



Gene Editing in Trypanosomatids: Tips and Tricks in the CRISPR-Cas9 Era

Akila Yagoubat, Rosa Corrales, Patrick Bastien, Maude Lévêque, Yvon Sterkers

► To cite this version:

Akila Yagoubat, Rosa Corrales, Patrick Bastien, Maude Lévêque, Yvon Sterkers. Gene Editing in Trypanosomatids: Tips and Tricks in the CRISPR-Cas9 Era. Trends in Parasitology, 2020, 36 (9), pp.745-760. <10.1016/j.pt.2020.06.005>. <hal-02920067>

HAL Id: hal-02920067

<https://hal.science/hal-02920067v1>

Submitted on 24 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Gene editing in trypanosomatids: tips and tricks in the CRISPR-Cas9 era

Akila Yagoubat, Rosa M Corrales, Patrick Bastien, Maude F L  v  que and Yvon Sterkers*

MiVEGEC, Univ. Montpellier, CNRS, IRD, CHU, Montpellier, France

*Correspondence: yvon.sterkers@umontpellier.fr (Y. Sterkers)

This manuscript has been published, please cite: Yagoubat A, Corrales RM, Bastien P, L  v  que MF, Sterkers Y. Gene Editing in Trypanosomatids: Tips and Tricks in the CRISPR-Cas9 Era. Trends Parasitol. 2020 Jul 20:S1471-4922(20)30180-X. doi: 10.1016/j.pt.2020.06.005.

Highlights

- CRISPR-Cas9 enables hitherto impossible procedures, such as, marker-free editing, point mutations and genome-wide loss-of-function analyses, thus opening a new era in the study of biological processes
- Transient PCR-based CRISPR-Cas9 is a straightforward method of gene editing in trypanosomatids
- Constant strategies allow single nucleotide modification, editing multicopy genes, and 'marker-free' tagging
- Modification of endogenous UTRs may cause failure of the expected gene editing, therefore strategies allowing UTR preservation are important to improve genetic manipulation in trypanosomatids
- Inducible CRISPR-Cas9 systems are now available in trypanosomatids, which can be useful tools to study essential as well as non-essential genes.

Outstanding Questions Box

- What are the advantages of using CRISPR-Cas9 for gene tagging as compared with classical methods? Is the tagging of all alleles necessary in all cases?
- How convincing is protein localization when endogenous UTRs are modified?
- Is inducible KO useful for non-essential genes?
- In *T. brucei*, what is the place for inducible KO vs RNAi knockdown?
- Toxicity due to constitutive expression of Cas9 and off-targets should be checked systematically. Should whole genome sequence become a gold standard for controlling gene editing?

Abstract

Gene editing in trypanosomatids has long been proven difficult. The development of CRISPR-Cas9 has improved this issue, opening the way to a better understanding of biological processes and drug resistance mechanisms, and screening of drug targets. Different strategies have now been developed: either PCR- or plasmid-based, mainly differing by the nature of the donor DNA and the single guide RNA transcription. Here we review the main genetic tools available in *Leishmania* spp., *Trypanosoma cruzi* and *Trypanosoma brucei* for gene tagging, single base editing and deletion of non-essential and essential genes. We discuss the main advantages and challenges of different strategies and how to choose 'the right cut' depending on the importance of untranslated regions. These considerations allow selection of the most accurate gene editing approach for a given functional analysis.

The CRISPR-Cas9 revolution in trypanosomatids

Trypanosomatids are highly divergent eukaryotes characterized by unique genomic and biological features. The human pathogens *Leishmania* spp., *Trypanosoma brucei* and *Trypanosoma cruzi* are the causative agents of neglected diseases: leishmaniasis, sleeping sickness and Chagas' disease, respectively. Elucidating divergent biological processes in these parasites may contribute to the discovery of novel preventative and therapeutic approaches. Genetic manipulations in *Leishmania* and *T. cruzi* have remained challenging and time consuming due to their particularly high degree of genomic plasticity

[1-3]. Thus, attempts to generate *Leishmania* null mutants may result in the amplification of the target gene. This can include the generation of supernumerary chromosomes or extrachromosomal elements retaining a copy of the gene of interest (GOI), even after several rounds of transfection [2, 4]. In *T. brucei*, genetic studies have been facilitated by highly efficient **homologous recombination** (see Glossary) and the presence of versatile and robust tools for gene tagging and knockdown using RNAi machinery [5, 6]. This situation has been evolving since 2014 when **CRISPR-Cas9** technology was first adapted in *T. cruzi*

[7], *Leishmania* [8, 9], and then *T. brucei* [10]. The CRISPR-Cas9 technique relies on two/three molecular components: (i) the Cas9 endonuclease, (ii) the **single guide RNA (sgRNA)** which contains a complementary sequence of the target gene and (iii) optionally the **donor DNA** to be inserted in the genome. The Cas9 and sgRNA form a ribonucleoprotein complex that recognizes the target DNA and generates a precise double-strand break (DSB) followed by DNA repair. In trypanosomatids, the key factors for **non-homologous end joining (NHEJ)** pathway are absent [11-13]. Consequently, DSBs are repaired through **microhomology-mediated end joining (MMEJ)**, **single strand annealing (SSA)** or **homology-directed repair (HDR)**. In the presence of donor DNA, DSBs are mainly repaired through HDR or SSA mechanisms [7, 9, 14-17](Key Table). Thanks to this technology, in less than five years, gene replacement in trypanosomatids has evolved from being cumbersome and time consuming to relatively straightforward; it has contributed to dissecting key biological traits such as flagellum composition, histone biology, and studying potential drug targets [18-20]. In 2018, Jeremy Mottram estimated that gene deletion attempts had been made for “200 *Leishmania* genes and 36 *T. cruzi* genes since the first published study in 1990” using a 'classical' approach [21]. We estimate that, in 2019 alone, more than 120 *Leishmania* genes were edited and published by different groups using different CRISPR-Cas9 strategies. Yet, the increasing number of papers in the literature describing different approaches for genome editing in trypanosomatids using CRISPR-Cas9 suggests that a single efficient approach is not readily available [4, 7-10, 15-17, 19, 20, 22-24]. The efficiency of the different systems depends on the optimal amount/quality of the sgRNA and of donor DNA, as well as the efficiency of DSB repair. The CRISPR-Cas9 strategies can be classified into 'stable' and 'transient' systems, depending on the expression of sgRNA and/or Cas9 (Key Table 1). In the transient systems (Figure 1), the sgRNA alone or together with Cas9 are transiently expressed in the cell, while in the stable systems (Figure 2), the Cas9 and sgRNA are constitutively expressed. In the different systems, the donor DNA can be provided either as a plasmid, a PCR product or a singlestranded oligonucleotide. In this review, we present the strategies commonly used for gene manipulation prior to and after the CRISPR-Cas9 era in *Leishmania*, *T. brucei* and

T. cruzi. Our goal is to provide guidelines for the selection of the most appropriate gene editing approach for a given analysis among the expanding repertoire of CRISPR-Cas9 strategies developed for trypanosomatids. We discuss the advantages and challenges of each system and provide an overview of the ongoing revolution of CRISPR-Cas9 technology in the trypanosomatids.

Box 1. Failure of CRISPR-Cas9 and definition of essentiality

Despite the enormous progress made by CRISPR-Cas9 in genome editing, a number of failures persist. If gene editing fails, it is important to verify potential mismatches in the sequences of the **seed**, the HRs, and the **protospacer adjacent motif (PAM)**. One should keep in mind that some discrepancies may exist between the reference genome sequence and the genome sequence of the strain actually used in the laboratory. Several sgRNAs should be designed, including a sgRNA targeting the 'neutral' intergenic regions described above to preserve the endogenous UTRs [27]. Beyond the field of trypanosomatids, online tools to design sgRNAs are available, including EuPaGDTⁱⁱⁱ, CRISPRscan^{iv}, Sequence Scan for CRISPR-cistrome^v, CCTop - CRISPR/Cas9 target online predictor^{vi} and CRISPRater [10, 89-93]. SgRNAs designed using these different tools may be different and present varying editing efficiency for a given target gene [15]. Further studies are needed to determine the best online tool for trypanosomatids. Finally, in case a deletion attempt fails, one should be cautious before concluding that the gene is 'essential'. Was the locus editable? *e.g.* was N- and C- terminal tagging possible without yielding a loss-of-function or an unexpected phenotype? Was gene inactivation only partial, leading to a recovering cell population made of heterozygous cells bearing edited and non-edited alleles? Other ways to assess essentiality is to directly observe cells dying, as shown for TOR1, RAD51 and DNA polymerase theta gene inactivation [9, 16], or to use an inducible system.

Acknowledgements

This work was supported by the Agence Nationale de la Recherche (ANR) within the frame of the “Investissements d'avenir” program [ANR 11 - LABX - 0024 - 01 “ParaFrap”]; the Centre National de la Recherche Scientifique (CNRS); the University of Montpellier, and the Centre Hospitalier Universitaire of Montpellier. AY was a recipient of a grant from the French Parasitology Consortium “ParaFrap” [ANR 11-LABX-0024-01]. We thank Philip Agnew for correcting our English.

Resources

ⁱ <http://tryptag.org/> ⁱⁱ <http://www.leishgedit.net/Home.html> ⁱⁱⁱ <http://gma.ctegd.uga.edu/> ^{iv} www.crisprscan.org ^v <http://cistrome.org/SSC/> ^{vi} <https://crispr.cos.uni-heidelberg.de/>

Keywords: CRISPR-Cas9, CRE-Lox, UTR modification, *Leishmania*, *Trypanosoma*, inducible genome editing

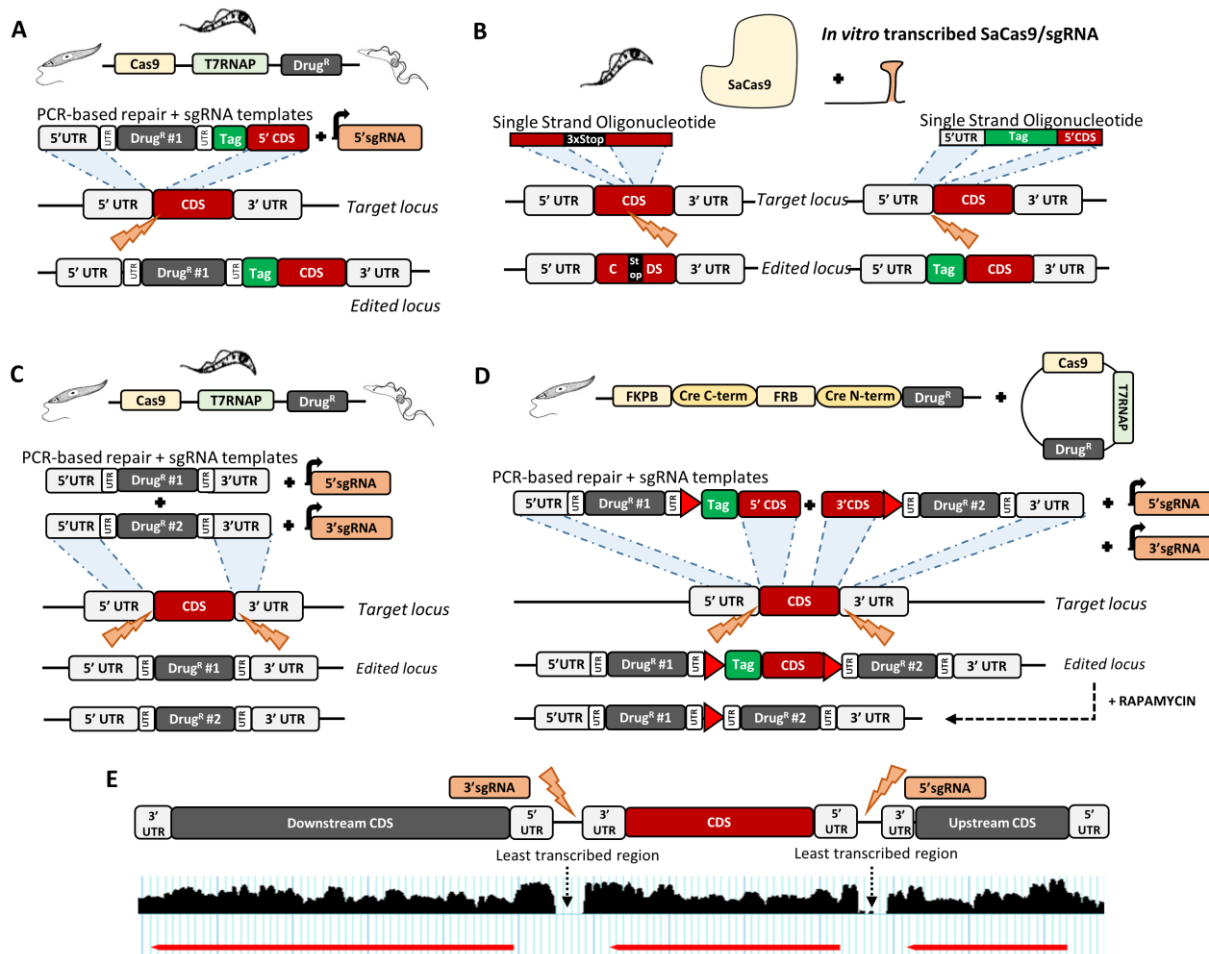


Figure 1. CRISPR-Cas9 transient strategies in trypanosomatids.

(A) PCR-based strategy for gene tagging in *Leishmania* spp., *T. brucei* and *T. cruzi* [10, 39]. Orange lightning symbols indicate DSB sites. SgRNA and repair DNA templates are delivered as PCR fragments encoding a selection marker and a tag flanked by HRs in cell lines constitutively expressing Cas9 and T7 polymerase. (B) Ribonucleoprotein complex-based strategy for gene disruption and tagging in *T. cruzi* [23]. The SaCas9 recombinant protein and *in vitro* transcribed sgRNA are delivered as a ribonucleoprotein complex. For gene disruption and gene tagging, the donor DNA is delivered as a single-stranded oligonucleotide containing 3 stop codons or a tag, respectively, flanked by HRs. (C) PCR-based strategy for gene deletion in *T. brucei*, *T. cruzi* and *Leishmania* [10, 39]. Two sgRNA templates targeting the 5'UTR and 3'UTR of the target gene and two repair DNA templates encoding different selection markers flanked by HRs are generated as PCR fragments and transfected in strains constitutively expressing Cas9 and T7 polymerase. (D) Inducible PCR-based strategy for gene deletion in *Leishmania* [27]. Two sgRNA templates and two repair DNA templates encoding (i) a drug resistance cassette and a loxP sequence (red arrowhead) upstream of a tag, flanked by HRs and (ii) a loxP sequence and a second selection marker, flanked by HRs, are generated as PCR fragments and transfected into *Leishmania* parasites constitutively expressing Cas9, T7 polymerase and DiCre subunits. Addition of rapamycin induces dimerization of the Cre subunits and subsequent excision of the floxed gene. (E) Alternative sgRNA targeting regions to preserve the UTRs of a target gene [27]. Screenshot from TriTrypDB [106] showing the RNA transcripts from the target, upstream and downstream genes. For gene deletion, two sgRNAs are designed in the 'least transcribed' regions surrounding the target gene to preserve the UTRs adjacent to the gene.

Gene functional analysis in the pre- and post-CRISPR-Cas9 eras

Gene Tagging

Protein localization provides functional clues to biological processes by delimiting the analysis to a given cell compartment and sometimes giving dynamic insights during the cell cycle. Endogenous proteins may be visualized using specific antibodies and immunofluorescence in fixed cells. However, the fixation and permeabilization of the cells can potentially cause artefacts and the production of specific antibodies may be difficult and time-consuming. Moreover, commercial antibodies against

protozoan parasite proteins are not readily available due to low consumer demand and antibodies developed for other eukaryotes rarely cross-react with trypanosomatid proteins. One alternative strategy is to express the target protein fused to a tag, such as GFP, enabling temporal subcellular localization studies directly in living cells. However, the engineering of a native protein involves the risks of artefactual localization and defects in cell fitness, partly depending on the position and size of the tag used [25-27]. The key point is to identify the

technical limitations of the chosen method and design appropriate controls to enable its interpretation. Several strategies were used for gene tagging in trypanosomatids in the pre-CRISPR-Cas9 era. Historically in *Leishmania* and *T. cruzi*, tagged copies of a protein of interest were constitutively expressed from plasmids that can be readily transfected and maintained as episomes [28, 29]. However, this approach has several drawbacks, among which are heterogeneous levels of expression of the tagged protein within cell populations [28]. This makes it difficult to be certain of the protein's subcellular localization or even causing it to be wrongly located. Several tools for protein tagging are available and widely used in *T. brucei* since the late 1990s (reviewed in [30]). In 2007, plasmids were developed which allowed endogenous tagging or integration in ribosomal RNA (rRNA) loci [31]. It is notable that ectopic expression in rRNA loci may be constitutive or inducible. As homologous recombination is efficient even with short homology regions (HRs) in *T. brucei*, this methodology was successfully applied for epitope tagging to a large number of proteins and is still widely used in this parasite. It has been shown that only 50 bp of homology are required for efficient HDR in *T. brucei* [32]. These results led to the development of a large toolkit of plasmids used as templates for long PCR primer strategies in *T. brucei* [33, 34]. Furthermore, an efficient and scalable long PCR primer approach [35] is currently used for whole genome tagging of *T. brucei* procyclic forms (TrypTag.orgⁱ: A Trypanosome Genome-wide Protein Localisation Resource [36]). Attempts to apply the long PCR primer tagging strategy used in *T. brucei* in *L. mexicana* were unsuccessful due to low recombination efficiency using short homology regions (<100bp) [35]. However, a strategy based on 500-bp homology regions and using PCR fusion of three fragments enabled efficient gene tagging in *Leishmania* [35, 37, 38]. Importantly, this strategy can be used in genetically unmodified *Leishmania* cell lines. The CRISPR-Cas9 approach developed by Beneke *et al.* [10] allowed a further reduction in the length of the homology arms to 30bp for efficient epitope tagging in both *Leishmania* and *T. brucei*. This methodology used genetically modified cell lines expressing the endonuclease *Streptococcus pyogenes* Cas9 (SpCas9) and T7 RNA polymerase (T7 RNAP). Briefly, Cas9 and

T7 RNAP are constitutively expressed and the sgRNA and donor DNA are PCR-generated without the need for molecular cloning (Figure 1A). PCR-generated donor DNA and sgRNA are maintained for only 48h in the transfected parasites. In *Leishmania*, transfection yielded edited cells at a rate of 1×10^{-5} per transfected cell, which is >10-fold more transfectants than when a non CRIPR-Cas9-based strategy is used. The authors also created LeishGEditⁱⁱ, a primer design tool to generate donor DNA and guide RNA sequences, which greatly facilitates the design of the tagging strategy. A similar system has recently been implemented in *T. cruzi* [39]. Here, the activity of the Cas9 combined with the PCR-generated donor DNA improves **editing efficiency** through the selective pressure to repair the DSB using the provided donor DNA. Efficient epitope tagging with integration of one drug resistance marker is achieved within non-clonal populations after a single round of transfection. This PCR-based strategy is fast, reliable and scalable, to a point where whole genome tagging and knock-out screens are now feasible in *Leishmania* [10, 18]. The main drawback of this strategy is the need for integration of selectable markers, which are limited in trypanosomatids, thus limiting further manipulation of the tagged strain. More importantly, when using this strategy, endogenous **untranslated regions (UTRs)** are actually interrupted by exogenous ones. As discussed in the last section of this paper, modification of UTRs of the target gene may result in nonfunctional products that may in turn affect cell fitness or the subcellular localization of the corresponding protein. To overcome this limitation, the so-called "stable plasmid-based" strategies prevent the modification of endogenous UTRs, thus avoiding the potential perturbation of gene expression. To tag the miltefosine transporter gene in *L. donovani*, cells expressing an episomal Cas9 and sgRNA (Table 1) were transfected with a PCR-generated donor DNA containing an eGFP tag flanked by 25bp HRs (Figure 2B). Transfected cells were subsequently subjected to drug selection with miltefosine; 1% of the MLF-resistant cells showed fluorescence by flow cytometry [9]. In *T. brucei*, a plasmid bearing donor DNA and encoding the sgRNA template was transfected into a strain expressing an episomal Cas9 (Figure 2A) [19]. The donor DNA was made of a tag flanked by long HRs (>200bp). This donor DNA carried a shield mutation to prevent further

recognition and editing of the targeted gene. Notably, this strategy is considered 'marker-free' with the selectable marker maintained on the episome and not integrated in the genome. Delayed dilution of transfected parasites resulted in 90% GFP-positive clones, as compared to ~10% after immediate dilution. In *T. cruzi*, endogenous gene tagging has been reported for the first time with the implementation of a stable CRISPR-Cas9 system [40]. Briefly, parasites were transfected with a single plasmid harboring the Cas9 and sgRNA that were constitutively expressed from an episome (Figure 2C). The donor DNA was provided as a PCR product containing 100bp HRs and a drug selection marker for DSB repair by HDR. Selection with a single drug resistance marker allowed the generation of homogenous cell populations after 4-5 weeks without cloning [40]. Furthermore, this system has been used for the generation of gene knock-ins, a result never achieved prior to the CRISPR-Cas9 era in *T. cruzi* [41]. Although the main advantage of this system is the high gene editing efficiency, it still requires the integration of a selection marker in the genome and constitutive expression of Cas9, which can be toxic in *T. cruzi* [7]. To overcome these limitations, a transient system has been developed involving the delivery by nucleofection of ribonucleoprotein (RNP) complexes, made of the *Staphylococcus aureus* Cas9 (SaCas9) and *in vitro*-transcribed sgRNAs, into different stages of *T. cruzi* [23] (Figure 1B and Table 1). This transient system allows endogenous tagging of the GP72 gene with 35% editing efficiency with donor DNA provided as a single-stranded oligonucleotide containing a 20nt sequence for DSB repair. To date, this is the only transient system in trypanosomatids based on the delivery of RNP complexes for CRISPR-Cas9 editing without the need of molecular cloning and drug selection. Since drug selection markers are limited in trypanosomatids, this transient selection-free system offers the potential to simultaneously edit multiple genes with a single transfection and to simply test several sgRNAs for each targeted gene. Given that editing efficiency is variable for this system (Table 1), which may be in part explained by the transient nature of the delivered Cas9-sgRNA complex, it is important to test several sgRNAs to achieve efficient genome editing. Overall, while the transient T7-PCR based strategy requires a selectable marker for the tag

donor, it eliminates the molecular cloning steps and provides a rapid and efficient tool particularly advantageous for large scale gene tagging. Yet, since the stable plasmid-based strategies are also capable of inserting marker-free tags into the genome, they are useful to ascertain the subcellular localization of proteins without modification of endogenous UTRs.

Box 2. Inducible knockout strategy using DiCre

Cre recombinase and LoxP (locus of X-over P1) are an enzyme and a sequence, respectively, naturally present in the bacteriophage P1. The Cre recombinase is able to catalyze site-specific DNA recombination between two LoxP sequences: either gene excision, if LoxP sequences are oriented in the same orientation, or gene flipping, if LoxP sequences are oriented in opposite direction. LoxP are 34bp sequences made of two 13bp palindromic sequences separated by an 8bp spacer. Several LoxP sequences deriving from the natural one have been produced; the sequence [ATAACTTCGTATAgcATagTATACGAAGTTAT] was used in *Leishmania* [27, 66]. In the DiCre strategy, the activity of the CRE enzyme is conditionally controlled by being split into two inactive subunits (CRE59 and CRE60) fused to a FK506-binding protein (FKBP-12) and to the binding domain of the FKBP12-rapamycin associated protein (FRB), respectively. The addition of rapamycin reconstitutes the enzymatic recombinase activity.

Box 3. CRISPR-Cas9 toxicity and off-target effects

Potential toxicity and off-target effects due to Cas9 expression may be a concern [7, 17, 84, 85]. Toxicity of Cas9 was reported for *T. cruzi* [7] and *T. brucei* [17]. So far, no study has reported off-target genome modification in trypanosomatid parasites expressing an active Cas9. The use of online tools to design sgRNA^{ii-vi} should in principle reduce the probability of off-target events. Softwares such as Protospacer Workbench software suite [94] can also be used to find and score putative off-targets and to tabulate the top 10 potential off-target sites. Since these precautions are based solely on *in silico* tests, the detection of off-target events should ideally be completed by whole genome sequencing of the mutant lines. Indeed, off-target effects may be dependent or independent of the sgRNA. Both are detected in whole genome sequencing as insertion/deletions (INDELS) but only sgRNA-dependent off-target events may be predicted by software. The sgRNA-dependent off-targets are due to the mispairing of the sgRNA and Cas9 complex in genomic regions displaying 12 bases identical to the target sequence. In total, preventing and detecting off-target events is a key challenge which should be taken into account (reviewed in [95]). Whole genome sequences should become a gold standard for controlling gene editing. The additional precaution of complementation (add-back) in knockout experiments, leading to at least partial phenotype recovery, is also highly recommended.

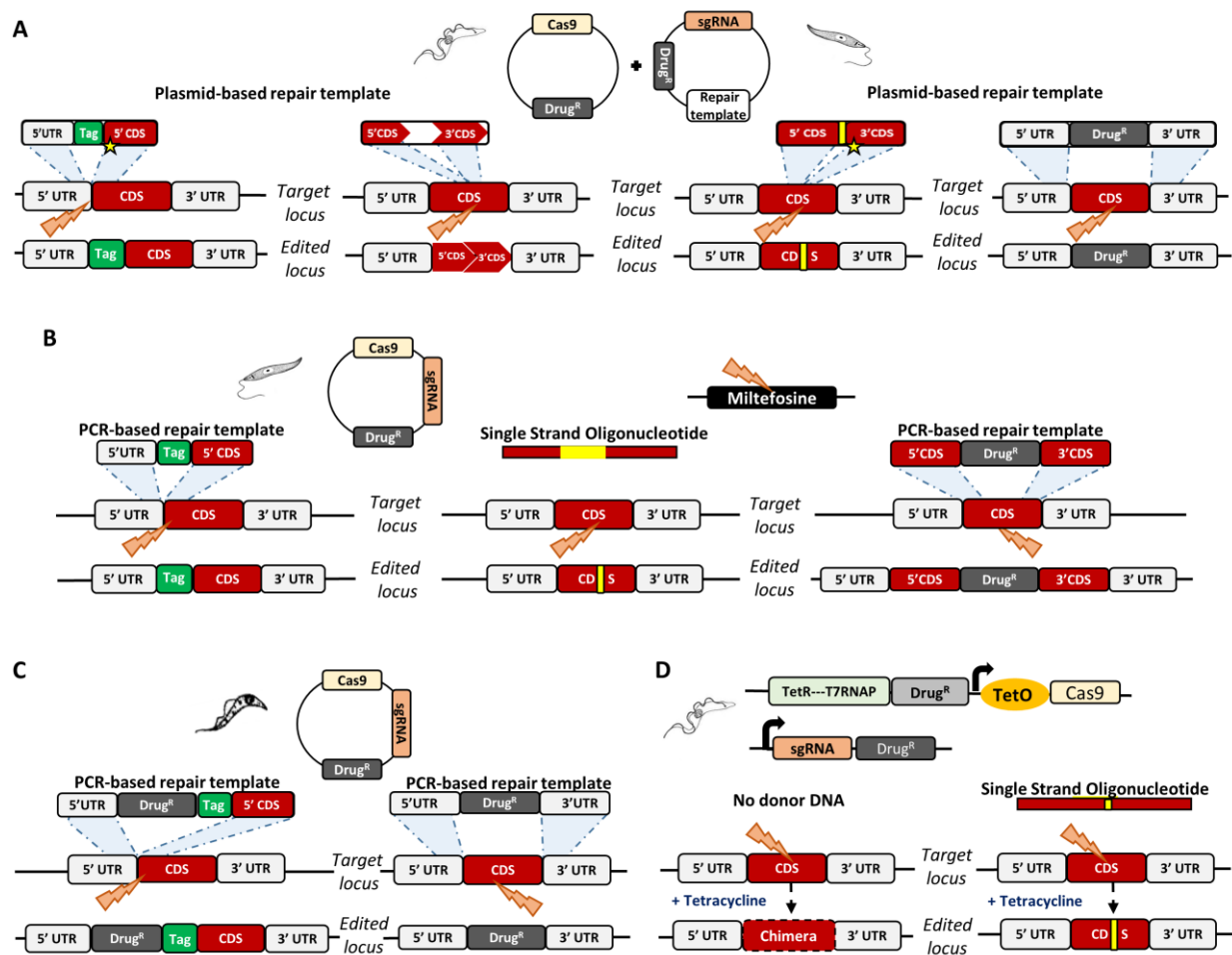


Figure 2. CRISPR-Cas9 stable strategies in trypanosomatids

(A) Plasmid-based strategy for gene tagging, gene disruption, gene deletion and single base editing in *T. brucei* and *Leishmania* [8, 19, 50]. Two plasmids encoding (i) Cas9 and (ii) sgRNA and the repair DNA templates are sequentially transfected into unmodified cell lines, resulting in gene editing without disturbing the endogenous UTRs. For single base editing, the repair template carries the mutation (yellow box) and a shield mutation (yellow star) to prevent further recognition of the edited gene. (B) Stable strategies for gene tagging and knockout in *Leishmania* [9, 15, 16, 107]. A plasmid encoding Cas9 and sgRNA is transfected in *Leishmania* cell lines. For gene tagging, the repair template is delivered as a PCR fragment containing a tag flanked by HRs. For single base editing, the repair template is provided as a single-stranded oligonucleotide encoding the intended mutation (yellow box) flanked by HRs. For gene knockout, the repair template is delivered as a PCR fragment encoding a selection marker flanked by HRs. Co-targeting the miltefosine transporter (MT) allows selection of cells with CRISPR-Cas9 activity. (C) PCR-based strategy for gene tagging and deletion in *T. cruzi* [108]. A plasmid-encoding Cas9 and sgRNA is co-transfected with a PCR-generated donor DNA encoding either a resistance marker with a tag, flanked by HRs (gene tagging), or a resistance marker flanked by HRs (gene deletion). (D) Inducible Cas9-based strategy for gene disruption and single base editing in *T. brucei*. Cas9, under the control of a Tet operator and T7 polymerase, is expressed upon addition of tetracycline. The sgRNA is constitutively expressed. For gene disruption, no donor DNA is provided, resulting in a chimeric gene likely produced by SSA. For single base editing, the repair template is provided as a single-stranded oligonucleotide encoding the intended mutation (yellow box) flanked with short HRs.

Single base genome editing

CRISPR-Cas9 technology has revolutionized the analysis of DNA point mutations. Single base genome editing is a key technique providing insights into biological processes, including post-translational modifications and enzymatic activities, the functional analysis of specific DNA/protein motifs, and improving our understanding of drug resistance mechanisms. In protozoan parasites, single-nucleotide substitution was achieved in *Plasmodium falciparum* using a CRISPR-Cas9 plasmid-based system with constitutive episomal expression of Cas9 and sgRNA [42]; the

authors were able to generate an artemisinin-resistant strain by introducing a single base polymorphism previously shown to be involved in this type of resistance. Drug resistance is also a concern for *T. cruzi* [43-45] and *Leishmania* [46-50]. In *Leishmania*, CRISPR-Cas9 mediating single base genome editing has been reported by two groups [9, 50]. The first one developed a single plasmid for co-expression of sgRNA and Cas9 by using the ribosomal RNA promoter to drive sgRNA expression; this can be used in several *Leishmania* species (Figure 2B)[9, 16]. By providing a 60 nt single-stranded oligonucleotide donor containing the desired

mutation and 25 nt-long homology arms (Figure 2B), stop codons and point mutations were introduced into the target gene, which resulted in disruption of the miltefosine transporter gene and a single conserved amino acid substitution (A189G) in the *L. donovani* RAD51 gene, a key recombinase in DNA repair [16]. Interestingly, likely resulting from an atypical DSB repair of MMEJ, a single point mutation (M381T) near the Cas9 cleavage site was detected in the 3-kb miltefosine transporter gene where no donor was provided [9]. The second group used a stable plasmid-based strategy with a donor DNA template encoding the intended mutation to explore drug resistance in *L. major* (Figure 2A) [50]. This strategy allowed the validation of a single amino acid substitution (H451Y) in the calcium-binding site of the orphan kinase SCAMK as a marker of antimony tolerance and resistance. This was subsequently confirmed by others in large chemical mutagenesis screens coupled with next-generation sequencing [51]. In *T. brucei*, a tetracycline-inducible Cas9 system was used to introduce single base editing in the *AQP2* gene encoding an aquaglyceroporin conferring resistance to the drug pentamidine with 100% efficiency (Figure 2D) (detailed in the Inducible gene editing section below)[17]. Overall, these strategies have proven efficient for the study of drug resistance mechanisms without the need to integrate a selectable marker into the genome; but these precise genome editing will be more difficult to detect when a drug cannot be used as a phenotypic read-out, particularly if multicopy genes are targeted. In this context, a promising tool for single base genome editing is the plasmid-based strategy seen above in *T. brucei* (Figure 2A) [19]. This approach allowed for the first time in *T. brucei* a precise single codon change in 90% of the multicopy histone H4 gene, a genetic manipulation hitherto impossible (Key Table 1).

Knockout of non-essential genes

A reverse genetic approach is crucial to address gene function. Prior to the CRISPR-Cas9 era, gene disruption in *Leishmania* was challenging due to low recombination efficiency and the presence of supernumerary chromosomes resulting from its unique genomic plasticity [1-3, 52]. Gene replacement using homologous recombination in *Leishmania* requires long homology arms of >500bp and several rounds of transfection to achieve editing

of all alleles if genes are present in supernumerary chromosomes [53]. In 2015, CRISPR-Cas9-mediated DSB using a stable strategy and HDR in *L. major* enabled the deletion of the multicopy *pfr2* gene in a single round of transfection with 25% editing efficiency (Figure 2A) [8]. In the same year, another proof of concept with another stable strategy was achieved by disrupting the miltefosine transporter in *L. donovani* [9]. In absence of donor DNA, large deletions ranging from 10 bp to more than 30 kb were obtained. To specifically insert stop codons and disrupt the target gene, a 60 nt single-stranded oligonucleotide was provided as a donor to repair DSB (Figure 2B). To facilitate isolation of the CRISPR disrupted mutants, a multidrug resistance protein-like gene was also disrupted by inserting a bleomycin resistance marker. This yielded 0.0025 to 0.005% resistant cells. In 2017, this stable strategy was used to delete 11 copies of the A2 gene in four months without any selection, and in six weeks by co-targeting the *L. donovani* miltefosine transporter gene followed by miltefosine selection to co-select for cells with CRISPR-Cas9 activity [16]. This stable strategy can also be used to express double or triple sgRNAs to improve gene targeting efficiency, to generate precise gene deletion mutants and to create specific chromosome translocations [9,16]. Although stable strategies [8, 9] (Table 1) have provided the proof of concept for CRISPR-Cas9 editing in *Leishmania*, they require cellular cloning and are time-consuming. They are also now challenged by the transient PCR-based strategy described above for tagging [10] (Figure 1A). This system requires constitutive expression of Cas-9 and T7 RNAP; sgRNA and donor DNAs are PCR-generated (Figure 1C). Recently, the same group further applied their strategy to a high-throughput knockout screen for the study of the flagellar proteome of *Leishmania* [18]. This was the first large-scale knockout screen in *Leishmania* (> 100 genes) demonstrating the speed and efficiency of this approach for the study of loss-of-function phenotypes for large groups of non-essential *Leishmania* genes. However, this transient system requires two selectable markers to achieve 100% editing efficiency in a single round of transfection. Although cellular cloning may save one selectable marker, this approach renders additional genetic manipulation difficult because of the need to express T7 polymerase and Cas9.

Moreover, this system is probably inefficient for editing multicopy genes dispersed over multiple loci. Regarding *T. brucei*, a CRISPR-Cas9 plasmid-based strategy mediating gene deletion in *T. brucei* without integration of a selectable marker was set up in 2018 [19] (Figure 2A). It allowed the deletion of both alleles of histone variants H3.V and H4.V in 100% and 66% of the clones, respectively. Interestingly, episomes can be removed from *T. brucei* through the generation of a DSB in the selectable marker, leading to its loss rather than repair by microhomology. Nevertheless, as with protein tagging, the transient PCR-based system provides a rapid and molecular cloning-free strategy with highly efficient editing of non-essential genes in both *T. brucei* and *Leishmania* spp.

Concerning *T. cruzi*, the first work allowing CRISPR-Cas9 mediated gene disruption was reported with the implementation of a transient system [7]. Based on the transfection of *in vitro* transcribed sgRNA in parasites constitutively expressing SpCas9, this system allowed downregulation of single and multicopy genes in the absence of donor DNA for DSB repair [7]. In this transient system, the constitutive expression of Cas9 was toxic to transfected parasites. To overcome these limitations, the same group improved this system with the delivery of RNP complexes [23] (Figure 1B). As explained in the Gene Tagging section, this cloning- and selection-free system is based on the transient expression of both Cas9 and sgRNA (Table 1). This strategy, providing donor DNA as a single-stranded oligonucleotide, allowed endogenous and exogenous reporter gene disruption with higher editing efficiency and no Cas9 toxicity. However, as explained above, depending on the choice of the sgRNA and the stability of the provided RNP complex, the editing efficiency of this transient system is variable, highlighting the importance of testing several sgRNAs for each targeted gene. A stable strategy involving the constitutive expression of both sgRNA and Cas9 also exists for *T. cruzi* [22]. Parasites can be transfected with either two vectors separately encoding Cas9 and sgRNA or a single vector co-expressing both Cas9 and sgRNA driven by the ribosomal promoter (Figure 2C). This strategy yielded 100% editing efficiency in non-clonal populations within 4-5 weeks for both gene deletion and tagging with donor DNA provided as a PCR product and without any Cas9 toxicity

[22, 40]. This system is widely used for gene editing in *T. cruzi* [41, 54-58] and was adapted for different strains and parasite stages [59-61] (Table 1). In addition to the transient system based on *in vitro* transcribed sgRNA [7, 23], the transient PCR-based system [10] (Figure 1C) was further adapted to *T. cruzi* [39]. Similarly to *Leishmania* and *T. brucei*, this system produced high editing efficiency for the deletion of non-essential genes in this parasite.

Tackling essential genes: inducible gene editing

Molecular biologists cannot limit themselves to the study of non-essential genes. Essential genes underpin core biological processes required for parasite viability. Functional gene analyses in trypanosomatids may contribute to the discovery and validation of new drug targets [21] and to the study of basic aspects of eukaryotic biology and evolution. The definition of gene essentiality is subject to debate: for example, whether a gene is required for one or all life cycle stages, or necessary only in specific growth conditions (see also Box 1). To get a better insight into the function and essentiality of a protein, robust and efficient inducible tools are required to carry out phenotypic analyses. Apart from essential genes, since genome plasticity is a hallmark in trypanosomatids, efficient inducible systems may also be useful to ascertain gene function by avoiding: (i) the recovery of ‘adapted’ parasites after the knockout of non-essential genes; (ii) phenotypes arising from a combination of summed up deleterious events observed after the selection of stable transfectants. *T. brucei* studies have largely benefited from the possibility of silencing genes by RNAi. This technology indeed allowed high-throughput phenotyping using genome-scale RNAi screens in both *T. brucei* bloodstream and insect stages [62]. However, target mRNA depletion is often partial, and may underestimate the functional importance of a gene for which complete depletion could induce more deleterious phenotypes. Recently, a CRISPR-Cas9 system using tetracycline-inducible expression of Cas9 was developed to study essential genes in *T. brucei* [17]. The sequence of Cas9 was codon-optimized to increase expression in *T. brucei* and transcription of sgRNA and Cas9 was driven by T7 and rRNA promoters, respectively (Figure 2D). Notably, all these elements were integrated in non-transcribed rDNA spacer loci. The authors disrupted an amino acid transporter

(AAT6) and the aquaglyceroporin (*AQP2*) genes conferring resistance to eflornithine and pentamidine, respectively, without genomic integration of a selectable marker. To date, this is the only system using inducible expression of Cas9 in trypanosomatids (Table 1). Several attempts to develop an efficient inducible system have been reported in *Leishmania*. Tetracycline-inducible expression of tagged genes has been reported for various *Leishmania* species [63, 64]. Although compatible with *in vivo* studies, the decreased expression of T7 polymerase in metacyclic promastigotes and amastigotes hinders the correct interpretation of results [64]. However, this technical issue may be resolved by using suitable UTRs for expression in these lifecycle stages. In addition, attempts to conditionally regulate gene expression using T7 RNA polymerase were unsuccessful due to high background activity and heterogeneous expression levels [65]. The first system allowing an inducible deletion of essential genes in *Leishmania* is a DiCre-based inducible system [66]. This strategy is based on the constitutive expression of two inactive subunits of the Cre recombinase which promotes the deletion of a target DNA sequence when activated by rapamycin (see Box 2). The replacement of one allele by the DiCre subunits or by a selectable marker gene, coupled with the integration of LoxP sequences flanking the second allele, allowed several essential genes to be studied: CRK3 [66], Rad9 [67], Hus1 [68] and pNT1 cysteine peptidase [20]. Still, this system is time-consuming and cumbersome, requiring molecular cloning and several rounds of transfection. In order to improve the DiCre system and to accelerate the characterization of *Leishmania* essential genes, we have recently generated a 'universal' system combining the DiCre strategy with the advantages of CRISPR-Cas9 technology [27]. Such a combination has already been successfully proposed in *P. falciparum* [69]. In *Leishmania*, the strategy derives from the transient PCR-based strategy for gene tagging [10] and uses *Leishmania* cell lines constitutively expressing Cas9, T7 polymerase and the DiCre subunits (Figure 1D). Cas9 endonuclease activity accelerates the process of LoxP integration (floxed target gene) in a single round of transfection. As a proof of concept, an essential gene, CRK3 [66] and the non-essential gene PF16 [10] were deleted. Floxed target gene-edited parasites were obtained 8-10 days after transfection and cell

cloning, representing a significant reduction in time compared to the previous multistep process [66]. Introducing a tag along with the LoxP sequences can be useful to determine a protein's localization before rapamycin induction and to follow its depletion by using immunofluorescence and western blot. Nevertheless, caution should be exercised with rapamycin induction. In our experience with *Leishmania*, whereas complete deletion of non-essential genes, such as PF16, could be readily achieved, for essential genes, such as CRK3, at least one copy of the floxed target gene persisted in approximately 10% of the cell population, even after several days of rapamycin induction. The nature of this 'resistance mechanism' is still unclear and has been observed by others [20, 66, 69]. A long-term rapamycin mediated gene deletion could thus lead to the recovery of parasites in the cell population still carrying the floxed target gene [20]. However, phenotypic analysis can be done early after rapamycin induction when 90% of the population has already lost the target gene. In addition, if the floxed gene has been tagged, phenotypic analysis may focus on negative fluorescent cells.

Inducible systems for the study of essential genes are scarce in *T. cruzi*. The destabilization domain of dihydrofolate reductase was only used to develop inducible suicidal vectors. Parasites bearing this system proved to be attenuated in mice, providing protection against lethal infection [70]. The DiCre system was adapted for the removal of exogenous selectable markers integrated in the genome, but with limited success: after rapamycin induction, removal of the selectable marker was 40-75% efficient in clonal populations [71]. Recently, a CRISPR-Cas9-Riboswitch system was developed for the downregulation of gene expression in *T. cruzi* [72]. Using CRISPR-Cas9 mediated endogenous tagging, the *glmS* ribozyme from *Bacillus subtilis* was introduced in both alleles of the glycoprotein GP72 and vacuolar proton pyrophosphatase VP1 genes after one round of transfection. Downregulation was achieved for both genes, making this strategy potentially useful for knocking down essential genes and for the validation of potential drug targets in *T. cruzi*.

Getting the right cut: Importance of the UTRs in trypanosomatids

Trypanosomatids have an unconventional gene expression from large **polycistronic**

transcription units and generally without any 'classical' promoters [73-75]. Consequently, in contrast to higher eukaryotes, gene expression in trypanosomatids is regulated predominantly by post-transcriptional mechanisms (Reviewed in [76]). Regulatory elements act on the stability of mRNAs, determining transcript abundance, and hence the amount of protein [74, 75, 77]. In *T. brucei*, the 3'-UTR is considered the main regulator of gene expression [58,61-64]; but regulatory elements have also been described in the 5' UTRs [78]. Likewise, several studies in *Leishmania* and *T. cruzi* have described the presence of regulatory elements in the UTRs of several genes influencing mRNA stability and translation efficiency [79-81]. Moreover, putative non-coding RNAs (ncRNAs) which are generally regulatory elements known in other eukaryotic models have been detected in UTRs of *Leishmania* genes [82]. Taken together, these works demonstrate the importance of UTRs for gene expression in trypanosomatids. This may potentially explain cases of genome editing

failures or unexplained cell fitness phenotypes when the endogenous UTRs of a target gene are modified [25, 26, 83] (Box 1). As an example, while developing the inducible DiCre-CRISPR-Cas9 (Di3C) system [27], modification of the endogenous 5'UTR of the PF16 gene affected the localization and function of this flagellar protein. As previously explained, the issue of UTR modification during gene tagging can be avoided using plasmid-based strategies or the ribonucleoprotein system, allowing 'marker-free' editing which preserves endogenous UTRs, in *Leishmania* [8, 9, 16, 23], *T. cruzi* [22, 23, 40] and *T. brucei* [19]. Furthermore, to preserve UTR sequences of the GOI, we have developed an alternative strategy using transcriptomic data for targeting regions with the least abundant transcript level in the intergenic regions (Fig. 1E) [27], to integrate the selectable marker and the LoxP sequences. This targeting strategy has increased editing success in the Di3C inducible system, but it may be adapted for any transient PCR-based CRISPR-Cas9 knockout strategy

Concluding Remarks

Implementation of CRISPR-Cas9 in trypanosomatids has remarkably improved gene editing in these important pathogens, particularly in *Leishmania* and *T. cruzi* where genetic manipulation has been historically challenging. Since 2014, several transient and stable strategies have been developed yielding heterogeneous gene editing efficiencies (Table 1). Several factors are critical to achieving highly efficient gene editing, including the selection, delivery and expression of the sgRNA/Cas9 complex and the repair mechanism. Based on the current literature (Key Table 1), increased editing efficiency in non-clonal populations is obtained with systems involving the integration of a selective marker in the genome. The transient PCR-based system [10] and the stable plasmid-based strategy in *T. cruzi* [22, 40] provide efficient and rapid tools for endogenous gene tagging and deletion of single copy non-essential genes. 'Marker-free' strategies avoiding modification of endogenous UTRs, such as the stable plasmid-based strategies, are suitable for the study of multicopy genes, gene families, and single base genome editing [9, 16, 19, 50]. Nevertheless, without selective markers, cell sorting or selection of clonal populations are required for a completely edited cell population. In this regard, the Cas9/sgRNA ribonucleoprotein complexes may be further explored in trypanosomatids (i) to improve editing efficiency and (ii) to overcome potential Cas9 toxicity and off-target editing in the genome [7, 17, 84, 85] (See Box 3 and Key Table1). Moreover, careful selection of sgRNA is a critical step for efficient genome editing and several sgRNAs should be tested since current algorithms may be inefficient to predict sgRNA activities in trypanosomatids [15]. Inducible strategies are now available for the study of essential genes in *Leishmania* and *T. brucei*. Although a CRISPR-Cas9 inducible system is now available in *T. brucei*, essential genes have not yet been studied with this system [17]. Further improvements are needed to expand the toolkit for the study of essential genes, particularly in *T. cruzi*. In this regard, the development of a Di3C strategy may represent an alternative approach to the RNAi system in *T. brucei* and to the riboswitch system in *T. cruzi*. The ability of CRISPR-Cas9 to precisely edit trypanosomatid genomes has largely opened the field to deciphering mechanisms of host-pathogen interactions, key cellular pathways, drug targets and drug resistance mechanisms. So far, CRISPR-Cas9 has already allowed large-scale knockout screening in *Leishmania* [18] and the study of potential drug targets [20, 50]. However, the CRISPR-Cas revolution also has many other applications in other distinct fields, such as RNA editing, applications of deactivated Cas9 (dCas9), and development of accurate and portable diagnostic tests. A novel application such as the CRISPR chromatin affinity purification with mass spectrometry may also shed light into the dynamics of proteins involved in DNA

damage and repair after a DSB in *T. brucei* [86]. If most of the other applications of CRISPR-Cas have not yet been used in trypanosomatids [87, 88], it is clear that the CRISPR-Cas-directed approaches will continue to expand in these divergent organisms.

References

- Bastien, P. et al. (1990) Interclonal Variations in Molecular Karyotype in *Leishmania Infantum* Imply a 'Mosaic' Strain Structure. *Mol Biochem Parasitol* 40 (1), 53-61.
- Cruz, A.K. et al. (1993) Plasticity in chromosome number and testing of essential genes in *Leishmania* by targeting. *Proc Natl Acad Sci U S A* 90 (4), 1599-603.
- Martinez-Calvillo, S. et al. (2005) Ploidy changes associated with disruption of two adjacent genes on *Leishmania major* chromosome 1. *Int J Parasitol* 35 (4), 419-29.
- Ishemgulova, A. et al. (2018) CRISPR/Cas9 in *Leishmania mexicana*: A case study of LmxBTN1. *PLoS One* 13 (2), e0192723.
- Kolev, N.G. et al. (2011) RNA interference in protozoan parasites: achievements and challenges. *Eukaryot Cell* 10 (9), 1156-63.
- Matthews, K.R. (2015) 25 Years of African Trypanosome Research: From Description to Molecular Dissection and New Drug Discovery. *Mol Biochem Parasitol* 200 (1-2), 30-40.
- Peng, D. et al. (2014) CRISPR-Cas9-mediated single-gene and gene family disruption in *Trypanosoma cruzi*. *mBio* 6 (1), e02097-14.
- Sollesis, L. et al. (2015) First efficient CRISPR-Cas9-mediated genome editing in *Leishmania* parasites. *Cell Microbiol* 17 (10), 1405-12.
- Zhang, W.W. and Matlashewski, G. (2015) CRISPR-Cas9-Mediated Genome Editing in *Leishmania donovani*. *mBio* 6 (4), e00861.
- Beneke, T. et al. (2017) A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. *R Soc Open Sci* 4 (5), 170095.
- Kelso, A.A. et al. (2017) Homologous Recombination in Protozoan Parasites and Recombinase Inhibitors. *Front Microbiol* 8 (01716).
- Burton, P. et al. (2007) Ku Heterodimer-Independent End Joining in *Trypanosoma Brucei* Cell Extracts Relies Upon Sequence Microhomology. *Eukaryot Cell* 6 (10), 1773-81.
- Glover, L. et al. (2011) Microhomology-mediated deletion and gene conversion in African trypanosomes. *Nucleic Acids Res* 39 (4), 1372-80.
- Glover, L. et al. (2008) Sequence Homology and Microhomology Dominate Chromosomal Double-Strand Break Repair in African Trypanosomes. *Nucleic Acids Res* 36 (8), 2608-18.
- Zhang, W.W. and Matlashewski, G. (2019) Single-Strand Annealing Plays a Major Role in Double-Strand DNA Break Repair Following CRISPR-Cas9 Cleavage in *Leishmania*. *mSphere* 4 (4), e00408-19.
- Zhang, W.W. et al. (2017) 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 (1).
- Rico, E. et al. (2018) Inducible High-Efficiency CRISPR-Cas9-targeted Gene Editing and Precision Base Editing in African Trypanosomes. *Scientific Rep* 8 (1), 7960.
- Beneke, T. et al. (2019) Genetic dissection of a *Leishmania* flagellar proteome demonstrates requirement for directional motility in sand fly infections. *PLoS Pathog* 15 (6), e1007828.
- Vasquez, J.J. et al. (2018) Exploiting CRISPR-Cas9 technology to investigate individual histone modifications. *Nucleic Acids Res* 46 (18), e106.
- Grewal, J.S. et al. (2019) Evaluation of Clan CD C11 Peptidase PNT1 and Other *Leishmania Mexicana* Cysteine Peptidases as Potential Drug Targets. *Biochimie* 166, 150-160.
- Jones, N.G. et al. (2018) Genetically Validated Drug Targets in *Leishmania*: Current Knowledge and Future Prospects. *ACS Infect Dis* 4 (4), 467-477.
- Lander, N. et al. (2015) CRISPR/Cas9-Induced Disruption of Paraflagellar Rod Protein 1 and 2 Genes in *Trypanosoma Cruzi* Reveals Their Role in Flagellar Attachment. *mBio* 6 (4), e01012.
- Soares Medeiros, L.C. et al. (2017) Rapid, Selection-Free, High-Efficiency Genome Editing in Protozoan Parasites Using CRISPR-Cas9 Ribonucleoproteins. *mBio* 8 (6).
- Fernandez-Prada, C. et al. (2018) High-throughput Cos-Seq screen with intracellular *Leishmania infantum* for the discovery of novel drug-resistance mechanisms. *Int J Parasitol Drugs Drug Resist* 8 (2), 165-173.
- Albisetti, A. et al. (2017) Interaction between the flagellar pocket collar and the hook complex via a novel microtubule-binding protein in *Trypanosoma brucei*. *PLoS Pathog* 13 (11), e1006710.
- Ahuja, K. et al. (2018) A Novel Signal Sequence Negative Multimeric Glycosomal Protein Required for Cell Cycle Progression of *Leishmania Donovanii* Parasites. *Biochim Biophys Acta Mol Cell Res* 1865 (8), 1148-59.
- Yagoubat, A. et al. (2020) Universal Highly Efficient Conditional Knockout System in *Leishmania*, With a Focus on Untranscribed Region Preservation. *Cell Microbiol* 22 (5), e13159.
- Tetaud, E. et al. (2002) A New Expression Vector for *Crithidia Fasciculata* and *Leishmania*. *Mol Biochem Parasitol* 120 (2), 195-204.
- Burle-Caldas, G.d.A. et al. (2015) Expanding the Tool Box for Genetic Manipulation of *Trypanosoma Cruzi*. *Mol Biochem Parasitol* 203 (1-2), 25-33.
- Clayton, C. (1999) Genetic Manipulation of Kinetoplastida. *Parasitology Today* 15 (9).
- Kelly, S. et al. (2007) Functional genomics in *Trypanosoma brucei*: a collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. *Mol Biochem Parasitol* 154 (1), 103-9.
- Barnes, R.L. and McCulloch, R. (2007) *Trypanosoma Brucei* Homologous Recombination Is Dependent on Substrate Length and Homology, Though Displays a Differential Dependence on Mismatch Repair as Substrate Length Decreases. *Nucleic Acids Res* 35 (10), 3478-93.
- Oberholzer, M. et al. (2006) A vector series for rapid PCR-mediated C-terminal in situ tagging of *Trypanosoma brucei* genes. *Mol Biochem Parasitol* 145 (1), 117-20.
- Arhin, G.K. et al. (2004) A PCR-based Method for Gene Deletion and Protein Tagging in *Trypanosoma Brucei*. *Methods Mol Biol* 270, 277-86.
- Dean, S. et al. (2015) A Toolkit Enabling Efficient, Scalable and Reproducible Gene Tagging in Trypanosomatids. *Open Biol* 5 (1), 140197.
- Dyer, P. et al. (2016) High-throughput Gene Tagging in *Trypanosoma Brucei*. *J Vis Exp* (114), 54342.
- Wheeler, R. et al. (2016) Flagellar Pocket Restructuring Through the *Leishmania* Life Cycle Involves a Discrete Flagellum Attachment Zone. *J Cell Sci* 129 (4), 854-67.
- Sunter, J. et al. (2019) *Leishmania* Flagellum Attachment Zone Is Critical for Flagellar Pocket Shape, Development in the Sand Fly, and Pathogenicity in the Host. *Proc Natl Acad Sci U S A* 116 (13), 6352_6360.
- Costa, F.C. et al. (2018) Expanding the Toolbox for *Trypanosoma Cruzi*: A Parasite Line Incorporating a Bioluminescence-Fluorescence Dual Reporter and Streamlined CRISPR/Cas9 Functionality for Rapid in Vivo Localisation and Phenotyping. *PLoS Negl Trop Dis* 12 (4), e0006388.
- Lander, N. et al. (2016) CRISPR/Cas9-mediated Endogenous C-terminal Tagging of *Trypanosoma Cruzi* Genes Reveals the Acidocalcisome Localization of the Inositol 1,4,5-Trisphosphate Receptor. *J Biol Chem* 291 (49), 25505-25515.
- Chiurillo, M.A. et al. (2019) Functional Analysis and Importance for Host Cell Infection of the Ca²⁺-conducting Subunits of the Mitochondrial Calcium Uniporter of *Trypanosoma cruzi*. *Mol Biol Cell* 30 (14), 1676-1690.
- Ghorbal, M. et al. (2014) Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-

- Cas9 system. *Nat Biotechnol* 32 (8), 819-21.
43. Campos, M. et al. (2015) "Benznidazole-resistance in *Trypanosoma Cruzi*: Evidence That Distinct Mechanisms Can Act in Concert" [*Mol. Biochem. Parasit.* (2014) 193, 17-19]. *Mol. Biochem. Parasit.* 201 (1), 83.
44. Zingales, B. et al. (2015) A Novel ABCG-like Transporter of *Trypanosoma Cruzi* Is Involved in Natural Resistance to Benznidazole. *Mem Inst Oswaldo Cruz* 110 (3), 433-34.
45. García-Huertas, P. et al. (2017) Prostaglandin F2 α Synthase in *Trypanosoma cruzi* Plays Critical Roles in Oxidative Stress and Susceptibility to Benznidazole. *R Soc Open Sci* 4 (9), 170773.
46. Das, V. et al. (2005) Magnitude of Unresponsiveness to Sodium Stibogluconate in the Treatment of Visceral Leishmaniasis in Bihar. *Natl Med J India* 18 (3), 131-3.
47. Rijal, S. et al. (2013) Increasing Failure of Miltefosine in the Treatment of Kala-azar in Nepal and the Potential Role of Parasite Drug Resistance, Reinfection, or Noncompliance. *Clin Infect Dis* 56 (11), 1530-8.
48. Srivastava, S. et al. (2017) Laboratory Confirmed Miltefosine Resistant Cases of Visceral Leishmaniasis From India. *Parasit Vectors* 10 (1), 49.
49. Dumetz, F. et al. (2018) Molecular Preadaptation to Antimony Resistance in *Leishmania Donovanii* on the Indian Subcontinent. *mSphere* 3 (2), e00548-17.
50. Vergnes, B. et al. (2019) A Single Amino Acid Substitution (H451Y) in *Leishmania* Calcium-Dependent Kinase SCAMK Confers High Tolerance and Resistance to Antimony. *J Antimicrob Chemother* 74 (11), 3231-3239.
51. Bhattacharya, A. et al. (2019) Coupling Chemical Mutagenesis to Next Generation Sequencing for the Identification of Drug Resistance Mutations in *Leishmania*. *Nature Commun* 10 (1), 5627.
52. Sterkers, Y. et al. (2011) FISH analysis reveals aneuploidy and continual generation of chromosomal mosaicism in *Leishmania major*. *Cell Microbiol* 13 (2), 274-83.
53. Dubessay, P. et al. (2002) The switch region on *Leishmania major* chromosome 1 is not required for mitotic stability or gene expression, but appears to be essential. *Nucleic Acids Res* 30, 3692-7.
54. Chiurillo, M. et al. (2017) Different Roles of Mitochondrial Calcium Uniporter Complex Subunits in Growth and Infectivity of *Trypanosoma cruzi*. *mBio* 8 (3), e00574-17.
55. Lander, N. et al. (2018) Calcium-sensitive Pyruvate Dehydrogenase Phosphatase Is Required for Energy Metabolism, Growth, Differentiation, and Infectivity of *Trypanosoma cruzi*. *J Biol Chem* 293 (45), 17402-17417.
56. Lander, N. et al. (2020) CRISPR/Cas9 Technology Applied to the Study of Proteins Involved in Calcium Signaling in *Trypanosoma Cruzi*. *Methods Mol Biol* 2116, 177-197.
57. Cruz-Bustos, T. et al. (2018) An Intracellular Ammonium Transporter Is Necessary for Replication, Differentiation, and Resistance to Starvation and Osmotic Stress in *Trypanosoma cruzi*. *mSphere* 3 (1), e00377-17.
58. Bertolini, M. et al. (2019) MICU1 and MICU2 Play an Essential Role in Mitochondrial Ca²⁺ Uptake, Growth, and Infectivity of the Human Pathogen *Trypanosoma Cruzi*. *mBio* 10 (3), e00348-19.
59. Romagnoli, B. et al. (2018) Improvements in the CRISPR/Cas9 System for High Efficiency Gene Disruption in *Trypanosoma Cruzi*. *Acta Trop* 178, 190-195.
60. Burle-Caldas, G. et al. (2018) Assessment of Two CRISPR-Cas9 Genome Editing Protocols for Rapid Generation of *Trypanosoma Cruzi* Gene Knockout Mutants. *Int J Parasitol* 48 (8), 591-596.
61. Akutsu, Y. et al. (2019) Introducing a Gene Knockout Directly Into the Amastigote Stage of *Trypanosoma Cruzi* Using the CRISPR/Cas9 System. *Journal of visualized experiments : JoVE* (149).
62. Alsford, S. et al. (2011) High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res* 21 (6), 915-24.
63. Yan, S. et al. (2001) Tetracycline Regulated Gene Expression in *Leishmania* *Donovani*. *Mol Biochem Parasitol* 112 (1), 61-9.
64. Ishemgulova, A. et al. (2016) T7 Polymerase-Driven Transcription Is Downregulated in Metacyclic Promastigotes and Amastigotes of *Leishmania Mexicana*. *Folia Parasitol* 63, 2016.016.
65. Yan, S. et al. (2002) A Low-Background Inducible Promoter System in *Leishmania Donovanii*. *Mol Biochem Parasitol* 119 (2), 217-23.
66. Duncan, S.M. et al. (2016) Conditional gene deletion with DiCre demonstrates an essential role for CRK3 in *Leishmania mexicana* cell cycle regulation. *Mol Microbiol* 100 (6), 931-44.
67. Santos, R. et al. (2017) A DiCre recombinase-based system for inducible expression in *Leishmania major*. *Mol Biochem Parasitol* 216, 45-48.
68. Damasceno, J.D. et al. (2018) Conditional genome engineering reveals canonical and divergent roles for the Hus1 component of the 9-1-1 complex in the maintenance of the plastic genome of *Leishmania*. *Nucleic Acids Res* 46 (22), 11835-11846.
69. Knuepfer, E. et al. (2017) Generating Conditional Gene Knockouts in *Plasmodium* - A Toolkit to Produce Stable DiCre Recombinase-Expressing Parasite Lines Using CRISPR/Cas9. *Sci Rep* 7 (1), 3881.
70. Ma, Y. et al. (2015) Inducible Suicide Vector Systems for *Trypanosoma Cruzi*. *Microbes Infect* 17 (6), 440-50.
71. Kangussu-Marcolino, M. et al. (2014) Conditional Removal of Selectable Markers in *Trypanosoma Cruzi* Using a Site-Specific Recombination Tool: Proof of Concept. *Mol Biochem Parasitol* 198 (2), 71-4.
72. Lander, N. et al. (2020) A CRISPR/Cas9-riboswitch-Based Method for Downregulation of Gene Expression in *Trypanosoma cruzi*. *Front Cell Infect Microbiol* 10, 68.
73. Vanhamme, L. and Pays, E. (1995) Control of Gene Expression in Trypanosomes. *Microbiol Rev* 59 (2), 223-40.
74. Clayton, C.E. (2002) Life without transcriptional control? From fly to man and back again. *EMBO J* 21 (8), 1881-8.
75. Cohen-Freue, G. et al. (2007) Global gene expression in *Leishmania*. *Int J Parasitol* 37 (10), 1077-86.
76. Clayton, C. (2016) Gene Expression in Kinetoplastids. *Current opinion in microbiology* 32, 46-51.
77. Quijada, L. et al. (2002) Expression of the Human RNA-binding Protein HuR in *Trypanosoma Brucei* Increases the Abundance of mRNAs Containing AU-rich Regulatory Elements. *Nucleic Acids Res* 30 (20), 4414-24.
78. Pasion, S.G. et al. (1996) Sequences within the 5' untranslated region regulate the levels of a kinetoplast DNA topoisomerase mRNA during the cell cycle. *Mol Cell Biol* 16 (12), 6724-35.
79. Campbell, D. and Sturm, N. (2007) The Untranslated Regions of Genes From *Trypanosoma Cruzi*: Perspectives for Functional Characterization of Strains and Isolates. *Mem Inst Oswaldo Cruz* 102 (1), 125-6.
80. Saxena, A. et al. (2007) Analysis of the *Leishmania Donovanii* Transcriptome Reveals an Ordered Progression of Transient and Permanent Changes in Gene Expression During Differentiation. *Mol Biochem Parasitol* 152 (1), 53-65.
81. Queiroz, R. et al. (2009) Transcriptome Analysis of Differentiating Trypanosomes Reveals the Existence of Multiple Post-Transcriptional Regulons. *BMC Genomics* 10 (495).
82. Freitas Castro, F. et al. (2017) Evidence of putative non-coding RNAs from *Leishmania* untranslated regions. *Mol Biochem Parasitol* 214, 69-74.
83. Florimond, C. et al. (2015) Correction: BILBO1 Is a Scaffold Protein of the Flagellar Pocket Collar in the Pathogen *Trypanosoma Brucei*. *PLoS Pathog* 11 (4), e1004844.
84. Jiang, W. et al. (2014) Successful Transient Expression of Cas9 and Single Guide RNA Genes in *Chlamydomonas reinhardtii*. *Eukaryot Cell* 13 (11), 1465-9.
85. Kim, S. et al. (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res* 24 (6), 1012-9.
86. Glover, L. (2020) mSphere of Influence: Expanding the CRISPR Sphere With Single-Locus Proteomics. *mSphere* 5 (1), e00001-20.
87. Bryant, J. et al. (2019) CRISPR in Parasitology: Not Exactly Cut and Dried! *Trends Parasitol* 35 (6), 409-422.
88. Gupta, D. et al. (2019) CRISPR-Cas9 System: A New-Fangled Dawn in Gene Editing. *Life Sci* 232, 116636.
89. Peng, D. and Tarleton, R. (2015)

- EuPaGDT: A Web Tool Tailored to Design CRISPR Guide RNAs for Eukaryotic Pathogens. *Microb Genom* 1 (4), e000033.
90. Moreno-Mateos, M. et al. (2015) CRISPRscan: Designing Highly Efficient sgRNAs for CRISPR-Cas9 Targeting in Vivo. *Nat Methods* 12 (10), 982-8.
91. Xu, H. et al. (2015) Sequence Determinants of Improved CRISPR sgRNA Design. *Genome Res* 25 (8), 1147-57.
92. Stemmer, M. et al. (2015) CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One* 10 (4), e0124633.
93. Labuhn, M. et al. (2018) Refined sgRNA Efficacy Prediction Improves Large- And Small-Scale CRISPR-Cas9 Applications. *Nucleic Acids Res* 46 (3), 1375-1385.
94. MacPherson, C.R. and Scherf, A. (2015) Protospacer Workbench: flexible guide-RNA design for CRISPR applications. *Nat. Biotechnol* 33(8), 805-6..
95. Li, J. et al. (2019) Advances in Detecting and Reducing Off-Target Effects Generated by CRISPR-mediated Genome Editing. *J Genet Genomics* 46 (11), 513-521.
96. Martel, D. et al. (2017) Characterisation of Casein Kinase 1.1 in *Leishmania donovani* Using the CRISPR Cas9 Toolkit. *Biomed Res Int* 2017, 4635605.
97. Halliday, C. et al. (2019) Cellular Landmarks of *Trypanosoma Brucei* and *Leishmania Mexicana*. *Mol Biochem Parasitol* 230, 24-36.
98. Schädeli, D. et al. (2019) Cardiolipin Depletion-Induced Changes in the *Trypanosoma brucei* Proteome. *FASEB J* 33 (12), 13161-13175.
99. Shrivastava, R. et al. (2019) Deletion of a Single LeishIF4E-3 Allele by the CRISPR-Cas9 System Alters Cell Morphology and Infectivity of *Leishmania*. *mSphere* 4 (5), e00450-19.
100. Beneke, T. et al. (2020) LAX28 Is Required for the Stable Assembly of the Inner Dynein Arm F Complex, and the Tether and Tether Head Complex in *Leishmania* Flagella. *J Cell Sci* 133 (2), jcs239855.
101. van der Laan, S. et al. (2019) Evolutionary Divergence of Enzymatic Mechanisms for Tubulin Detyrosination. *Cell Rep* 29 (12), 4159-4171.e6.
102. Shaw, S. et al. (2019) Flagellar cAMP Signaling Controls Trypanosome Progression Through Host Tissues. *Nat Commun* 10 (1), 803.
103. Rojas, F. et al. (2019) Oligopeptide Signaling Through TbGPR89 Drives Trypanosome Quorum Sensing. *Cell* 176 (1-2), 306-317.
104. Alves, A. et al. (2020) Control of Assembly of Extra-Axonemal Structures: The Paraflagellar Rod of Trypanosomes. *J Cell Sci* 133 (10), jcs242271.
105. Takagi, Y. et al. (2019) Utilization of Proliferable Extracellular Amastigotes for Transient Gene Expression, Drug Sensitivity Assay, and CRISPR/Cas9-mediated Gene Knockout in *Trypanosoma Cruzi*. *PLoS Negl Trop Dis* 13 (1), e0007088.
106. Aslett, M. et al. (2010) TriTrypDB: a functional genomic resource for the *Trypanosomatidae*. *Nucleic Acids Res* 38 (Database issue), D457-62.
107. Giri, S. and Shaha, C. (2019) *Leishmania Donovanii* Parasite Requires Atg8 Protein for Infectivity and Survival Under Stress. *Cell Death Dis* 10 (11), 808.
108. Lander, N. and Chiurillo, M. (2019) State-of-the-art CRISPR/Cas9 Technology for Genome Editing in Trypanosomatids. *J Eukaryot Microbiol* 66 (6), 981-990.

Glossary

CRISPR-Cas9: A bacterial derived immunity system providing protection against foreign nucleic acids (DNA, RNA). The endonuclease Cas9 forms an active complex with the sgRNA to target complementary DNA and generate a double strand break (DSB). This tool is now widely used and has been adapted in several 'divergent' organisms, including apicomplexan and trypanosomatid parasites.

Donor DNA: DNA sequence which serves as a template to repair the DSB induced by Cas9. It can be delivered to the cell cloned into a plasmid, as a purified PCR product, or as a single-stranded oligonucleotide. Depending on the application, donor DNA may include a selectable marker, a tag or a single nucleotide modification. In the absence of genomic integration of a selectable marker, the strategy is called 'marker-free' and the edition is scarless.

Editing efficiency: We define efficiency as the number of surviving cells that are correctly edited divided by the number of cells originally transfected. This efficiency is rarely documented and most authors report the number of edited clones out of the total number of clones screened (See Key Table).

Homologous recombination or homology directed recombination (HDR): A type of DNA repair process providing high-fidelity homology sequence-dependent repair of DSBs. It is catalyzed by specific recombinases, like RAD51. HDR is widely used to introduce genetic modifications into cells. In trypanosomatids, when donor DNA is provided, HDR prevails over other mechanisms.

Microhomology-mediated end joining (MMEJ): A pathway for repairing DSBs. Contrary to HDR, it is considered as highly error-prone. It involves alignment of micro homology sequences on either side of the break to guide DSB repair, thus leading to deletion of the DNA sequence between the microhomology regions. Trypanosomatids use MMEJ or SSA to repair DSBs in the absence of donor DNA.

Non-homologous end joining pathway (NHEJ): A DSB repair pathway mediating direct ligation of DNA breaks without the need for a homologous sequence and often leading to small sequence insertions or deletions (INDELs). NHEJ is the most efficient DSB repair pathway in mammalian cells but is absent in trypanosomatids.

Polycistronic transcription unit: Gene clusters of two or more members transcribed from the same transcription initiation point. Since trypanosomatids almost completely lack polymerase II promoters, genes are usually transcribed as large polycistronic units of functionally non-related genes. The resulting polycistronic mRNAs are processed post-transcriptionally to generate monocistronic mature mRNAs.

Protospacer adjacent motif (PAM): See sgRNA definition below.

Seed: See sgRNA definition below

Single guide RNAs (sgRNAs): RNA molecules made of two non-coding parts: (i) the 'seed' made of 20 nts specific to the target gene and (ii) the scaffold RNA at the 3' end which serves as a base for Cas9 binding. For its endonuclease activity, Cas9 recognizes a protospacer adjacent motif (PAM, here a NGG for *SpCas9*) downstream of the seed sequence in the genome. Several RNA polymerases, *i.e.* T7RNAP, Pol I or III, can be used for transcription of sgRNA, driven by a T7, U6 or PARP (procyclic acidic repetitive protein) promoter. In addition to the amount of sgRNA, a critical parameter for an efficient system is the accurate start and termination of transcription of the sgRNA. This can be achieved (i) when the sgRNA is PCR-generated, (ii) when a terminator is added downstream of the sgRNA and (iii) when the sgRNA is flanked with ribozyme sequences.

Single strand annealing (SSA): A DSB repair pathway using homologous repeat sequences (26 to 500 bp) flanking the DSB to bridge DSB ends, thereby resulting in large deletion mutations between the tandem repeats. Trypanosomatids may repair DSBs by SSA in the absence or presence of donor DNA.

Untranslated regions (UTRs): Sequences located upstream of the start codon and downstream of the stop codon of a gene. These regions are transcribed but not translated. As discussed in the "UTR importance" section, UTRs may play diverse and key roles in regulating gene expression, including mRNA stability and translational efficiency.

Key Table. Essential features of different CRISPR-Cas9 strategies used in trypanosomatids.

Applications	Parasite stage	Cas9 expression	sgRNA	Donor DNA	UTRs preservation	Selection-free	Efficiency (Nb of edited clones/total clones screened)	Refs
Leishmania spp.								
Non-essential gene KO; precise base editing; gene tagging	promastigote	Constitutive expression from a plasmid DHFR-TS promoter	Constitutive episomal expression; LmU6 promoter	Constant, from an episome selection marker with HRs of 0.5-1kb	Yes	No	25% edited parasites after 4 weeks selection and cellular cloning	[8, 50]
Non-essential gene KO (single and multicopy genes); gene tagging; precise base editing	promastigote	Constitutive expression from a plasmid expressing either SpCas9 or SaCas9	Constitutive expression from a plasmid	No donor DNA or transient: either Single-strandedOligonucleotide or PCR with HRs of 25pb	Yes	Yes	10-50% to 100% depending on the presence of donor DNA or not; 6 weeks culture with cellular cloning and repeated transfections in absence of selection marker	[9, 15, 16]
		or Constitutive expression from plasmid co-expressing Cas9 and sgRNA; rRNA promoter with/without HDV ribozyme sequences						
Non-essential gene KO; gene tagging	promastigote	Constitutive expression from a plasmid	Transient expression as PCR product; T7 RNAP promoter	Transient: PCR product containing selection marker with HRs of 30bp	No	No	100% edited total population obtained after 1 week without cellular cloning	[4, 10, 18, 20, 27, 39, 96-100]
Rapamycin inducible KO combining CRISPR-Cas9 and DiCre	promastigote	Constitutive expression from a plasmid	Transient expression as PCR product; T7 RNAP promoter	Transient: PCR product containing selection marker; LoxP sequences with HRs of 30bp	Yes	No	25-100% edited parasites depending on the target gene after 3 weeks including cellular cloning	[27]
Non-essential gene KO	promastigote	Transient expression as a Ribonucleoprotein (RNP) complex transcribed in vitro		No donor DNA or transient: Single-stranded Oligonucleotide	Yes	Yes	>55% edited parasites	[23]
Trypanosoma brucei								
Non-essential gene KO; gene tagging	procyclic, BSF	Constitutive expression from a plasmid	Transient expression as PCR product; T7 RNAP promoter	Transient: PCR product containing selection marker with HRs of 30bp	No	No	100% after 1 week without cellular cloning	[10, 101-104]
Tetracycline-inducible gene KO (non-essential and essential genes);	BSF	Constitutive expression from integrative plasmid, integrated at rDNA	Constitutive expression from integrative plasmid integrated in the rDNA	No donor DNA or transient donor DNA: Single-stranded Oligonucleotide of 50bp	No	No	100% after 1week without integration of selection marker (co-targeting drug	[17]

precise base editing		locus. Tetracycline-inducible rRNA promoter	locus; T7 RNAP promoter including HDV ribozyme				resistance selection)	
Non-essential gene KO (single and multicopy genes); gene tagging: precise base editing	procyclic	Constitutive expression from an episome	Constitutive expression from an episome PARP promoter with ribozyme	Constant from an episome; 'marker free' with HRs ~200bp	Yes	Yes	60-100% edited parasites after 3-5 months with cellular cloning.	[19]
Tested for KO of exogenous reporter	BSF	Transient expression: Ribonucleoprotein complex RNP transcribed <i>in vitro</i> (SaCas9/sgRNA)		No donor DNA or transient donor DNA as Single-stranded Oligonucleotide	N/A	Yes	55% editing for a reporter gene	[23]
Trypanosoma cruzi								
Gene disruption, gene KO; knockdown of multicopy genes	epimastigote, trypomastigote	Constitutive expression from a plasmid	Transient expression in vitro transcribed sgRNA	No donor DNA	N/A	Yes	1-70% for gene replacement	[7]
Non-essential gene KO; gene knock-in; gene tagging	epimastigote	Constitutive expression from a plasmid or co-expression with sgRNA. ribosomal promoter		Transient: PCR product: selection marker with HRs of 100bp	No	No	95-100% edited parasites after 5 weeks of selection and cellular cloning	[22, 40, 55]
Non-essential gene KO; gene knock in; gene tagging	epimastigote	Constitutive expression from an integrative plasmid (tubulin locus)	Transient expression as PCR product; T7 RNAP promoter	Transient: PCR product containing selection marker with HRs of 30bp	No	No	variable edition rates; 100% edited parasites after 3 weeks of selection	[39]
Non-essential gene KO, knockdown; gene tagging	epimastigote, trypomastigote	Transient expression: Ribonucleoprotein (RNP) complex transcribed <i>in vitro</i> (SaCas9/sgRNA)		No donor DNA or transient: Single-stranded Oligonucleotide 100bp or a tag with HRs of 20bp	Yes	Yes	35 -93% from 2 days to 2 weeks after transfection depending on the sgRNA and RNP complex amount	[23]
Knockdown combining CRISPR-Cas9 and Riboswitch	epimastigote	Constitutive expression from a plasmid or co-expression with sgRNA. ribosomal promoter		Transient: PCR product: selection marker with HRs of 100bp	No	No	95-100% edited parasites after 5 weeks of selection and cellular cloning	[72]
Non-essential gene KO	amastigote	Constitutive expression from a plasmid	Transient expression of <i>in vitro</i> transcribed sgRNA	No donor DNA	N/A	Yes	ND	[105]