

# SCIENTIFIC REPORTS



OPEN

## Inducible high-efficiency CRISPR-Cas9-targeted gene editing and precision base editing in African trypanosomes

Eva Rico, Laura Jeacock, Julie Kovářová & David Horn 

The Cas9 endonuclease can be programmed by guide RNA to introduce sequence-specific breaks in genomic DNA. Thus, Cas9-based approaches present a range of novel options for genome manipulation and precision editing. African trypanosomes are parasites that cause lethal human and animal diseases. They also serve as models for studies on eukaryotic biology, including 'divergent' biology. Genome modification, exploiting the native homologous recombination machinery, has been important for studies on trypanosomes but often requires multiple rounds of transfection using selectable markers that integrate at low efficiency. We report a system for delivering tetracycline inducible Cas9 and guide RNA to *Trypanosoma brucei*. In these cells, targeted DNA cleavage and gene disruption can be achieved at close to 100% efficiency without further selection. Disruption of aquaglyceroporin (*AQP2*) or amino acid transporter genes confers resistance to the clinical drugs pentamidine or eflornithine, respectively, providing simple and robust assays for editing efficiency. We also use the new system for homology-directed, precision base editing; a single-stranded oligodeoxyribonucleotide repair template was delivered to introduce a single *AQP2* - T<sup>791</sup>G/L<sup>264</sup>R mutation in this case. The technology we describe now enables a range of novel programmed genome-editing approaches in *T. brucei* that would benefit from temporal control, high-efficiency and precision.

The African trypanosomes of the *Trypanosoma brucei* group are transmitted by tsetse flies and cause lethal diseases in humans and livestock, known as sleeping sickness and nagana, respectively. Research on these protozoan parasites, and related parasitic trypanosomatids, often focuses on virulence mechanisms and druggable biology<sup>1</sup>. However, their divergent position within the eukaryotic phylogeny, and the experimental tractability of *T. brucei* in particular, means they are also of interest as model Excavates for studies on various aspects of eukaryotic biology, including novel biology not seen in other eukaryotic model systems. Unusual features include an almost complete lack of introns, widespread polycistronic transcription and *trans*-splicing of every mRNA<sup>2</sup>.

The ability to manipulate parasite genomes has had a huge impact on this field of research<sup>3</sup>. In *T. brucei*, DNA constructs, typically introduced by electroporation, can be used for gene knockout, gene knockdown exploiting the native RNA interference (RNAi) machinery, tagging native alleles or inducible expression of recombinant products; robust inducible expression<sup>4</sup> has been of particular value. Sequences towards the ends of the linearised DNA constructs serve as templates for homologous recombination reactions. Thus, targeted integration depends upon the native homologous recombination machinery; microhomology-mediated end-joining (MMEJ)<sup>5</sup> and single-strand annealing<sup>6</sup> also operate in *T. brucei*, while non-homologous end-joining appears to be absent. Targeted integration occurs at low frequency, however (typically  $<1 \times 10^{-5}$  electroporated cells), meaning that antibiotic selectable marker genes must be integrated at a modified site followed by selection for the desired recombinant cells. Since the *T. brucei* core genome is diploid<sup>7</sup>, gene disruption approaches typically require two rounds of electroporation and selection.

Chromosomal site-specific DNA double strand breaks (DSBs) can increase recombination efficiency in *T. brucei*, as demonstrated following inducible expression of the I-SceI meganuclease<sup>8</sup>. The same endonuclease can efficiently trigger inter-chromosomal homologous recombination<sup>5</sup> or antigenic variation when breaks are targeted

Wellcome Trust Centre for Anti-Infectives Research, School of Life Sciences, University of Dundee, Dow Street, Dundee, DD1 5EH, UK. Correspondence and requests for materials should be addressed to D.H. (email: [d.horn@dundee.ac.uk](mailto:d.horn@dundee.ac.uk))

close to the expressed Variant Surface Glycoprotein gene<sup>9,10</sup>. For this approach, however, a specific 18-bp I-SceI cleavage site must first be engineered at the desired target site, which limits its versatility. Cas9 is an RNA-guided DNA endonuclease that can be programmed in a sequence-specific manner using a short (17–20 b) synthetic guide RNA sequence; specific targeting involves RNA:DNA base-pairing. The only additional requirement is that this 'protospacer' sequence must immediately precede a protospacer adjacent motif (PAM; at the DNA target site but not to be included in the gRNA). The PAM recognized by the most widely used Cas9 is 5'-NGG, for the type II system from *Streptococcus pyogenes*<sup>11</sup>. It is this tremendous versatility that has made Cas9 repurposing so popular for genome manipulation and editing<sup>12</sup>. Cas9 systems function naturally as bacterial adaptive immune systems. In the bacterial cells, Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), derived from invading phage DNA, produce CRISPR RNA (crRNA) which, together with a *trans*-activating crRNA (tracrRNA) program Cas9 (CRISPR associated protein 9) to cut both the RNA-complementary and non-complementary strands, producing DSBs in the target DNA, 3 bp upstream of the PAM. The crRNA and tracrRNA can be fused to generate a chimeric single guide RNA (sgRNA)<sup>11</sup>, such that a two-component system requires only Cas9 and an sgRNA for programmed editing in a heterologous system.

The application of Cas9-based genome editing to trypanosomatids has progressed rapidly in the last few years. For example, SpCas9 has been used to edit genes in *T. brucei* with a cloning-free approach, allowing for one-step gene knockout, but still requiring integration of antibiotic selectable markers at the target sites<sup>13</sup>. Another recent report describes the delivery of *in vitro* assembled ribonucleoproteins containing the smaller *Staphylococcus aureus* Cas9 for editing in trypanosomatid parasites; *T. brucei*, *Trypanosoma cruzi* and *Leishmania major*. In this case, a GFP-reporter was disrupted in >50% of targeted *T. brucei* cells<sup>14</sup>; the PAM recognized by SaCas9 is 5'-NNGRRT, where R represents A or G. SpCas9 has also been used to edit genes in the other trypanosomatids<sup>15</sup>, including those for which loss-of-function studies have previously been hampered, notably due to the absence of a native RNAi machinery<sup>16</sup>. Thus, these advances will have a major impact on tractability in terms of genetic manipulation in *T. cruzi*<sup>17–19</sup>, and *Leishmania spp*<sup>13,20–23</sup>.

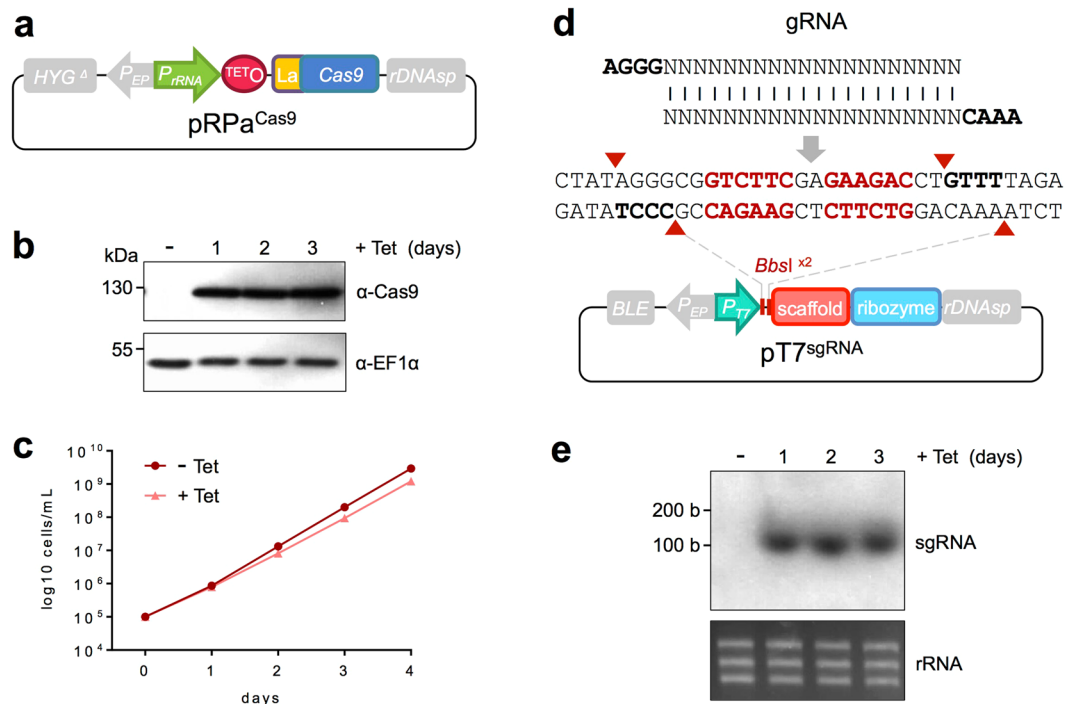
Cas9-based approaches will also enable a range of new approaches in *T. brucei*. To date, however, an inducible Cas9 editing system has not yet been reported for any trypanosomatid. In *T. brucei*, double-allele editing without selectable markers has not yet been reported, double-allele editing efficiencies have not been determined and precision base editing is not readily achievable. We sought to address three specific issues to advance Cas9-based editing tools and approaches for *T. brucei*. Since current systems in protozoan parasites often use constitutive Cas9 expression, which can be toxic and continually risks off-target editing<sup>18,24</sup>, we first wanted to establish an inducible system. Second, we wanted to develop a system that allowed for DSB formation (and repair) in close to 100% of cells in the population in a week or less. Third, we wanted to use the system for template-directed precision base editing. By manipulating genes associated with resistance to clinical drugs, we report the successful attainment of all three goals.

## Results

**A system for inducible Cas9 expression and sgRNA expression in *T. brucei*.** We assembled an expression construct with a tetracycline (Tet)-inducible ribosomal RNA (*rRNA*) promoter to drive SpCas9 expression in *T. brucei*. These pRPa constructs can be used in combination with *T. brucei* bloodstream-form 2T1 strains<sup>25</sup>, which express the Tet-repressor and have a partial hygromycin (*HYG*) drug-selectable marker gene at a non-transcribed *rDNA* spacer locus on chromosome 2; the expression construct contains a complementing portion of the *HYG* gene<sup>26</sup>. This allows for selection of hygromycin-resistant clones with expression cassettes reproducibly integrated at the same locus, a locus that supports robust and uniform Tet-inducible expression<sup>26</sup>. A synthetic SpCas9 gene with a La nuclear localisation signal<sup>27</sup> at the N-terminus was codon-optimised for expression in *T. brucei*<sup>28</sup>. This gene was cloned in the pRPa expression construct and the resulting pRPa<sup>Cas9</sup> plasmid (Fig. 1a) was linearised and introduced into 2T1<sup>T7</sup> cells; these cells also express the phage T7 polymerase to drive sgRNA expression (see below). Analysis of the resulting 2T1<sup>T7-Cas9</sup> clones revealed tightly regulated and robust Cas9 expression (Fig. 1b, Supplementary Fig. 1). We observed a moderate reduction in growth rate associated with Cas9 induction (Fig. 1c), which could reflect off-target editing in the absence of sgRNA<sup>24</sup>.

For expression of the sgRNA, a phage T7 promoter, a cloning site, the sgRNA scaffold and a hepatitis delta virus (HDV) ribozyme<sup>29</sup> were synthesised in tandem (Fig. 1d). The cloning site comprises dual *BbsI* restriction enzyme cleavage sites, which allows for high-efficiency cloning of annealed oligonucleotide pairs (AGGG[N]<sub>20</sub>/AAAC[N]<sub>20</sub>, Fig. 1d) without introducing additional unnecessary flanking sequences. The HDV ribozyme is a self-cleaving RNA that liberates the upstream RNA. Thus, the ribozyme generates the 3'-end of the sgRNA. This synthetic cassette was used to generate pT7<sup>sgRNA</sup> that integrates at a second non-transcribed *rDNA* spacer locus (Fig. 1d); there are estimated to be 10–20 *rDNA* spacer loci in the *T. brucei* genome<sup>7</sup>.

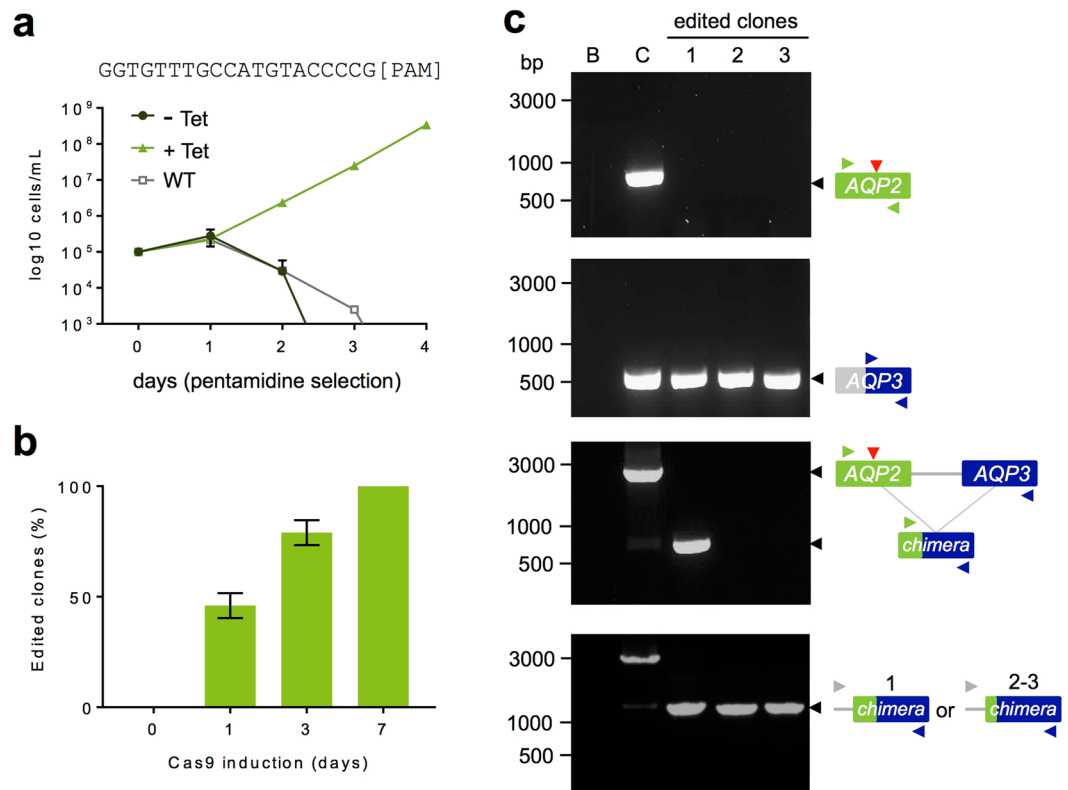
We selected the aquaglyceroporin gene, *AQP2*, (Tb927.10.14170) as a suitable target for editing since this gene is involved in drug resistance and parasites are resistant to pentamidine when *AQP2* is defective<sup>30</sup>. The full complement of >9,000 *T. brucei* protein-coding genes comprises 51% G + C content<sup>7</sup> so suitable PAMs can be found approximately once every 8 bp in these sequences. We identified potential PAMs in *AQP2*, selected a suitable 20 b *AQP2*-specific gRNA sequence, annealed the cognate oligonucleotide pair and ligated the product to *BbsI*-digested pT7<sup>sgRNA</sup>. The resulting construct was linearised and introduced into 2T1<sup>T7-Cas9</sup> cells. We analysed two independent Cas9/sgRNA<sup>AQP2</sup> clones and observed robust and reproducible expression of sgRNA of the expected size (96 b), which was dependent upon Cas9 expression (Fig. 1e, Supplementary Fig. 1). This suggests that sgRNA is unstable in the absence of Cas9. Thus, sgRNA is efficiently processed by the HDV ribozyme and is specifically stabilised by Cas9, providing evidence for physical interaction of these two components and assembly of Cas9 ribonucleoproteins in *T. brucei*.



**Figure 1.** An inducible Cas9 system. **(a)** Schematic map of the pRPa<sup>Cas9</sup> construct. P<sub>rRNA</sub>, rRNA promoter; TET<sub>O</sub>, tetracycline operator; La, nuclear localisation signal. Other components are from the pRPa parent vector (see the text for more details). **(b)** The protein blots show tightly regulated and inducible Cas9 expression (predicted mass: 159.4 kDa). The EF1α panel provides a loading control. Multiple independent clones yielded similar results. **(c)** Cell growth with (+Tet) and without (-Tet) Cas9 induction. Error bars (SD) from triplicate counts are obscured by the data symbols. A second independent clone gave similar results. **(d)** Schematic map of the pT7<sup>sgRNA</sup> construct showing the gRNA cloning strategy using tandem BbsI sites. The sgRNA scaffold and HDV ribozyme are also shown. P<sub>T7</sub>, phage T7 promoter. Other components are from the pRP<sup>c6MYCn</sup> parent vector. **(e)** The RNA blot shows expression of the predicted 96 b sgRNA<sup>AQP2</sup> that is dependent upon Cas9 expression. The rRNA panel provides a loading control. A second independent clone yielded similar results.

**Tightly regulated and high-efficiency AQP2 editing.** The *T. brucei* genome is diploid, meaning that Cas9-based editing for the vast majority of genes will involve targeting two identical sequences, ideally simultaneously in many cases. When AQP2 function is lost, the effective concentration of pentamidine to reduce growth by 50%, or the EC<sub>50</sub>, is approximately 15-fold greater than for wild-type cells<sup>30</sup>. Thus, pentamidine can be used to provide a robust phenotypic read-out and for quantitative analysis of bi-allelic editing. Using three independent Cas9/sgRNA<sup>AQP2</sup> clones, we pre-induced Cas9 programmed to target AQP2 for three days, and then selected with pentamidine. The growth profiles indicated an absence of drug-resistance without induction and robust drug-resistance with induction (Fig. 2a).

Encouraged by these results, using two independent Cas9/sgRNA<sup>AQP2</sup> clones, we induced editing in the absence of pentamidine selection and sub-cloned the population either prior to induction, or one, three and seven days later. Under these conditions, un-edited, drug-sensitive cells remain viable and edited cells have no selective advantage. Twelve sub-clones from each clone and at each time-point were expanded under conditions where Cas9 was no longer induced and then tested for pentamidine resistance. This revealed remarkably tightly regulated editing (Fig. 2b). The absence of pentamidine-resistant sub-clones (0/24) from the uninduced population provided strong evidence for efficient temporal control, and this 'baseline' data allowed us to accurately assess editing during the induction time-course. This revealed efficiently induced editing, which increased from ~45% resistant clones after 1-day to ~80% after 3-days and reached the maximum 100% (24/24 sub-clones) after 7-days (Fig. 2b). This temporal response reflects the kinetics of Cas9 induction<sup>4</sup>, the Cas9 target search<sup>31</sup> and DNA repair in *T. brucei*<sup>32</sup>, and may also be impacted by multiple rounds of single-allele break and error-free repair by homologous recombination. A series of PCR-based assays were used to assess the AQP2-3 locus in three independent resistant clones (Fig. 2c, Supplementary Fig. 1). AQP2 had indeed been disrupted (Fig. 2c, top panel) and replaced by AQP2-3 chimeras in all three clones (Fig. 2c, lower panels), similar to what has been seen previously in drug-resistant isolates from relapsed patients<sup>33,34</sup>. Cloning and sequencing of the products revealed recombination within a block of identity from 168–222 nt in AQP2 in clone 1, meaning that from 223 nt into the CDS is AQP3 sequence in the chimera, and recombination within a block of identity from 1–91 nt in AQP2 in clones 2–3, meaning that from 92 nt into the CDS is AQP3 sequence in these chimeras. The chimeras likely arise by single-strand annealing, whereby single-stranded regions extend to adjacent repeated sequences and complementary strands anneal, leading to repair and the loss of one repeat<sup>35</sup>; the PCR-assays indicate loss of ~1.8 kbp in each case (Fig. 2c, lower panels), which is the size of the semi-repetitive AQP2-3 region. Microhomology-mediated

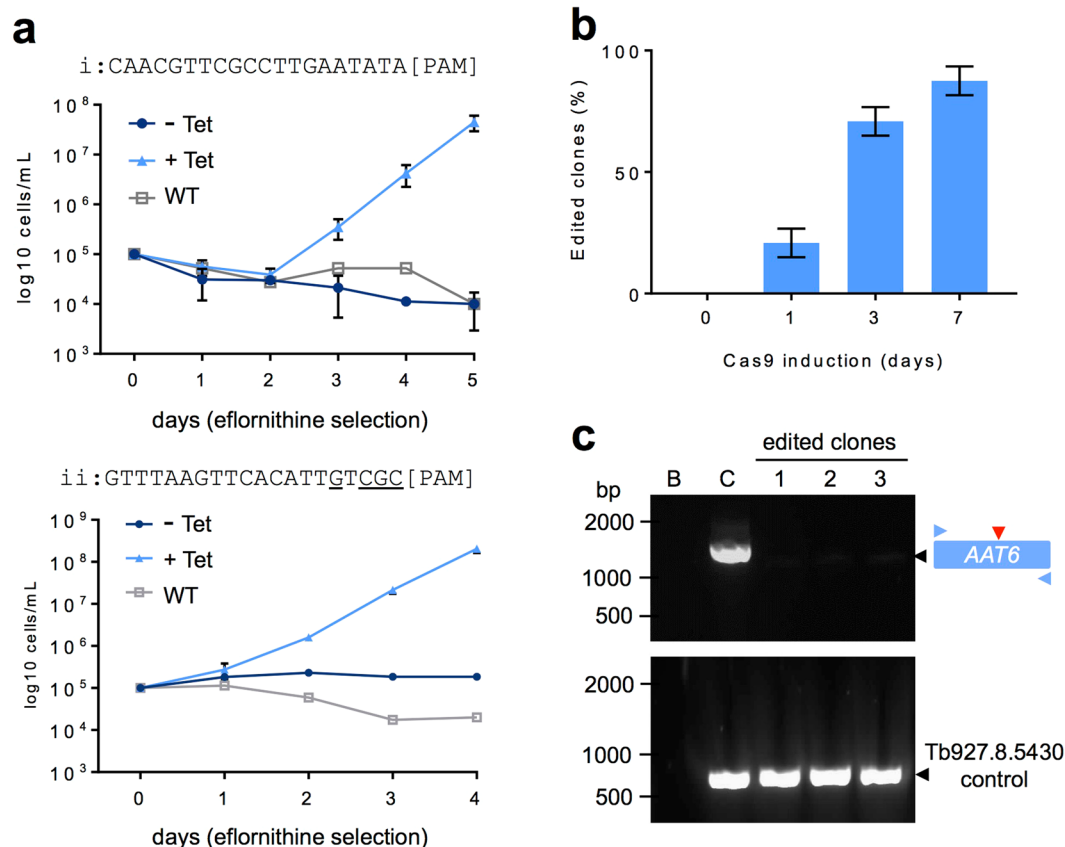


**Figure 2.** High-efficiency Cas9-based editing of the *AQP2* gene. **(a)** Cumulative cell growth in the presence of 6 nM pentamidine (estimated  $2 \times EC_{50}$ ) following 3-days with (+Tet) or without (–Tet) Cas9 induced editing of *AQP2*. Three independent Cas9/sgRNA<sup>AQP2</sup> clones and wild-type (WT) cells were analysed. Error bars, SD. The gRNA sequence is shown above the plot. **(b)** A cloning assay reveals efficient Cas9 editing. Cas9 was induced as indicated followed by sub-cloning in the absence of Tet. The sub-clones were then transferred to medium with or without 6 nM pentamidine and considered ‘edited’ if drug-resistant; twelve sub-clones were analysed from each of two independent Cas9/sgRNA<sup>AQP2</sup> clones ( $n = 24$  sub-clones) for each time-point. **(c)** The PCR assays reveal *AQP2* editing in three independent drug-resistant clones. B, blank – no genomic DNA; C, control – wild-type genomic DNA. The *AQP3* gene is adjacent to *AQP2*. The vertical (red) arrowheads indicate the Cas9 target site in *AQP2*. The horizontal (green, blue and grey) arrowheads indicate the primers used for the PCR-assays. Predicted wild-type *AQP2* fragment, 724 bp; predicted *AQP3* fragment, 508 bp; predicted wild-type *AQP2*-3 fragments in bottom two panels, 2,440 and 3,200 bp, respectively.

end-joining is more likely to be involved in repair at loci that lack extensive blocks of adjacent homology<sup>5</sup>. Thus, our inducible Cas9 editing system is tightly regulated and achieves temporally controlled and remarkably efficient bi-allelic editing.

**Tightly regulated and high-efficiency *AAT6* editing.** To determine whether the efficiency of our system could be extended to another native locus, and to further test sgRNA design for *T. brucei*, we selected a second gene, on a different chromosome, for analysis. Again exploiting a known drug-resistance mechanism, we selected the amino acid transporter, or *AAT6* gene (Tb927.8.5450). Trypanosomes become resistant to the amino acid analogue, eflornithine when *AAT6* is defective and the eflornithine  $EC_{50}$  is >40-fold greater than for wild-type cells in this case<sup>36</sup>. Thus, eflornithine can also be used to provide a robust phenotypic read-out for bi-allelic editing. Since a PAM-proximal GC-content of >50% (at least 4 of 6 bases) yields optimal efficiency gRNAs in other cell-types<sup>37</sup>, we cloned two 20b *AAT6*-specific gRNA sequences in pT7<sup>sgRNA</sup>. One gRNA (i) has a PAM-adjacent GC content of 0/6 and the other (ii) has a PAM-adjacent GC content of 4/6; both had an overall GC-content of 40% (Fig. 3a). The resulting constructs were introduced into 2T1<sup>T7-Cas9</sup> cells as above. Using two independent Cas9/sgRNA<sup>AAT6</sup> clones for each sgRNA, we induced Cas9 programmed to target *AAT6* for three days, followed by selection with eflornithine. In both cases, the growth profiles indicated an absence of resistance without induction and robust resistance with induction (Fig. 3a). We observed a striking difference between sgRNA-i and sgRNA-ii, however. As predicted by previous studies<sup>37</sup>, the sgRNA with the optimised PAM-proximal GC-content was associated with more efficient editing; with sgRNA-ii yielding drug-resistant cells more rapidly than sgRNA-i during selection (Fig. 3a).

As above, and using the optimised sgRNA (*AAT6*-ii) we next induced editing in the absence of eflornithine and sub-cloned the populations. The proportions of eflornithine-resistant sub-clones revealed remarkably efficient editing, reaching ~90% after 7-days (Fig. 3b). PCR-based assays confirmed disruption of the *AAT6* gene in three independent resistant clones (Fig. 3c, Supplementary Fig. 1). The process, from gRNA design to validation

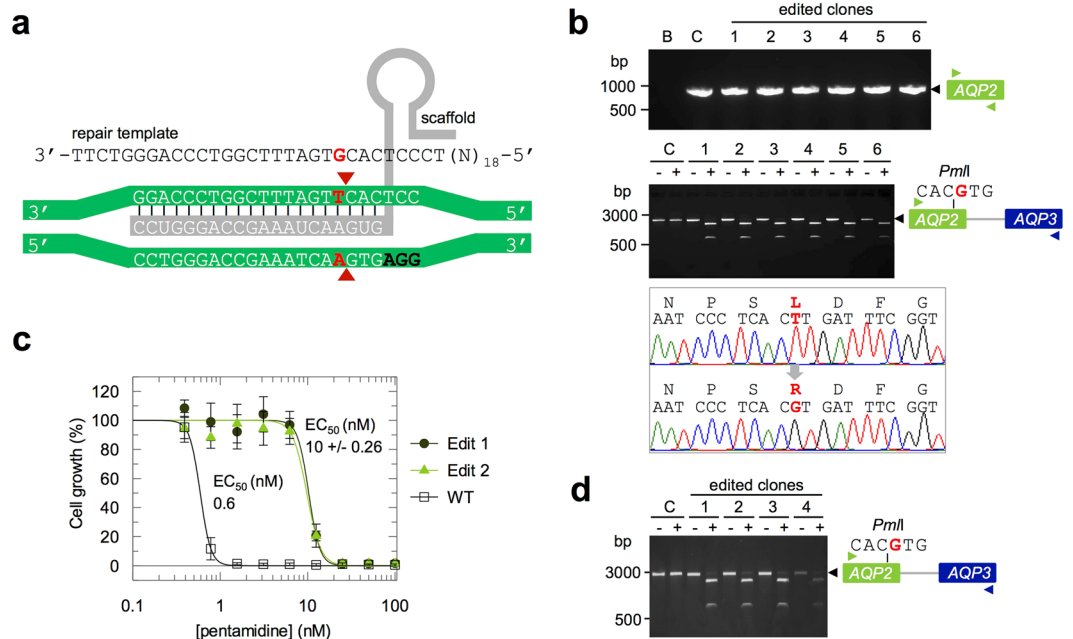


**Figure 3.** High efficiency Cas9-based editing of the *AAT6* gene. **(a)** Cumulative cell growth in the presence of 270  $\mu$ M eflornithine (estimated  $10 \times EC_{50}$ ) following 3-days with (+Tet) or without (–Tet) Cas9 induced editing of *AAT6*. Two independent Cas9/sgRNA<sup>*AAT6*</sup> clones and wild-type (WT) cells were analysed for each gRNA (i and ii). Error bars, SD. The gRNA sequences are shown above the plot with PAM-adjacent G/C residues underlined (see text for details). As can be seen, eflornithine is a cytostatic rather than cytotoxic drug. **(b)** A cloning assay reveals efficient Cas9 editing by gRNA-ii. Cas9 was induced as indicated followed by sub-cloning in the absence of Tet. The sub-clones were then transferred to medium with or without 270  $\mu$ M eflornithine and considered ‘edited’ if drug-resistant. Twelve sub-clones were analysed from each of two independent clones ( $n = 24$  sub-clones) for each time-point. **(c)** The PCR assays reveal *AAT6* editing in three independent drug-resistant clones. B, blank – no genomic DNA; C, control – wild-type genomic DNA. The Tb927.8.5430 gene is close to *AAT6* and serves as an unedited control. The vertical (red) arrowhead indicates the Cas9 target site. The horizontal (blue) arrowheads indicate the primers used for the *AAT6* PCR-assay. Predicted wild-type *AAT6* fragment, 1,467 bp; predicted wild-type Tb927.8.5430 fragment, 648 bp.

of edited clones can be performed in approximately three weeks. Thus, our tightly regulated inducible Cas9 editing system achieves rapid, high-efficiency, bi-allelic editing at native chromosomal loci.

**Precise and ‘scar-less’ single base editing.** The analyses above demonstrate robust and efficient inducible editing. The edits described are site-specific and are sufficiently efficient that neither antibiotic selectable markers, nor exhaustive screening, are required to isolate edited clones. However, these edits typically result in deletions that may also affect neighbouring sequences or genes. We next sought to use this system for precision editing by co-opting the native homologous recombination machinery. This requires the introduction of a DSB adjacent to the edit site and the introduction of a DNA template containing the desired site-specific edit (Fig. 4a). We once again chose *AQP2* as the target gene for this experiment. *AQP2* is an atypical aquaglyceroporin that lacks many of the conserved pore-lining residues, including the ‘aromatic arginine’ motif<sup>30</sup>. A leucine residue is present at this position and a single L<sup>264</sup>R change generates a mutant *AQP2* that fails to reverse pentamidine resistance when over-expressed in *aqp2*-null trypanosomes<sup>38</sup>. Cas9 editing technology now presents opportunities to test the behaviour of such mutant genes in their native context. Short homology flanks are sufficient for recombination in *T. brucei*<sup>13</sup> and single-stranded oligonucleotides can be used to edit trypanosomatid<sup>14,22</sup> and metazoan genomes<sup>39</sup>. We, therefore, used a new gRNA, cloned in the pT7<sup>sgRNA</sup> construct, that targets a sequence adjacent to the L<sup>264</sup> codon (Fig. 4a). In this case, although the overall GC-content is 55%, the PAM-adjacent GC content of the new gRNA is 3/6 and may be suboptimal (Fig. 4a, see above).

We used this new ‘AQP2.2’ gRNA to assemble Cas9/sgRNA<sup>*AQP2.2*</sup> clones. We next delivered the short (48 bp) single-stranded oligonucleotide repair-template (Fig. 4a) immediately followed by induction of Cas9-based



**Figure 4.** Precision base editing with Cas9. **(a)** The schematic illustrates the AQP2-editing strategy (one allele shown). The genomic DNA sequence is on a green background, the PAM is in bold black text, the gRNA is on a grey background and the arrowheads indicate the expected Cas9-induced break sites. The single-stranded repair template is centred on the editing site and shows the T-G edit in red letters. **(b)** Analysis of sub-clones from an edited and pentamidine-selected population. The PCR assays are as described in Fig. 2c. B, blank – no genomic DNA; C, control – wild-type genomic DNA. The AQP2-3 PCR products were also digested with *PmlI* (lower panel) and sequencing confirmed that all six clones had the templated T<sup>791</sup>G/L<sup>264</sup>R edit (one example shown, lower trace). **(c)** Dose-response curves for pentamidine. Both edited clones analysed display a drug-resistant phenotype. WT, wild-type cells. **(d)** Analysis of the four pentamidine-resistant sub-clones among 48 sub-clones from an edited population. The PCR assay and *PmlI* digest are as described in b.

cleavage. The repair template carries the single edited base-change (Fig. 4a, red) and was delivered by electroporation while Cas9 was induced for only 24 h in this case, to maximise template-directed repair. Note that, despite the base-change, the edited alleles may remain substrates for re-cleavage, albeit at reduced efficiency.

We expected a number of possible outcomes within the resulting population of cells; no editing, non-templated deletions (as described above) or the desired edit. Un-edited cells remain pentamidine-sensitive and these were selected against by growth in pentamidine for three days, immediately following the 24 h induction period. The resulting pentamidine-resistant population was sub-cloned and six clones were subjected to an AQP2 PCR assay. In this case, all six clones analysed yielded an AQP2 PCR-product (Fig. 4b, upper panel, Supplementary Fig. 1), in contrast to the clones above that underwent non-templated repair (Fig. 2c). Using a second PCR assay, all six clones also yielded an AQP2-3 PCR-product that contained a new *PmlI* restriction site introduced by the single-base edit (Fig. 4b, middle panel, Supplementary Fig. 1). Sequencing of these products confirmed the desired T<sup>791</sup>G/L<sup>264</sup>R precision edit in every clone (Fig. 4b, lower panel; Supplementary Fig. 2). This suggests that the electroporated short single-stranded oligonucleotide is used remarkably efficiently as a template and that this particular homology-directed repair pathway is favoured over the error-prone single-strand annealing pathway. To assess the impact of expression of the AQP2 - L<sup>264</sup>R mutant in the native context, we determined the EC<sub>50</sub> for pentamidine. This revealed robust resistance and an EC<sub>50</sub> that was increased approximately 17-fold relative to the wild-type control (Fig. 4c).

The remarkable efficiency of precision base editing above suggested that we may be able to isolate precision edited clones without selection to enrich for the edited cells. We, therefore, repeated the experiment above but, instead of selecting against unedited cells by growth in pentamidine, we simply sub-cloned the population immediately following the 24 h induction period. Forty-eight clones were expanded for several days without further Cas9-induction and were then tested for pentamidine-resistance. We found four clones (~8%) to be resistant and, again using a PCR-assay and diagnostic digest (Fig. 4b), found that all four were precision-edited (Fig. 4d, Supplementary Fig. 1). These results suggest that, using the Cas9-editing system described here, it will be possible to isolate clones with precision-edits at different loci, following screening of a readily manageable number of clones.

## Discussion

The versatility and accessibility of tools for genome editing have improved rapidly in recent years and Cas9-based approaches are central to this transition. Substantial progress has been made in developing these systems in trypanosomatids but an inducible Cas9 editing system has not yet been reported. Additionally, for *T. brucei*,

double-allele editing has not been achieved without selectable markers nor have efficiencies been determined. Given these limitations, precision base editing is not readily achievable. To advance Cas9-based editing tools and approaches for *T. brucei*, we have established a highly efficient inducible system and have used the system for templated, homology-directed, precision base editing.

The system we describe incorporates several features that improve convenience and editing efficiency. In terms of convenience, the current strategy involves annealing a pair of oligonucleotides and ligating the product to the *BbsI*-digested pT7<sup>sgRNA</sup> construct prior to transfection. This cloning step is highly efficient, but sgRNA delivery can also be achieved in a cloning-free format. Such a cloning-free format, involving transient transfection of the sgRNA or an sgRNA-encoding PCR-product, may be suitable for strategies that do not require maximal-efficiency editing, such as when integrated antibiotic resistance markers are used to select for correctly edited cells<sup>13</sup>. In terms of the editing efficiency of our system, several features likely contribute; first, Cas9 was codon-optimised to increase expression in *T. brucei*<sup>28</sup>. Second, Cas9 and the sgRNA are transcribed from strong *rDNA* and T7 promoters, respectively. Third, the T7 polymerase, *BbsI*-cloning approach and ribozyme all help to produce sgRNAs with minimal superfluous sequence, likely enhancing assembly with Cas9. In addition, tightly regulated temporal control of Cas9 expression increases the versatility of the system, also minimising off-target edits and allowing for phenotypic analysis of edited cells without potential interference from continuously expressed Cas9. Finally, and as in metazoan cells<sup>37</sup>, our findings support the view that gRNA sequences achieve optimal efficiency cleavage in *T. brucei* when the PAM-adjacent GC-content is >50%.

A major source of off-target edits may involve the non-specific association of Cas9 with native RNA. Indeed, expression of Cas9 in the absence of sgRNA has been reported to be toxic in *T. cruzi*<sup>18</sup> and *Toxoplasma gondii*, and a “decoy” sgRNA has been used in *Toxoplasma* to reduce these effects<sup>24</sup>. The current system should minimise such problems in *T. brucei* since we constitutively transcribe the sgRNA and induce a tightly regulated Cas9 gene, which stabilises the sgRNA. Thus, cells are not exposed to Cas9 in the absence of sgRNA. Titration of Cas9-dosage can dramatically increase cleavage specificity and reduce off-target effects<sup>40</sup> and the inducible system now presents an opportunity to titrate Cas9 expression in *T. brucei* simply by reducing the tetracycline concentration in the growth medium<sup>4</sup> or by inducing Cas9 for only a short period of time. Temporally constrained exposure to Cas9 is also important for precision base editing since precision-edited sequences may still be potential substrates for re-cleavage<sup>37,40–42</sup>. It will be worthwhile to test other versions of Cas9 in *T. brucei*. For example, HypaCas9 (*SpCas9*: N692A/M694A/Q695A/H698A) displays reduced off-target editing<sup>43</sup> while xCas9 variants display broader PAM compatibility<sup>44</sup>.

Further optimisation of specific editing and minimising off-target edits can also be achieved by modifying the gRNA. In metazoan cells, gRNA sequences can be truncated by 2–3 nt to reduce off-target cleavage without substantially reducing on-target efficiency<sup>45</sup> and these gRNAs can be further optimised to improve specific activity<sup>37,40,41</sup>. On the other hand, trypanosomatid genomes are relatively small at ~30 Mbp, likely presenting relatively few potential off-target sites. In addition, homologous recombination is the dominant repair pathway in *T. brucei*<sup>5</sup> such that low-frequency off-target edits in one allele will typically be reversed, since allelic templated, homology-directed repair will be favoured over error-prone repair by end-joining. In circumstances where off-target edits must be avoided, genome sequencing and profiling of potential off-target sites will remain an option.

Established tools for the genetic dissection of *T. brucei* already present a range of options. Cas9-based approaches present alternative strategies and, perhaps more importantly, present completely new opportunities, for increased throughput and for marker-free and precision editing, in particular. In terms of alternative strategies, RNAi is currently the favoured approach for conditional loss-of-function studies in *T. brucei*, particularly for many likely essential genes. Similarly, native, single-allele tagging for protein localisation studies is readily achievable using selectable marker based constructs in *T. brucei*<sup>46</sup> and can now be considered alongside (double-allele) Cas9-based editing options. Clearly, editing efficiency need not be high if selectable markers are to be used to enrich for cells with a desired, non-lethal edit. When an edit is lethal on the other hand, the desired editing will fail to yield long-term viable recombinant cells. In these circumstances, editing that is both inducible and close to 100% efficiency, will be a major advantage since it will facilitate the study of phenotypes associated with lethal edits that precede cell death. High efficiency will also be important for other marker-free or high-throughput approaches. In these cases, high validation rates for a desired edit will substantially reduce the number of clones that must be screened and will improve coverage across large numbers of editing sites, respectively.

Key advantages of precision editing without selectable markers are that the (double-allele) edit is generated in one step in a native context, retaining native flanking sequences that may control expression level or timing during the cell cycle or the life cycle. A recombinant AQP2 - L<sup>264</sup>R mutant was previously ectopically overexpressed in an *aqp2* null background, for example<sup>38</sup>. The pentamidine resistance phenotype associated with AQP2 - L<sup>264</sup>R expression is confirmed here but the prior ectopic expression approach required three genetic manipulation steps and was also associated with risks of potential artefacts associated with non-physiological expression. Thus, the current editing approach adds an important new ‘one-step, precision’ option to the genetic manipulation ‘toolbox’ for *T. brucei*. The efficiency reported here indicates that it is now feasible to obtain cells with edited loci without selection, by screening manageable numbers of clones.

The Cas9-based editing system we describe for *T. brucei* incorporates key features, including robust, temporal control of Cas9 and sgRNA expression. We show that null-clones (both alleles disrupted) can be generated at exceptionally high frequency without selection. Future applications will likely extend beyond precision edits, deletions and insertions to high-throughput approaches and the repurposing of nuclease-dead Cas9 for targeting other functionalities to specific genomic loci<sup>12</sup>.

## Materials and Methods

***T. brucei* strains and growth conditions.** Bloodstream form *T. brucei*, Lister 427 (MITat 1.2), clone 221a cells, 2T1 and 2T1<sup>T7</sup> cells<sup>25</sup> and their derivatives were grown in HMI-11 medium. Cumulative growth rates were monitored by splitting to  $1 \times 10^4$  cells/ml and by counting daily. Tetracycline (Sigma) was applied at 1 µg/mL. Pentamidine isethionate was from Sigma and eflornithine was from Sanofi Aventis.

***T. brucei* genetic manipulation and plasmid construction.** *T. brucei* were genetically manipulated using electroporation as described<sup>25</sup>. To generate the pRPa<sup>Cas9</sup> construct, the SpCas9 sequence was codon optimised and synthesised (Genscript) with a La nuclear localisation signal<sup>27</sup> at the N-terminus. The synthetic Cas9 cassette, digested with *HindIII*/*BglII*, was ligated to pRPa<sup>c6MYCn26</sup> [Addgene ID:69242] digested with *HindIII*/*BamHI*. The resulting pRPa<sup>Cas9</sup> construct was linearised with *AscI* prior to electroporation; 2T1<sup>T7-Cas9</sup> parasites were selected with hygromycin at 2.5 µg/mL and checked for sensitivity to puromycin to indicate correct construct integration. To generate the pT7<sup>sgRNA</sup> construct, the *BLA* resistance cassette in pRPa<sup>c6MYCn26</sup> [Addgene ID:69243] was replaced by a *BLE* cassette following *NgoMIV*/*BstZ171* digestion. A synthetic sequence comprising a T7 promoter, tandem *BbsI* sites, the sgRNA scaffold and an HDV ribozyme was then used to replace the expression cassette from pRPa<sup>c6MYCn</sup>, following *NheI*/*ApaI* digestion. Potential gRNA sequences were identified either manually or using the LeishGEdit.net online tool<sup>13</sup>, and then selected based on PAM-adjacent GC-content<sup>37</sup>. The gRNA oligonucleotide pairs were FwgAQP2 (**AGGGGGTGTGGCCATGTACCCCG**) and RvgAQP2 (**AAACCGGGGTACATGGCAAACACC**); FwgAAT6i (**AGGGGTTTAAAGTTCACATTGTTCGC**) and RvgAAT6i (**AAACGCGACAATGTGAACCTAAC**); FwgAAT6ii (**AGGGCAACGTTCCGCTTGAATATA**) and RvgAAT6ii (**AAACTATATTCAAGGCGAACGTTG**) and, for the 'AQP2.2' precision edit, FwgAQP2c (**AGGGCCTGGGACCGAAATCAAGTG**) and RvgAQP2c (**AAACCACTTGATTTCCGGTCCCAGG**); the overhanging ends that facilitate cloning are in bold. Oligonucleotide pairs were annealed by heating to 70°C for 3 min followed by slow cooling. Annealed pairs were ligated to *BbsI*-digested pT7<sup>sgRNA</sup> and constructs were confirmed by sequencing. pT7<sup>sgRNA</sup> constructs containing specific gRNA sequences were linearised with *NotI* prior to electroporation and parasites were selected with phleomycin at 2 µg/mL. 40 µg of single-stranded oligonucleotide (GTCTCCCCTGCGATGAATCCCTCAGTGATTTCCGGTCCCAGGGTCTT) was used as a repair template, which was resuspended in 10 µL of H<sub>2</sub>O and delivered by electroporation as described above; the edited base is in bold.

**DNA analysis.** DNA was extracted from *T. brucei* using a DNeasy blood and tissue Kit (Qiagen). The primer pairs used for the PCR assays were FwAQP2only (GAGCGGTGGGATGCAGATGTAC) and RvAQP2only (CCCCGAGAAGGATCGCACCG), FwAQP3only (GTGTGGTCCGACGGTTATC) and RvAQP3only (GCGTAACCCGTTGAGTAACCG), FwAAT6only (ATGAGAGAGCCGATACAAACTTCAAC) and RvAAT6only (TCAGAGTTCAGCAATGACGCTG), Fw5430ATG (ATGGGCAACAACCGGAAGTAGC) and Rv5430STOP (CTATGCCGCCAGTCGAACAC). The FwAQP2only or Fw14180 (CCAAAATCAGCGGGTTCCTGA) primers were combined with RvAQP3only for the AQP2-3 PCR assays (lower panels in Fig. 2c).

**RNA analysis.** For northern blotting, RNA was extracted from *T. brucei* using an RNeasy Mini Kit (Qiagen). 1.5 µg of RNA was resolved on formaldehyde agarose gels in MOPS buffer and transferred to nylon membranes by capillary blotting. A digoxigenin (DIG)-labelled riboprobe was hybridized to the membrane and the transcript was detected using an anti-DIG antibody (Roche) with a chemiluminescent CDP-star substrate (Roche). rRNA stained with ethidium-bromide was used as loading control.

**Protein analysis.** Total *T. brucei* extracts equivalent to  $1 \times 10^6$  cells were separated on SDS-polyacrylamide gels and subjected to western blotting analysis according to the manufacturers' instructions (BioRad). Blots were blocked in 5% milk in PBS-T and washes were performed in PBS-T (0.05% Tween). Blots were then cut in two, and the upper part (above 70KDa according to the pre-stained protein ladder) was probed with 1/1000 α-Cas9 primary antibody (Abcam) and 1/5000 α-mouse secondary antibody (Bio-Rad). As a loading control, the lower part of each blot was incubated with 1/20000 α-EF1α (Merck Millipore) and 1/5000 α-mouse secondary antibody (Bio-Rad).

**Drug sensitivity analysis.** EC<sub>50</sub> analysis was performed using the AlamarBlue method as described<sup>47</sup>; drug exposure was for 72 h and incubation with resazurin (Sigma-Aldrich) was for 5–6 h. Plates were read on an Infinite 200 Pro plate-reader (Tecan).

**Data Availability.** The datasets generated during the current study are available from the corresponding author on reasonable request.

## References

1. Field, M. C. *et al.* Anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need. *Nat Rev Microbiol* **15**, 217–231 (2017).
2. Clayton, C. E. Gene expression in Kinetoplastids. *Curr Opin Microbiol* **32**, 46–51 (2016).
3. Meissner, M., Agop-Nersesian, C. & Sullivan, W. J. Jr. Molecular tools for analysis of gene function in parasitic microorganisms. *Appl Microbiol Biotechnol* **75**, 963–975 (2007).
4. Wirtz, E. & Clayton, C. Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. *Science* **268**, 1179–1183 (1995).
5. Glover, L., McCulloch, R. & Horn, D. Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. *Nucleic Acids Res* **36**, 2608–2618 (2008).
6. Glover, L. & Horn, D. Locus-specific control of DNA resection and suppression of subtelomeric VSG recombination by HAT3 in the African trypanosome. *Nucleic Acids Res* **42**, 12600–12613 (2014).



7. Berriman, M. *et al.* The genome of the African trypanosome *Trypanosoma brucei*. *Science* **309**, 416–422 (2005).
8. Glover, L. & Horn, D. Site-specific DNA double-strand breaks greatly increase stable transformation efficiency in *Trypanosoma brucei*. *Mol Biochem Parasitol* **166**, 194–197 (2009).
9. Boothroyd, C. E. *et al.* A yeast-endonuclease-generated DNA break induces antigenic switching in *Trypanosoma brucei*. *Nature* **459**, 278–281 (2009).
10. Glover, L., Alsford, S. & Horn, D. DNA break site at fragile subtelomeres determines probability and mechanism of antigenic variation in African trypanosomes. *PLoS Pathog* **9**, e1003260 (2013).
11. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
12. Barrangou, R. & Doudna, J. A. Applications of CRISPR technologies in research and beyond. *Nat Biotechnol* **34**, 933–941 (2016).
13. Beneke, T. *et al.* A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplasts. *R Soc Open Sci* **4**, 170095 (2017).
14. Soares Medeiros, L. C. *et al.* Rapid, selection-free, high-efficiency genome editing in protozoan parasites using CRISPR-Cas9 ribonucleoproteins. *MBio* **8**, e01788–17 (2017).
15. Lander, N., Chiurillo, M. A. & Docampo, R. Genome editing by CRISPR/Cas9: A game change in the genetic manipulation of protists. *J Eukaryot Microbiol* **63**, 679–690 (2016).
16. Lye, L. F. *et al.* Retention and loss of RNA interference pathways in trypanosomatid protozoans. *PLoS Pathog* **6**, e1001161 (2010).
17. Lander, N., Li, Z. H., Niyogi, S. & Docampo, R. CRISPR/Cas9-induced disruption of paraflagellar rod protein 1 and 2 genes in *Trypanosoma cruzi* reveals their role in flagellar attachment. *MBio* **6**, e01012 (2015).
18. Peng, D., Kurup, S. P., Yao, P. Y., Minning, T. A. & Tarleton, R. L. CRISPR-Cas9-mediated single-gene and gene family disruption in *Trypanosoma cruzi*. *MBio* **6**, e02097–02014 (2014).
19. Romagnoli, B. A. A. *et al.* Improvements in the CRISPR/Cas9 system for high efficiency gene disruption in *Trypanosoma cruzi*. *Acta Trop* **178**, 190–195 (2018).
20. Duncan, S. M., Jones, N. G. & Mottram, J. C. Recent advances in *Leishmania* reverse genetics: Manipulating a manipulative parasite. *Mol Biochem Parasitol* **216**, 30–38 (2017).
21. Sollis, L. *et al.* First efficient CRISPR-Cas9-mediated genome editing in *Leishmania* parasites. *Cell Microbiol* **17**, 1405–1412 (2015).
22. Zhang, W. W., Lypaczewski, P. & Matlashewski, G. Optimized CRISPR-Cas9 genome editing for *Leishmania* and its use to target a multigene family, induce chromosomal translocation, and study DNA break repair mechanisms. *mSphere* **2**, e00340–16 (2017).
23. Zhang, W. W. & Matlashewski, G. CRISPR-Cas9-mediated genome editing in *Leishmania donovani*. *MBio* **6**, e00861 (2015).
24. Sidik, S. M. *et al.* A genome-wide CRISPR screen in *Toxoplasma* identifies essential apicomplexan genes. *Cell* **166**, 1423–1435 e1412 (2016).
25. Alsford, S., Kawahara, T., Glover, L. & Horn, D. Tagging a *T. brucei* *RRNA* locus improves stable transfection efficiency and circumvents inducible expression position effects. *Mol Biochem Parasitol* **144**, 142–148 (2005).
26. Alsford, S. & Horn, D. Single-locus targeting constructs for reliable regulated RNAi and transgene expression in *Trypanosoma brucei*. *Mol Biochem Parasitol* **161**, 76–79 (2008).
27. Marchetti, M. A., Tschudi, C., Kwon, H., Wolin, S. L. & Ullu, E. Import of proteins into the trypanosome nucleus and their distribution at karyokinesis. *J Cell Sci* **113**, 899–906 (2000).
28. Horn, D. Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids. *BMC Genomics* **9**, 2 (2008).
29. Gao, Y. & Zhao, Y. Self-processing of ribozyme-flanked RNAs into guide RNAs *in vitro* and *in vivo* for CRISPR-mediated genome editing. *J Integr Plant Biol* **56**, 343–349 (2014).
30. Baker, N. *et al.* Aquaglyceroporin 2 controls susceptibility to melarsoprol and pentamidine in African trypanosomes. *Proc Natl Acad Sci USA* **109**, 10996–11001 (2012).
31. Jones, D. L. *et al.* Kinetics of dCas9 target search in *Escherichia coli*. *Science* **357**, 1420–1424 (2017).
32. Marin, P. A., da Silva, M. S., Pavani, R. S., Machado, C. R. & Elias, M. C. Recruitment kinetics of the homologous recombination pathway in procyclic forms of *Trypanosoma brucei* after ionizing radiation treatment. *Sci Rep* **8**, 5405 (2018).
33. Graf, F. E. *et al.* Chimerization at the *AQP2-AQP3* locus is the genetic basis of melarsoprol-pentamidine cross-resistance in clinical *Trypanosoma brucei gambiense* isolates. *Int J Parasitol Drugs Drug Resist* **5**, 65–68 (2015).
34. Graf, F. E. *et al.* Aquaporin 2 mutations in *Trypanosoma brucei gambiense* field isolates correlate with decreased susceptibility to pentamidine and melarsoprol. *PLoS Negl Trop Dis* **7**, e2475 (2013).
35. Ivanov, E. L., Sugawara, N., Fishman-Lobell, J. & Haber, J. E. Genetic requirements for the single-strand annealing pathway of double-strand break repair in *Saccharomyces cerevisiae*. *Genetics* **142**, 693–704 (1996).
36. Vincent, I. M. *et al.* A molecular mechanism for efloornithine resistance in African trypanosomes. *PLoS Pathog* **6**, e1001204 (2010).
37. Ren, X. *et al.* Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in *Drosophila*. *Cell Rep* **9**, 1151–1162 (2014).
38. Song, J. *et al.* Pentamidine is not a permeant but a nanomolar inhibitor of the *Trypanosoma brucei* aquaglyceroporin-2. *PLoS Pathog* **12**, e1005436 (2016).
39. Paix, A. *et al.* Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. *Proc Natl Acad Sci USA* **114**, E10745–E10754 (2017).
40. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827–832 (2013).
41. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**, 184–191 (2016).
42. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822–826 (2013).
43. Chen, J. S. *et al.* Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* **550**, 407–410 (2017).
44. Hu, J. H. *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018).
45. Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* **32**, 279–284 (2014).
46. Dean, S., Sunter, J. D. & Wheeler, R. J. TrypTag.org: A trypanosome genome-wide protein localisation resource. *Trends Parasitol* **33**, 80–82 (2017).
47. Raz, B., Iten, M., Grether-Buhler, Y., Kaminsky, R. & Brun, R. The Alamar Blue assay to determine drug sensitivity of African trypanosomes (*T.b. rhodesiense* and *T.b. gambiense*) *in vitro*. *Acta Trop* **68**, 139–147 (1997).

## Acknowledgements

We thank M. Pawlowic for comments on the draft manuscript. This work was funded by a Wellcome Trust Senior Investigator Award to D.H. (100320/Z/12/Z) and a Wellcome Trust Centre Award (203134/Z/16/Z).

## Author Contributions

E.R. performed the experiments and prepared Figures 1–4. E.R. and D.H. designed the experiments and analysed the data. L.J. performed preliminary experiments. J.K. contributed to plasmid construction. D.H. supervised the study and wrote the original draft. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-26303-w>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018