-V1-2-MovieS3_20Apr2020.mov

Movie 3. Serial optical sections of a *S. mansoni* female adult worm transfected with fluorescently labelled Cas9-gRNA (ATTO™ 550 signal in red), fixed and DAPI-stained (DAPI signal in aqua blue). In these series of optical sections, the anterior end of the worm is observed. Scale bar: 100 µm.

- 16031-V1-2-MovieS4_20Apr2020.mov

Movie 4. Serial optical sections of a *S. mansoni* sporocyst transfected with fluorescently labelled Cas9-gRNA (ATTO™ 550 signal in red), fixed and DAPI-stained (DAPI signal in aqua blue). Scale bar: 25 µm.

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Author information

Patrick Driguez is currently affiliated with King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia.

References

1. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators: **Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015.** *Lancet.* 2016; **388**(10053): 1545–1602.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

2. Gryseels B, Polman K, Clerinx J, *et al.*: **Human schistosomiasis.** *Lancet.* 2006; **368**(9541): 1106–1118.
   PubMed Abstract | **Publisher Full Text**

3. Crellen T, Walker M, Lamberton PHL, *et al.*: **Reduced Efficacy of Praziquantel Against *Schistosoma mansoni* Is Associated With Multiple Rounds of Mass Drug Administration.** *Clin Infect Dis.* 2016; **63**(9): 1151–1159.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

4. Swain MT, Larkin DM, Caffrey CR, *et al.*: ***Schistosoma* comparative genomics:**

integrating genome structure, parasite biology and anthelmintic discovery.** *Trends Parasitol.* 2011; **27**(12): 555–64.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. Protasio AV, Tsai IJ, Babbage A, *et al.*: **A systematically improved high quality genome and transcriptome of the human blood fluke *Schistosoma mansoni*.** *PLoS Negl Trop Dis.* 2012; **6**(1): e1455.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

6. Stroehlein AJ, Korhonen PK, Chong TM, *et al.*: **High-quality *Schistosoma haematobium* genome achieved by single-molecule and long-range sequencing.** *Gigascience.* 2019; **8**(9): giz108.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

7. Luo F, Yin M, Mo X, *et al.*: **An improved genome assembly of the fluke *Schistosoma japonicum*.** *PLoS Negl Trop Dis.* 2019; **13**(8): e0007612.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

8. Wangwiwatsin A, Protasio AV, Wilson S, *et al.*: **Transcriptome of the parasitic flatworm *Schistosoma mansoni* during intra-mammalian development.** *bioRxiv.*
Publisher Full Text

9. Diaz Soria CL, Lee J, Chong T, *et al.*: **Single-cell atlas of the first intra-mammalian developmental stage of the human parasite *Schistosoma mansoni*.** *bioRxiv.* 2019; 754713.
Publisher Full Text

10. Wendt G, Zhao L, Chen R, *et al.*: **A single-cell RNAseq atlas of the pathogenic stage of *Schistosoma mansoni* identifies a key regulator of blood feeding**. Microbiology. *bioRxiv.* 2020; 69.
Publisher Full Text

11. Wang J, Paz C, Padalino G, *et al.*: **Large-scale RNAi screening uncovers new therapeutic targets in the human parasite *Schistosoma mansoni*.** *bioRxiv.* 2020; 2020.02.05.935833.
Publisher Full Text

12. Boutros M, Ahringer J: **The art and design of genetic screens: RNA interference.** *Nat Rev Genet.* 2008; **9**(7): 554–566.
PubMed Abstract | **Publisher Full Text**

13. Farboud B, Severson AF, Meyer BJ: **Strategies for Efficient Genome Editing Using CRISPR-Cas9.** *Genetics.* 2019; **211**(2): 431–457.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

14. Dickinson DJ, Goldstein B: **CRISPR-Based Methods for *Caenorhabditis elegans* Genome Engineering.** *Genetics.* 2016; **202**(3): 885–901.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

15. Suttiprapa S, Rinaldi G, Tsai IJ, *et al.*: **HIV-1 Integrates Widely throughout the Genome of the Human Blood Fluke *Schistosoma mansoni*.** *PLoS Pathog.* 2016; **12**(10): e1005931.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

16. Mann VH, Suttiprapa S, Skinner DE, *et al.*: **Pseudotyped murine leukemia virus for schistosome transgenesis: approaches, methods and perspectives.** *Transgenic Res.* 2014; **23**(3): 539–556.
PubMed Abstract | **Publisher Full Text**

17. Rinaldi G, Eckert SE, Tsai IJ, *et al.*: **Germline transgenesis and insertional mutagenesis in *Schistosoma mansoni* mediated by murine leukemia virus.** *PLoS Pathog.* 2012; **8**(7): e1002820.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Hagen J, Young ND, Every AL, *et al.*: **Omega-1 knockdown in *Schistosoma mansoni* eggs by lentivirus transduction reduces granuloma size *in vivo*.** *Nat Commun.* 2014; **5**: 5375.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

19. Jiang F, Doudna JA: **CRISPR-Cas9 Structures and Mechanisms.** *Annu Rev Biophys.* 2017; **46**: 505–529.
PubMed Abstract | **Publisher Full Text**

20. Gang SS, Castelletto ML, Bryant AS, *et al.*: **Targeted mutagenesis in a human-parasitic nematode.** *PLoS Pathog.* 2017; **13**(10): e1006675.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

21. Ittiprasert W, Mann VH, Karinshak SE, *et al.*: **Programmed genome editing of the omega-1 ribonuclease of the blood fluke, *Schistosoma mansoni*.** *eLife.* 2019; **8**: e41337.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

22. Arunsan P, Ittiprasert W, Smout MJ, *et al.*: **Programmed knockout mutation of liver fluke granulin attenuates virulence of infection-induced hepatobiliary morbidity.** *eLife.* 2019; **8**: e41463.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

23. Valentim CLL, Cioli D, Chevalier FD, *et al.*: **Genetic and molecular basis of drug resistance and species-specific drug action in schistosome parasites.** *Science.* 2013; **342**(6164): 1385–1389.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

24. Pica-Mattoccia L, Dias LC, Moroni R, *et al.*: ***Schistosoma mansoni*: genetic complementation analysis shows that two independent hycanthone/oxamniquine-resistant strains are mutated in the same gene.** *Exp Parasitol.* 1993; **77**(4): 445–449.
PubMed Abstract | **Publisher Full Text**

25. Lu Z, Zhang Y, Berriman M: **A web portal for gene expression across all life stages of *Schistosoma mansoni*.** *bioRxiv.* 2018; 308213.
Publisher Full Text

26. Tucker MS, Karunaratne LB, Lewis FA, *et al.*: **Schistosomiasis.** *Curr Protoc Immunol.* 2013; **103**(3): 19 11 11–19 11 58.
PubMed Abstract | **Publisher Full Text**

27. Mann VH, Morales ME, Rinaldi G, *et al.*: **Culture for genetic manipulation of developmental stages of *Schistosoma mansoni*.** *Parasitology.* 2010; **137**(3): 451–462.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

28. Dalton JP, Day SR, Drew AC, *et al.*: **A method for the isolation of schistosome eggs and miracidia free of contaminating host tissues.** *Parasitology.* 1997; **115**(Pt 1): 29–32.
PubMed Abstract | **Publisher Full Text**

29. Rinaldi G, Morales ME, Alrefaei YN, *et al.*: **RNA interference targeting leucine aminopeptidase blocks hatching of *Schistosoma mansoni* eggs.** *Mol Biochem Parasitol.* 2009; **167**(2): 118–126.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

30. Bruntraeger M, Byrne M, Long K, *et al.*: **Editing the Genome of Human Induced Pluripotent Stem Cells Using CRISPR/Cas9 Ribonucleoprotein Complexes.** *Methods Mol Biol.* 2019; **1961**: 153–183.
PubMed Abstract | **Publisher Full Text**

31. Bolt BJ, Rodgers FH, Shafie M, *et al.*: **Using WormBase ParaSite: An Integrated Platform for Exploring Helminth Genomic Data.** *Methods Mol Biol.* 2018; **1757**: 471–491.
PubMed Abstract | **Publisher Full Text**

32. Kines KJ, Rinaldi G, Okatcha TI, *et al.*: **Electroporation facilitates introduction of reporter transgenes and virions into schistosome eggs.** *PLoS Negl Trop Dis.* 2010; **4**(2): e593.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

33. Rinaldi G, Okatcha TI, Popratiloff A, *et al.*: **Genetic manipulation of *Schistosoma haematobium*, the neglected schistosome.** *PLoS Negl Trop Dis.* 2011; **5**(10): e1348.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

34. Rinaldi G: **Large CRISPR-Cas-induced deletions in the oxamniquine resistance locus of the human parasite *Schistosoma mansoni*.** 2020.
http://www.doi.org/10.17605/OSF.IO/Z45BG

35. Bolger AM, Lohse M, Usadel B: **Trimmomatic: A Flexible Trimmer for Illumina Sequence Data.** *Bioinformatics.* 2014; **30**(15): 2114–2120.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

36. Ewing B, Green P: **Base-calling of automated sequencer traces using *phred*. II. Error probabilities.** *Genome Res.* 1998; **8**(3): 186–94.
PubMed Abstract | **Publisher Full Text**

37. Canver MC, Haeussler M, Bauer DE, *et al.*: **Integrated Design, Execution, and Analysis of Arrayed and Pooled CRISPR Genome-Editing Experiments.** *Nat Protoc.* 2018; **13**(5): 946–986.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

38. Pinello L, Canver MC, Hoban MD, *et al.*: **Analyzing CRISPR Genome-Editing Experiments With CRISPResso.** *Nat Biotechnol.* 2016; **34**(7): 695–697.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

39. Ginzinger DG: **Gene Quantification Using Real-Time Quantitative PCR: An Emerging Technology Hits the Mainstream.** *Exp Hematol.* 2002; **30**(6): 503–512.
PubMed Abstract | **Publisher Full Text**

40. Livak KJ, Schmittgen TD: **Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2(-Delta Delta C^T) Method.** *Methods.* 2001; **25**(4): 402–408.
PubMed Abstract | **Publisher Full Text**

41. International Helminth Genomes Consortium: **Comparative genomics of the major parasitic worms.** *Nat Genet.* 2019; **51**(1): 163–174.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

42. Fincher CT, Wurtzel O, de Hoog T, *et al.*: **Cell type transcriptome atlas for the planarian *Schmidtea mediterranea*.** *Science.* 2018; **360**(6391): eaaq1736.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

43. Zheng T, Hou Y, Zhang P, *et al.*: **Profiling Single-Guide RNA Specificity Reveals a Mismatch Sensitive Core Sequence.** *Sci Rep.* 2017; **7**: 40638.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

44. Sankaranarayanan G, Coghlan A, Driguez P, *et al.*: **Video files for 'Large CRISPR-Cas-induced deletions in the oxamniquine resistance locus of the human parasite *Schistosoma mansoni* '.** *Wellcome Open Research Media.* 2020.
http://www.doi.org/10.6084/m9.figshare.12631670.v1

45. Shin HY, Wang C, Lee HK, *et al.*: **CRISPR/Cas9 Targeting Events Cause Complex Deletions and Insertions at 17 Sites in the Mouse Genome.** *Nat Commun.* 2017; **8**: 15464.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

46. Chiu H, Schwartz HT, Antoshechkin I, *et al.*: **Transgene-free genome editing in *Caenorhabditis elegans* using CRISPR-Cas.** *Genetics.* 2013; **195**(3): 1167–1171.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

47. Kosicki M, Tomberg K, Bradley A: **Repair of Double-Strand Breaks Induced by CRISPR-Cas9 Leads to Large Deletions and Complex Rearrangements.** *Nat Biotechnol.* 2018; **36**(8): 765–771.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

48. Friedland AE, Tzur YB, Esvelt KM, *et al.*: **Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system.** *Nat Methods.* 2013; **10**(8): 741–3.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

49. Kosicki M, Allen F, Bradley A: **Cas9-induced large deletions and small indels are controlled in a convergent fashion.** *bioRxiv.* 2020.
Publisher Full Text

50. Vu GTH, Cao HX, Fauser F, *et al.*: **Endogenous Sequence Patterns Predispose the Repair Modes of CRISPR/Cas9-induced DNA Double-Stranded Breaks in *Arabidopsis Thaliana*.** *Plant J.* 2017; **92**(1): 57–67.
PubMed Abstract | **Publisher Full Text**

51. Ashton PD, Harrop R, Shah B, *et al.*: **The Schistosome Egg: Development and Secretions.** *Parasitology.* 2001; **122**(Pt 3): 329–338.
PubMed Abstract

52. Cao HM, Wang YF, Long S: **A Study of Ultrastructure of Egg Shell of *Schistosoma Japonicum*. I. Transmission Electron Microscopic Observation of *S. Japonicum* Egg.** *Ann Parasitol Hum Comp.* 1982; **57**(4): 345–352.
PubMed Abstract | **Publisher Full Text**

53. Erickson HP: **Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy.** *Biol Proced Online.* 2009; **11**: 32–51.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

54. Kragelund BB, Weterings E, Hartmann-Petersen R, *et al.*: **The Ku70/80 Ring in Non-Homologous End-Joining: Easy to Slip On, Hard to Remove.** *Front Biosci (Landmark Ed).* 2016; **21**: 514–527.
PubMed Abstract | **Publisher Full Text**

55. Verkuijl SA, Rots MG: **The Influence of Eukaryotic Chromatin State on CRISPR-Cas9 Editing Efficiencies.** *Curr Opin Biotechnol.* 2019; **55**: 68–73.
PubMed Abstract | **Publisher Full Text**

56. Uusi-Mäkelä MIE, Barker HR, Bäuerlein CA, *et al.*: **Chromatin Accessibility Is Associated With CRISPR-Cas9 Efficiency in the Zebrafish (*Danio rerio*).** *PLoS One.* 2018; **13**(4): e0196238.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

57. Jensen KT, Fløe L, Petersen TS, *et al.*: **Chromatin Accessibility and Guide Sequence Secondary Structure Affect CRISPR-Cas9 Gene Editing Efficiency.** *FEBS Lett.* 2017; **591**(13): 1892–1901.
PubMed Abstract | **Publisher Full Text**

58. Beckmann S, Wippersteg V, El-Bahay A, *et al.*: ***Schistosoma mansoni*: germ-line transformation approaches and actin-promoter analysis.** *Exp Parasitol.* 2007; **117**(3): 292–303.
PubMed Abstract | **Publisher Full Text**

59. Rinaldi G, Morales ME, Cancela M, *et al.*: **Development of Functional Genomic Tools in Trematodes: RNA Interference and Luciferase Reporter Gene Activity in *Fasciola hepatica*.** *PLoS Negl Trop Dis.* 2008; **2**(7): e260.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

60. Dell'Oca N, Basika T, Corvo I, *et al.*: **RNA interference in *Fasciola hepatica* newly excysted juveniles: Long dsRNA induces more persistent silencing than siRNA.** *Mol Biochem Parasitol.* 2014; **197**(1–2): 28–35.
PubMed Abstract | **Publisher Full Text**

61. van Schendel R, Roerink SF, Portegijs V, *et al.*: **Polymerase Θ is a key driver of genome evolution and of CRISPR/Cas9-mediated mutagenesis.** *Nat Commun.* 2015; **6**: 7394.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Open Peer Review

Current Peer Review Status: ✓ ✓ ✓ ✓

Version 1

Reviewer Report 01 September 2020

https://doi.org/10.21956/wellcomeopenres.17586.r39712

✓   **Patrick Skelly** (iD)

Molecular Helminthology Laboratory, Department of Infectious Disease and Global Health, Cummings School of Veterinary Medicine, Tufts University, North Grafton, MA, USA

This study by Sankaranarayanan & Coghlan *et al.* employed CRISPR-Cas9 to target for disruption the SULT-OR sulfotransferase gene in *Schistosoma mansoni* using electroporation to deliver a ribonucleoprotein (RNP) complex consisting of gRNA, tracrRNA and Cas9 protein. Mutations induced by this treatment were subsequently detected by high-throughput sequencing. The results provide independent confirmation that CRISPR/Cas delivery to *S. mansoni* can indeed induce mutations in these parasites, albeit (so far) at a low level. Three developmental stages were examined, and mutations were most commonly detected in adult worms (0.3-2.0% of aligned sequence reads), followed by sporocysts (0.1-0.2%), and were extremely rare in eggs. The work provides hope that ongoing research will permit researchers to use refined methods to target the germ line to create stable knock-out or knock-in strains of any schistosome gene of interest.

Some comments and questions:

Animal Procedures. Why were snails "moved into dark cupboards at 28°C when they start shedding cercariae" and is there a reference showing that this is helpful? Page 4: How was the water "conditioned"? (At least in my version) the "Welfare assessments...." sentence needs correcting.

Parasite Material. Useful, detailed methods are reported. So, for consistency, give specifics of the Percoll-sucrose solution.

CRISPR-Cas9 ribonucleoprotein complex assembly. Since we expect the gRNA and the nuclease to associate in a 1:1 ratio, why not mix these in equimolar amounts? How can the authors confirm that the gRNA and nuclease actually formed the RNP complex? Is there a way to determine how efficiently this occurred? Low efficiency might provide some explanation for the relatively low level of mutation detected here?

Bioinformatic analysis. Consider adding a reference to Phred Quality Scoring so that interested readers can make sense of the <23 score cut off. It may also be useful to define here "SNP" and "NHEJ".

Results. CRISPR-Cas9 machinery.....The authors should report if ALL of the adults (males and females) yielded a similar staining pattern. In Figure 2A and Movie 1, surface staining is not uniform along the length of the worm; did all males exhibit such staining in a similar region and to a similar extent (and in all replicates)? Does any of the RNP staining co-localize with DAPI-stained nuclei? If yes, point examples out. If not, how well is the nuclear localization signal working?

Discussion. It is reported that "no evidence for CRISPR-Cas9-induced insertions or substitutions in the SULT-OR gene was observed." For context, it would be useful to know how this equates (if possible, under comparable experimental conditions) with work in *Strongyloides*, *C. elegans*, and mammalian cell-lines.

I disagree that because mutations were detected in 2% of aligned reads, this necessarily means that ~2% of cells were impacted.

The control data shown in e.g. Figure 3A – were they derived from the mock treated worms or from worms treated with Cas9 only or from those treated with gRNA only? Or were data from all controls combined?

Report how the tree shown in Figure S1A was generated. What does "PRJEA36577" etc. refer to in the figure? What is the value of showing anything outside of the area bounded by the red dashed line?

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular parasitology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 07 Dec 2020
Gabriel Rinaldi, Wellcome Sanger Institute, Hinxton, UK

07 December 2020,
Wellcome Open Research,

On behalf of all the authors, we acknowledge and appreciate the efforts of the reviewers with the review of our manuscript. We have been able to revise along the lines of the recommendations, which has certainly contributed to improve our paper. Point-by-point responses are provided below: **Response (R)**

30 Jul 2020 | for Version 1
Phillip LoVerde, Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

Schistosomiasis remains a serious global health problem. The toolbox available to scientists continues to improve. The ability to manipulate genes is a key technique to the continued success of the field to address questions of biology, host-parasite interactions, vaccines, drugs and diagnostics among others. The manuscript by the Wellcome Sanger Group is a welcome addition as it provides very detailed methods, state of the art approach and detailed analysis of the results. In fact, it provides a roadmap for others to follow. This is in spite of the fact they were not successful in producing a drug resistant transgenic strain. Clearly part of the reason was that only somatic cells were affected by the CRISPR-Cas 9 construct. They recognize that affecting germ line cells will likely make a difference. Another plausible explanation is that the cells in which CRISPR-Cas9 did induce mutations represented a small fraction of the cells expressing SmSULT-OR. They also acknowledge that the resistance trait is double recessive inferring that both alleles must be interrupted. I may have missed it but did not see where they attempted to demonstrate that both alleles of SmSULT-OR were interrupted.

R. We acknowledge the reviewer for the positive comments, and as indicated, the resistance trait is double recessive inferring that both alleles must be interrupted in the germline to generate a transgenic line with a fully OXA-resistant phenotype. In this study, we did not demonstrate that both alleles of SmSULT-OR were interrupted, given that the sequence data were generated from a large number of cells, probably including a mixture of wild type and mutant cells. In order to identify allele-specific mutations single cell DNA PCR and/ or sequencing approaches need to be applied (PMID: 31827197; PMID: 30992375). This can be optimised in the future, in particular after having shown that single cell sequencing approaches are feasible in schistosomes (PMID: 32973030; bioRxiv 754713; doi: https://doi.org/10.1101/754713)

The manuscript is well-written and follows a logical progression with well-designed experiments. Suggest use the term intramolluscan instead of intrasnail.

R. Edited as recommended

An issue out of their control is the difficulty in navigating the Extended data Figures and Tables. However, all the data is available to the reader.

R. We agree and have also found difficulties in navigating the Extended Data files. We will raise a comment to the editorial team in this regard.

04 Aug 2020 | for Version 1
Arnon Jurberg, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

The work by Sankaranarayanan & Coghlan et al. sought to deploy CRISPR/Cas9 by electroporation at distinct developmental stages of schistosomes (more specifically, eggs, sporocysts and adult worms). A successful implementation of this technique in schistosomes will open up numerous possibilities for the study of gene function in these flatworms and is likely to revolutionize the field in the same manner as gene knockout by homologous recombination in mice and rats.

The paper is well written and provided a detailed description of their methods and findings. The figures are well conceived and the paper is likely to provide grounds for others to come, although CRISPR efficiency rate was below 5%.

1. Considering the low editing efficiency, did the authors evaluate the activity of other gRNAs against SULT-OR?

R. This is a good question, and indeed we have originally designed three gRNAs in total, but have only tested one of them so far. The gRNA employed in this study is the closest to the start codon of the exon 1 (Figure 1B), and predicted to be SmSULT-OR specific; therefore, reducing the chances of off-target mutations in the other members of the sulfotransferase family (as shown in Figure S3 A). Additionally, it is expected that an indel mutation (that is not a multiple of 3 bp in size) at the 5' end of the gene could disrupt the reading frame of nearly the whole gene and so nearly the whole protein, while such a mutation near the 3' end of the gene would only affect the sequence of the end of the protein.

2. It also caught my attention the apparent lack of nuclear staining for fluorescently labelled Cas9-gRNA. Did the authors address this?

R. We agree with the reviewer that nuclear staining, i.e. co-staining of the nuclei with DAPI and ATTO^{TM} 550 is not evident. Probably the confocal imaging is not sensitive enough to detect the fluorescent ribonucleoprotein complex within the cell nucleus. However, the Cas9 nuclease contains a nuclear localization sequence, and we identified site-specific deletions across the expected double-strand break site in the three developmental stages tested herein. This strongly suggests that even though no nuclear signal was detected (as above, probably due to sensitivity of the confocal assay), the ribonucleoprotein complex did indeed reach the genomic DNA of the cells. In this regard, we have now incorporated the following

statement in Discussion (4th paragraph):" *The absence of nuclear localisation of the RNP complex in the confocal images may indicate that the 4-hour timeframe between the molecule delivery and parasite fixation was too short, and which may explain, at least partially, the low CRISPR-induced deletion efficiency."*

3. It was also unclear to me whether the staining pattern between adult males and females were similar, and I would like to suggest the addition of further details. An alternative approach that can assist in achieving greater CRISPR efficiencies in schistosomes is the use of microinjection (perhaps in the ovary of females).

R. No evident differences were observed in the staining pattern between adult males and females. Accordingly, we have incorporated the following sentence in the Results (section "CRISPR-Cas9 machinery successfully delivered into schistosome developmental stages"): *"…no evident differences in the staining pattern were evident between sexes".* Regarding the use of microinjection to introduce the CRISPR machinery into the germ line, it is a very interesting idea worth testing in future experiments.

4. Once mutations in the target gene are predicted to induce oxamniquine resistance, did the authors evaluate whether incubation of electroporated parasites with this drug could improve CRISPR efficiency?

R. We acknowledge this suggestion, and indeed in one experiment we incubated CRISPR-mutated worms in the presence of oxamniquine (OXA), but no obvious resistant phenotype was detected. These findings were not unexpected, given the intrinsic low efficiency of CRISPR and that only a few (presumably somatic) cells, were mutated in the whole parasite. In the current study, we decided to focus on describing the induced deletions as a *proof-of-principle* that the technology can be applied to several developmental stages, in this case as a "somatic transgenesis" approach. Further experiments will be performed in the future to investigate the development of an OXA-resistance phenotype.

24 Aug 2020 | for Version 1
Mattie Pawlowic, Biological Chemistry and Drug Discovery, Wellcome Centre for Anti-Infectives Research, University of Dundee, Dundee, UK

Sankaranarayanan and colleagues report their work to mutate *Schitosoma mansoni* SULT-OR, a gene related to oxamniquine resistance, using transfection of ribonucleoprotein (RNP) complex to induce CRISPR/Cas9 double-stranded breaks. They use this approach on worm adults, sporocysts, and eggs, and deep sequence the resulting parasites to determine which mutations are present. This represents a new approach for genetic modification of Schistosoma.

The authors find that adult and sporocysts (not eggs) take up fluorescently labelled RNPs. These bind non-specifically on the surface of adult worms and also concentrate in the gut. In the future, the authors should visualise what uptake may occur in adult worms without electroporation, as it appears a significant amount of uptake may be due to feeding. It is curious that the Cas9, which contains a nuclear localisation sequence, does not localise the

RNP complex to the nucleus of worm cells. It could be that the 4-hour time frame is too short to observe this, or it could be that lack of nuclear localisation may explain the low editing efficiency.

R. Although the fluorescence signal is faint in eggs, in particular compared to adult worms and sporocysts, Figures 2F and H show it is inside the eggshell and in the miracidium surface. Given the high concentration of RNP in cells lining the gut, as suggested by the reviewer, in future experiments we may compare the RNP delivery by soaking and electroporation. Regarding the absence of co-localisation of DAPI and ATTO^{TM} 550, please see the answer to question 2 of the previous reviewer. The 4-hour timeframe being too short is a good hypothesis and worth testing it in future experiments. In this regard, we added the following statement in the 4th paragraph of Discussion: *"Future experiments comparing the delivery of the RNP complex by soaking versus electroporation would inform about the best delivery approach. The absence of nuclear localisation of the RNP complex in the confocal images may indicate that the 4-hour timeframe between the molecule delivery and parasite fixation was too short, and which may explain, at least partially, the low CRISPR-induced deletion efficiency."*

After 4 days of culture, the authors extracted DNA from transfected parasites, PCR amplified the region of interest, and deep sequenced to identify mutations. I understand that a larger PCR fragment is required to identify deletions, however the authors designed the ends of the PCR products to overlap at the Cas9 cut site. This led to difficulties in identifying true mutations vs sequencing errors at the cut site. Therefore, I think indels may be under-reported as they were not well captured.

R. In Figure 1C, we show the positions of the PCR primers in green and the amplicon in pale blue. The amplicon length is 262bp and it covers the predicted Cas9 site (position 137), the PCR products do not overlap at the Cas9 cut site. On the other hand, the forward and reverse sequence reads do overlap the cut site, increasing the accuracy to detect CRISPR-induced mutations. To make the Figure clearer, we have now incorporated more information in the Figure 1C legend as follows: *"(C) Reference PCR amplicon (pale blue), showing the positions of the gRNA, PAM, DSB, forward and reverse PCR primers (green), and forward and reverse sequence reads (orange), as well as a SNP site found in many sequence reads. The diagrams are drawn to scale."*

The authors find that deletions are the most common mutations the occur at the SULT-OR locus. I disagree that deletions in the range of 24-102 bp should be called "large". Consistent with use of CRISPR in other systems, deletions occur upstream of the cut site. Although adult worms took up RNPs the best, the mutation rate was found to be only <2%. This is likely due to the RNPs not transfecting every cell in the adult worm.

R. As such, the term "large" is relative, and we agree that deletions in the range of 24-102 are not large compared to deletions of hundreds and even thousands of bases. However, given that in the earlier CRISPR studies the majority of identified deletions were in the range of 1-3bp around the double-stranded break site (see https://www.biorxiv.org/content/10.1101/2020.08.05.216739v1), and our study is the second in schistosomes, and the first one in this parasite showing deletions up to 102bp, we have

decided to keep the term "large" to describe them. Finally, even longer deletions (~ 500bp) as shown for *Strongyloides* spp (PMID: 29016680) but not detectable by our PCR-based approach, cannot be ruled out in our system.

Despite SULT-OR being at its highest expression levels in adult worms, mutation rates were too low to see an impact on SULT-OR expression levels.
Overall this work reports new methods for genetic modification of *Schistosoma mansoni*. While the efficiency is low, further optimisation makes this approach promising.

R. We completely agree with the reviewer, and further optimisation, not only to increase the CRISPR-induced mutations, but also to introduce them into the germline, is planned.

01 Sep 2020 | for Version 1
Patrick Skelly, Molecular Helminthology Laboratory, Department of Infectious Disease and Global Health, Cummings School of Veterinary Medicine, Tufts University, North Grafton, MA, USA

This study by Sankaranarayanan & Coghlan *et al.* employed CRISPR-Cas9 to target for disruption the SULT-OR sulfotransferase gene in *Schistosoma mansoni* using electroporation to deliver a ribonucleoprotein (RNP) complex consisting of gRNA, tracrRNA and Cas9 protein. Mutations induced by this treatment were subsequently detected by high-throughput sequencing. The results provide independent confirmation that CRISPR/Cas delivery to *S. mansoni* can indeed induce mutations in these parasites, albeit (so far) at a low level. Three developmental stages were examined, and mutations were most commonly detected in adult worms (0.3-2.0% of aligned sequence reads), followed by sporocysts (0.1-0.2%), and were extremely rare in eggs. The work provides hope that ongoing research will permit researchers to use refined methods to target the germ line to create stable knock-out or knock-in strains of any schistosome gene of interest.

Some comments and questions:

<u>Animal Procedures</u>. Why were snails "moved into dark cupboards at 28°C when they start shedding cercariae" and is there a reference showing that this is helpful? Page 4: How was the water "conditioned"? (At least in my version) the "Welfare assessments...." sentence needs correcting.

R. Regarding the question "why the snails were moved into dark cupboards at 28°C when they start shedding cercariae", there is not a reference showing this is helpful to obtain more cercariae. In order to experimentally obtain the highest number of cercariae, as far as we are aware, there are mainly two approaches used by most of the laboratories working with schistosomes. One approach involves keeping the patent snails under 12-hour light/dark cycles and 24 hours before shedding the snails are moved to the dark as described (PMID: 24510597). A second approach, that originally was followed in David Dunne laboratory at the Department of Pathology, University of Cambridge (personal communication by Anna Protasio), and that we inherited involves moving the snails into dark incubators when they start to shed cercariae, and keeping them at dark all the time.

Even though, as far as we are aware, no studies comparing the two approaches have been conducted, the number of cercariae obtained by the latter approach, currently used by us, has been high enough to carry out all the experiments described in the current manuscript, in addition to maintaining the whole life cycle independently at the Wellcome Sanger Institute and supporting the rest of the schistosoma-related projects.

R. In the Methods section "Animal Procedures" we have now included the reference for the conditioned water recipe (i.e. PMID: 24510597)

R. We have now edited the sentence '*Welfare assessments are carried out daily abnormal signs of behaviour or clinical signs of concern are reported.*' as follows: '*Welfare assessments are carried out daily, and abnormal signs of behaviour or clinical signs of concern are reported.*'

<u>Parasite Material</u>. Useful, detailed methods are reported. So, for consistency, give specifics of the Percoll-sucrose solution.

R. Following the reviewer's suggestion we have now included more details about the Percoll-sucrose solution (in Methods, "Parasite Material"), as follows: '*The filtrate was passed through a Percoll-sucrose gradient prepared by mixing 8 ml of Percoll with 32 ml of sterile-filtered 0.25M sucrose...*'

<u>CRISPR-Cas9 ribonucleoprotein complex assembly</u>. Since we expect the gRNA and the nuclease to associate in a 1:1 ratio, why not mix these in equimolar amounts? How can the authors confirm that the gRNA and nuclease actually formed the RNP complex? Is there a way to determine how efficiently this occurred? Low efficiency might provide some explanation for the relatively low level of mutation detected here?

R. Indeed, as the reviewer pointed out we did not use equimolar amounts of gRNA and nuclease. For the RNP complex assembly we followed the protocol suggested by the manufacturer IDT, slightly modified based on a well-optimised protocol to introduce CRISPR-Cas9 mutations by RNP complex in pluripotent stem cells (PMID: 30912046). In these protocols an excess of gRNA compared to Cas9 nuclease is employed, presumably to ensure that all the nuclease molecules are combined with the gRNA. Having said that, future experiments could be conducted to test the CRISPR efficiency by changing the relative concentrations of the RNP complex components. We have now provided in the manuscript further information about the RNP assembly protocol. In the Methods (section, "CRISPR-Cas 9 ribonucleoprotein complex assembly") we have included the following statement: '*The CRISPR-Cas9 ribonucleoprotein complex (RNP) was assembled in vitro following the manufacturer recommendations slightly modified based on (Bruntraeger et al. 2019) by combining...*'

R. We have not confirmed the actual RNP complex was properly assembled, but the presence of definitive CRISPR-Cas -induced deletions around the predicted double-stranded break site strongly suggests that the complex was, at least partially, active. Protocols to evaluate the assembly of Cas9/gRNA complexes by using a fluorometric molecular beacon-like assay have been developed (PMID: 26945042), and could be applied in our future studies. We agree with the reviewer that low efficiency of assembly of RNP complexes could,

at least partially, explain the low efficiency of the system to induce site-specific mutations. However, the CRISPR efficiency reported by Ittiprasert et al (PMID: 30644357), where RNP complex and lentivirus were (separately) employed to induce site-specific mutations, was even lower than that reported here. This suggests that the low efficiency may be related to the species rather than the approach employed to deliver the CRISPR-Cas cargo.

R. In the first paragraph of Discussion we have now included the following sentence to address this point raised by the reviewer: *'The low CRISPR-Cas efficiency in our study may be explained, at least partially, by low efficiency of RNP complex assembly; however, as discussed above, a low CRISPR-Cas efficiency was also described when lentivirus (rather than RNP complex) was used to deliver the CRISPR-Cas9 cargo into the parasite (Ittiprasert et al. 2019)'.*

<u>Bioinformatic analysis.</u> Consider adding a reference to Phred Quality Scoring so that interested readers can make sense of the <23 score cut off. It may also be useful to define here "SNP" and "NHEJ".

R. Edited as suggested.

<u>Results. CRISPR-Cas9 machinery.....</u>The authors should report if ALL of the adults (males and females) yielded a similar staining pattern. In Figure 2A and Movie 1, surface staining is not uniform along the length of the worm; did all males exhibit such staining in a similar region and to a similar extent (and in all replicates)? Does any of the RNP staining co-localize with DAPI-stained nuclei? If yes, point examples out. If not, how well is the nuclear localization signal working?

R. Following the reviewer's suggestion we have now included in Results (section "CRISPR-Cas9 machinery successfully delivered into schistosome developmental stages") the following statement: *'The staining pattern was similar in all observed specimens, and no evident differences in the staining pattern were evident between sexes.'* Regarding the absence of co-localisation of DAPI and ATTO^{TM} 550, please see the answer to question 2 of the previous reviewer (Arnon Jurberg). We have not evaluated how well the nuclear localisation signal worked in our system. However, the presence of definitive CRISPR-Cas -induced deletions around the predicted double-stranded breaking points strongly suggests the system worked.

<u>Discussion.</u> It is reported that "no evidence for CRISPR-Cas9-induced insertions or substitutions in the SULT-OR gene was observed." For context, it would be useful to know how this equates (if possible, under comparable experimental conditions) with work in *Strongyloides*, *C. elegans*, and mammalian cell-lines.

R. Following the reviewer's suggestion we have now included in the first paragraph of Discussion the following statement: *' In Strongyloides spp, Gang et al (Gang et al. 2017) showed that in the absence of a repair template, small insertions or deletions (indels) or substitutions were not observed, but instead the authors found deletions of >500 bp at the target site in the unc-22 gene. On the other hand, in C. elegans, one study in the absence of a repair template, detected deletions ranging from 7 bp to >2 kb in the dpy-11 and unc-4 genes (Chiu et al. 2013); while another study found only small insertions and deletions of <20 bp (Friedland et al.*

2013)'. In mammalian cell lines, the majority of indels are relatively small (1-50 bp), but larger deletions of kilobases in size are also sometimes observed (Kosicki et al, https://www.biorxiv.org/content/10.1101/2020.08.05.216739v1).'

I disagree that because mutations were detected in 2% of aligned reads, this necessarily means that ~2% of cells were impacted.

R. We agree that in order to claim that 2% of aligned reads means that ~2% of cells were impacted, we have to assume an even distribution of mutations, and this is not necessarily the case. In Discussion (4th paragraph) we have now edited the text accordingly: *'The large deletions spanning the predicted Cas9 cut site were found in 0.3–2.0% of aligned reads from CRISPR-treated adult worms, so our best estimate of the fraction of adult cells in which CRISPR worked is 0.3–2.0% assuming an even distribution of mutations across the transfected parasites. However, since a pool of five adult worms were transfected with the RNP complex, the efficiency of CRISPR (and the amount of knock-down at the mRNA level) may have varied between worms, as well as between cells of an individual worm.'*

The control data shown in e.g. Figure 3A – were they derived from the mock treated worms or from worms treated with Cas9 only or from those treated with gRNA only? Or were data from all controls combined?

R. For both experimental samples and controls in all the figures we provide the sample identifiers. The information for all these identifiers and samples can be found in Supplementary Table 1. In particular for Figure 3A the controls include worm-only controls for the adult samples, worm+Cas9-only and worm-only controls for the sporocyst samples, egg+Cas9-only and egg-only controls for the egg samples. To clarify this point, we have now added this information in the Figure 3A legend as follows: *'(A) Frequency of deletions in NGS sequencing data, identified with the assistance of CRISPResso in three biological replicates from adults, two from sporocysts, and three from eggs, as indicated (sample identifiers at the bottom). The controls include worm-only controls for the adult samples, worm treated with Cas9-only and worm-only controls for the sporocyst samples, eggs treated with Cas9-only and egg-only controls for the egg samples.'*

Report how the tree shown in Figure S1A was generated. What does "PRJEA36577" etc. refer to in the figure? What is the value of showing anything outside of the area bounded by the red dashed line?

R. To clarify this point we have now included two sentences in the figure legend describing how the tree shown in Figure S1A (now in Extended data V2) was generated and the meaning of identifiers next to the species names, as follows: *'The phylogenetic tree was generated by WormBase ParaSite using the EnsemblCompara pipeline [34]. Identifiers beside the species names in the tree indicate the NCBI BioProject accession numbers for the sequencing projects for each species or species strain.'*

R. The red dashed line shows a clade of the phylogenetic tree that includes paralogues of SULT-OR that lie in a region about 3.2 Mb along *S. mansoni* chromosome 6. The rest of the phylogenetic tree, outside the red dashed line, includes other more distant paralogues of

SULT-OR, that are scattered around the *S. mansoni* genome. We think it could be of interest to readers that, as well as several closely related paralogues of SULT-OR that are found nearby to SULT-OR in the genome, the SULT-OR gene also has more distantly related paralogues scattered around the genome (these are the ones outside the red dashed line).

Competing Interests: No competing interests were disclosed.

Reviewer Report 24 August 2020

https://doi.org/10.21956/wellcomeopenres.17586.r39705

✔   **Mattie Pawlowic** (iD)

Biological Chemistry and Drug Discovery, Wellcome Centre for Anti-Infectives Research, University of Dundee, Dundee, UK

Sankaranarayanan and colleagues report their work to mutate *Schitosoma mansoni* SULT-OR, a gene related to oxamniquine resistance, using transfection of ribonucleoprotein (RNP) complex to induce CRISPR/Cas9 double-stranded breaks. They use this approach on worm adults, sporocysts, and eggs, and deep sequence the resulting parasites to determine which mutations are present. This represents a new approach for genetic modification of Schistosoma.

The authors find that adult and sporocysts (not eggs) take up fluorescently labelled RNPs. These bind non-specifically on the surface of adult worms and also concentrate in the gut. In the future, the authors should visualise what uptake may occur in adult worms without electroporation, as it appears a significant amount of uptake may be due to feeding. It is curious that the Cas9, which contains a nuclear localisation sequence, does not localise the RNP complex to the nucleus of worm cells. It could be that the 4-hour time frame is too short to observe this, or it could be that lack of nuclear localisation may explain the low editing efficiency.

After 4 days of culture, the authors extracted DNA from transfected parasites, PCR amplified the region of interest, and deep sequenced to identify mutations. I understand that a larger PCR fragment is required to identify deletions, however the authors designed the ends of the PCR products to overlap at the Cas9 cut site. This led to difficulties in identifying true mutations vs sequencing errors at the cut site. Therefore, I think indels may be under-reported as they were not well captured.

The authors find that deletions are the most common mutations the occur at the SULT-OR locus. I disagree that deletions in the range of 24-102 bp should be called "large". Consistent with use of CRISPR in other systems, deletions occur upstream of the cut site. Although adult worms took up RNPs the best, the mutation rate was found to be only <2%. This is likely due to the RNPs not transfecting every cell in the adult worm.

Despite SULT-OR being at its highest expression levels in adult worms, mutation rates were too low to see a impact on SULT-OR expression levels.

Overall this work reports new methods for genetic modification of *Schistosoma mansoni*. While the efficiency is low, further optimisation makes this approach promising.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Parasite genetics and biochemistry, Cryptosporidium.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 07 Dec 2020
Gabriel Rinaldi, Wellcome Sanger Institute, Hinxton, UK

07 December 2020,
Wellcome Open Research,

On behalf of all the authors, we acknowledge and appreciate the efforts of the reviewers with the review of our manuscript. We have been able to revise along the lines of the recommendations, which has certainly contributed to improve our paper. Point-by-point responses are provided below: **Response (R)**

30 Jul 2020 | for Version 1
Phillip LoVerde, Department of Biochemistry and Structural Biology, University of Texas

Health Science Center at San Antonio, San Antonio, TX, USA

Schistosomiasis remains a serious global health problem. The toolbox available to scientists continues to improve. The ability to manipulate genes is a key technique to the continued success of the field to address questions of biology, host-parasite interactions, vaccines, drugs and diagnostics among others. The manuscript by the Wellcome Sanger Group is a welcome addition as it provides very detailed methods, state of the art approach and detailed analysis of the results. In fact, it provides a roadmap for others to follow. This is in spite of the fact they were not successful in producing a drug resistant transgenic strain. Clearly part of the reason was that only somatic cells were affected by the CRISPR-Cas 9 construct. They recognize that affecting germ line cells will likely make a difference. Another plausible explanation is that the cells in which CRISPR-Cas9 did induce mutations represented a small fraction of the cells expressing SmSULT-OR. They also acknowledge that the resistance trait is double recessive inferring that both alleles must be interrupted. I may have missed it but did not see where they attempted to demonstrate that both alleles of SmSULT-OR were interrupted.

R. We acknowledge the reviewer for the positive comments, and as indicated, the resistance trait is double recessive inferring that both alleles must be interrupted in the germline to generate a transgenic line with a fully OXA-resistant phenotype. In this study, we did not demonstrate that both alleles of SmSULT-OR were interrupted, given that the sequence data were generated from a large number of cells, probably including a mixture of wild type and mutant cells. In order to identify allele-specific mutations single cell DNA PCR and/ or sequencing approaches need to be applied (PMID: 31827197; PMID: 30992375). This can be optimised in the future, in particular after having shown that single cell sequencing approaches are feasible in schistosomes (PMID: 32973030; bioRxiv 754713; doi: https://doi.org/10.1101/754713)

The manuscript is well-written and follows a logical progression with well-designed experiments. Suggest use the term intramolluscan instead of intrasnail.

R. Edited as recommended

An issue out of their control is the difficulty in navigating the Extended data Figures and Tables. However, all the data is available to the reader.

R. We agree and have also found difficulties in navigating the Extended Data files. We will raise a comment to the editorial team in this regard.

04 Aug 2020 | for Version 1
Arnon Jurberg, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

The work by Sankaranarayanan & Coghlan et al. sought to deploy CRISPR/Cas9 by electroporation at distinct developmental stages of schistosomes (more specifically, eggs, sporocysts and adult worms). A successful implementation of this technique in schistosomes will open up numerous possibilities for the study of gene function in these

flatworms and is likely to revolutionize the field in the same manner as gene knockout by homologous recombination in mice and rats.

The paper is well written and provided a detailed description of their methods and findings. The figures are well conceived and the paper is likely to provide grounds for others to come, although CRISPR efficiency rate was below 5%.

1. Considering the low editing efficiency, did the authors evaluate the activity of other gRNAs against SULT-OR?

R. This is a good question, and indeed we have originally designed three gRNAs in total, but have only tested one of them so far. The gRNA employed in this study is the closest to the start codon of the exon 1 (Figure 1B), and predicted to be SmSULT-OR specific; therefore, reducing the chances of off-target mutations in the other members of the sulfotransferase family (as shown in Figure S3 A). Additionally, it is expected that an indel mutation (that is not a multiple of 3 bp in size) at the 5' end of the gene could disrupt the reading frame of nearly the whole gene and so nearly the whole protein, while such a mutation near the 3' end of the gene would only affect the sequence of the end of the protein.

2. It also caught my attention the apparent lack of nuclear staining for fluorescently labelled Cas9-gRNA. Did the authors address this?

R. We agree with the reviewer that nuclear staining, i.e. co-staining of the nuclei with DAPI and ATTO^{TM} 550 is not evident. Probably the confocal imaging is not sensitive enough to detect the fluorescent ribonucleoprotein complex within the cell nucleus. However, the Cas9 nuclease contains a nuclear localization sequence, and we identified site-specific deletions across the expected double-strand break site in the three developmental stages tested herein. This strongly suggests that even though no nuclear signal was detected (as above, probably due to sensitivity of the confocal assay), the ribonucleoprotein complex did indeed reach the genomic DNA of the cells. In this regard, we have now incorporated the following statement in Discussion (4th paragraph):" *The absence of nuclear localisation of the RNP complex in the confocal images may indicate that the 4-hour timeframe between the molecule delivery and parasite fixation was too short, and which may explain, at least partially, the low CRISPR-induced deletion efficiency."*

3. It was also unclear to me whether the staining pattern between adult males and females were similar, and I would like to suggest the addition of further details. An alternative approach that can assist in achieving greater CRISPR efficiencies in schistosomes is the use of microinjection (perhaps in the ovary of females).

R. No evident differences were observed in the staining pattern between adult males and females. Accordingly, we have incorporated the following sentence in the Results (section "CRISPR-Cas9 machinery successfully delivered into schistosome developmental stages"): *"...no evident differences in the staining pattern were evident between sexes".* Regarding the use of microinjection to introduce the CRISPR machinery into the germ line, it is a very interesting idea worth testing in future experiments.

4. Once mutations in the target gene are predicted to induce oxamniquine resistance, did the authors evaluate whether incubation of electroporated parasites with this drug could improve CRISPR efficiency?

R. We acknowledge this suggestion, and indeed in one experiment we incubated CRISPR-mutated worms in the presence of oxamniquine (OXA), but no obvious resistant phenotype was detected. These findings were not unexpected, given the intrinsic low efficiency of CRISPR and that only a few (presumably somatic) cells, were mutated in the whole parasite. In the current study, we decided to focus on describing the induced deletions as a *proof-of-principle* that the technology can be applied to several developmental stages, in this case as a "somatic transgenesis" approach. Further experiments will be performed in the future to investigate the development of an OXA-resistance phenotype.

24 Aug 2020 | for Version 1
Mattie Pawlowic, Biological Chemistry and Drug Discovery, Wellcome Centre for Anti-Infectives Research, University of Dundee, Dundee, UK

Sankaranarayanan and colleagues report their work to mutate *Schitosoma mansoni* SULT-OR, a gene related to oxamniquine resistance, using transfection of ribonucleoprotein (RNP) complex to induce CRISPR/Cas9 double-stranded breaks. They use this approach on worm adults, sporocysts, and eggs, and deep sequence the resulting parasites to determine which mutations are present. This represents a new approach for genetic modification of Schistosoma.

The authors find that adult and sporocysts (not eggs) take up fluorescently labelled RNPs. These bind non-specifically on the surface of adult worms and also concentrate in the gut. In the future, the authors should visualise what uptake may occur in adult worms without electroporation, as it appears a significant amount of uptake may be due to feeding. It is curious that the Cas9, which contains a nuclear localisation sequence, does not localise the RNP complex to the nucleus of worm cells. It could be that the 4-hour time frame is too short to observe this, or it could be that lack of nuclear localisation may explain the low editing efficiency.

R. Although the fluorescence signal is faint in eggs, in particular compared to adult worms and sporocysts, Figures 2F and H show it is inside the eggshell and in the miracidium surface. Given the high concentration of RNP in cells lining the gut, as suggested by the reviewer, in future experiments we may compare the RNP delivery by soaking and electroporation. Regarding the absence of co-localisation of DAPI and ATTO^{TM} 550, please see the answer to question 2 of the previous reviewer. The 4-hour timeframe being too short is a good hypothesis and worth testing it in future experiments. In this regard, we added the following statement in the 4th paragraph of Discussion: *"Future experiments comparing the delivery of the RNP complex by soaking versus electroporation would inform about the best delivery approach. The absence of nuclear localisation of the RNP complex in the confocal images may indicate that the 4-hour timeframe between the molecule delivery and parasite fixation was too short, and which may explain, at least partially, the low CRISPR-induced deletion efficiency."*

After 4 days of culture, the authors extracted DNA from transfected parasites, PCR amplified the region of interest, and deep sequenced to identify mutations. I understand that a larger PCR fragment is required to identify deletions, however the authors designed the ends of the PCR products to overlap at the Cas9 cut site. This led to difficulties in identifying true mutations vs sequencing errors at the cut site. Therefore, I think indels may be under-reported as they were not well captured.

R. In Figure 1C, we show the positions of the PCR primers in green and the amplicon in pale blue. The amplicon length is 262bp and it covers the predicted Cas9 site (position 137), the PCR products do not overlap at the Cas9 cut site. On the other hand, the forward and reverse sequence reads do overlap the cut site, increasing the accuracy to detect CRISPR-induced mutations. To make the Figure clearer, we have now incorporated more information in the Figure 1C legend as follows: *"(C) Reference PCR amplicon (pale blue), showing the positions of the gRNA, PAM, DSB, forward and reverse PCR primers (green), and forward and reverse sequence reads (orange), as well as a SNP site found in many sequence reads. The diagrams are drawn to scale."*

The authors find that deletions are the most common mutations the occur at the SULT-OR locus. I disagree that deletions in the range of 24-102 bp should be called "large". Consistent with use of CRISPR in other systems, deletions occur upstream of the cut site. Although adult worms took up RNPs the best, the mutation rate was found to be only <2%. This is likely due to the RNPs not transfecting every cell in the adult worm.

R. As such, the term "large" is relative, and we agree that deletions in the range of 24-102 are not large compared to deletions of hundreds and even thousands of bases. However, given that in the earlier CRISPR studies the majority of identified deletions were in the range of 1-3bp around the double-stranded break site (see https://www.biorxiv.org/content/10.1101/2020.08.05.216739v1), and our study is the second in schistosomes, and the first one in this parasite showing deletions up to 102bp, we have decided to keep the term "large" to describe them. Finally, even longer deletions (~ 500bp) as shown for *Strongyloides* spp (PMID: 29016680) but not detectable by our PCR-based approach, cannot be ruled out in our system.

Despite SULT-OR being at its highest expression levels in adult worms, mutation rates were too low to see an impact on SULT-OR expression levels.
Overall this work reports new methods for genetic modification of *Schistosoma mansoni*. While the efficiency is low, further optimisation makes this approach promising.

R. We completely agree with the reviewer, and further optimisation, not only to increase the CRISPR-induced mutations, but also to introduce them into the germline, is planned.

01 Sep 2020 | for Version 1
Patrick Skelly, Molecular Helminthology Laboratory, Department of Infectious Disease and Global Health, Cummings School of Veterinary Medicine, Tufts University, North Grafton, MA, USA

This study by Sankaranarayanan & Coghlan *et al.* employed CRISPR-Cas9 to target for disruption the SULT-OR sulfotransferase gene in *Schistosoma mansoni* using electroporation to deliver a ribonucleoprotein (RNP) complex consisting of gRNA, tracrRNA and Cas9 protein. Mutations induced by this treatment were subsequently detected by high-throughput sequencing. The results provide independent confirmation that CRISPR/Cas delivery to *S. mansoni* can indeed induce mutations in these parasites, albeit (so far) at a low level. Three developmental stages were examined, and mutations were most commonly detected in adult worms (0.3-2.0% of aligned sequence reads), followed by sporocysts (0.1-0.2%), and were extremely rare in eggs. The work provides hope that ongoing research will permit researchers to use refined methods to target the germ line to create stable knock-out or knock-in strains of any schistosome gene of interest.

Some comments and questions:

Animal Procedures. Why were snails "moved into dark cupboards at 28°C when they start shedding cercariae" and is there a reference showing that this is helpful? Page 4: How was the water "conditioned"? (At least in my version) the "Welfare assessments…." sentence needs correcting.

R. Regarding the question "why the snails were moved into dark cupboards at 28°C when they start shedding cercariae", there is not a reference showing this is helpful to obtain more cercariae. In order to experimentally obtain the highest number of cercariae, as far as we are aware, there are mainly two approaches used by most of the laboratories working with schistosomes. One approach involves keeping the patent snails under 12-hour light/dark cycles and 24 hours before shedding the snails are moved to the dark as described (PMID: 24510597). A second approach, that originally was followed in David Dunne laboratory at the Department of Pathology, University of Cambridge (personal communication by Anna Protasio), and that we inherited involves moving the snails into dark incubators when they start to shed cercariae, and keeping them at dark all the time. Even though, as far as we are aware, no studies comparing the two approaches have been conducted, the number of cercariae obtained by the latter approach, currently used by us, has been high enough to carry out all the experiments described in the current manuscript, in addition to maintaining the whole life cycle independently at the Wellcome Sanger Institute and supporting the rest of the schistosoma-related projects.

R. In the Methods section "Animal Procedures" we have now included the reference for the conditioned water recipe (i.e. PMID: 24510597)

R. We have now edited the sentence '*Welfare assessments are carried out daily abnormal signs of behaviour or clinical signs of concern are reported*.' as follows: '*Welfare assessments are carried out daily, and abnormal signs of behaviour or clinical signs of concern are reported*.'

Parasite Material. Useful, detailed methods are reported. So, for consistency, give specifics of the Percoll-sucrose solution.

R. Following the reviewer's suggestion we have now included more details about the

Percoll-sucrose solution (in Methods, "Parasite Material"), as follows: *'The filtrate was passed through a Percoll-sucrose gradient prepared by mixing 8 ml of Percoll with 32 ml of sterile-filtered 0.25M sucrose...'*

<u>CRISPR-Cas9 ribonucleoprotein complex assembly</u>. Since we expect the gRNA and the nuclease to associate in a 1:1 ratio, why not mix these in equimolar amounts? How can the authors confirm that the gRNA and nuclease actually formed the RNP complex? Is there a way to determine how efficiently this occurred? Low efficiency might provide some explanation for the relatively low level of mutation detected here?

R. Indeed, as the reviewer pointed out we did not use equimolar amounts of gRNA and nuclease. For the RNP complex assembly we followed the protocol suggested by the manufacturer IDT, slightly modified based on a well-optimised protocol to introduce CRISPR-Cas9 mutations by RNP complex in pluripotent stem cells (PMID: 30912046). In these protocols an excess of gRNA compared to Cas9 nuclease is employed, presumably to ensure that all the nuclease molecules are combined with the gRNA. Having said that, future experiments could be conducted to test the CRISPR efficiency by changing the relative concentrations of the RNP complex components. We have now provided in the manuscript further information about the RNP assembly protocol. In the Methods (section, "CRISPR-Cas 9 ribonucleoprotein complex assembly") we have included the following statement: '*The CRISPR-Cas9 ribonucleoprotein complex (RNP) was assembled in vitro following the manufacturer recommendations slightly modified based on* (Bruntraeger et al. 2019) *by combining...'*

R. We have not confirmed the actual RNP complex was properly assembled, but the presence of definitive CRISPR-Cas -induced deletions around the predicted double-stranded break site strongly suggests that the complex was, at least partially, active. Protocols to evaluate the assembly of Cas9/gRNA complexes by using a fluorometric molecular beacon-like assay have been developed (PMID: 26945042), and could be applied in our future studies. We agree with the reviewer that low efficiency of assembly of RNP complexes could, at least partially, explain the low efficiency of the system to induce site-specific mutations. However, the CRISPR efficiency reported by Ittiprasert et al (PMID: 30644357), where RNP complex and lentivirus were (separately) employed to induce site-specific mutations, was even lower than that reported here. This suggests that the low efficiency may be related to the species rather than the approach employed to deliver the CRISPR-Cas cargo.

R. In the first paragraph of Discussion we have now included the following sentence to address this point raised by the reviewer: *'The low CRISPR-Cas efficiency in our study may be explained, at least partially, by low efficiency of RNP complex assembly; however, as discussed above, a low CRISPR-Cas efficiency was also described when lentivirus (rather than RNP complex) was used to deliver the CRISPR-Cas9 cargo into the parasite* (Ittiprasert et al. 2019)'.

<u>Bioinformatic analysis.</u> Consider adding a reference to Phred Quality Scoring so that interested readers can make sense of the <23 score cut off. It may also be useful to define here "SNP" and "NHEJ".

R. Edited as suggested.

Results. CRISPR-Cas9 machinery.....The authors should report if ALL of the adults (males and females) yielded a similar staining pattern. In Figure 2A and Movie 1, surface staining is not uniform along the length of the worm; did all males exhibit such staining in a similar region and to a similar extent (and in all replicates)? Does any of the RNP staining co-localize with DAPI-stained nuclei? If yes, point examples out. If not, how well is the nuclear localization signal working?

R. Following the reviewer's suggestion we have now included in Results (section "CRISPR-Cas9 machinery successfully delivered into schistosome developmental stages") the following statement: *'The staining pattern was similar in all observed specimens, and no evident differences in the staining pattern were evident between sexes.'* Regarding the absence of co-localisation of DAPI and ATTO^TM 550, please see the answer to question 2 of the previous reviewer (Arnon Jurberg). We have not evaluated how well the nuclear localisation signal worked in our system. However, the presence of definitive CRISPR-Cas -induced deletions around the predicted double-stranded breaking points strongly suggests the system worked.

Discussion. It is reported that "no evidence for CRISPR-Cas9-induced insertions or substitutions in the SULT-OR gene was observed." For context, it would be useful to know how this equates (if possible, under comparable experimental conditions) with work in *Strongyloides*, *C. elegans*, and mammalian cell-lines.

R. Following the reviewer's suggestion we have now included in the first paragraph of Discussion the following statement: ' *In Strongyloides spp, Gang et al (Gang et al. 2017) showed that in the absence of a repair template, small insertions or deletions (indels) or substitutions were not observed, but instead the authors found deletions of >500 bp at the target site in the unc-22 gene. On the other hand, in C. elegans, one study in the absence of a repair template, detected deletions ranging from 7 bp to >2 kb in the dpy-11 and unc-4 genes (Chiu et al. 2013); while another study found only small insertions and deletions of <20 bp (Friedland et al. 2013)'. In mammalian cell lines, the majority of indels are relatively small (1-50 bp), but larger deletions of kilobases in size are also sometimes observed (Kosicki et al, https://www.biorxiv.org/content/10.1101/2020.08.05.216739v1).'*

I disagree that because mutations were detected in 2% of aligned reads, this necessarily means that ~2% of cells were impacted.

R. We agree that in order to claim that 2% of aligned reads means that ~2% of cells were impacted, we have to assume an even distribution of mutations, and this is not necessarily the case. In Discussion (4th paragraph) we have now edited the text accordingly: *'The large deletions spanning the predicted Cas9 cut site were found in 0.3–2.0% of aligned reads from CRISPR-treated adult worms, so our best estimate of the fraction of adult cells in which CRISPR worked is 0.3–2.0% assuming an even distribution of mutations across the transfected parasites. However, since a pool of five adult worms were transfected with the RNP complex, the efficiency of CRISPR (and the amount of knock-down at the mRNA level) may have varied between worms, as well as between cells of an individual worm.'*

The control data shown in e.g. Figure 3A – were they derived from the mock treated worms or from worms treated with Cas9 only or from those treated with gRNA only? Or were data from all controls combined?

R. For both experimental samples and controls in all the figures we provide the sample identifiers. The information for all these identifiers and samples can be found in Supplementary Table 1. In particular for Figure 3A the controls include worm-only controls for the adult samples, worm+Cas9-only and worm-only controls for the sporocyst samples, egg+Cas9-only and egg-only controls for the egg samples. To clarify this point, we have now added this information in the Figure 3A legend as follows: *'(A) Frequency of deletions in NGS sequencing data, identified with the assistance of CRISPResso in three biological replicates from adults, two from sporocysts, and three from eggs, as indicated (sample identifiers at the bottom). The controls include worm-only controls for the adult samples, worm treated with Cas9-only and worm-only controls for the sporocyst samples, eggs treated with Cas9-only and egg-only controls for the egg samples.'*

Report how the tree shown in Figure S1A was generated. What does "PRJEA36577" etc. refer to in the figure? What is the value of showing anything outside of the area bounded by the red dashed line?

R. To clarify this point we have now included two sentences in the figure legend describing how the tree shown in Figure S1A (now in Extended data V2) was generated and the meaning of identifiers next to the species names, as follows: *'The phylogenetic tree was generated by WormBase ParaSite using the EnsemblCompara pipeline [34]. Identifiers beside the species names in the tree indicate the NCBI BioProject accession numbers for the sequencing projects for each species or species strain.'*

R. The red dashed line shows a clade of the phylogenetic tree that includes paralogues of SULT-OR that lie in a region about 3.2 Mb along *S. mansoni* chromosome 6. The rest of the phylogenetic tree, outside the red dashed line, includes other more distant paralogues of SULT-OR, that are scattered around the *S. mansoni* genome. We think it could be of interest to readers that, as well as several closely related paralogues of SULT-OR that are found nearby to SULT-OR in the genome, the SULT-OR gene also has more distantly related paralogues scattered around the genome (these are the ones outside the red dashed line).

Competing Interests: No competing interests were disclosed.

Reviewer Report 04 August 2020

✔ **Arnon Jurberg**

Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

The work by Sankaranarayanan & Coghlan *et al.* sought to deploy CRISPR/Cas9 by electroporation at distinct developmental stages of schistosomes (more specifically, eggs, sporocysts and adult worms). A successful implementation of this technique in schistosomes will open up numerous possibilities for the study of gene function in these flatworms and is likely to revolutionize the field in the same manner as gene knockout by homologous recombination in mice and rats.

The paper is well written and provided a detailed description of their methods and findings. The figures are well conceived and the paper is likely to provide grounds for others to come, although CRISPR efficiency rate was below 5%.

1. Considering the low editing efficiency, did the authors evaluate the activity of other gRNAs against *SULT-OR*?

2. It also caught my attention the apparent lack of nuclear staining for fluorescently labelled Cas9-gRNA. Did the authors address this?

3. It was also unclear to me whether the staining pattern between adult males and females were similar, and I would like to suggest the addition of further details. An alternative approach that can assist in achieving greater CRISPR efficiencies in schistosomes is the use of microinjection (perhaps in the ovary of females).

4. Once mutations in the target gene are predicted to induce oxamniquine resistance, did the authors evaluate whether incubation of electroporated parasites with this drug could improve CRISPR efficiency?

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Developmental biology, molecular biology and gene editing, cell biology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 07 Dec 2020

Gabriel Rinaldi, Wellcome Sanger Institute, Hinxton, UK

07 December 2020,
Wellcome Open Research,

On behalf of all the authors, we acknowledge and appreciate the efforts of the reviewers with the review of our manuscript. We have been able to revise along the lines of the recommendations, which has certainly contributed to improve our paper. Point-by-point responses are provided below: **Response (R)**

30 Jul 2020 | for Version 1
Phillip LoVerde, Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

Schistosomiasis remains a serious global health problem. The toolbox available to scientists continues to improve. The ability to manipulate genes is a key technique to the continued success of the field to address questions of biology, host-parasite interactions, vaccines, drugs and diagnostics among others. The manuscript by the Wellcome Sanger Group is a welcome addition as it provides very detailed methods, state of the art approach and detailed analysis of the results. In fact, it provides a roadmap for others to follow. This is in spite of the fact they were not successful in producing a drug resistant transgenic strain. Clearly part of the reason was that only somatic cells were affected by the CRISPR-Cas 9 construct. They recognize that affecting germ line cells will likely make a difference. Another plausible explanation is that the cells in which CRISPR-Cas9 did induce mutations represented a small fraction of the cells expressing SmSULT-OR. They also acknowledge that the resistance trait is double recessive inferring that both alleles must be interrupted. I may have missed it but did not see where they attempted to demonstrate that both alleles of SmSULT-OR were interrupted.

R. We acknowledge the reviewer for the positive comments, and as indicated, the resistance trait is double recessive inferring that both alleles must be interrupted in the germline to generate a transgenic line with a fully OXA-resistant phenotype. In this study, we did not demonstrate that both alleles of SmSULT-OR were interrupted, given that the sequence data were generated from a large number of cells, probably including a mixture of wild type and mutant cells. In order to identify allele-specific mutations single cell DNA PCR and/ or sequencing approaches need to be applied (PMID: 31827197; PMID: 30992375). This can be optimised in the future, in particular after having shown that single cell sequencing approaches are feasible in schistosomes (PMID: 32973030; bioRxiv 754713; doi: https://doi.org/10.1101/754713)

The manuscript is well-written and follows a logical progression with well-designed experiments. Suggest use the term intramolluscan instead of intrasnail.

R. Edited as recommended

An issue out of their control is the difficulty in navigating the Extended data Figures and Tables. However, all the data is available to the reader.

R. We agree and have also found difficulties in navigating the Extended Data files. We will raise a comment to the editorial team in this regard.

04 Aug 2020 | for Version 1
Arnon Jurberg, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

The work by Sankaranarayanan & Coghlan et al. sought to deploy CRISPR/Cas9 by electroporation at distinct developmental stages of schistosomes (more specifically, eggs, sporocysts and adult worms). A successful implementation of this technique in schistosomes will open up numerous possibilities for the study of gene function in these flatworms and is likely to revolutionize the field in the same manner as gene knockout by homologous recombination in mice and rats.

The paper is well written and provided a detailed description of their methods and findings. The figures are well conceived and the paper is likely to provide grounds for others to come, although CRISPR efficiency rate was below 5%.

1. Considering the low editing efficiency, did the authors evaluate the activity of other gRNAs against SULT-OR?

R. This is a good question, and indeed we have originally designed three gRNAs in total, but have only tested one of them so far. The gRNA employed in this study is the closest to the start codon of the exon 1 (Figure 1B), and predicted to be SmSULT-OR specific; therefore, reducing the chances of off-target mutations in the other members of the sulfotransferase family (as shown in Figure S3 A). Additionally, it is expected that an indel mutation (that is not a multiple of 3 bp in size) at the 5' end of the gene could disrupt the reading frame of nearly the whole gene and so nearly the whole protein, while such a mutation near the 3' end of the gene would only affect the sequence of the end of the protein.

2. It also caught my attention the apparent lack of nuclear staining for fluorescently labelled Cas9-gRNA. Did the authors address this?

R. We agree with the reviewer that nuclear staining, i.e. co-staining of the nuclei with DAPI and ATTO^{TM} 550 is not evident. Probably the confocal imaging is not sensitive enough to detect the fluorescent ribonucleoprotein complex within the cell nucleus. However, the Cas9 nuclease contains a nuclear localization sequence, and we identified site-specific deletions across the expected double-strand break site in the three developmental stages tested herein. This strongly suggests that even though no nuclear signal was detected (as above,

probably due to sensitivity of the confocal assay), the ribonucleoprotein complex did indeed reach the genomic DNA of the cells. In this regard, we have now incorporated the following statement in Discussion (4th paragraph):" *The absence of nuclear localisation of the RNP complex in the confocal images may indicate that the 4-hour timeframe between the molecule delivery and parasite fixation was too short, and which may explain, at least partially, the low CRISPR-induced deletion efficiency."*

3. It was also unclear to me whether the staining pattern between adult males and females were similar, and I would like to suggest the addition of further details. An alternative approach that can assist in achieving greater CRISPR efficiencies in schistosomes is the use of microinjection (perhaps in the ovary of females).

R. No evident differences were observed in the staining pattern between adult males and females. Accordingly, we have incorporated the following sentence in the Results (section "CRISPR-Cas9 machinery successfully delivered into schistosome developmental stages"): *"...no evident differences in the staining pattern were evident between sexes".* Regarding the use of microinjection to introduce the CRISPR machinery into the germ line, it is a very interesting idea worth testing in future experiments.

4. Once mutations in the target gene are predicted to induce oxamniquine resistance, did the authors evaluate whether incubation of electroporated parasites with this drug could improve CRISPR efficiency?

R. We acknowledge this suggestion, and indeed in one experiment we incubated CRISPR-mutated worms in the presence of oxamniquine (OXA), but no obvious resistant phenotype was detected. These findings were not unexpected, given the intrinsic low efficiency of CRISPR and that only a few (presumably somatic) cells, were mutated in the whole parasite. In the current study, we decided to focus on describing the induced deletions as a *proof-of-principle* that the technology can be applied to several developmental stages, in this case as a "somatic transgenesis" approach. Further experiments will be performed in the future to investigate the development of an OXA-resistance phenotype.

24 Aug 2020 | for Version 1
Mattie Pawlowic, Biological Chemistry and Drug Discovery, Wellcome Centre for Anti-Infectives Research, University of Dundee, Dundee, UK

Sankaranarayanan and colleagues report their work to mutate *Schitosoma mansoni* SULT-OR, a gene related to oxamniquine resistance, using transfection of ribonucleoprotein (RNP) complex to induce CRISPR/Cas9 double-stranded breaks. They use this approach on worm adults, sporocysts, and eggs, and deep sequence the resulting parasites to determine which mutations are present. This represents a new approach for genetic modification of Schistosoma.

The authors find that adult and sporocysts (not eggs) take up fluorescently labelled RNPs. These bind non-specifically on the surface of adult worms and also concentrate in the gut. In the future, the authors should visualise what uptake may occur in adult worms without

electroporation, as it appears a significant amount of uptake may be due to feeding. It is curious that the Cas9, which contains a nuclear localisation sequence, does not localise the RNP complex to the nucleus of worm cells. It could be that the 4-hour time frame is too short to observe this, or it could be that lack of nuclear localisation may explain the low editing efficiency.

R. Although the fluorescence signal is faint in eggs, in particular compared to adult worms and sporocysts, Figures 2F and H show it is inside the eggshell and in the miracidium surface. Given the high concentration of RNP in cells lining the gut, as suggested by the reviewer, in future experiments we may compare the RNP delivery by soaking and electroporation. Regarding the absence of co-localisation of DAPI and ATTO^TM 550, please see the answer to question 2 of the previous reviewer. The 4-hour timeframe being too short is a good hypothesis and worth testing it in future experiments. In this regard, we added the following statement in the 4th paragraph of Discussion: *"Future experiments comparing the delivery of the RNP complex by soaking versus electroporation would inform about the best delivery approach. The absence of nuclear localisation of the RNP complex in the confocal images may indicate that the 4-hour timeframe between the molecule delivery and parasite fixation was too short, and which may explain, at least partially, the low CRISPR-induced deletion efficiency."*

After 4 days of culture, the authors extracted DNA from transfected parasites, PCR amplified the region of interest, and deep sequenced to identify mutations. I understand that a larger PCR fragment is required to identify deletions, however the authors designed the ends of the PCR products to overlap at the Cas9 cut site. This led to difficulties in identifying true mutations vs sequencing errors at the cut site. Therefore, I think indels may be under-reported as they were not well captured.

R. In Figure 1C, we show the positions of the PCR primers in green and the amplicon in pale blue. The amplicon length is 262bp and it covers the predicted Cas9 site (position 137), the PCR products do not overlap at the Cas9 cut site. On the other hand, the forward and reverse sequence reads do overlap the cut site, increasing the accuracy to detect CRISPR-induced mutations. To make the Figure clearer, we have now incorporated more information in the Figure 1C legend as follows: *"(C) Reference PCR amplicon (pale blue), showing the positions of the gRNA, PAM, DSB, forward and reverse PCR primers (green), and forward and reverse sequence reads (orange), as well as a SNP site found in many sequence reads. The diagrams are drawn to scale."*

The authors find that deletions are the most common mutations the occur at the SULT-OR locus. I disagree that deletions in the range of 24-102 bp should be called "large". Consistent with use of CRISPR in other systems, deletions occur upstream of the cut site. Although adult worms took up RNPs the best, the mutation rate was found to be only <2%. This is likely due to the RNPs not transfecting every cell in the adult worm.

R. As such, the term "large" is relative, and we agree that deletions in the range of 24-102 are not large compared to deletions of hundreds and even thousands of bases. However, given that in the earlier CRISPR studies the majority of identified deletions were in the range of 1-3bp around the double-stranded break site (see

https://www.biorxiv.org/content/10.1101/2020.08.05.216739v1), and our study is the second in schistosomes, and the first one in this parasite showing deletions up to 102bp, we have decided to keep the term "large" to describe them. Finally, even longer deletions (~ 500bp) as shown for *Strongyloides* spp (PMID: 29016680) but not detectable by our PCR-based approach, cannot be ruled out in our system.

Despite SULT-OR being at its highest expression levels in adult worms, mutation rates were too low to see an impact on SULT-OR expression levels.
Overall this work reports new methods for genetic modification of *Schistosoma mansoni*. While the efficiency is low, further optimisation makes this approach promising.

R. We completely agree with the reviewer, and further optimisation, not only to increase the CRISPR-induced mutations, but also to introduce them into the germline, is planned.

01 Sep 2020 | for Version 1
Patrick Skelly, Molecular Helminthology Laboratory, Department of Infectious Disease and Global Health, Cummings School of Veterinary Medicine, Tufts University, North Grafton, MA, USA

This study by Sankaranarayanan & Coghlan *et al.* employed CRISPR-Cas9 to target for disruption the SULT-OR sulfotransferase gene in *Schistosoma mansoni* using electroporation to deliver a ribonucleoprotein (RNP) complex consisting of gRNA, tracrRNA and Cas9 protein. Mutations induced by this treatment were subsequently detected by high-throughput sequencing. The results provide independent confirmation that CRISPR/Cas delivery to *S. mansoni* can indeed induce mutations in these parasites, albeit (so far) at a low level. Three developmental stages were examined, and mutations were most commonly detected in adult worms (0.3-2.0% of aligned sequence reads), followed by sporocysts (0.1-0.2%), and were extremely rare in eggs. The work provides hope that ongoing research will permit researchers to use refined methods to target the germ line to create stable knock-out or knock-in strains of any schistosome gene of interest.

Some comments and questions:

<u>Animal Procedures</u>. Why were snails "moved into dark cupboards at 28°C when they start shedding cercariae" and is there a reference showing that this is helpful? Page 4: How was the water "conditioned"? (At least in my version) the "Welfare assessments...." sentence needs correcting.

R. Regarding the question "why the snails were moved into dark cupboards at 28°C when they start shedding cercariae", there is not a reference showing this is helpful to obtain more cercariae. In order to experimentally obtain the highest number of cercariae, as far as we are aware, there are mainly two approaches used by most of the laboratories working with schistosomes. One approach involves keeping the patent snails under 12-hour light/dark cycles and 24 hours before shedding the snails are moved to the dark as described (PMID: 24510597). A second approach, that originally was followed in David Dunne laboratory at the Department of Pathology, University of Cambridge (personal

communication by Anna Protasio), and that we inherited involves moving the snails into dark incubators when they start to shed cercariae, and keeping them at dark all the time. Even though, as far as we are aware, no studies comparing the two approaches have been conducted, the number of cercariae obtained by the latter approach, currently used by us, has been high enough to carry out all the experiments described in the current manuscript, in addition to maintaining the whole life cycle independently at the Wellcome Sanger Institute and supporting the rest of the schistosoma-related projects.

R. In the Methods section "Animal Procedures" we have now included the reference for the conditioned water recipe (i.e. PMID: 24510597)

R. We have now edited the sentence '*Welfare assessments are carried out daily abnormal signs of behaviour or clinical signs of concern are reported*.' as follows: '*Welfare assessments are carried out daily, and abnormal signs of behaviour or clinical signs of concern are reported.*'

<u>Parasite Material</u>. Useful, detailed methods are reported. So, for consistency, give specifics of the Percoll-sucrose solution.

R. Following the reviewer's suggestion we have now included more details about the Percoll-sucrose solution (in Methods, "Parasite Material"), as follows: '*The filtrate was passed through a Percoll-sucrose gradient prepared by mixing 8 ml of Percoll with 32 ml of sterile-filtered 0.25M sucrose...*'

<u>CRISPR-Cas9 ribonucleoprotein complex assembly</u>. Since we expect the gRNA and the nuclease to associate in a 1:1 ratio, why not mix these in equimolar amounts? How can the authors confirm that the gRNA and nuclease actually formed the RNP complex? Is there a way to determine how efficiently this occurred? Low efficiency might provide some explanation for the relatively low level of mutation detected here?

R. Indeed, as the reviewer pointed out we did not use equimolar amounts of gRNA and nuclease. For the RNP complex assembly we followed the protocol suggested by the manufacturer IDT, slightly modified based on a well-optimised protocol to introduce CRISPR-Cas9 mutations by RNP complex in pluripotent stem cells (PMID: 30912046). In these protocols an excess of gRNA compared to Cas9 nuclease is employed, presumably to ensure that all the nuclease molecules are combined with the gRNA. Having said that, future experiments could be conducted to test the CRISPR efficiency by changing the relative concentrations of the RNP complex components. We have now provided in the manuscript further information about the RNP assembly protocol. In the Methods (section, "CRISPR-Cas 9 ribonucleoprotein complex assembly") we have included the following statement: '*The CRISPR-Cas9 ribonucleoprotein complex (RNP) was assembled in vitro following the manufacturer recommendations slightly modified based on* (Bruntraeger et al. 2019) *by combining...*'

R. We have not confirmed the actual RNP complex was properly assembled, but the presence of definitive CRISPR-Cas -induced deletions around the predicted double-stranded break site strongly suggests that the complex was, at least partially, active. Protocols to evaluate the assembly of Cas9/gRNA complexes by using a fluorometric molecular beacon-

like assay have been developed (PMID: 26945042), and could be applied in our future studies. We agree with the reviewer that low efficiency of assembly of RNP complexes could, at least partially, explain the low efficiency of the system to induce site-specific mutations. However, the CRISPR efficiency reported by Ittiprasert et al (PMID: 30644357), where RNP complex and lentivirus were (separately) employed to induce site-specific mutations, was even lower than that reported here. This suggests that the low efficiency may be related to the species rather than the approach employed to deliver the CRISPR-Cas cargo.

R. In the first paragraph of Discussion we have now included the following sentence to address this point raised by the reviewer: *'The low CRISPR-Cas efficiency in our study may be explained, at least partially, by low efficiency of RNP complex assembly; however, as discussed above, a low CRISPR-Cas efficiency was also described when lentivirus (rather than RNP complex) was used to deliver the CRISPR-Cas9 cargo into the parasite* (Ittiprasert et al. 2019)*'.*

Bioinformatic analysis. Consider adding a reference to Phred Quality Scoring so that interested readers can make sense of the <23 score cut off. It may also be useful to define here "SNP" and "NHEJ".

R. Edited as suggested.

Results. CRISPR-Cas9 machinery…..The authors should report if ALL of the adults (males and females) yielded a similar staining pattern. In Figure 2A and Movie 1, surface staining is not uniform along the length of the worm; did all males exhibit such staining in a similar region and to a similar extent (and in all replicates)? Does any of the RNP staining co-localize with DAPI-stained nuclei? If yes, point examples out. If not, how well is the nuclear localization signal working?

R. Following the reviewer's suggestion we have now included in Results (section "CRISPR-Cas9 machinery successfully delivered into schistosome developmental stages") the following statement: *'The staining pattern was similar in all observed specimens, and no evident differences in the staining pattern were evident between sexes.'* Regarding the absence of co-localisation of DAPI and ATTO^{TM} 550, please see the answer to question 2 of the previous reviewer (Arnon Jurberg). We have not evaluated how well the nuclear localisation signal worked in our system. However, the presence of definitive CRISPR-Cas -induced deletions around the predicted double-stranded breaking points strongly suggests the system worked.

Discussion. It is reported that "no evidence for CRISPR-Cas9-induced insertions or substitutions in the SULT-OR gene was observed." For context, it would be useful to know how this equates (if possible, under comparable experimental conditions) with work in *Strongyloides*, *C. elegans*, and mammalian cell-lines.

R. Following the reviewer's suggestion we have now included in the first paragraph of Discussion the following statement: ' *In Strongyloides spp, Gang et al (Gang et al. 2017) showed that in the absence of a repair template, small insertions or deletions (indels) or substitutions were not observed, but instead the authors found deletions of >500 bp at the target site in the unc-22 gene. On the other hand, in C. elegans, one study in the absence of a repair*

template, detected deletions ranging from 7 bp to >2 kb in the dpy-11 and unc-4 genes (Chiu et al. 2013); while another study found only small insertions and deletions of <20 bp (Friedland et al. 2013)'. In mammalian cell lines, the majority of indels are relatively small (1-50 bp), but larger deletions of kilobases in size are also sometimes observed (Kosicki et al, https://www.biorxiv.org/content/10.1101/2020.08.05.216739v1).'

I disagree that because mutations were detected in 2% of aligned reads, this necessarily means that ~2% of cells were impacted.

R. We agree that in order to claim that 2% of aligned reads means that ~2% of cells were impacted, we have to assume an even distribution of mutations, and this is not necessarily the case. In Discussion (4th paragraph) we have now edited the text accordingly: *'The large deletions spanning the predicted Cas9 cut site were found in 0.3–2.0% of aligned reads from CRISPR-treated adult worms, so our best estimate of the fraction of adult cells in which CRISPR worked is 0.3–2.0% assuming an even distribution of mutations across the transfected parasites. However, since a pool of five adult worms were transfected with the RNP complex, the efficiency of CRISPR (and the amount of knock-down at the mRNA level) may have varied between worms, as well as between cells of an individual worm.'*

The control data shown in e.g. Figure 3A – were they derived from the mock treated worms or from worms treated with Cas9 only or from those treated with gRNA only? Or were data from all controls combined?

R. For both experimental samples and controls in all the figures we provide the sample identifiers. The information for all these identifiers and samples can be found in Supplementary Table 1. In particular for Figure 3A the controls include worm-only controls for the adult samples, worm+Cas9-only and worm-only controls for the sporocyst samples, egg+Cas9-only and egg-only controls for the egg samples. To clarify this point, we have now added this information in the Figure 3A legend as follows: *'(A) Frequency of deletions in NGS sequencing data, identified with the assistance of CRISPResso in three biological replicates from adults, two from sporocysts, and three from eggs, as indicated (sample identifiers at the bottom). The controls include worm-only controls for the adult samples, worm treated with Cas9-only and worm-only controls for the sporocyst samples, eggs treated with Cas9-only and egg-only controls for the egg samples.'*

Report how the tree shown in Figure S1A was generated. What does "PRJEA36577" etc. refer to in the figure? What is the value of showing anything outside of the area bounded by the red dashed line?

R. To clarify this point we have now included two sentences in the figure legend describing how the tree shown in Figure S1A (now in Extended data V2) was generated and the meaning of identifiers next to the species names, as follows: *'The phylogenetic tree was generated by WormBase ParaSite using the EnsemblCompara pipeline [34]. Identifiers beside the species names in the tree indicate the NCBI BioProject accession numbers for the sequencing projects for each species or species strain.'*

R. The red dashed line shows a clade of the phylogenetic tree that includes paralogues of

SULT-OR that lie in a region about 3.2 Mb along *S. mansoni* chromosome 6. The rest of the phylogenetic tree, outside the red dashed line, includes other more distant paralogues of SULT-OR, that are scattered around the *S. mansoni* genome. We think it could be of interest to readers that, as well as several closely related paralogues of SULT-OR that are found nearby to SULT-OR in the genome, the SULT-OR gene also has more distantly related paralogues scattered around the genome (these are the ones outside the red dashed line).

Competing Interests: No competing interests were disclosed.

Reviewer Report 30 July 2020

https://doi.org/10.21956/wellcomeopenres.17586.r39710

Phillip LoVerde [iD]

Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

Schistosomiasis remains a serious global health problem. The toolbox available to scientists continues to improve. The ability to manipulate genes is a key technique to the continued success of the field to address questions of biology, host-parasite interactions, vaccines, drugs and diagnostics among others. The manuscript by the Wellcome Sanger Group is a welcome addition as it provides very detailed methods, state of the art approach and detailed analysis of the results. In fact, it provides a roadmap for others to follow. This is in spite of the fact they were not successful in producing a drug resistant transgenic strain. Clearly part of the reason was that only somatic cells were affected by the CRISPR-Cas 9 construct. They recognize that affecting germ line cells will likely make a difference. Another plausible explanation is that the cells in which CRISPR-Cas9 did induce mutations represented a small fraction of the cells expressing SmSULT-OR. They also acknowledge that the resistance trait is double recessive inferring that both alleles must be interrupted. I may have missed it but did not see where they attempted to demonstrate that both alleles of SmSULT-OR were interrupted.

The manuscript is well-written and follows a logical progression with well-designed experiments. Suggest use the term intramolluscan instead of intrasnail.
An issue out of their control is the difficulty in navigating the Extended data Figures and Tables. However, all the data is available to the reader.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Schistosomiasis host-parasite interactions involving molecular, immunological and genetic approaches. Drug Discovery.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 07 Dec 2020
Gabriel Rinaldi, Wellcome Sanger Institute, Hinxton, UK

07 December 2020,
Wellcome Open Research,

On behalf of all the authors, we acknowledge and appreciate the efforts of the reviewers with the review of our manuscript. We have been able to revise along the lines of the recommendations, which has certainly contributed to improve our paper. Point-by-point responses are provided below: **Response (R)**

30 Jul 2020 | for Version 1
Phillip LoVerde, Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

Schistosomiasis remains a serious global health problem. The toolbox available to scientists continues to improve. The ability to manipulate genes is a key technique to the continued success of the field to address questions of biology, host-parasite interactions, vaccines, drugs and diagnostics among others. The manuscript by the Wellcome Sanger Group is a welcome addition as it provides very detailed methods, state of the art approach and detailed analysis of the results. In fact, it provides a roadmap for others to follow. This is in spite of the fact they were not successful in producing a drug resistant transgenic strain. Clearly part of the reason was that only somatic cells were affected by the CRISPR-Cas 9 construct. They recognize that affecting germ line cells will likely make a difference. Another

plausible explanation is that the cells in which CRISPR-Cas9 did induce mutations represented a small fraction of the cells expressing SmSULT-OR. They also acknowledge that the resistance trait is double recessive inferring that both alleles must be interrupted. I may have missed it but did not see where they attempted to demonstrate that both alleles of SmSULT-OR were interrupted.

R. We acknowledge the reviewer for the positive comments, and as indicated, the resistance trait is double recessive inferring that both alleles must be interrupted in the germline to generate a transgenic line with a fully OXA-resistant phenotype. In this study, we did not demonstrate that both alleles of SmSULT-OR were interrupted, given that the sequence data were generated from a large number of cells, probably including a mixture of wild type and mutant cells. In order to identify allele-specific mutations single cell DNA PCR and/ or sequencing approaches need to be applied (PMID: 31827197; PMID: 30992375). This can be optimised in the future, in particular after having shown that single cell sequencing approaches are feasible in schistosomes (PMID: 32973030; bioRxiv 754713; doi: https://doi.org/10.1101/754713)

The manuscript is well-written and follows a logical progression with well-designed experiments. Suggest use the term intramolluscan instead of intrasnail.

R. Edited as recommended

An issue out of their control is the difficulty in navigating the Extended data Figures and Tables. However, all the data is available to the reader.

R. We agree and have also found difficulties in navigating the Extended Data files. We will raise a comment to the editorial team in this regard.

04 Aug 2020 | for Version 1
Arnon Jurberg, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

The work by Sankaranarayanan & Coghlan et al. sought to deploy CRISPR/Cas9 by electroporation at distinct developmental stages of schistosomes (more specifically, eggs, sporocysts and adult worms). A successful implementation of this technique in schistosomes will open up numerous possibilities for the study of gene function in these flatworms and is likely to revolutionize the field in the same manner as gene knockout by homologous recombination in mice and rats.

The paper is well written and provided a detailed description of their methods and findings. The figures are well conceived and the paper is likely to provide grounds for others to come, although CRISPR efficiency rate was below 5%.

1. Considering the low editing efficiency, did the authors evaluate the activity of other gRNAs against SULT-OR?

R. This is a good question, and indeed we have originally designed three gRNAs in total, but

have only tested one of them so far. The gRNA employed in this study is the closest to the start codon of the exon 1 (Figure 1B), and predicted to be SmSULT-OR specific; therefore, reducing the chances of off-target mutations in the other members of the sulfotransferase family (as shown in Figure S3 A). Additionally, it is expected that an indel mutation (that is not a multiple of 3 bp in size) at the 5' end of the gene could disrupt the reading frame of nearly the whole gene and so nearly the whole protein, while such a mutation near the 3' end of the gene would only affect the sequence of the end of the protein.

2. It also caught my attention the apparent lack of nuclear staining for fluorescently labelled Cas9-gRNA. Did the authors address this?

R. We agree with the reviewer that nuclear staining, i.e. co-staining of the nuclei with DAPI and ATTO^{TM} 550 is not evident. Probably the confocal imaging is not sensitive enough to detect the fluorescent ribonucleoprotein complex within the cell nucleus. However, the Cas9 nuclease contains a nuclear localization sequence, and we identified site-specific deletions across the expected double-strand break site in the three developmental stages tested herein. This strongly suggests that even though no nuclear signal was detected (as above, probably due to sensitivity of the confocal assay), the ribonucleoprotein complex did indeed reach the genomic DNA of the cells. In this regard, we have now incorporated the following statement in Discussion (4th paragraph):" *The absence of nuclear localisation of the RNP complex in the confocal images may indicate that the 4-hour timeframe between the molecule delivery and parasite fixation was too short, and which may explain, at least partially, the low CRISPR-induced deletion efficiency."*

3. It was also unclear to me whether the staining pattern between adult males and females were similar, and I would like to suggest the addition of further details. An alternative approach that can assist in achieving greater CRISPR efficiencies in schistosomes is the use of microinjection (perhaps in the ovary of females).

R. No evident differences were observed in the staining pattern between adult males and females. Accordingly, we have incorporated the following sentence in the Results (section "CRISPR-Cas9 machinery successfully delivered into schistosome developmental stages"): *"...no evident differences in the staining pattern were evident between sexes".* Regarding the use of microinjection to introduce the CRISPR machinery into the germ line, it is a very interesting idea worth testing in future experiments.

4. Once mutations in the target gene are predicted to induce oxamniquine resistance, did the authors evaluate whether incubation of electroporated parasites with this drug could improve CRISPR efficiency?

R. We acknowledge this suggestion, and indeed in one experiment we incubated CRISPR-mutated worms in the presence of oxamniquine (OXA), but no obvious resistant phenotype was detected. These findings were not unexpected, given the intrinsic low efficiency of CRISPR and that only a few (presumably somatic) cells, were mutated in the whole parasite. In the current study, we decided to focus on describing the induced deletions as a *proof-of-principle* that the technology can be applied to several developmental stages, in this case as a "somatic transgenesis" approach. Further experiments will be performed in the future to

investigate the development of an OXA-resistance phenotype.

24 Aug 2020 | for Version 1
Mattie Pawlowic, Biological Chemistry and Drug Discovery, Wellcome Centre for Anti-Infectives Research, University of Dundee, Dundee, UK

Sankaranarayanan and colleagues report their work to mutate *Schitosoma mansoni* SULT-OR, a gene related to oxamniquine resistance, using transfection of ribonucleoprotein (RNP) complex to induce CRISPR/Cas9 double-stranded breaks. They use this approach on worm adults, sporocysts, and eggs, and deep sequence the resulting parasites to determine which mutations are present. This represents a new approach for genetic modification of Schistosoma.

The authors find that adult and sporocysts (not eggs) take up fluorescently labelled RNPs. These bind non-specifically on the surface of adult worms and also concentrate in the gut. In the future, the authors should visualise what uptake may occur in adult worms without electroporation, as it appears a significant amount of uptake may be due to feeding. It is curious that the Cas9, which contains a nuclear localisation sequence, does not localise the RNP complex to the nucleus of worm cells. It could be that the 4-hour time frame is too short to observe this, or it could be that lack of nuclear localisation may explain the low editing efficiency.

R. Although the fluorescence signal is faint in eggs, in particular compared to adult worms and sporocysts, Figures 2F and H show it is inside the eggshell and in the miracidium surface. Given the high concentration of RNP in cells lining the gut, as suggested by the reviewer, in future experiments we may compare the RNP delivery by soaking and electroporation. Regarding the absence of co-localisation of DAPI and ATTO^{TM} 550, please see the answer to question 2 of the previous reviewer. The 4-hour timeframe being too short is a good hypothesis and worth testing it in future experiments. In this regard, we added the following statement in the 4th paragraph of Discussion: *"Future experiments comparing the delivery of the RNP complex by soaking versus electroporation would inform about the best delivery approach. The absence of nuclear localisation of the RNP complex in the confocal images may indicate that the 4-hour timeframe between the molecule delivery and parasite fixation was too short, and which may explain, at least partially, the low CRISPR-induced deletion efficiency."*

After 4 days of culture, the authors extracted DNA from transfected parasites, PCR amplified the region of interest, and deep sequenced to identify mutations. I understand that a larger PCR fragment is required to identify deletions, however the authors designed the ends of the PCR products to overlap at the Cas9 cut site. This led to difficulties in identifying true mutations vs sequencing errors at the cut site. Therefore, I think indels may be under-reported as they were not well captured.

R. In Figure 1C, we show the positions of the PCR primers in green and the amplicon in pale blue. The amplicon length is 262bp and it covers the predicted Cas9 site (position 137), the PCR products do not overlap at the Cas9 cut site. On the other hand, the forward and

reverse sequence reads do overlap the cut site, increasing the accuracy to detect CRISPR-induced mutations. To make the Figure clearer, we have now incorporated more information in the Figure 1C legend as follows: *"(C) Reference PCR amplicon (pale blue), showing the positions of the gRNA, PAM, DSB, forward and reverse PCR primers (green), and forward and reverse sequence reads (orange), as well as a SNP site found in many sequence reads. The diagrams are drawn to scale."*

The authors find that deletions are the most common mutations the occur at the SULT-OR locus. I disagree that deletions in the range of 24-102 bp should be called "large". Consistent with use of CRISPR in other systems, deletions occur upstream of the cut site. Although adult worms took up RNPs the best, the mutation rate was found to be only <2%. This is likely due to the RNPs not transfecting every cell in the adult worm.

R. As such, the term "large" is relative, and we agree that deletions in the range of 24-102 are not large compared to deletions of hundreds and even thousands of bases. However, given that in the earlier CRISPR studies the majority of identified deletions were in the range of 1-3bp around the double-stranded break site (see https://www.biorxiv.org/content/10.1101/2020.08.05.216739v1), and our study is the second in schistosomes, and the first one in this parasite showing deletions up to 102bp, we have decided to keep the term "large" to describe them. Finally, even longer deletions (~ 500bp) as shown for *Strongyloides* spp (PMID: 29016680) but not detectable by our PCR-based approach, cannot be ruled out in our system.

Despite SULT-OR being at its highest expression levels in adult worms, mutation rates were too low to see an impact on SULT-OR expression levels.
Overall this work reports new methods for genetic modification of *Schistosoma mansoni*. While the efficiency is low, further optimisation makes this approach promising.

R. We completely agree with the reviewer, and further optimisation, not only to increase the CRISPR-induced mutations, but also to introduce them into the germline, is planned.

01 Sep 2020 | for Version 1
Patrick Skelly, Molecular Helminthology Laboratory, Department of Infectious Disease and Global Health, Cummings School of Veterinary Medicine, Tufts University, North Grafton, MA, USA

This study by Sankaranarayanan & Coghlan *et al.* employed CRISPR-Cas9 to target for disruption the SULT-OR sulfotransferase gene in *Schistosoma mansoni* using electroporation to deliver a ribonucleoprotein (RNP) complex consisting of gRNA, tracrRNA and Cas9 protein. Mutations induced by this treatment were subsequently detected by high-throughput sequencing. The results provide independent confirmation that CRISPR/Cas delivery to *S. mansoni* can indeed induce mutations in these parasites, albeit (so far) at a low level. Three developmental stages were examined, and mutations were most commonly detected in adult worms (0.3-2.0% of aligned sequence reads), followed by sporocysts (0.1-0.2%), and were extremely rare in eggs. The work provides hope that ongoing research will permit researchers to use refined methods to target the germ line to create stable knock-

out or knock-in strains of any schistosome gene of interest.

Some comments and questions:

Animal Procedures. Why were snails "moved into dark cupboards at 28°C when they start shedding cercariae" and is there a reference showing that this is helpful? Page 4: How was the water "conditioned"? (At least in my version) the "Welfare assessments...." sentence needs correcting.

R. Regarding the question "why the snails were moved into dark cupboards at 28°C when they start shedding cercariae", there is not a reference showing this is helpful to obtain more cercariae. In order to experimentally obtain the highest number of cercariae, as far as we are aware, there are mainly two approaches used by most of the laboratories working with schistosomes. One approach involves keeping the patent snails under 12-hour light/dark cycles and 24 hours before shedding the snails are moved to the dark as described (PMID: 24510597). A second approach, that originally was followed in David Dunne laboratory at the Department of Pathology, University of Cambridge (personal communication by Anna Protasio), and that we inherited involves moving the snails into dark incubators when they start to shed cercariae, and keeping them at dark all the time. Even though, as far as we are aware, no studies comparing the two approaches have been conducted, the number of cercariae obtained by the latter approach, currently used by us, has been high enough to carry out all the experiments described in the current manuscript, in addition to maintaining the whole life cycle independently at the Wellcome Sanger Institute and supporting the rest of the schistosoma-related projects.

R. In the Methods section "Animal Procedures" we have now included the reference for the conditioned water recipe (i.e. PMID: 24510597)

R. We have now edited the sentence '*Welfare assessments are carried out daily abnormal signs of behaviour or clinical signs of concern are reported*.' as follows: '*Welfare assessments are carried out daily, and abnormal signs of behaviour or clinical signs of concern are reported.*'

Parasite Material. Useful, detailed methods are reported. So, for consistency, give specifics of the Percoll-sucrose solution.

R. Following the reviewer's suggestion we have now included more details about the Percoll-sucrose solution (in Methods, "Parasite Material"), as follows: *'The filtrate was passed through a Percoll-sucrose gradient prepared by mixing 8 ml of Percoll with 32 ml of sterile-filtered 0.25M sucrose...'*

CRISPR-Cas9 ribonucleoprotein complex assembly. Since we expect the gRNA and the nuclease to associate in a 1:1 ratio, why not mix these in equimolar amounts? How can the authors confirm that the gRNA and nuclease actually formed the RNP complex? Is there a way to determine how efficiently this occurred? Low efficiency might provide some explanation for the relatively low level of mutation detected here?

R. Indeed, as the reviewer pointed out we did not use equimolar amounts of gRNA and

nuclease. For the RNP complex assembly we followed the protocol suggested by the manufacturer IDT, slightly modified based on a well-optimised protocol to introduce CRISPR-Cas9 mutations by RNP complex in pluripotent stem cells (PMID: 30912046). In these protocols an excess of gRNA compared to Cas9 nuclease is employed, presumably to ensure that all the nuclease molecules are combined with the gRNA. Having said that, future experiments could be conducted to test the CRISPR efficiency by changing the relative concentrations of the RNP complex components. We have now provided in the manuscript further information about the RNP assembly protocol. In the Methods (section, "CRISPR-Cas 9 ribonucleoprotein complex assembly") we have included the following statement: '*The CRISPR-Cas9 ribonucleoprotein complex (RNP) was assembled in vitro following the manufacturer recommendations slightly modified based on* *(Bruntraeger et al. 2019)* *by combining...'*

R. We have not confirmed the actual RNP complex was properly assembled, but the presence of definitive CRISPR-Cas -induced deletions around the predicted double-stranded break site strongly suggests that the complex was, at least partially, active. Protocols to evaluate the assembly of Cas9/gRNA complexes by using a fluorometric molecular beacon-like assay have been developed (PMID: 26945042), and could be applied in our future studies. We agree with the reviewer that low efficiency of assembly of RNP complexes could, at least partially, explain the low efficiency of the system to induce site-specific mutations. However, the CRISPR efficiency reported by Ittiprasert et al (PMID: 30644357), where RNP complex and lentivirus were (separately) employed to induce site-specific mutations, was even lower than that reported here. This suggests that the low efficiency may be related to the species rather than the approach employed to deliver the CRISPR-Cas cargo.

R. In the first paragraph of Discussion we have now included the following sentence to address this point raised by the reviewer: '*The low CRISPR-Cas efficiency in our study may be explained, at least partially, by low efficiency of RNP complex assembly; however, as discussed above, a low CRISPR-Cas efficiency was also described when lentivirus (rather than RNP complex) was used to deliver the CRISPR-Cas9 cargo into the parasite* *(Ittiprasert et al. 2019)'*.

Bioinformatic analysis. Consider adding a reference to Phred Quality Scoring so that interested readers can make sense of the <23 score cut off. It may also be useful to define here "SNP" and "NHEJ".

R. Edited as suggested.

Results. CRISPR-Cas9 machinery.....The authors should report if ALL of the adults (males and females) yielded a similar staining pattern. In Figure 2A and Movie 1, surface staining is not uniform along the length of the worm; did all males exhibit such staining in a similar region and to a similar extent (and in all replicates)? Does any of the RNP staining co-localize with DAPI-stained nuclei? If yes, point examples out. If not, how well is the nuclear localization signal working?

R. Following the reviewer's suggestion we have now included in Results (section "CRISPR-Cas9 machinery successfully delivered into schistosome developmental stages") the following statement: '*The staining pattern was similar in all observed specimens, and no evident*

differences in the staining pattern were evident between sexes.' Regarding the absence of co-localisation of DAPI and ATTO^{TM} 550, please see the answer to question 2 of the previous reviewer (Arnon Jurberg). We have not evaluated how well the nuclear localisation signal worked in our system. However, the presence of definitive CRISPR-Cas -induced deletions around the predicted double-stranded breaking points strongly suggests the system worked.

Discussion. It is reported that "no evidence for CRISPR-Cas9-induced insertions or substitutions in the SULT-OR gene was observed." For context, it would be useful to know how this equates (if possible, under comparable experimental conditions) with work in *Strongyloides*, *C. elegans*, and mammalian cell-lines.

R. Following the reviewer's suggestion we have now included in the first paragraph of Discussion the following statement: '*In Strongyloides spp, Gang et al (Gang et al. 2017) showed that in the absence of a repair template, small insertions or deletions (indels) or substitutions were not observed, but instead the authors found deletions of >500 bp at the target site in the unc-22 gene. On the other hand, in C. elegans, one study in the absence of a repair template, detected deletions ranging from 7 bp to >2 kb in the dpy-11 and unc-4 genes (Chiu et al. 2013); while another study found only small insertions and deletions of <20 bp (Friedland et al. 2013)'. In mammalian cell lines, the majority of indels are relatively small (1-50 bp), but larger deletions of kilobases in size are also sometimes observed (Kosicki et al, https://www.biorxiv.org/content/10.1101/2020.08.05.216739v1).'*

I disagree that because mutations were detected in 2% of aligned reads, this necessarily means that ~2% of cells were impacted.

R. We agree that in order to claim that 2% of aligned reads means that ~2% of cells were impacted, we have to assume an even distribution of mutations, and this is not necessarily the case. In Discussion (4th paragraph) we have now edited the text accordingly: *'The large deletions spanning the predicted Cas9 cut site were found in 0.3–2.0% of aligned reads from CRISPR-treated adult worms, so our best estimate of the fraction of adult cells in which CRISPR worked is 0.3–2.0% assuming an even distribution of mutations across the transfected parasites. However, since a pool of five adult worms were transfected with the RNP complex, the efficiency of CRISPR (and the amount of knock-down at the mRNA level) may have varied between worms, as well as between cells of an individual worm.'*

The control data shown in e.g. Figure 3A – were they derived from the mock treated worms or from worms treated with Cas9 only or from those treated with gRNA only? Or were data from all controls combined?

R. For both experimental samples and controls in all the figures we provide the sample identifiers. The information for all these identifiers and samples can be found in Supplementary Table 1. In particular for Figure 3A the controls include worm-only controls for the adult samples, worm+Cas9-only and worm-only controls for the sporocyst samples, egg+Cas9-only and egg-only controls for the egg samples. To clarify this point, we have now added this information in the Figure 3A legend as follows: *'(A) Frequency of deletions in NGS sequencing data, identified with the assistance of CRISPResso in three biological replicates from*

adults, two from sporocysts, and three from eggs, as indicated (sample identifiers at the bottom). The controls include worm-only controls for the adult samples, worm treated with Cas9-only and worm-only controls for the sporocyst samples, eggs treated with Cas9-only and egg-only controls for the egg samples.'

Report how the tree shown in Figure S1A was generated. What does "PRJEA36577" etc. refer to in the figure? What is the value of showing anything outside of the area bounded by the red dashed line?

R. To clarify this point we have now included two sentences in the figure legend describing how the tree shown in Figure S1A (now in Extended data V2) was generated and the meaning of identifiers next to the species names, as follows: *'The phylogenetic tree was generated by WormBase ParaSite using the EnsemblCompara pipeline [34]. Identifiers beside the species names in the tree indicate the NCBI BioProject accession numbers for the sequencing projects for each species or species strain.'*

R. The red dashed line shows a clade of the phylogenetic tree that includes paralogues of SULT-OR that lie in a region about 3.2 Mb along *S. mansoni* chromosome 6. The rest of the phylogenetic tree, outside the red dashed line, includes other more distant paralogues of SULT-OR, that are scattered around the *S. mansoni* genome. We think it could be of interest to readers that, as well as several closely related paralogues of SULT-OR that are found nearby to SULT-OR in the genome, the SULT-OR gene also has more distantly related paralogues scattered around the genome (these are the ones outside the red dashed line).

Competing Interests: No competing interests were disclosed.