
Article:
Jones, Nathaniel orcid.org/0000-0001-7328-4487, Catta Preta, Carolina, Lima, Ana Paula C.A. et al. (1 more author) (2018) Genetically Validated Drug Targets in Leishmania: Current Knowledge and Future Prospects. ACS Infectious Diseases. pp. 467-477. ISSN 2373-8227

https://doi.org/10.1021/acsinfecdis.7b00244

Perspective.

Title: Genetically Validated Drug Targets in *Leishmania*; Current Knowledge and Future Prospects.

Author Names: Nathaniel G. Jones[1], Carolina M.C. Catta-Preta[1], Ana Paula C.A. Lima[2], Jeremy C. Mottram[1*].

Author Affiliations:

[1]Centre for Immunology and Infection, Department of Biology, University of York, Wentworth Way, Heslington, York, YO10 5DD, UK.

[2] Instituto de Biofisica Carlos Chagas Filho, Centro de Ciências da Saúde, Universidade Federal do Rio de Janeiro, Cidade Universitária, Rio de Janeiro, RJ, 21949-900, Brazil.

Corresponding Author Email Address:

* Email: jeremy.mottram@york.ac.uk

ORCID:

Nathaniel G. Jones 0000-0001-7328-4487

Carolina M.C. Catta-Preta 0000-0001-8983-3487

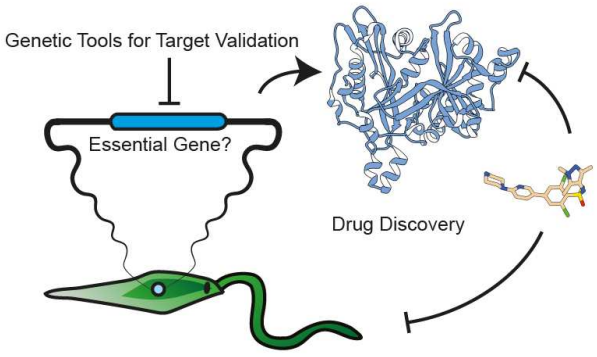
Ana Paula C.A. Lima

Jeremy C. Mottram 0000-0001-5574-3766

Abstract: There has been a very limited number of high throughput screening campaigns carried out with *Leishmania* drug targets. In part, this is due to the small number of suitable target genes that have been shown by genetic or chemical methods to be essential for the parasite. In this perspective, we discuss the state of genetic target validation in the field of *Leishmania* research and review the 200 *Leishmania* genes and 36 *Trypanosoma cruzi* genes for which gene deletion attempts have been made since the first published case in 1990. We define a quality score for the different genetic deletion techniques that can be used to identify potential drug targets. We also discuss how the advances in genome-scale gene disruption techniques have been used to assist target based and phenotypic based drug development in other parasitic protozoa and why Leishmania has lacked a similar approach so far. The prospects for this scale of work are considered in the context of the application of CRISPR/Cas9 gene editing as a useful tool in *Leishmania*.

Keywords: *Leishmania*, Gene Knockouts, Null, CRISPR/Cas9, Target Validation, Pathogen, *Trypanosoma cruzi, Drug Discovery.*

Table of Contents Graphic:



- **Introduction to *Leishmania* and need for new drugs**

Leishmaniasis is a disease that mainly affects those burdened by extreme poverty, across large swathes of the tropics and sub-tropics. It consists of a spectrum of human and animal diseases caused by at least 20 species of parasites in the genus *Leishmania*. The symptoms range from self-limiting ulcers, to destruction of muco-cutaneous surfaces, to fatal disease caused by visceral leishmaniasis; dermal sequelae can also occur following cure of primary visceral disease. There are up to 1.2 million cases of cutaneous leishmaniasis and up to 0.4 million cases of visceral leishmaniasis each year – resulting in up to 30, 000 deaths[1]. Collectively, 350 million people live at risk of catching leishmaniasis, due to the wide geographic range of the vector insects, phlebotomine sand flies. There are medicines available to treat these diseases but they have notable deficiencies due to emerging resistance, poor safety profiles, and long duration of treatment; so more effective treatments with improved safety and dosing profiles are desirable. The complex picture of a spectrum of diseases, in several organs, caused by different parasite species, with differing biology mean that several new medicines may be required to satisfy quite stringent target product profiles (TPPs) for visceral and cutaneous leishmaniasis[2,3]. New chemical entities will probably require different chemical features for achieving desired bioavailability in skin versus viscera coupled with the challenge of drugging an intracellular parasite that resides in a phagolysosomal compartment. There is a growing appreciation that *Leishmania* infections can also form quiescent, persistent forms[4]. Targeting these stages is desirable, though may also prove challenging. There are a number of ongoing drug discovery programmes that aim to address this need, typically working to defined phenotypic or target-based screening strategies to best achieve success[2,3,5-8]. Phenotypic or whole organism screening selects for compounds that cause loss of parasite fitness, whereas the target-based approach involves

screening against a selected target of interest thought to be essential for parasite survival. We will assess these approaches and the role that genetic target validation can play within them.

- **Target validation in drug discovery programmes.**

Drug target validation for anti-parasitic compounds consists of acquiring the evidence that defines whether a target molecule (usually a protein/enzyme) is selectively inhibited by a chemical entity leading to the death of the parasite; as well as linkage of the target molecule to an essential parasite process. The evidence used to validate targets can arise from genetic manipulations of a pathogen, or from interrogation of the organism with specific chemical probes. The strength of validation will depend on how much evidence is accrued and well validated targets will be supported by both genetic and chemical evidence[5,9]. For non-infectious diseases, it is estimated that selecting genetically supported targets doubles the success rate in later stages of clinical development[10], highlighting the importance of genetic validation. The best validation of all comes when the molecular target for a compound is known and the compound is available on the market for clinical use – the one example for trypanosomatids being the treatment of Human African Trypanosomiasis with Eflornithine (difluoromethylornithine -DFMO), a compound that inhibits ornithine decarboxylase[11]. It is notable that for most anti-leishmanial drugs (miltefosine, paromomycin, antimonials or pentamidine) that there are no target proteins identified despite extensive research, although amphotericin B is reported to specifically target ergosterol-containing membranes[12,13].

  - **Phenotypic Screening**

For *Leishmania* several hit compounds have been identified by phenotypic based drug discovery, which involves screening drug-like molecules against *Leishmania* axenic amastigotes or amastigotes in macrophages to identify compounds that kill the parasites. This approach has the advantage in that it can identify bioactive compounds that have appropriate cell permeability characteristics to kill the parasite within the parasitophorous vacuole. One disadvantage is that the screening is highly stringent and consequently few bioactive compounds are identified[5]. Another disadvantage is that one does not immediately know the molecule being targeted by the bioactive compounds. If a large (>1 million compounds) screen is performed it may conceivably identify diverse target-class inhibitors whereas sub-libraries can be used to increase the chances of hits within a certain target-class, for example by using a protein kinase inhibitor focussed library[14]. Optimisation of hit compounds by medicinal chemistry involves a level of randomness as it may not be immediately apparent which functional groups of the molecule are the important pharmacophores. Structural activity analyses are performed in whole cell assays and it is not possible to know if increases in activity against parasites are concurrent with activity against a specific target. Therefore, a campaign of target deconvolution is often initiated to identify the molecular target of a hit compound. This can involve a combination of forward genetic screens such as the generation of resistant parasites and the identification of mutant alleles in the target gene by whole genome sequencing[15]; affinity purification of the target using compound conjugated to beads[16]; the use of overexpression libraries[17,18], or metabolomics to identify blocks in biochemical pathways[19]. It is also important to have an understanding an experimental compound's mode-of-action as this can enable an assessment into the likelihood of resistance mechanisms evolving in the parasite. Recent, successful examples of phenotypic screening of large compound libraries, or focussed sub-set chemical libraries,

against *Leishmania* parasites include the identification of the ”Leish-Box” of inhibitors by GSK[20], the identification of a selective proteasome inhibitor (GNF6702) by Novartis[15], and the identification of natural product inhibitors[21].

- **Target Based Drug Discovery**

The alternative strategy of target-based drug discovery is used extensively by the pharmaceutical industry and has been applied to *Leishmania* – this consists of the selection of a protein target, the development of a biochemical or biophysical assay that can be used to identify inhibitors of that target, followed by a high-throughput screening (HTS) campaign. Hit compounds from such screens are validated then chemically optimised into lead compounds that have improved properties in terms of potency, selectivity and bioavailability. This is a resource-intensive and costly process and needs to be built on a solid understanding of why the selected target is an appropriate choice, a process defined as target assessment. Drug targets are evaluated based on gene essentiality, in addition to other criteria, for example the presence of close homologues in the parasite genome that may allow for easier evolution of resistance through functional redundancy. Alternatively, single point mutations in a gene may lead to drug resistance with little associated cost to parasite fitness. These examples emphasise the importance of a good understanding of parasite genetics for drug target prioritisation. Other parameters for target assessment include the potential druggability of the protein, the availability of *in vitro* assays, the availability of structural information, and the presence of structurally similar proteins in the host proteome that may present a potential risk for host toxicity. The assessment of a target against each of these criteria can be scored, generating a “traffic lights” rating – red ratings will typically stop a target from progression, but those scoring green should progress more

easily through development[22]. The level of rigour in building a strong evidence base to support target focused approaches means that many targets are shown to be invalid. For these reasons, as well as due to the limitation of financial and material resources, there have been notably few HTS campaigns against *Leishmania* proteins, namely CRK3 (cyclin dependent kinase)[23], NMT (*N*-myristoyltransferase)[24], and PTR1 (pteridine reductase)[25]. Casein Kinase (LmCK1.2) has also been the subject of a HTS campaign[14] after being chemically validated as a target[26,27].

- **Drug repurposing and piggybacking**

Alternatives to developing or identifying new chemical entities to target *Leishmania* exist in the form of drug repurposing, a strategy where a compound that has been developed to treat one disease is used to treat another, different disease. The benefit of drug repurposing is that the treatment can be used in patients relatively quickly and with lower cost compared to *de novo* drug discovery. Notable examples for leishmaniasis include the use of miltefosine, amphotericin B and pentamidine, which were all designed or approved for other indications[6]. Piggybacking is the concept of identifying existing pharmaceutical material, investment and knowledge around a given target or inhibitor class and applying it to infectious disease targets, but this usually stops short of direct drug repurposing.

- **Genetic Target Validation in *Leishmania***

Leishmania has some biological features that have made it somewhat difficult to genetically manipulate; it is highly prone to aneuploidy meaning that genetic modification of essential genes can be challenging or unstable[28-30] . RNA-interference does not function in most *Leishmania* species[31]. In those species where RNAi does function it has not been exploited

to the same extent as *Trypanosoma brucei*, primarily due to the lack of an inducible system to investigate the function of essential genes. *Leishmania* parasites are easily cultured and genetically manipulated as promastigote stages, those found in the insect vector, but these are not the relevant stage for human disease. It is therefore very important to confirm the essentiality of a potential target using the amastigote stages of the parasite life-cycle. Some species of *Leishmania*, such as *L. mexicana* are readily converted to axenic amastigotes *in vitro*[32] and this can facilitate study of gene function in the pathogenic stages, but phenotypic analysis is perhaps best carried out using macrophages to perform intracellular amastigote growth and survival assays and/or mouse infections. Amastigotes can be genetically manipulated but this is not a common practice[33]. Typically, an *in vivo* model is used to demonstrate a fitness or virulence defect of a genetically engineered mutant.

Leishmania possess a very efficient homologous recombination pathway for DNA repair that has been exploited as a basis for generating gene deletions, integrating epitope tags or the expression of transgenes[34] (see Table 1 for definition of terms). As the technology to manipulate *Leishmania* has improved over time so has the ability to achieve more confidence in a gene's essentiality. Based on the literature we used a scoring system for the strength of validation resulting from gene deletion studies, on a 1-5-star scale (Table 1 and Figure 1), and discuss these below, with examples. We propose that the higher scoring methods provide better quality evidence of a genes essentiality but acknowledge they are also associated with increased time and effort to generate the mutants. Despite this, we hope to use this guide to inform good practices for *Leishmania* genetic manipulation.

- **Approaches for genetic manipulation of *Leishmania*.**
    - **Deletion by double homologous replacement and facilitated null mutants**

In 1990 it was demonstrated that genes could be deleted from *Leishmania* promastigotes by homologous replacement with linear dsDNA[35]. The authors demonstrated proof of concept, which involved deleting a single allele of *DHFR-TS* in wild-type *L. major*. This was then repeated this in a strain that was already identified as heterozygous for *DHFR-TS,* which permitted the deletion of the remaining allele using a homology flanked neomycin resistance cassette. The generation of a *DHFR-TS* null mutant led to thymidine auxotrophy and could only be achieved in the presence of thymidine nutritional supplementation; removal of thymidine supplementation stalled cell growth, demonstrating that *DHFR-TS* was essential for parasite survival. Complementation of the null mutant with an episome expressing *DHFR-TS* restored wild-type growth levels, indicating that the function of this gene could be reconstituted genetically as well as chemically. This set the stage for other researchers to begin performing reverse genetic analyses in *Leishmania*. Development of new resistance markers allowed for gene deletions of two alleles of a gene[36], given that *Leishmania* was considered at the time to have a diploid genome this allowed the generation of null mutants from wild-type backgrounds. In the following years a number of papers demonstrated that upon deletion of both alleles of a suspected essential gene the system would fail. Either no transfectants would survive the drug selections, parasites would survive would but had duplicated their genome to retain the gene of interest as well as the drug resistance markers[36], or that there were extra copies of the gene remaining on supernumerary chromosomes or ectopic elements (Figure 1A). At the time, this was considered evidence for gene essentiality, but with the inability to perform any further this is classified as the weakest form of evidence (one star) as the failure to generate a null mutant may result from a simple technical failure. For example, as the technology for performing electroporation became more sophisticated, transfection efficiencies increased,

allowing some previously intractable genes to be revealed as non-essential. The Metacaspase gene, *MCA*, initially could not be deleted from *L. major* [37] and overexpression caused a growth defect, which was interpreted to mean that the gene played an essential role in cellular proliferation. However, *MCA* null mutants could be generated in both *L. major*[38] and *L. mexicana*[39] using a newer, high-efficiency electroporation system that causes minimal cellular damage. Generating this deletion allowed for the characterisation of the role of MCA in amastigotes and mice where the null mutant exhibited an increase in virulence[39].

Several studies have performed deletions that resulted in parasites auxotrophic for a given metabolite, and required chemical supplementation to allow null mutants to be recovered[35,40] (Table S1). The advantage of a chemical complementation is that it allows for rapid withdrawal and the resulting phenotype investigated. Where nutritional supplementation could not be used to facilitate deletion of essential genes, researchers began to introduce an episomal copy of the gene of interest into the parasite prior to deletion of the two chromosomal alleles[41] (Figure 1B). This approach increases the confidence in the validation, as it demonstrates that the failure to delete both alleles of the gene of interest is not a technical failure or due to the locus being refractory to genetic manipulation (Fig 2B, two stars).

An episome expressing a gene in a null mutant background will be retained if it confers a selective advantage for growth, even if the gene is dispensable for parasite viability. In a facilitated null mutant of a non-essential gene, the episome can be lost in the absence of drug selection and a null mutant is generated. For essential genes, even after prolonged

culturing of the facilitated null mutant in medium that lacks drug selection for the episome, or by passaging the parasites through mice, the episome will be retained. This is considered indirect evidence that the gene is essential (Figure 1B, 3 stars). However, most episomes cannot be forced from the parasite so this still does not provide direct evidence for gene essentiality or allow the phenotype analysis of loss-of-function mutants.

- **Facilitated Null with Negative Selection/Forced Plasmid Shuffle**

As an improvement in the ability to select for loss of the episome in a facilitated null mutant, negative selection or forced plasmid shuffle can be achieved using the negative selectable marker, *herpes simplex thymidine kinase* (TK). Ganciclovir is used to impose a fitness cost on parasites that retain the plasmid, selecting for parasites that lose the episome. If it is complementing an essential gene then the parasites must retain the plasmid to remain viable and so any loss will prove fatal. Inclusion of a Green Fluorescent Protein (*GFP*) reporter gene helps to distinguish cells containing the episome, which is useful for fluorescence activated cell sorting (FACS) of different populations of cells. (Figure 1C).
This strategy was first applied to *Leishmania major* to investigate the role of 5,10-methylene tetrahydrofolate dehydrogenase (DHCH) in 10-Formyl tetrahydrofolate (10-CHO-THF) metabolism[42] following the failure of classic gene deletion and nutritional supplementation strategies. *DHCH1* was expressed from an episome (also containing TK and GFP), allowing a facilitated null mutant to be generated. The episome was selected against using ganciclovir and parasites were then flow sorted into GFP bright or dim populations and cloned into 96 well plates. The dim clones, lacking the DHCH1 episome were not viable, and failed to grow. This system has also been used to examine the essentiality of MAP Kinase 4 (*MPK4*) in *L. major* promastigotes in a manner allowing more detailed functional genetic analysis[43]. A

facilitated null mutant was achieved by episomal complementation, followed by forced plasmid shuffle to assess gene essentiality. After ascertaining that MPK4 was essential for promastigote viability the authors complemented with a panel of secondary episomes (lacking TK) containing functionally mutated versions of MPK4. This allowed for exploration of the importance of protein motifs and residues for MPK4 ATP binding, activity, and activation by upstream MKKs. This study highlighted the power of plasmid shuffle in that a MPK4 mutant with altered ATP binding properties, predicted to have reduced protein kinase activity, could replace wild-type MPK4, resulting in resistance to acidic stress. The results from this study support previous facilitated null and unforced plasmid shuffle data suggesting MPK4 from *L. mexicana*[44] is essential for the mammalian infection cycle. Profiling of null mutant phenotypes in the amastigote stage is critical to the validation of drug targets, so the compatibility of plasmid shuffle with this is welcome. However, a limitation of the plasmid shuffle approach is that null mutants of essential genes are never generated in the absence of ganciclovir (which acts as a cytostatic but not cytocidal drug), so direct analysis of the null mutant phenotypes is not possible. In reference to the suitability of MPK4 as a drug target , it has so far been impossible to identify a biochemical assay for screening of small molecule inhibitors, mainly due to its lack of protein kinase activity against generic substrates[43]. Currently this lack of an enzymatic activity assay renders MPK4 a genetically well-validated but a poorly tractable target for drug discovery programmes.

- **DiCre Inducible Deletion**

In order to directly measure loss of gene function, a population of cells would ideally be subjected to an inducible gene deletion approach, allowing synchronous, direct analysis of emerging phenotypes in real time. A technique that has recently been applied to *Leishmania*

to directly study gene function is the DiCre inducible gene deletion strategy[45] (Figure 1D). This was first used in *Leishmania mexicana* to inducibly delete the cyclin dependent protein kinase, *CRK3*, a gene that had previous indirect evidence for essentiality in *Leishmania*[46]. Prior to the study by Duncan *et al*. there was a significant body of indirect genetic evidence for CRK3 essentiality, such as changes in ploidy associated with attempts to delete both alleles and that it was only possible to delete both chromosomal alleles in the presence of an episome expressing *CRK3*, with the episome being retained in absence of drug selection[46]. Other chemical evidence existed for CRK3 to be a potential drug target as known inhibitors of CDKs were shown to induce a G2/M cell cycle block and halt parasite growth[46-48].

This DiCre mediated deletion of *CRK3* involved replacing a genomic allele of the gene with a version flanked by *loxP* sites (floxed), the second copy replaced by a cassette encoding two subunits of the Cre recombinase, fused to rapamycin binding domains[45] (Figure 1D). Once this strain was generated, the dimerization of Cre subunits was initiated by the addition of rapamycin, forming a functional Cre recombinase that excised the remaining floxed copy of the gene of interest, resulting in a null mutant (Figure 1D). The advantage of this system is that phenotypes emerging from the gene deletion can be studied over several days, and the mechanism of cell death or dysregulation can be elucidated. The disadvantage is that RNA and protein remain in the cell after DiCre induced deletion of the gene and the phenotype may take several rounds of cell division before becoming apparent.

By performing the DiCre analysis, Duncan et al. were able to delete the *CRK3*^{flx} gene from promastigote cells and observe precisely the phenotype resulting in *CRK3* null mutants – cell

cycle (G2/M) arrest, the initiation of G1 phase in the absence of cytokinesis, and impaired cytokinesis leading to an increase in both multinucleated cells and "zoids" (cells lacking nuclear DNA content) and associated cell death. Cell lines were complemented with both wild-type *CRK3* and *CRK3* ^{T178E} (a mutation in the activation loop) genes, to demonstrate the phenotype was specific to *CRK3* deletion, that the presence of threonine 178 in the activation loop of the protein kinase domain was important for proper function and that *CRK3* ^{T178E} could not complement the loss of wild-type *CRK3*. Not only did this study show the importance of *CRK3* in promastigote stages, the gene was deleted in stationary phase cultures (metacyclic enriched), allowing the parasites to be inoculated into mice and the progress of the infection monitored by non-invasive, bioluminescent imaging using a strain engineered to express red-shifted luciferase. *CRK3* was shown to be important for the establishment and maintenance of infection by reduction in bioluminescence from mice infected with the *CRK3*-deficient cells. This was ranked as a 5-star validation (Table1, Figure 1D). Despite the now extremely good genetic validation of *L. mexicana CRK3* as an essential gene, a previous HTS campaign (against LmxCRK3/CYC6) failed to identify compounds that had bioactivity against the parasite in macrophages[49]. The compounds identified failed to inhibit parasite growth in macrophages despite having good physiochemical properties and having different scaffolds, suggesting that these compounds failed because of some unknown aspect of CRK3 biology in *Leishmania*, or unknown bioavailability within the parasitophorous vacuole, thus CRK3/CYC6 has not been fully chemically validated.

A more definitive assessment of essentiality would be to induce gene deletions in amastigotes during animal infections. This experiment is not possible with the DiCre system due to the lack of bioavailability of rapamycin/rapalogues and the toxicity of rapamycin to

amastigotes. Such a model would enable researchers to observe that an experimental infection was not compromised during the establishment phase, but upon inducible deletion of an essential gene the infection would fail and be resolved by the host immune system.

- **Overview of existing gene deletion attempts.**

To assess the current state of target validation in *Leishmania*, a literature search was performed to identify all instances where these genetic techniques have been used to attempt to generate null mutants in human pathogenic *Leishmania* (Figure 2A, Table S1). We focussed on identifying the first instance where researchers had attempted to generate a null mutant for a given gene and do not provide a comprehensive review of every genetic study on the same gene in multiple species of *Leishmania*. The strategy used to attempt the genetic deletion was identified, whether this was successful or not, and a broad categorisation of the phenotype; i.e. essential for promastigotes, conditionally essential, or non-essential. If a complementation was required, it was defined if this was genetic or a nutritional supplementation. Two hundred genes were identified where attempts have been made to make null mutants in promastigotes of various *Leishmania* species; 65 were classified as essential and the remaining 135 were classified as non-essential. The rate of genetic deletions studies being published appeared to accelerate after 2008, possibly due to the publication of the *L. major* genome in 2005[50] and the *L. braziliensis* and *L. infantum* genomes in 2007[51], but has remained fairly constant between 2008-2017. Our literature search identified a range of different species of *Leishmania* under investigation, including those responsible for both cutaneous and visceral diseases (60% vs. 40%, respectively), with

L. major (37%) being the most commonly studied, followed by *L. donovani (*29%), (Figure 2B).

A Gene Ontology (GO) analysis was performed on the 200 genes using TriTrypDB[52] and REVIGO[53] (Figure 3, Table S1). GO were analysed by clustering to describe the biological processes that are performed by the 200 gene products. Clear clusters emerge for processes involved in nutrient transport and stress response, broader clusters emerge for metabolic processes including nucleotide salvage and synthesis. There are many instances of GO terms associated with oxidation-reduction processes and protein modification such as protein folding, proteolysis and protein phosphorylation – demonstrating the strength of research on metabolism, peptidases and protein kinases. Important single gene processes were also identified, such as N-terminal protein myristoylation.

Due to the high variability in the assay protocols used for phenotypic analyses in each study it is challenging to allocate each gene deletion mutant into a defined phenotypic category, other than essential or non-essential to promastigotes. For example, gene deletions that confer altered virulence may be essential in a given context such as mammalian infection, but the classification of these phenotypes can be quite subtle or dependent on the species, infectious dose or mouse strain used and is outside the scope of this perspective. In this regard, there is a clear need for a comparative analysis on *Leishmania* null mutants, where the same genetic approaches and phenotype analysis can be carried out on a large scale, which is discussed in the next section.

By way of comparison and to extend the analysis, the literature was also reviewed to identify *Trypanosoma cruzi* genes for which creation of null mutants had been attempted (Table S2). The genetic tools available for *Leishmania* are broadly applicable to *Trypanosoma cruzi*, in that this organism also has a functional homologous recombination system and can support expression of genes from episomes[54,55]. As for *Leishmania*, we focussed on identifying the first study to report the deletion attempt for any given gene. Based on our five stars scoring system, studies in *T. cruzi* did not reach a high-quality level of drug target validation. Since 1993 only 36 genes had deletion attempts reported in the literature, and 16 were considered essential in epimastigotes (Figure 4, Table S2). Only one study, of *NMT*[56], used a variant of the facilitated null approach to validate the target gene; deletion of both chromosomal alleles was not possible until a "rescue" allele was integrated into the ribosomal locus of the parasite. This demonstrated that the generation of chromosomal allelic deletions was possible, but only after genetic complementation; therefore the gene is likely to be essential. However, the because the "rescue" allele was integrated into the genomic DNA there was no possibility to test that it could be lost to create a true facilitated null cell line. Recently, DiCre[57] systems were applied to *T. cruzi* opening the possibility for better target validation.

- **Prospects for large sale genetic manipulations of *Leishmania;* lessons from other protozoa**

When compared to other, more genetically tractable, protozoan parasites the situation in *Leishmania* looks limited, mainly due to the low throughput of the prevailing techniques used to generate *Leishmania* gene deletions. Through a brief description of tools and

screens that have been used in other kinetoplastids and apicomplexans, despite the unique biology and technical challenges each organism and method present, desirable features for future genetic screens in *Leishmania* can be identified. To date genetic manipulation of *Leishmania* has been on a gene-by-gene basis, yet several other parasitic protozoa such as *Trypanosoma* brucei and *Toxoplasma gondii* have been amenable to genome-wide screens[58-60](Figure 4). Combining these massively paralleled genetic depletion resources with various selection pressures can yield sophisticated experiments that are able to interrogate, identify and validate drug targets and elucidate resistance mechanisms.

Most relevant to *Leishmania* target validation is the use of RNAi screening in *T. brucei*, which evolved from gene-by-gene projects to genome-wide RNAi screens[58,59] and medium-throughput gene family screens[61,62]. These massively paralleled RIT-seq (RNA interference target–sequencing) experiments have led to the identification of genes essential for , life cycle progression, drug resistance and response to stress[63,64]. Of particular relevance is the way this technology was used to identify organelles, proteins and metabolic pathways responsible for the mode of action of trypanocidal drugs[65,66] as this may be used to shed light on pan-kinetoplastid biology and drug targets.

Other model pathogens such as *Plasmodium berghei* have also been subject to high throughput gene deletion screens. Although less amenable to genetic manipulation, a herculean effort has allowed half of the *P. berhei* genome to be assessed for essentiality[67,68](Figure 4). The PlasmoGem project used recombineering to build libraries of long-homology flanked deletion vectors that can be transfected into parasites prior to infecting mice[67,68]. Pooling these libraries allowed for more than half of the parasite's genes to be assessed for their role in parasite fitness by sequencing DNA barcodes that were

included in the null mutants. This analysis suggested a wealth of drug targets as ~ 60% of the *P. berghei* genes were considered essential. This compares to ~30% in *T. brucei*, and based on our literature search we expect *Leishmania* to be similar as 32.6% of the *Leishmania* genes investigated so far are essential.

Increased-throughput and efficiency for genetic manipulation of parasitic protozoa has occurred by the adoption of CRISPR/Cas9 mediated gene disruption and editing, with the first genome wide screen in a parasite performed in *Toxoplasma gondii* [60]. *T. gondii* has a propensity for repairing double stranded DNA breaks by error-prone, non-homologous end joining (NHEJ), allowing for inactivation of genes by frame-shifting insertions or deletions. Next-generation sequencing was used to quantify the remaining sgRNAs in the population and thus calculate a score for how important a given gene is for parasite fitness. When combined with comparative genomic assessment of essential *T. gondii* genes to conserved genes in other apicomplexans the authors identified potential pan-apicomplexan targets, some of which were then investigated and confirmed as essential in *Plasmodium falciparum*. The recent application of CRISPR/Cas9 to *Leishmania* may enable such genome wide screening, but this requires further development[69-73]. This is due to *Leishmania* lacking a functional NHEJ pathway, although they can repair double strand break using microhomology-mediated end joining (MMEJ). Despite this, the medium-throughput analysis of gene families in *Leishmania* is now a reality using the highly efficient *L. mexicana* T7/Cas9 system. This method uses PCR amplicons for both sgRNA and repair templates with no cloning steps, both alleles can be targeted in a single transfection, and drug selection in pools has a high efficiency[73]. These factors permit easy and fast de-prioritisation of non-essential genes from a family of potential targets.

Combined with the ability to generate facilitated null mutants, CRISPR/Cas9 approaches can begin to provide indirect evidence for a gene's essentiality. Indeed, we expect that soon the number of CRISPR/Cas9 genetic deletion attempts in *Leishmania* will rapidly exceed the total number of gene deletions generated between 1990 – 2017, possibly even genome-wide. Because the other genetic techniques, such as DiCre, are more labour intensive, it would be beneficial to focus them on prioritised genes that cannot be deleted by CRISPR/Cas9, therefore CRISPR/Cas9 should facilitate faster and more detailed studies on genes that have interesting or essential functions. CRISPR/Cas9 systems have also been developed for *T. cruzi,* further expanding the toolbox for genetic manipulation[72,74,75].

While constructing the spreadsheet of attempted gene deletions, we were struck by the range of different phenotypic analyses that were performed, which restricted placing of genes into neat categories based on the type or strength of phenotype observed. One of the benefits of cell library-screens or massively-paralleled approaches is the ability to simultaneously analyse all of the mutants in the same assay, which allows for easier comparison of phenotypes. Integration of such data into EuPathDB, would make an exceptionally useful resource[52]. Target selection for *Leishmania* can be guided by mining the datasets available in resources such as TriTrypDB so we encourage fellow researchers to contribute to this community resource with both historical and current data.

- **Conclusion**

The discovery of new drugs for intracellular pathogens is a complex process that usually has low chance of success. In order to give it the best chance the evidence needed to support the process must be of a high standard. The optimisation of existing genetic techniques, the

development of new techniques, and the application of best practises is critical for identifying and validating novel druggable targets. A balanced approach between target based drug development and phenotypic screening, supported by genetic tools for molecular target validation is desirable. High-quality genetic tools are now available for conducting research on potential drug targets in *Leishmania* is optimal and the new technologies available will facilitate and improve the number of well genetically validated targets suitable for entry into HTS campaigns.

Supporting Information

Genetically Validated Drug Targets in *Leishmania*; Current Knowledge and Future Perspectives.

Nathaniel G. Jones[1], Carolina M.C. Catta-Preta[1], Ana Paula C.A. Lima[2], Jeremy C. Mottram[1*].

Pages: 0

Figures: 0

Tables: 2

Supplemental Table 1:

List of genes where deletion attempts in human pathogenic *Leishmania* species have been published. Includes information for: Gene ID, name, product, species, year of publication, originating lab, reference, method and result of deletion, and quality score. Panel 1 for essential genes, Panel 2 for non-essential genes.

Supplemental Table 2:

List of genes where deletion attempts in *Trypanosoma cruzi* have been published. Includes information for: Gene ID, name, product, year of publication, originating lab, reference, method and result of deletion and quality score. Panel 1 for essential genes, Panel 2 for non-essential genes.

Acknowledgements.

Research in the JCM laboratory is supported by a Wellcome Trust Investigator Award (R1758501). We apologise for any published gene deletions that have been inadvertently omitted during our search.

References :

(1) Alvar, J.; Vélez, I. D.; Bern, C.; Herrero, M.; Desjeux, P.; Cano, J.; Jannin, J.; Boer, den, M.; WHO Leishmaniasis Control Team. Leishmaniasis worldwide and global estimates of its incidence. *PLoS ONE* **2012**, *7* (5), e35671 DOI: 10.1371/journal.pone.0035671.

(2) DNDi. https://www.dndi.org/diseases-projects/leishmaniasis/.

(3) WHO. http://www.who.int/leishmaniasis/en/.

(4) Mandell, M. A.; Beverley, S. M. Continual renewal and replication of persistent Leishmania major parasites in concomitantly immune hosts. *Proceedings of the National Academy of Sciences of the United States of America* **2017**, *114* (5), E801–E810 DOI: 10.1073/pnas.1619265114.

(5) Field, M. C.; Horn, D.; Fairlamb, A. H.; Ferguson, M. A. J.; Gray, D. W.; Read, K. D.; De Rycker, M.; Torrie, L. S.; Wyatt, P. G.; Wyllie, S.; et al. Anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need. *Nature reviews. Microbiology* **2017**, *15* (4), 217–231 DOI: 10.1038/nrmicro.2016.193.

(6) Nagle, A. S.; Khare, S.; Kumar, A. B.; Supek, F.; Buchynskyy, A.; Mathison, C. J. N.; Chennamaneni, N. K.; Pendem, N.; Buckner, F. S.; Gelb, M. H.; et al. Recent developments in drug discovery for leishmaniasis and human African trypanosomiasis. *Chemical Reviews* **2014**, *114* (22), 11305–11347 DOI: 10.1021/cr500365f.

(7) Zulfiqar, B.; Shelper, T. B.; Avery, V. M. Leishmaniasis drug discovery: recent progress and challenges in assay development. *Drug discovery today* **2017**, *22* (10), 1516–1531 DOI: 10.1016/j.drudis.2017.06.004.

(8) Rajasekaran, R.; Chen, Y.-P. P. Potential therapeutic targets and the role of technology in developing novel antileishmanial drugs. *Drug discovery today* **2015**, *20* (8), 958–968 DOI: 10.1016/j.drudis.2015.04.006.

(9) Gilbert, I. H. Drug discovery for neglected diseases: molecular target-based and phenotypic approaches. *Journal of medicinal chemistry* **2013**, *56* (20), 7719–7726 DOI: 10.1021/jm400362b.

(10) Nelson, M. R.; Tipney, H.; Painter, J. L.; Shen, J.; Nicoletti, P.; Shen, Y.; Floratos, A.; Sham, P. C.; Li, M. J.; Wang, J.; et al. The support of human genetic evidence for approved drug indications. *Nat. Genet.* **2015**, *47* (8), 856–860 DOI: 10.1038/ng.3314.

(11) Priotto, G.; Pinoges, L.; Fursa, I. B.; Burke, B.; Nicolay, N.; Grillet, G.; Hewison, C.; Balasegaram, M. Safety and effectiveness of first line eflornithine for Trypanosoma brucei gambiense sleeping sickness in Sudan: cohort study. *BMJ : British Medical Journal* **2008**, *336* (7646), 705–708 DOI: 10.1136/bmj.39485.592674.BE.

(12) Saha, A. K.; Mukherjee, T.; Bhaduri, A. Mechanism of action of amphotericin B on Leishmania donovani promastigotes. *Mol. Biochem. Parasitol.* **1986**, *19* (3), 195–200.

(13) Mbongo, N.; Loiseau, P. M.; Billion, M. A.; Robert-Gero, M. Mechanism of amphotericin B resistance in Leishmania donovani promastigotes. *Antimicrobial agents and chemotherapy* **1998**, *42* (2), 352–357.

(14) Durieu, E.; Prina, E.; Leclercq, O.; Oumata, N.; Gaboriaud-Kolar, N.; Vougogiannopoulou, K.; Aulner, N.; Defontaine, A.; No, J. H.; Ruchaud, S.; et al. From Drug Screening to Target Deconvolution: a Target-Based Drug Discovery Pipeline Using Leishmania Casein Kinase 1 Isoform 2 To Identify Compounds with Antileishmanial Activity. *Antimicrobial agents and chemotherapy* **2016**, *60* (5), 2822–2833 DOI: 10.1128/AAC.00021-16.

(15) Khare, S.; Nagle, A. S.; Biggart, A.; Lai, Y. H.; Liang, F.; Davis, L. C.; Barnes, S. W.; Mathison, C. J. N.; Myburgh, E.; Gao, M.-Y.; et al. Proteasome inhibition for treatment of leishmaniasis, Chagas disease and sleeping sickness. *Nature* **2016**, *537* (7619), 229–233 DOI: 10.1038/nature19339.

(16) Nishino, M.; Choy, J. W.; Gushwa, N. N.; Oses-Prieto, J. A.; Koupparis, K.; Burlingame, A. L.; Renslo, A. R.; McKerrow, J. H.; Taunton, J. Hypothemycin, a fungal natural product, identifies therapeutic targets in Trypanosoma brucei. *eLife* **2013**, *2*, e00712 DOI: 10.7554/eLife.00712.

(17) Begolo, D.; Erben, E.; Clayton, C. Drug target identification using a trypanosome overexpression library. *Antimicrobial agents and chemotherapy* **2014**, *58* (10), 6260–6264 DOI: 10.1128/AAC.03338-14.

(18) Gazanion, É.; Fernández-Prada, C.; Papadopoulou, B.; Leprohon, P.; Ouellette, M. Cos-Seq

for high-throughput identification of drug target and resistance mechanisms in the protozoan parasite Leishmania. *Proceedings of the National Academy of Sciences of the United States of America* **2016**, *113* (21), E3012–E3021 DOI: 10.1073/pnas.1520693113.

(19) Vincent, I. M.; Barrett, M. P. Metabolomic-based strategies for anti-parasite drug discovery. *Journal of biomolecular screening : the official journal of the Society for Biomolecular Screening* **2015**, *20* (1), 44–55 DOI: 10.1177/1087057114551519.

(20) Peña, I.; Pilar Manzano, M.; Cantizani, J.; Kessler, A.; Alonso-Padilla, J.; Bardera, A. I.; Alvarez, E.; Colmenarejo, G.; Cotillo, I.; Roquero, I.; et al. New compound sets identified from high throughput phenotypic screening against three kinetoplastid parasites: an open resource. *Scientific Reports* **2015**, *5*, 8771 DOI: 10.1038/srep08771.

(21) Zulfiqar, B.; Jones, A. J.; Sykes, M. L.; Shelper, T. B.; Davis, R. A.; Avery, V. M. Screening a Natural Product-Based Library against Kinetoplastid Parasites. *Molecules* **2017**, *22* (10), 1715 DOI: 10.3390/molecules22101715.

(22) Frearson, J. A.; Wyatt, P. G.; Gilbert, I. H.; Fairlamb, A. H. Target assessment for antiparasitic drug discovery. *Trends Parasitol.* **2007**, *23* (12), 589–595 DOI: 10.1016/j.pt.2007.08.019.

(23) Walker, R. G.; Thomson, G.; Malone, K.; Nowicki, M. W.; Brown, E.; Blake, D. G.; Turner, N. J.; Walkinshaw, M. D.; Grant, K. M.; Mottram, J. C. High throughput screens yield small molecule inhibitors of Leishmania CRK3:CYC6 cyclin-dependent kinase. *PLoS neglected tropical diseases* **2011**, *5* (4), e1033 DOI: 10.1371/journal.pntd.0001033.

(24) Bell, A. S.; Mills, J. E.; Williams, G. P.; Brannigan, J. A.; Wilkinson, A. J.; Parkinson, T.; Leatherbarrow, R. J.; Tate, E. W.; Holder, A. A.; Smith, D. F. Selective inhibitors of protozoan protein N-myristoyltransferases as starting points for tropical disease medicinal chemistry programs. *PLoS neglected tropical diseases* **2012**, *6* (4), e1625 DOI: 10.1371/journal.pntd.0001625.

(25) Cavazzuti, A.; Paglietti, G.; Hunter, W. N.; Gamarro, F.; Piras, S.; Loriga, M.; Allecca, S.; Corona, P.; McLuskey, K.; Tulloch, L.; et al. Discovery of potent pteridine reductase inhibitors to guide antiparasite drug development. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105* (5), 1448–1453 DOI: 10.1073/pnas.0704384105.

(26) Rachidi, N.; Taly, J. F.; Durieu, E.; Leclercq, O.; Aulner, N.; Prina, E.; Pescher, P.; Notredame, C.; Meijer, L.; Späth, G. F. Pharmacological assessment defines Leishmania donovani casein kinase 1 as a drug target and reveals important functions in parasite viability and intracellular infection. *Antimicrobial agents and chemotherapy* **2014**, *58* (3), 1501–1515 DOI: 10.1128/AAC.02022-13.

(27) Allocco, J. J.; Donald, R.; Zhong, T.; Lee, A.; Tang, Y. S.; Hendrickson, R. C.; Liberator, P.; Nare, B. Inhibitors of casein kinase 1 block the growth of Leishmania major promastigotes in vitro. *International Journal for Parasitology* **2006**, *36* (12), 1249–1259 DOI: 10.1016/j.ijpara.2006.06.013.

(28) Sterkers, Y.; Lachaud, L.; Crobu, L.; Bastien, P.; Pagès, M. FISH analysis reveals aneuploidy and continual generation of chromosomal mosaicism in Leishmania major. *Cellular Microbiology* **2011**, *13* (2), 274–283 DOI: 10.1111/j.1462-5822.2010.01534.x.

(29) Dumetz, F.; Imamura, H.; Sanders, M.; Seblova, V.; Myskova, J.; Pescher, P.; Vanaerschot, M.; Meehan, C. J.; Cuypers, B.; De Muylder, G.; et al. Modulation of Aneuploidy in Leishmania donovani during Adaptation to Different In Vitro and In Vivo Environments and Its Impact on Gene Expression. *mBio* **2017**, *8* (3), e00599–17 DOI: 10.1128/mBio.00599-17.

(30) Rogers, M. B.; Hilley, J. D.; Dickens, N. J.; Wilkes, J.; Bates, P. A.; Depledge, D. P.; Harris, D.; Her, Y.; Herzyk, P.; Imamura, H.; et al. Chromosome and gene copy number variation allow major structural change between species and strains of Leishmania. *Genome research* **2011**, *21* (12), 2129–2142 DOI: 10.1101/gr.122945.111.

(31) Lye, L. F.; Owens, K.; Shi, H.; Murta, S. M. F.; Vieira, A. C.; Turco, S. J.; Tschudi, C.; Ullu, E.; Beverley, S. M. Retention and Loss of RNA Interference Pathways in Trypanosomatid Protozoans. *PLOS Pathogens* **2010**, *6* (10), e1001161 DOI: 10.1371/journal.ppat.1001161.

(32) Bates, P. A. Complete developmental cycle of Leishmania mexicana in axenic culture. *Parasitology* **1994**, *108 (Pt 1)*, 1–9.

(33) Sereno, D.; Roy, G.; Lemesre, J. L.; Papadopoulou, B.; Ouellette, M. DNA Transformation of Leishmania infantum Axenic Amastigotes and Their Use in Drug Screening. *Antimicrobial agents and chemotherapy* **2001**, *45* (4), 1168–1173 DOI: 10.1128/AAC.45.4.1168-

1173.2001.
(34) Duncan, S. M.; Jones, N. G.; Mottram, J. C. Molecular & Biochemical Parasitology Recent advances in Leishmania reverse genetics : Manipulating a manipulative parasite. *Mol. Biochem. Parasitol.* **2017**, *216* (June), 30–38 DOI: 10.1016/j.molbiopara.2017.06.005.
(35) Cruz, A.; Beverley, S. M. Gene replacement in parasitic protozoa. *Nature* **1990**, *348* (6297), 171–173 DOI: 10.1038/348171a0.
(36) Cruz, A.; Coburn, C. M.; Beverley, S. M. Double targeted gene replacement for creating null mutants. *Proceedings of the National Academy of Sciences* **1991**, *88* (16), 7170–7174.
(37) Ambit, A.; Fasel, N.; Coombs, G. H.; Mottram, J. C. An essential role for the Leishmania major metacaspase in cell cycle progression. *Cell Death Differ.* **2008**, *15* (1), 113–122 DOI: 10.1038/sj.cdd.4402232.
(38) Casanova, M.; Gonzalez, I. J.; Sprissler, C.; Zalila, H.; Dacher, M.; Basmaciyan, L.; Späth, G. F.; Azas, N.; Fasel, N. Implication of different domains of the Leishmania major metacaspase in cell death and autophagy. *Cell Death Dis* **2015**, *6* (10), cddis2015288 DOI: 10.1038/cddis.2015.288.
(39) Castanys-Muñoz, E.; Brown, E.; Coombs, G. H.; Mottram, J. C. Leishmania mexicana metacaspase is a negative regulator of amastigote proliferation in mammalian cells. *Cell Death Dis* **2012**, *3* (9), e385 DOI: 10.1038/cddis.2012.113.
(40) Jiang, Y.; Roberts, S. C.; Jardim, A.; Carter, N. S.; Shih, S.; Ariyanayagam, M.; Fairlamb, A. H.; Ullman, B. Ornithine decarboxylase gene deletion mutants of Leishmania donovani. *J. Biol. Chem.* **1999**, *274* (6), 3781–3788.
(41) Vergnes, B.; Sereno, D.; Tavares, J.; Cordeiro-da-Silva, A.; Vanhille, L.; Madjidian-Sereno, N.; Depoix, D.; Monte-Alegre, A.; Ouaissi, A. Targeted disruption of cytosolic SIR2 deacetylase discloses its essential role in Leishmania survival and proliferation. *Gene* **2005**, *363*, 85–96 DOI: 10.1016/j.gene.2005.06.047.
(42) Murta, S. M. F.; Vickers, T. J.; Scott, D. A.; Beverley, S. M. Methylene tetrahydrofolate dehydrogenase/cyclohydrolase and the synthesis of 10-CHO-THF are essential in Leishmania major. *Molecular Microbiology* **2009**, *71* (6), 1386–1401 DOI: 10.1111/j.1365-2958.2009.06610.x.
(43) Dacher, M.; Morales, M. A.; Pescher, P.; Leclercq, O.; Rachidi, N.; Prina, E.; Cayla, M.; Descoteaux, A.; Späth, G. F. Probing druggability and biological function of essential proteins in Leishmania combining facilitated null mutant and plasmid shuffle analyses. *Molecular Microbiology* **2014**, *93* (1), 146–166 DOI: 10.1111/mmi.12648.
(44) Wang, Q.; Melzer, I. M.; Kruse, M.; Sander-Juelch, C.; Wiese, M. LmxMPK4, a mitogen-activated protein (MAP) kinase homologue essential for promastigotes and amastigotes of Leishmania mexicana. *Kinetoplastid biology and disease* **2005**, *4*, 6 DOI: 10.1186/1475-9292-4-6.
(45) Duncan, S. M.; Myburgh, E.; Philipon, C.; Brown, E.; Meissner, M.; Brewer, J.; Mottram, J. C. Conditional gene deletion with DiCre demonstrates an essential role for CRK3 in Leishmania mexicana cell cycle regulation. *Molecular Microbiology* **2016**, *100* (6), 931–944 DOI: 10.1111/mmi.13375.
(46) Hassan, P.; Fergusson, D.; Grant, K. M.; Mottram, J. C. The CRK3 protein kinase is essential for cell cycle progression of Leishmania mexicana. *Mol. Biochem. Parasitol.* **2001**, *113* (2), 189–198.
(47) Grant, K. M.; Dunion, M. H.; Yardley, V.; Skaltsounis, A.-L.; Marko, D.; Eisenbrand, G.; Croft, S. L.; Meijer, L.; Mottram, J. C. Inhibitors of Leishmania mexicana CRK3 cyclin-dependent kinase: chemical library screen and antileishmanial activity. *Antimicrobial agents and chemotherapy* **2004**, *48* (8), 3033–3042 DOI: 10.1128/AAC.48.8.3033-3042.2004.
(48) Grant, K. M.; Hassan, P.; Anderson, J. S.; Mottram, J. C. The crk3 gene of Leishmania mexicana encodes a stage-regulated cdc2-related histone H1 kinase that associates with p12. *J. Biol. Chem.* **1998**, *273* (17), 10153–10159.
(49) Walker, R. G.; Thomson, G.; Malone, K.; Nowicki, M. W.; Brown, E.; Blake, D. G.; Turner, N. J.; Walkinshaw, M. D.; Grant, K. M.; Mottram, J. C. High throughput screens yield small molecule inhibitors of Leishmania CRK3:CYC6 cyclin-dependent kinase. *PLoS neglected tropical diseases* **2011**, *5* (4), e1033 DOI: 10.1371/journal.pntd.0001033.
(50) Ivens, A. C.; Peacock, C. S.; Worthey, E. A.; Murphy, L.; Aggarwal, G.; Berriman, M.; Sisk, E.; Rajandream, M. A.; Adlem, E.; Aert, R.; et al. The genome of the kinetoplastid parasite, Leishmania major. *Science (New York, N.Y.)* **2005**, *309* (5733), 436–442 DOI:

10.1126/science.1112680.
(51) Peacock, C. S.; Seeger, K.; Harris, D.; Murphy, L.; Ruiz, J. C.; Quail, M. A.; Peters, N.; Adlem, E.; Tivey, A.; Aslett, M.; et al. Comparative genomic analysis of three Leishmania species that cause diverse human disease. **2007**, *39* (7), 839–847.
(52) Aurrecoechea, C.; Barreto, A.; Basenko, E. Y.; Brestelli, J.; Brunk, B. P.; Cade, S.; Crouch, K.; Doherty, R.; Falke, D.; Fischer, S.; et al. EuPathDB: the eukaryotic pathogen genomics database resource. *Nucleic Acids Research* **2017**, *45* (D1), D581–D591 DOI: 10.1093/nar/gkw1105.
(53) Supek, F.; Bošnjak, M.; Škunca, N.; Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **2011**, *6* (7), e21800 DOI: 10.1371/journal.pone.0021800.
(54) Kelly, J. M.; Ward, H. M.; Miles, M. A.; Kendall, G. A shuttle vector which facilitates the expression of transfected genes in Trypanosoma cruzi and Leishmania. *Nucleic Acids Research* **1992**, *20* (15), 3963–3969.
(55) Burle-Caldas, G. de A.; Grazielle-Silva, V.; Laibida, L. A.; DaRocha, W. D.; Teixeira, S. M. R. Expanding the tool box for genetic manipulation of Trypanosoma cruzi. *Mol. Biochem. Parasitol.* **2015**, *203* (1-2), 25–33 DOI: 10.1016/j.molbiopara.2015.10.004.
(56) Roberts, A. J.; Torrie, L. S.; Wyllie, S.; Fairlamb, A. H. Biochemical and genetic characterization of Trypanosoma cruzi N-myristoyltransferase. *The Biochemical journal* **2014**, *459* (2), 323–332 DOI: 10.1042/BJ20131033.
(57) Kangussu-Marcolino, M. M.; Cunha, A. P.; Avila, A. R.; Herman, J.-P.; DaRocha, W. D. Conditional removal of selectable markers in Trypanosoma cruzi using a site-specific recombination tool: proof of concept. *Mol. Biochem. Parasitol.* **2014**, *198* (2), 71–74 DOI: 10.1016/j.molbiopara.2015.01.001.
(58) Schumann Burkard, G.; Jutzi, P.; Roditi, I. Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters. *Mol. Biochem. Parasitol.* **2011**, *175* (1), 91–94 DOI: 10.1016/j.molbiopara.2010.09.002.
(59) Alsford, S.; Turner, D. J.; Obado, S. O.; Sanchez-Flores, A.; Glover, L.; Berriman, M.; Hertz-Fowler, C.; Horn, D. High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome research* **2011**, *21* (6), 915–924 DOI: 10.1101/gr.115089.110.
(60) Sidik, S. M.; Huet, D.; Ganesan, S. M.; Huynh, M.-H.; Wang, T.; Nasamu, A. S.; Thiru, P.; Saeij, J. P. J.; Carruthers, V. B.; Niles, J. C.; et al. A Genome-wide CRISPR Screen in Toxoplasma Identifies Essential Apicomplexan Genes. *Cell* **2016**, *166* (6), 1423–1435.e12 DOI: 10.1016/j.cell.2016.08.019.
(61) Jones, N. G.; Thomas, E. B.; Brown, E.; Dickens, N. J.; Hammarton, T. C.; Mottram, J. C. Regulators of Trypanosoma brucei Cell Cycle Progression and Differentiation Identified Using a Kinome-Wide RNAi Screen. *PLOS Pathogens* **2014**, *10* (1), e1003886 DOI: 10.1371/journal.ppat.1003886.
(62) Fernandez-Cortes, F.; Serafim, T. D.; Wilkes, J. M.; Jones, N. G.; Ritchie, R.; Mcculloch, R.; Mottram, J. C. RNAi screening identifies Trypanosoma brucei stress response protein kinases required for survival in the mouse. *Scientific Reports* **2017**, No. May, 1–10 DOI: 10.1038/s41598-017-06501-8.
(63) Stortz, J. A.; Serafim, T. D.; Alsford, S.; Wilkes, J.; Fernandez-Cortes, F.; Hamilton, G.; Briggs, E.; Lemgruber, L.; Horn, D.; Mottram, J. C.; et al. Genome-wide and protein kinase-focused RNAi screens reveal conserved and novel damage response pathways in Trypanosoma brucei. *PLOS Pathogens* **2017**, *13* (7), e1006477 DOI: 10.1371/journal.ppat.1006477.
(64) Mony, B. M.; MacGregor, P.; Ivens, A.; Rojas, F.; Cowton, A.; Young, J.; Horn, D.; Matthews, K. Genome-wide dissection of the quorum sensing signalling pathway in Trypanosoma brucei. *Nature* **2014**, *505* (7485), 681–685 DOI: 10.1038/nature12864.
(65) Alsford, S.; Eckert, S.; Baker, N.; Glover, L.; Sanchez-Flores, A.; Leung, K. F.; Turner, D. J.; Field, M. C.; Berriman, M.; Horn, D. High-throughput decoding of antitrypanosomal drug efficacy and resistance. *Nature* **2012**, *482* (7384), 232–236 DOI: 10.1038/nature10771.
(66) Baker, N.; Hamilton, G.; Wilkes, J. M.; Hutchinson, S.; Barrett, M. P.; Horn, D. Vacuolar ATPase depletion affects mitochondrial ATPase function, kinetoplast dependency, and drug sensitivity in trypanosomes. *Proceedings of the National Academy of Sciences of the United States of America* **2015**, *112* (29), 9112–9117 DOI: 10.1073/pnas.1505411112.
(67) Bushell, E.; Gomes, A. R.; Sanderson, T.; Anar, B.; Girling, G.; Herd, C.; Metcalf, T.;

Modrzynska, K.; Schwach, F.; Martin, R. E.; et al. Functional Profiling of a Plasmodium Genome Reveals an Abundance of Essential Genes. *Cell* **2017**, *170* (2), 260–272.e268 DOI: 10.1016/j.cell.2017.06.030.

(68) Gomes, A. R.; Bushell, E.; Schwach, F.; Girling, G.; Anar, B.; Quail, M. A.; Herd, C.; Pfander, C.; Modrzynska, K.; Rayner, J. C.; et al. A genome-scale vector resource enables high-throughput reverse genetic screening in a malaria parasite. *Cell Host and Microbe* **2015**, *17* (3), 404–413 DOI: 10.1016/j.chom.2015.01.014.

(69) Zhang, W. W.; Matlashewski, G. CRISPR-Cas9-mediated genome editing in Leishmania donovani. *mBio* **2015**, *6* (4), 1–14 DOI: 10.1128/mBio.00861-15.

(70) Sollelis, L.; Ghorbal, M.; Macpherson, C. R.; Martins, R. M.; Kuk, N.; Crobu, L.; Bastien, P.; Scherf, A.; Lopez-Rubio, J. J.; Sterkers, Y. First efficient CRISPR-Cas9-mediated genome editing in Leishmania parasites. *Cellular Microbiology* **2015**, *17* (10), 1405–1412 DOI: 10.1111/cmi.12456.

(71) Wen-Wei Zhang, Patrick Lypaczewski; 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* **2017**, *Jan-Feb; 2* (e00340-16.), 1–15.

(72) Soares Medeiros, L. C.; South, L.; Peng, D.; Bustamante, J. M.; Wang, W.; Bunkofske, M.; Perumal, N.; Sanchez-Valdez, F.; Tarleton, R. L. Rapid, Selection-Free, High-Efficiency Genome Editing in Protozoan Parasites Using CRISPR-Cas9 Ribonucleoproteins. *mBio* **2017**, *8* (6), e01788–17 DOI: 10.1128/mBio.01788-17.

(73) Beneke, T.; Madden, R.; Makin, L.; Valli, J.; Sunter, J.; Gluenz, E. A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. *Royal Society Open Science* **2017**, *May* DOI: 10.1098/rsos.170095.

(74) 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* **2014**, *6* (1), e02097–14 DOI: 10.1128/mBio.02097-14.

(75) 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* **2015**, *6* (4), e01012 DOI: 10.1128/mBio.01012-15.

All other papers that were identified as part of the literature review are referenced in Table S1 and Table S2.

Figures and Tables

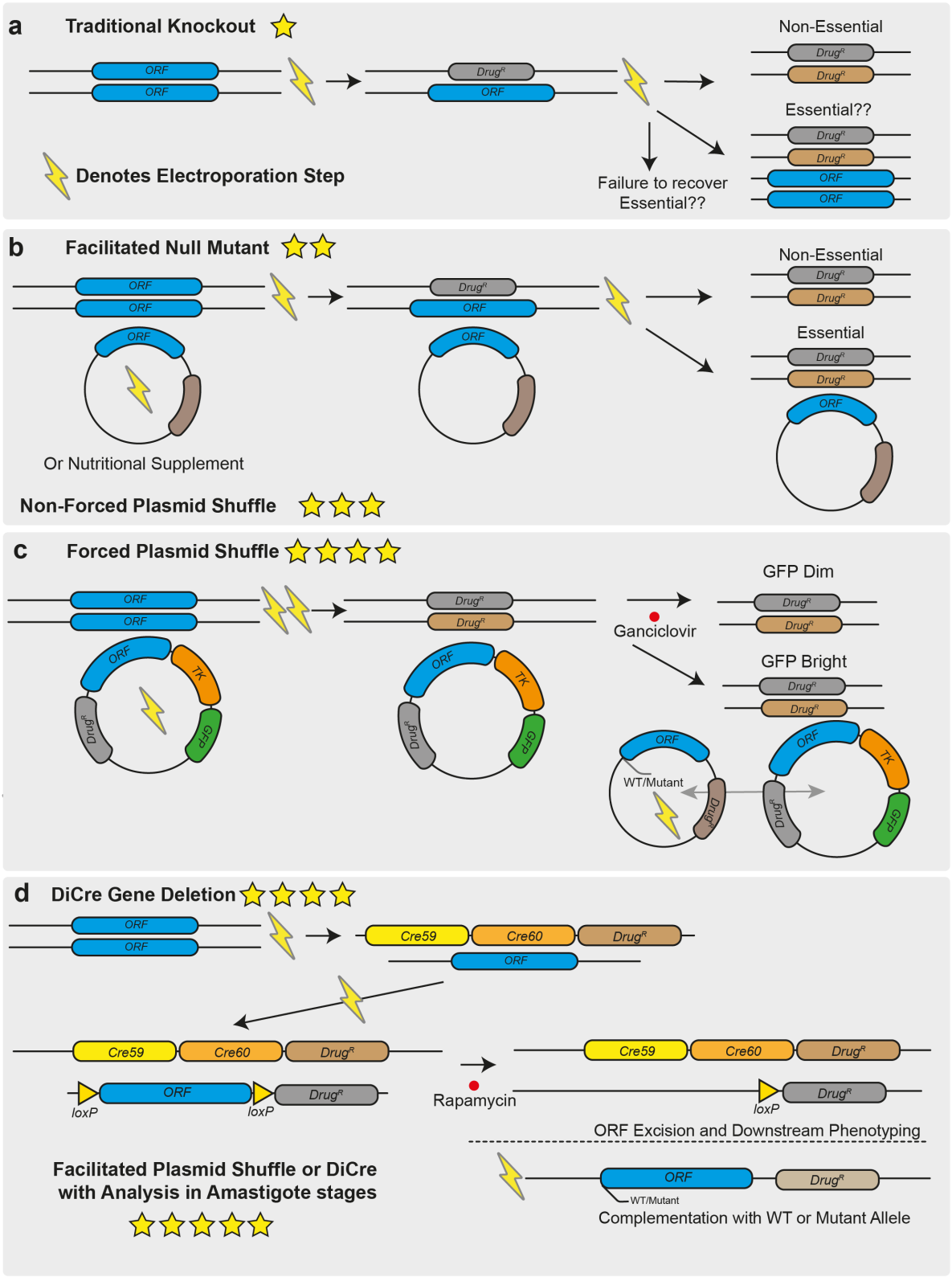


Figure 1: Overview of techniques that can be used for genetic target validation in *Leishmania.* A: Gene deletion by homologous replacement. Drug resistance markers are

targeted to the gene of interest by long homology flanks (0.5-1kb) in sequential transfections by electroporation. This process can now be facilitated using CRISPR/Cas9 and short homology flanked cassettes in a single transfection. Deletions targeting essential genes will result in cell death and failure to isolate null mutants, or in ploidy changes that allow the cell to retain alleles of the wild-type locus as well as drug resistance markers. B: Facilitated null mutant with unforced plasmid shuffle. An episome expressing the gene of interest is first transfected into the cell line, or nutritional supplement is provided, allowing it to survive the subsequent deletion of the chromosomal alleles of the gene of interest. The drug selection pressure for the episome can be removed and retention of the plasmid determined, if the gene is not essential then it will be possible to isolate parasites that lack the episome. C: Forced plasmid shuffle. As in B., an episome expressing the gene of interest is transfected into the parasite to allow deletion of the chromosomal alleles of the target gene. The episome also encodes a negative-selectable marker, herpes simplex virus thymidine kinase. Selection with ganciclovir favours survival of parasites that lack the episome, so if a gene is non-essential the episome will rapidly be lost from the population, but will be retained for an essential gene despite the associated costs. Addition of a second episome containing mutant versions of the gene of interest allow for exploration of the roles of specific domains and residues in the encoded protein for correct gene function, by assessing which plasmid of the two is preferentially retained. D: DiCre Inducible Gene Deletion. One allele of the target gene is replaced by a drug selectable cassette containing a “floxed” allele, in a second transfection stage the remaining allele is replaced by a second drug resistance marker. Addition of rapamycin induces DiCre dimerization, and excision of the floxed allele, the phenotypes that emerge in the induced null mutants can then be analysed. As in C., complementation allows for assessment of null mutant specificity and

functional assessment of defined domains or residues in the protein. In all panels, the number of stars indicate the quality of the genetic evidence for gene essentiality, with one star being the weakest, and 5 being the strongest.

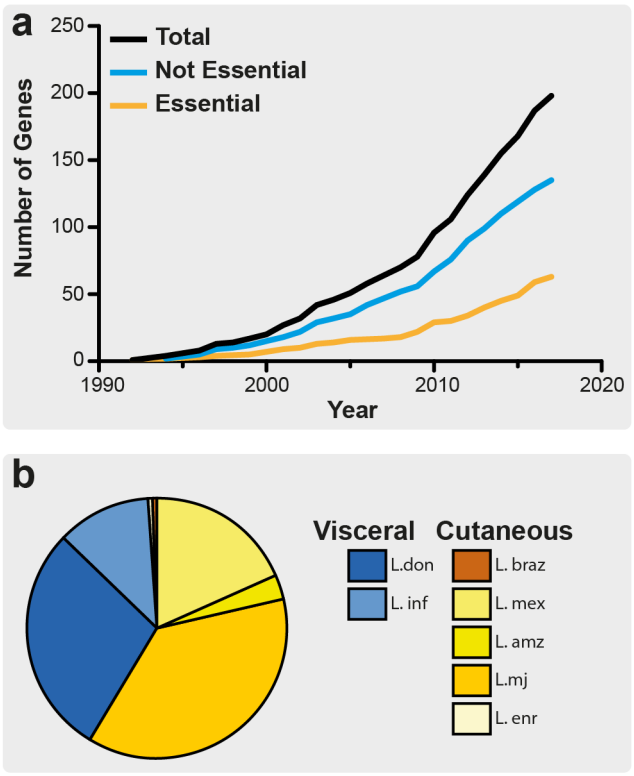


Figure 2: Overview of the number of *Leishmania* genes with published attempts at creation of a null mutant. A: Line graph depicting cumulative number of genes for which attempts have been made to generate null mutants, for human infective *Leishmania* species. Only the first attempt at a gene deletion for each individual gene was recorded. Data from this study were ordered by year and cumulative values of publications per year were derived, the total number of attempted gene deletions is shown as well as the number of essential genes identified. B: Pie chart showing proportion of unique gene deletion attempts by species of Leishmania, cutaneous species are shaded in yellows, visceral species in blues.

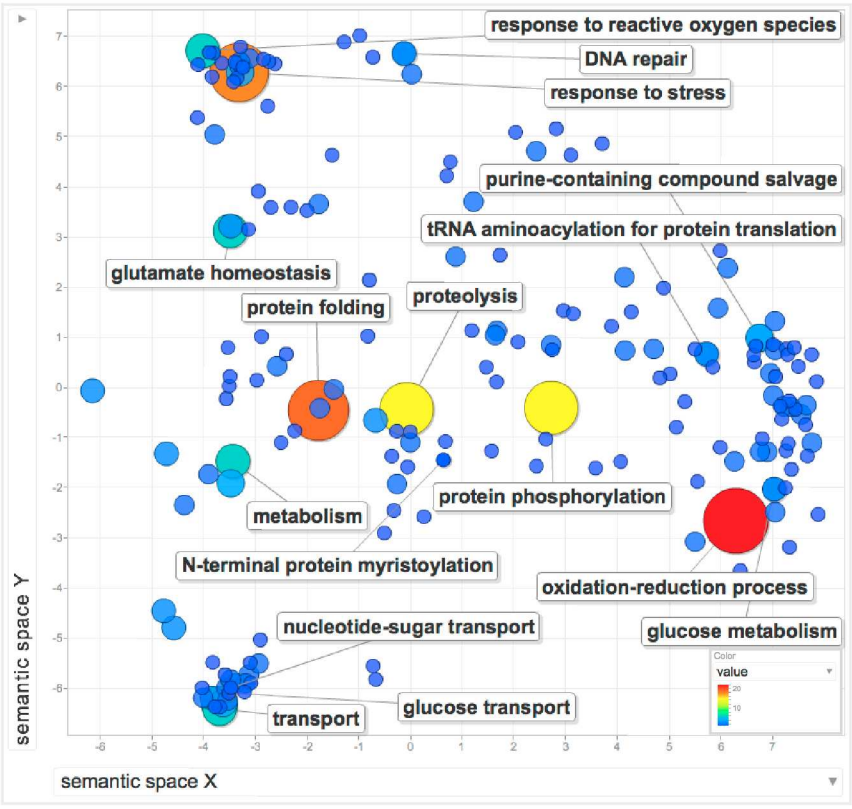


Figure 3: REVIGO analysis of GO Terms associated with targeted *Leishmania* genes.

L. major orthologues for all gene IDs in Table S1 were used to recover the associated GO terms for biological processes. The number of occurrences of each GO term was used to weight a REVIGO analysis, depicted as bigger circles and hotter colours. The more frequently occurring biological processes are annotated in the figure, as are key (but less frequent) GO terms such as N-terminal protein myristoylation.

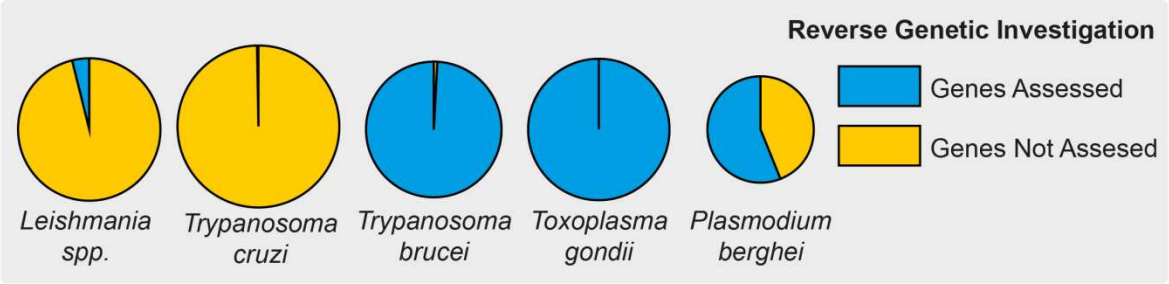


Figure 4: Number of reverse genetic manipulations of *Leishmania* in comparison to other model parasitic protozoans. A: Comparison of reverse genetic screening in *Leishmania* spp

(Table S1), *Trypanosoma cruzi* (Table S2), *Trypanosoma brucei*[59], *Toxoplasma gondii*[60] and *Plasmodium berghei*[67]. Pie chart segments depict the proportion of genes in each organism that have been targeted using reverse genetics, with the overall area of the pie depicting the relative sizes of the genomes.

Table 1: Definitions of key terms.

Term	Definition	Synonym
Allele	Variant form of a gene, for diploid chromosomes there are two alleles in the nuclear DNA (if identical these are homozygous, if different these are heterozygous), further alleles or copies of a gene can be complemented on an episome or integrated into the nuclear DNA in wild-type, mutant or floxed forms.	
Gene Deletion by Double Homologous Replacement	An allele of a gene is replaced by a drug resistance marker resulting in deletion of the coding DNA sequence of the allele. This is performed twice, in sequential steps, with two different drug resistance markers to result in replacement of two alleles of a gene, yielding a null mutant if the gene is not essential.	Knockout (verb)
Null Mutant	A parasite strain that lacks both alleles of a gene of interest after a deletion strategy.	Knockout (noun)
Facilitated Null Mutant	Deletion of both chromosomal alleles can be performed but requires prior genetic (episomal allele) or nutritional supplementation.	Episomal Rescue, Conditional Null Mutant
Unforced Plasmid Shuffle	Drug selection for plasmid episome that complements a chromosomal null mutant is removed, loss or retention of this plasmid is analysed.	Plasmid Cure
Forced Plasmid Shuffle	Complementing plasmid episome contains both a positive and negative selectable marker which allows a selection pressure to be imposed to favour parasites that lose the episome	
Inducible Gene Deletion	A parasite line is engineered to contain an allele that can be removed by addition of a trigger – for example, rapamycin induced DiCre recombinase deletion of a floxed allele	
Floxed	Flanked by *loxP* (locus of X-overP1) site, used to describe alleles of genes that can be inducibly deleted by a Cre or DiCre (Inducible Dimerizable Cre) recombinase system.	
Complementation	Restoration of null mutants by addition of an extra allele of the gene of interest (episomally or genomically) or by addition of a nutritional supplement in the case of gene that encodes a metabolic enzyme.	Add Back
Episome	Circular DNA molecule that can be replicated independently of the nuclear or kinetoplast DNA.	Plasmid

Table 2: Scoring system for the assessing the quality of evidence for target essentiality by gene deletion techniques.

Quality Score	Definition	Issues	*LEISHMANIA*		*TRYPANOSOMA CRUZI*	
			Number	Genes	Number	Genes
*	Attempt to perform gene deletion by double homologous replacement failed.	Changes in ploidy indicative of essentiality. May be a technical failure.	30	*TR, A2-A2REL, SODB1, LACK, PK, CRN12, TOR1, TOR2, SPASE I, MYO21, HSLU1, LHR1, HUS1, HEMAC, ENDOG, ASNA, ARP, RAD51-6, AIRK, LMIT1, RAD50, TYRRS, ABC3, RAB5A, RAB5B, TWF, SIR2RP2, SODA, ACECS, CPN10*	15	*TC52, DHOD, GALE, DHFR-TS, ECH1, SUB2, GPI8, GPI12, CRT, IP3R, RPA2, GALF, CYP51, STI1*
**	Facilitated null mutant. Gene of interest is complemented with an episome or nutritional supplement, allowing genomic alleles to be deleted.	Shows gene locus can be targeted for deletion.	20	*TUB, CRK1, NMT, SIR2, RPC2, TOPS, TXN1, GSH1, GLO1, SGT, DHS34, HSLV, NTR, TRYS, CYP51, EIF4E, RAD51-3, RPIB, RAD9, LYSRS-1*	1	*NMT*
***	Unforced plasmid shuffle. Plasmid retained in the absence of antibiotic selection	Indirect evidence for a gene being essential. Best carried out in vivo.	12	*PRT1, ODC, DHFR-TS, SPSDS, ADOMETDC, ARG, DHCH1, STI1, UMPS, CPS, UPRT*	-	-
****	Forced plasmid shuffle or DiCre inducible deletion	Death of parasite after induction is evidence of essentiality.	2	*H2A.Z, H2B.V, MPK4*	-	-
*****	As above but with analysis in amastigote stages and or mouse model.	Application to amastigotes provides best evidence for essentiality in vivo.	1	*CRK3*	-	-