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

Nathaniel G. Jones, Carolina M. C. Catta-Preta, Ana Paula C. A. Lima, and Jeremy C. Mottram

**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

*Leishmaniasis* is a disease that mainly affects those burdened by extreme poverty, across large swaths of the tropics and subtropics. 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 the destruction of mucocutaneous surfaces to fatal disease caused by visceral leishmaniasis; dermal sequelae can also occur following the 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.

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 the 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 means that several new medicines may be required to satisfy quite stringent target product profiles (TPPs) for visceral and cutaneous leishmaniasis. New chemical entities will probably require different chemical features for achieving the desired bioavailability in skin versus viscera coupled with the challenge of drugging an intracellular parasite that resides in a phagolysosomal compartment. There is growing appreciation that *Leishmania* infections can also form quiescent, persistent forms. Targeting these stages is desirable, though it may also prove challenging. There are a number of ongoing drug discovery programs that aim to address this need, typically working within defined phenotypic or target-based screening strategies to best achieve success. Phenotypic or whole organism screening selects for compounds that cause a loss of parasite fitness, whereas the target-based approach involves screening against a selected target of interest thought to be essential to parasite survival. We will assess these approaches and the role that genetic target validation can play within them.

**TARGET VALIDATION IN DRUG DISCOVERY PROGRAMS**

Drug target validation for antiparasitic 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 the 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 the 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 noninfectious 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, DFM0), a compound that inhibits ornithine decarboxylase.11 It is notable that for most antileishmanial drugs (miltefosine, paromomycin, antimonials, and pentamidine) 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 druglike 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 very stringent; consequently, few bioactive compounds are identified.9 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, then it may conceivably identify diverse target-class inhibitors whereas sublibraries can be used to increase the chances of hits within a certain target class, for example, by using a protein kinase inhibitor-focused library.14 The optimization of hit compounds by medicinal chemistry involves a level of randomness in which 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 a compound conjugated to beads,16 and the use of overexpression libraries17,18 or metabolomics to identify blocks in biochemical pathways.19 It is also important to have an understanding of an experimental compound’s mode of action as this can enable an assessment of the likelihood of resistance mechanisms evolving in the parasite. Recent successful examples of the phenotypic screening of large compound libraries or focused subset chemical libraries against Leishmania parasites include the identification of the “LeishBox” 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 and 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 and then chemically optimized 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 on the basis of gene essentiality in addition to other criteria, for example, the presence of close homologues in the parasite genome that may allow for an 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 emphasize the importance of a good understanding of parasite genetics for drug target prioritization. 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 rigor 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 an HTS campaign26 after being chemically validated as a target.

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.9 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 lifecycle. Some species of Leishmania, such as L. mexicana, are readily converted to axenic amastigotes in vitro32 and this can facilitate the study of gene function in the pathogenic stages, but phenotypic analysis is perhaps best carried out using macrophages to perform
in intracellular amastigote growth and survival assays and/or mouse infections. Amastigotes can be genetically manipulated, but this is not a common practice. Typically, in a model system 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 expressing transgenes. (See Table 1 for a 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. On the basis of the literature, we used a scoring system to rank the strength of validation resulting from gene deletion studies, on a 1–5-star scale (Table 1 and Figure 1), and we discuss these below, with examples. We propose that the higher scoring methods provide better-quality evidence of a gene’s essentiality but acknowledge that 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 TO THE 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. The authors demonstrated proof of concept, which involved deleting a single allele of *DHFR-TS* in wild-type *L. major*. This was then repeated 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. The 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*. The development of new resistance markers allowed for gene deletions of two alleles of a gene, given that *Leishmania* was considered at the time to have a diploid genome that allowed for 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 but had duplicated their genome to retain the gene of interest as well as the drug resistance markers, or 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 electro- poration became more sophisticated, transfection efficiencies increased, allowing some previously intractable genes to be revealed as nonessential. The Metacaspase gene, *MCA*, initially could not be deleted from *L. major* 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* and *L. mexicana* using a newer high-efficiency electroporation system that causes minimal cellular damage. Generating this deletion allowed for the characterization of the role of *MCA* in amastigotes and mice where the null mutant exhibited an increase in virulence.

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 (Table S1). The advantage of a chemical complementation is that it allows for rapid withdrawal and the resulting phenotype was investigated. Where nutritional supplementation could not be used to facilitate the deletion of essential genes, researchers began to introduce an episomal copy of the gene of interest into the parasite prior to the deletion of the two chromosomal alleles (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

### Table 1. Definitions of Key Terms

<table>
<thead>
<tr>
<th>term</th>
<th>definition</th>
<th>synonym</th>
</tr>
</thead>
<tbody>
<tr>
<td>allele</td>
<td>This is a 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). Additional 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.</td>
<td>knockout (verb)</td>
</tr>
<tr>
<td>gene deletion by double homologous replacement</td>
<td>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 the replacement of two alleles of a gene, yielding a null mutant if the gene is not essential.</td>
<td>knockout (noun)</td>
</tr>
<tr>
<td>null mutant</td>
<td>This is a parasite strain that lacks both alleles of a gene of interest after a deletion strategy.</td>
<td>episomal rescue, conditional null mutant</td>
</tr>
<tr>
<td>facilitated null mutant</td>
<td>Both chromosomal alleles can be deleted, but this requires prior genetic (episomal allele) or nutritional supplementation.</td>
<td></td>
</tr>
<tr>
<td>unforced plasmid shuffle</td>
<td>Drug selection for a plasmid episome that complements a chromosomal null mutant is removed, and the loss or retention of this episome is analyzed.</td>
<td>plasmid cure</td>
</tr>
<tr>
<td>forced plasmid shuffle</td>
<td>The complementing plasmid episome contains both a positive and negative selectable marker, which allows a selection pressure to be imposed to favor parasites that lose the episome.</td>
<td></td>
</tr>
<tr>
<td>inducible gene deletion</td>
<td>A parasite line is engineered to contain an allele that can be removed by the addition of a trigger, for example, the rapamycin-induced DiCre recombinase deletion of a floxed allele.</td>
<td></td>
</tr>
<tr>
<td>floxed</td>
<td>Flanked by theloxP (locus of X-overP1) site, this term is used to describe alleles of genes that can be inducibly deleted by a Cre or DnCre (inducible dimerizable Cre) recombinase system.</td>
<td></td>
</tr>
<tr>
<td>complementation</td>
<td>This involves the restoration of null mutants by the addition of an extra allele of the gene of interest (episomally or genomically) or by the addition of a nutritional supplement in the case of a gene that encodes a metabolic enzyme.</td>
<td>add back</td>
</tr>
<tr>
<td>episome</td>
<td>A circular DNA molecule that can be replicated independently of the nuclear or kinetoplast DNA.</td>
<td>plasmid</td>
</tr>
</tbody>
</table>
the locus being refractory to genetic manipulation (Figure 2B, two stars).
deletion attempts by species of Leishmania genes identified. Values of publications per year were derived, where the total number of recorded. Data from this study were ordered by year, and cumulative difficulty. Generated. For essential genes, even after prolonged culturing of the facilitated null mutant in a 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, three 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 the loss of the episome in a facilitated null mutant, negative selection or forced plasmid shuffle can be achieved using a negative selectable marker herpes simplex thymidine kinase (TK). Ganciclovir is used to impose a fitness cost on parasites that retain the episome, selecting for parasites that lose the episome. If it is complementing an essential gene, then the parasites must retain the episome to remain viable, so any loss will prove fatal. The 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 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 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. 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 shuffling in that an 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 enforced plasmid shuffling data suggesting that MPK4 from L. mexicana\(^{44}\) is essential to 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 the 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 a 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 the 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 poorly tractable target for drug discovery programs.

**DiCre Inducible Deletion.** In order to directly measure the 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 possible to delete both chromosomal alleles only in the presence of an epimeme expressing CRK3, with the epimeme being retained in the 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\(^{36}\) gene from promastigote cells and observe precisely the phenotype resulting in CRK3 null mutant—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 “zooids” (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 that 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, but the gene was deleted in stationary phase cultures (metacyclic-enriched), allowing the parasites to be inoculated into mice and the progress of the infection to be monitored by noninvasive, 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 a reduction in bioluminescence from mice infected with the CRK3-deficient cells. This was ranked as a five-star validation (1, 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 focused on identifying the first instance where researchers had attempted to generate a null mutant for a given gene, but we 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, along with a broad categorization of the phenotype, i.e. essential for promastigotes, conditionally essential, or nonessential. 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 nonessential. The rate of genetic deletion 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,\(^{71}\) but has remained fairly constant between 2008 and 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 and REVIGO (Figure 3, Table S1). GO was analyzed 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, and 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.

Because of the high variability in the assay protocols used for phenotypic analyses in each study, it is challenging to allocate each gene deletion mutant to a defined phenotypic category, other than essential or nonessential 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 of 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 the 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 the expression of genes from episomes. As for Leishmania, we focused on identifying the first study to report the deletion attempt for any given gene. On the basis of our five star scoring system, studies on 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 to be essential in epimastigotes (Figure 4, Table S2). Only one
study, of NMT, used a variant of the facilitated null approach to validate the target gene; the 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, 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 systems were applied to *T. cruzi*, opening the possibility for better target validation.

### PROSPECTS FOR LARGE-SCALE 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, the 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. 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 and medium-throughput gene family screens. 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 the response to stress. Of particular relevance is the way that this technology was used to identify organelles, proteins, and metabolic pathways responsible for the mode of action of trypanocidal drugs 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. berghei* genome to be assessed for essentiality. The PlasmoGem project used recombinating to build libraries of long-homology flanked deletion vectors that can be transfected into parasites prior to infecting mice. 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 approximately 60% of the *P. berghei* genes were considered essential. This compares to approximately 30% in *T. brucei*, and on the basis of our literature search, we expect *Leishmania* to be similar to 32.6% of the *Leishmania* genes investigated so far that are essential.

Increased throughput and efficiency for the 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*. *T. gondii* has a propensity for repairing double-stranded DNA breaks by error-prone, nonhomologous end joining (NHEJ), allowing for the 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 the 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. This is due to *Leishmania* lacking a functional NHEJ pathway, although it can repair a 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. These factors permit easy and fast deprioritization of nonessential 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 and 2017, possibly even genome-wide. Because the other genetic techniques, such as DiCre, are more labor-intensive, it would be beneficial to focus them on prioritized 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.

While constructing the spreadsheet of attempted gene deletions, we were struck by the range of different phenotypic analyses that were performed, which restricted placing 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 analyze all of the mutants in the same assay, which allows for an easier comparison of phenotypes. The integration of such data into EuPathDB would make an exceptionally useful resource. Target selection for *Leishmania* can be guided by mining the data sets available in resources such as TriTrypDB, so we encourage fellow researchers to contribute to this community resource with both historical and current data.

### CONCLUSIONS

The discovery of new drugs for intracellular pathogens is a complex process that usually has a 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 optimization of existing genetic techniques, the development of new techniques, and the application of best practices 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 now available for conducting research on potential drug targets in *Leishmania* are optimal, and the new technologies available will facilitate and improve the number of genetically well-validated targets suitable for entry into HTS campaigns.
Perspective

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsinfecdis.7b00244.

Lists of genes for Leishmania species, where gene deletions have been attempted. This table contains a tab for essential genes and a tab for nonessential genes.

(XLSX)

Lists of genes for Trypanosoma cruzi, where gene deletions have been attempted. This table contains a tab for essential genes and a tab for nonessential genes.

(XLSX)

Descriptions of Tables S1 and S2 (PDF)

AUTHOR INFORMATION

Corresponding Author
*E-mail: jeremy.mottram@york.ac.uk.

ORCID
Nathaniel G. Jones: 0000-0001-7328-4487
Jeremy C. Mottram: 0000-0001-5574-3766

Author Contributions
N.G.J., C.M.C.C.-P., and J.C.M. designed and developed the work. N.G.J., C.M.C.C.-P., A.P.C.A.L., and J.C.M. carried out data analysis. N.G.J., C.M.C.C.-P., and J.C.M. wrote the article.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Research in the J.C.M. laboratory is supported by a Wellcome Trust Investigator Award (200807/Z/16/Z). The research leading to these results has, in part, received funding form the Research Council United Kingdom Grand Challenges Research Funder under grant agreement ‘A Global Network for Neglected Tropical Diseases’ grant number MR/P027989/1. We apologize for any published gene deletions that have been inadvertently omitted during our search.

REFERENCES


