First efficient CRISPR-Cas9-mediated genome editing in *Leishmania* parasites

Lauriane Sollelis[1,2], Mehdi Ghorbal[1,2], Cameron Ross MacPherson[3], Rafael Miyazawa Martins[3], Nada Kuk[1], Lucien Crobu[2], Patrick Bastien[1,2,4], Artur Scherf[3], Jose-Juan Lopez-Rubio[1,2] and Yvon Sterkers[1,2,4#]

Running title: **CRISPR-Cas9 for genome editing in *Leishmania***

[1] University of Montpellier, Faculty of Medicine, Laboratory of Parasitology-Mycology, [2] CNRS 5290 - IRD 224 – University of Montpellier (UMR "MiVEGEC"), [3] Institut Pasteur - INSERM U1201 - CNRS ERL9195, "Biology of Host-Parasite Interactions" Unit, Paris, France, [4] CHRU (Centre Hospitalier Universitaire de Montpellier), Department of Parasitology-Mycology, Montpellier, France

[#]To whom correspondence should be addressed.

Postal address: Laboratoire de Parasitologie-Mycologie**,** CHRU de Montpellier, 39, Avenue Charles Flahault, 34295 Montpellier cedex 5, France.

Telephone: +33 467 63 55 13, Fax: +33 467 63 23 58, Email:yvon.sterkers@univ-montp1.fr

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/cmi.12456

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

Summary

Protozoan pathogens that cause leishmaniasis in humans are relatively refractory to genetic manipulation. In this work we implemented the CRISPR-Cas9 system in *Leishmania* parasites and demonstrated its efficient use for genome editing. The Cas9 endonuclease was expressed under the control of the DHFR-TS promoter and the single guide RNA was produced under the control of the U6snRNA promoter and terminator. As a proof of concept we chose to knockout a tandemly repeated gene family, the paraflagellar rod-2 locus. We were able to obtain null mutants in a single round of transfection. In addition, we confirmed the absence of off-target editions by whole genome sequencing of two independent clones. Our work demonstrates that CRISPR-Cas9-mediated gene knockout represents a major improvement in comparison with existing methods. Beyond gene knockout, this genome editing tool opens avenues for a multitude of functional studies to speed up research on leishmaniasis.

Introduction

The protozoan parasite *Leishmania* causes a large spectrum of diseases in 88 countries around the world, thus representing the second cause of parasite-related deaths after malaria. The absence of effective vaccines, the extremely limited available therapeutic arsenal and the development of drug resistances are major drawbacks for clinical case management and eradication (den Boer *et al.*, 2011).

In the post-genomic era, genome editing became a major tool used in a wide range of experimental systems (Gilles *et al.*, 2014). Gene knockouts were particularly used, aiming at the replacement of the targeted region by a selective marker. For this, classical molecular biology methods were first used, based on double cross-over events in homology regions (HRs) initiated by a random DNA double strand break (DSB). Linear DNA molecules bearing HRs at the ends were then used in order to increase efficiency, as the cellular machinery

This article is protected by copyright. All rights reserved.

recognizes the DNA free ends which trigger recombination, then using the endogenous locus as a template. DSB is one of the most deleterious DNA damages that can affect cell viability; and cells have developped different DNA repair pathways to overcome DSBs and survive: homologous recombination, non homologous end joining (NHEJ) and micro-homology mediated end joining (MMEJ). In *Leishmania*, homologous recombination is prevailing, and, in spite of the presense of KU proteins, the NHEJ pathway seems to be absent (Passos-Silva *et al*., 2010). More recently, it has been shown that *Trypanosoma cruzi*, a related parasite of the trypanosomatid familly, can use the MMEJ pathway in the absence of template to repair DSB (Peng *et al.*, 2015).

In order to enhance the probability of a DSB and to target it on a desired region of the genome, different genome editing systems using engineered nucleases have been developed, such as Zinc-finger nucleases, Talen and CRISPR-Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR associated endonuclease 9) (Gaj *et al.*, 2013). None of these have been implemented so far in *Leishmania*. The CRISPR-Cas9 editing system, which was engineered from a prokaryotic immune system (Jinek *et al.*, 2012), is constituted by a single guide RNA (sgRNA) which guides the Cas9 endonuclease to cause targeted DSBs. The sgRNA carries the Cas9 binding domain and a customizable 20 nt-sequence (hereafter referred to as the seed) which gives the specificity for the target DNA site. DSB occurs three nucleotides upstream of a protospacer-adjacent motif (PAM), within the guide RNA target region (Jinek *et al.*, 2012). CRISPR-Cas9 system has been recently developed as a powerful technique for genome editing and as a molecular genetic tool to study gene function in various organisms, including protozoan parasites like *Plasmodium*, *Toxoplasma* and *Trypanosoma cruzi* (Ghorbal *et al.*, 2014, Shen *et al.*, 2014, Wagner *et al.*, 2014, Peng *et al.*, 2015). The system appears particularly interesting in *Leishmania* sp. because classical genome manipulation methods experience serious limitations in this organism: (i) one of them is the difficulty in realizing gene knockouts due to its high genomic plasticity, and in particular to 'mosaic aneuploidy' (Sterkers *et al.*, 2011, Lachaud *et al.*,

This article is protected by copyright. All rights reserved.

2014); indeed, more than two rounds of transfection may be necessary, and in the case of essential genes, additional copies of the targeted gene/chromosome fragment/whole chromosome may arise at each round of transfection (Dubessay *et al.*, 2002); (ii) moreover, RNA interference is inefficient due to the absence of RNAi-related genes like Dicer and Argonaute in most species (Lye *et al.*, 2010). Here, we set up the CRISPR-Cas9 system in *Leishmania major* using the disruption of the paraflagellar rod (PFR)-2 genes as a proof of concept. In addition, we confirmed the absence of off-target edition by whole genome sequencing, as this is a major concern which is generally poorly addressed in this type of experiments.

Results and Discussion

Edition of the PFR2 locus

PFR2 is a major constituent of the PFR, which is a trypanosomatid-specific filamentous network attached to the axoneme of the flagellum (Maga *et al.*, 1999). It is encoded by three genes tandemly arrayed in a single locus on chromosome 16: LmjF16.1425 (from 591,295 to 593,094 bp), LmjF.16.1427 (594,664-596,463 bp), and LmjF.16.1430 (598,036-599,835 bp). We chose to edit the PFR2 locus because (i) it is not an essential gene since null mutants that divide normally in culture have been reported (Santrich *et al.*, 1997, Lye *et al.*, 2010) and (ii) specific antibodies targeting the PFR2 are available (Ismach *et al.*, 1989). To edit the PFR2 locus, we have chosen to use a two plasmids strategy: pTCAS9 carrying the Cas9 (Cong *et al.*, 2013) endonuclease gene, and pLS7 which encodes the sgRNA and the donor cassette constituted by the HRs and the resistance marker (Figure 1). We first generated *L. major* mutant parasites (termed LmjF_3171-CAS9 or M756) stably carrying the pTCAS9 episomal vector (Figure 1A). Expression of CAS9 was checked by Western blot using an anti-flag antibody (Figure 1B). The system requires sgRNA expression without RNA polII-associated modifications. RNA polIII promoters like U6 has been successfully used to implement the CRISPR-Cas9 system in other organisms, and in particular, U6 small nuclear (sn)RNA regulatory elements have been used in *P. falciparum* (Ghorbal *et al.*, 2014). The

This article is protected by copyright. All rights reserved.

Accepted Article

promoter and terminator of the U6snRNA (LmjF.24snRNA.01) have been reported in *Leishmania* and the minimal sequences to promote its transcription characterized (Nakaar *et al.*, 1994). The promoter sequence used here include the tRNAGln with its canonical BOX A and B, flanked by the 5'UTR of U6snRNA (215 bp). The terminator comprised 120 bp of the 3'UTR from U6snRNA. In pLS7, the sgRNA was inserted between the promotor and the terminator (Figure 1BC). To ensure maximal specificity and efficiency of the CRISPR-Cas9 editing system, we searched for (i) two large HRs to promote efficient homologous recombination and (ii) a seed close to one of the HRs. We designed two HRs of ~1kb flanking the complex locus PFR2 and a seed which targeted a 20-nt single copy sequence, preceding an NGG motif in the 5'UTR next to HR1. The HRs were cloned flanking the puromycin resistance gene (Figure 2A and Table S1). M756 (LmjF3171-CAS9) promastigotes were transfected with the linearized pLS7 plasmid and transfectant parasites (M790) were obtained at day 11. PCR experiments showed the presence of edited loci but also the persistence of wild type loci (Figure 2B). FISH analysis was then performed using two DNA probes from chromosome 16, one targeting the PFR2 locus and the second one another repeated gene (LmjF16.1030): PFR2-positive cells were reduced in M790 in comparison with the wild type (LmjF), whereas LmjF16.1030-positive cells were similar in both cell lines (Figure 2C and Table S2). FISH analysis thus detected a subpopulation of parasites with no signal for PFR2 (Figure 2C). This was confirmed by a specific immunofluorescence assay, showing both PFR2-positive and PFR2-negative cells in the M790 cell line. We then decided to clone this cell line by limiting dilution. Eight clones were screened by PCR and FISH. In two clones (M790_B4 and M790_D2), no remaining PFR2 sequence was found neither by PCR nor by FISH (Figure 2BC and Table S2); no PFR2 expression was found by Western blot and the flagella were not labeled in the immunofluorescence assay (Figure 2D and Table S2). Besides, when we transfected circular pLS7, the resulting transfectant cell line (M788) was found positive both for PFR2 and for puromycin integration using PCR; and a mixed population of PFR2-positive and -negative cells was observed using the immunofluorescence assay (Table S2). As negative controls,

This article is protected by copyright. All rights reserved.

Accepted Article

we transfected either linear or circular pLS6 plasmids, *i.e.* plasmids containing the HRs and the sgRNA with no seed: no edited genomes were obtained in the resulting transfectant cell lines (M809 and M810) (Table S2).

Our observation of mixed cell populations suggests that the efficiency of CRISPR-Cas9 edition was not 100%. This may be explained in two ways: (i) as commonly seen in *Leishmania* when using episomal expression vectors, Cas9 expression levels may be heterogeneous in the population, with high expression and low expression cells; (ii) sgRNA transcription driven by the U6snRNA promotor and terminator minimal sequences may be weak. A second concern is how the non-edited cells could resist to puromycin: for this, we suspected either the presence of some circular pLS7 or random integration in the genome.

Assessment of the CRISPR-Cas9 procedure and off-target mutation analysis

In order to verify the absence of off-target activity, we performed whole genome sequencing of the parental (LmjF_3171) and two edited clones (M790_B4 and M790_D2). The Protospacer Workbench software suite was used to find putative off-targets up to 10-nt different from the original seed (MacPherson and Scherf, 2015). The software scored each off-target according to the algorithm developed by Hsu *et al.*, 2013. The overall quality of the sgRNA design was visually inspected by ranking and then plotting the Hsu *et al.*'s score of each off-target site (Figure 3). These analyses showed a clear separation between the intended target and the next-best target, as well as rapidly decreasing scores as rank increases (Figure 3). The top-10 potential off-target sites are tabulated in Table S3. A hallmark of off-target cleavage by Cas9 is the formation of INDELs within or proximal to the off-target site. Here, using a p-value threshold of 0.05, the software VarScan2 detected 1271, 1996, and 1643 INDELs in M790_B4, _D2, and LmjF_3171, respectively. M790_B4 shared 70% of its parental INDELs, leaving 401 potentially novel INDELs caused by the CRISPR protocol. Likewise, the M790_D2 line contained 868 novel INDELs. However, none of the INDELs were found within 20-nt of any putative off-target site (up to 10 mismatches) ruling out the possibility that these were caused by Cas9 off-target cleavage. All the evidences

This article is protected by copyright. All rights reserved.

Accepted Article

therefore point toward successful modification of the target with no off-target cleavage within detectable limits. The data obtained from whole genome sequencing of the reference strain and the two edited clones were also used to map total genomic contents to a reference-plasmid hybrid genome and thereby confirm recombination. The targeted site (Figure 2E) appeared as a deletion in M790_B4 and _D2, but not in the wild-type. In summary, the target region appears exactly as one would expect from a successful homology-directed repair between HR1 and HR2 (Figure 2E).

Conclusions and Perspectives

To our knowledge, this is the first demonstration of CRISPR-Cas9 genome editing in *Leishmania*. We report here a specific gene knockout targeting a complex locus made of three tandemly arrayed genes. We were able to obtain null parasites in a single round of transfection, *i.e.* in a short time frame (<1 month), whereas two rounds of transfection were needed in the classical approach to knockout PFR2 (Santrich *et al.*, 1997). The CRISPR-Cas9 system may also overcome the remarkable genomic plasticity observed in *Leishmania* parasites which leads to additional copies of the targeted gene when it is essential for viability. Recovering transfectant lines when an essential gene is targeted remains cumbersome in spite of the adaptation of CRISPR-Cas9 system. This issue could be solved by transfecting an ectopic copy of the target gene. Yet, a major improvement of the system would be to develop an inducible Cas9 gene. The CRISPR-Cas9 genome editing system in *Leishmania* sp. should open opportunities for other applications, in particular, *in situ* tagging and single nucleotide substitution marker-free approaches, which, considering the limited number of selectable markers available for *Leishmania*, would be invaluable for consecutive genome manipulations. We therefore expect that the CRISPR-Cas9 system will rapidly become a routine laboratory genome editing tool in the field of *Leishmania*.

Experimental procedures

Parasites

This article is protected by copyright. All rights reserved.

Accepted Article

Promastigote forms of the *Leishmania major* genome reference strain 'Friedlin' (MHOM/IL/80/Friedlin, LEM3171), herein termed LmjF_3171, were grown in supplemented RPMI1640 medium as described (Dubessay *et al.*, 2004). Parasites were maintained in log phase by splitting the culture every second day. For FISH analysis, in order to minimize the number of dividing cells in the population, parasites were grown to late log phase and then harvested.

Construction of vectors

All cloning experiments were performed using the In-Fusion HD Cloning Kit (Clontech), the 15nt adaptors were added to the primers. The endonuclease Cas9 gene together with Nuclear Localization Signal (NLS) motifs and three Flag tags were amplified from pUF1-Cas9 (Ghorbal *et al.*, 2014) using the primer pair LS84/LS85. The PCR product was cloned into pTH6 (Dubessay *et al.*, 2006) which bears the hygromycin resistance gene and where Cas9 is expressed under the DHFR-TS promoter (pTCAS9) (Figure 1A). The design of pLS6 was based on the pL6 plasmid used for *P. falciparum* (Ghorbal *et al.*, 2014). The sgRNA cassette expression was ordered as a gene block (IDT Company®) to avoid extra bases from restriction sites; it was then inserted on a pGEMTeasy™ (Promega®) backbone between the *Apa*I and *Nco*I restriction sites. We then proceeded to cloning the homology regions boxes (HR1 and HR2): single-copy sequences on each side of the three copies of PFR2 genes were amplified using the primer pairs LS75/LS76 and LS77/LS79 for HR1 and HR2 respectively (Figure 1B). HR1 (1137 bp) and HR2 (1054 bp) were cloned between *Kpn*I and *Xba*I and between *Mfe*I and *Hpa*I, respectively, and flanked the selection marker (puromycin).

The 20-nt of the seed surrounded by the 15-nt adaptators necessary for InFusion® cloning were ordered as two oligonucleotides (PFR2Seed3For and PFR2Seed3Rev primers in Table S4) and annealed using the following program: 95 °C for 10 minutes; cooling from 95°C to 25°C using a -1.5°C/s ramp and 1min pause every 10°C and keeping at 4°C. Then, the seed was cloned into the pLS6 plasmid previously digested by *Bsg*I to obtain pLS7.

This article is protected by copyright. All rights reserved.

Accepted Article

Transfection and cloning

Parasites were transfected as previously described by electroporation (Dubessay *et al.*, 2006). For transfections with circular DNA, 100µg of each plasmid were used. For transfections with linear DNA, plasmids were linearized using the restriction enzyme ScaI, verified for linearization in agarose gel, and 100 µg were subsequently used for transfection. Prior to transfection, the appropriate DNA was ethanol-precipitated and resuspended in 50 µL of water. Drug pressure (30 µg/mL hygromycin and 30 µg/mL puromycin) was applied 15–20 h after transfection. For cloning by limit dilution, exponential Log phase culture was diluted to 5.10^3 cells/mL and several dilutions were performed to obtain 0.5 cell/ well in 96 culture plates. Positive wells were transferred into 24 wells plate to expand the culture.

Western blot

Parasites were centrifuged and washed with a protease inhibitor cocktail (Sigma®). Cells were then lysed for 5 min at 75°C in loading buffer (0.5 M Tris-HCl 10%glycerol, 2%SDS, 2%DTT, 0.01% bromophenol blue). 5.10^6 cell equivalents were loaded per well and the lysates were separated on a 10% SDS-PAGE gel and then transferred onto a PVDF membrane. Membranes were blocked using PBS, 6% skimmed milk and 0.05% Tween 20 for 1 h. For Cas9 expression, membranes were probed with a mouse anti-FLAG antibody (Sigma) diluted at 1:100 then by alkaline phosphatase-conjugated anti-mouse PA secondary antibodies diluted at 1:7500 and developed with NBT/BCIP (Promega®). To detect PFR2 expression, the antibodies used were a mouse monoclonal anti-PFR2 antibody (generously provided by Prof. Diane McMahon-Pratt, Yale) at 1:500 and an anti-mouse Phosphatase alcaline conjugated (S3721, Promega®) at 1:7500.

Analysis of mutant parasite strains by PCR.

In the mutants, target-gene disruption and puromycin-cassette integration (5' integration and 3' integration) were verified by performing PCR on genomic DNA (See Table S4 for primer pairs list). LmjF_3171 was used as a control. gDNA extraction was performed by DNeasy Blood and Tissue Kit (Qiagen®) according to the manufacturer's instructions.

This article is protected by copyright. All rights reserved.

Accepted Article

Fluorescence in situ hybridization (FISH)

The two DNA probes used for FISH were constituted by the PFR2 genes and the LmjF16.1030 loci which were PCR-amplified using the primer pairs LS110/LS111 and LS113/LS114, respectively (see Table S4). Probes were labelled with tetramethyl-rhodamine-5-dUTP (Roche Applied Sciences®) using the Nick Translation Mix (Roche Applied Sciences®). FISH analysis was performed as described elsewhere (Sterkers *et al.*, 2011). Slides were mounted in Vectashield with DAPI (Vector Laboratories®).

Immunofluorescence assay

Cells were fixed in 4% paraformaldehyde and air-dried on microscope immunofluorescence slides. Slides were treated with triton 0.2% in PBS1X and saturated with fetal bovine serum (FBS) 2%. Anti-PFR2 antibody (2E10B7, kindly provided by Prof. Diane McMahon-Pratt, Yale) diluted to 1:1000 was added for 1 hour. After 3 washes with PBS1X/2% FBS, goat anti-mouse AF488 secondary antibodies (Molecular Probes®, ref. A-11001) diluted to 1:1500 were added for 45min. Slides were finally mounted with Vectashield (Vector laboratories®) and Hoechst at 1:1000.

Microscopy and imaging

Leishmania cells were viewed by bright field, and fluorescence was visualized using appropriate filters on a Zeiss® Axioplan 2 microscope with a 100X objective. Digital images were captured using a Photometrics CoolSnap CDD camera (Roper Scientific®) and processed with MetaView (Universal Imaging®). In order to view the entire nucleus and to count the total number of labeled chromosomes, we used a Piezo controller for Z-Stack image acquisitions of 20 planes of 0.25 mm for each analyzed cell. Of note, for FISH analysis, mitotic cells were excluded from the analysis.

Whole genome sequencing

DNeasy Blood and Tissue Kit (Qiagen®) was used to prepare the genomic DNA from LmjF_3171, M790_B4 and M790_D2 cell lines. High-power sonication in a Bioruptor UCD-

This article is protected by copyright. All rights reserved.

Accepted Article

200 (Diagenode®), for three 10-min cycles, with 30 s ON pulses, followed by 30 s OFF, was used to fragment genomic DNA (1.25 µg) in 125 µL of Tris-EDTA buffer. DNA fragments (100 ng) were used to prepare indexed DNA libraries for next-generation sequencing (NGS) using the TruSeq Nano DNA LT Sample Prep Kit (Illumina®) following the manufacturer's instructions. After purification, fragment sizes and library quality were estimated in a Bioanalyzer 2100 (Agilent®) and the DNA concentration was determined in a Qubit Fluorometer (Life Technologies®). Then, 100 bp reads were sequenced in a HiSeq2500 sequencer (Illumina®) in the Transcriptome and Epigenome Platform (PF2) at Institut Pasteur.

Off-target analysis

The Protospacer Workbench software suite was used to find putative off-targets up to 10-nucleotides different from the original seed, “CAGGCACCCCACGGTATTCA”. The software scored each off-target according to the algorithm published elsewhere (Hsu *et al.*, 2013). The resulting score for each off-target is proportional to the chance of off-target cleavage actually occurring. VarScan2 software was used to detect INDELs and INDEL/off-target overlap occurring in B4, D2 and LmjF_3171 in comparison with the current version of the *L. major* genome (Friedlin strain) at GeneDB database (2013-01-16). The software's suitability and parameters were previously evaluated (Ghorbal *et al.*, 2014) and showed high sensitivity and specificity at a p-value threshold of 0.05.

This article is protected by copyright. All rights reserved.

Acknowledgements

We gratefully thank Michel Pagès for fruitful discussions and for critical reading of the manuscript. We thank Prof. Diane McMahon-Pratt (Yale School of Public Health) for the gift of anti-PFR2 antibodies. We also thank Odile Sismeiro, Caroline Proux and Jean-Yves Coppée from Institut Pasteur's Plateforme Transcriptome et Epigenome (PF2) for helping in the DNA sequencing. Finally, we gratefully acknowledge Vicki Diakou (Plateau MRI DBS-UM Optique, RIO Imaging Platform, Montpellier) and Laurence Berry (University of Montpellier/CNRS, UMR 5539) for expert advice and assistance in fluorescence microscopy, as well as Mauld Lamarque (University of Montpellier/CNRS, UMR 5539) for her fruitful help at a critical step.

This work was supported by the Agence Nationale de la Recherche within the frame of the "Investissements d'avenir" program (ANR-11-LABX-0024-01 "PARAFRAP"), ERC grant (PlasmoEscape 250320), ANR-13-ISV3-0003-01–NSFC (no. 81361130411), ANR 11 JSV3 004 01 PlasmoPiggyBAc, the Centre National de la Recherche Scientifique (CNRS), the French Ministry of Research and the Centre Hospitalier Universitaire of Montpellier. LS was funded by the French Parasitology consortium ParaFrap (ANR11-LABX0024), MG by ANR 11 JSV3 004 01 PlasmoPiggyBAc; CRM and RMM were funded by ANR-13-ISV3-0003-01–NSFC. JJLR was funded by the Institut National de la Santé et de la Recherche Médicale (INSERM).

Accepted Article

This article is protected by copyright. All rights reserved.

References

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N.*, et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339,** 819-823.

den Boer, M., Argaw, D., Jannin, J. and Alvar, J. (2011). Leishmaniasis impact and treatment access. *Clin Microbiol Infect* **17,** 1471-1477.

Dubessay, P., Blaineau, C., Bastien, P. and Pages, M. (2004). Chromosome fragmentation in leishmania. *Methods Mol Biol* **270,** 353-378.

Dubessay, P., Blaineau, C., Bastien, P., Tasse, L., Van Dijk, J., Crobu, L. and Pages, M. (2006). Cell cycle-dependent expression regulation by the proteasome pathway and characterization of the nuclear targeting signal of a Leishmania major Kin-13 kinesin. *Mol Microbiol* **59,** 1162-1174.

Dubessay, P., Ravel, C., Bastien, P., Crobu, L., Dedet, J.P., Pages, M. and Blaineau, C. (2002). The switch region on Leishmania major chromosome 1 is not required for mitotic stability or gene expression, but appears to be essential. *Nucleic Acids Res* **30,** 3692-3697.

Gaj, T., Gersbach, C.A. and Barbas, C.F., 3rd (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* **31,** 397-405.

Ghorbal, M., Gorman, M., Macpherson, C.R., Martins, R.M., Scherf, A. and Lopez-Rubio, J.J. (2014). Genome editing in the human malaria parasite Plasmodium falciparum using the CRISPR-Cas9 system. *Nat Biotechnol* **32,** 819-821.

Gilles, A.F. and Averof, M. (2014). Functional genetics for all: engineered nucleases, CRISPR and the gene editing revolution. *Evodevo* **5,** 43.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V.*, et al.* (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31,** 827-832.

Ismach, R., Cianci, C.M., Caulfield, J.P., Langer, P.J., Hein, A. and McMahon-Pratt, D. (1989). Flagellar membrane and paraxial rod proteins of Leishmania: characterization employing monoclonal antibodies. *J Protozool* **36,** 617-624.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337,** 816-821.

Lachaud, L., Bourgeois, N., Kuk, N., Morelle, C., Crobu, L., Merlin, G.*, et al.* (2014). Constitutive mosaic aneuploidy is a unique genetic feature widespread in the Leishmania genus. *Microbes Infect* **16,** 61-66.

Lye, L.F., Owens, K., Shi, H., Murta, S.M., Vieira, A.C., Turco, S.J.*, et al.* (2010). Retention and loss of RNA interference pathways in trypanosomatid protozoans. *PLoS Pathog* **6,** e1001161.

MacPherson, C.R. and Scherf, A. (2015). Protospacer Workbench: flexible guide-RNA design for CRISPR applications. *Nat. Biotechnol* **In Press**.

Maga, J.A. and LeBowitz, J.H. (1999). Unravelling the kinetoplastid paraflagellar rod. *Trends Cell Biol* **9,** 409-413.

Nakaar, V., Dare, A.O., Hong, D., Ullu, E. and Tschudi, C. (1994). Upstream tRNA genes are essential for expression of small nuclear and cytoplasmic RNA genes in trypanosomes. *Mol Cell Biol* **14,** 6736-6742.

Passos-Silva, D.G., Rajao, M.A., Nascimento de Aguiar, P.H., Vieira-da-Rocha, J.P., Machado, C.R. and Furtado, C. (2010). Overview of DNA Repair in Trypanosoma cruzi, Trypanosoma brucei, and Leishmania major. *J Nucleic Acids* **2010,** 840768.

Peng, D., Kurup, S.P., Yao, P.Y., Minning, T.A. and Tarleton, R.L. (2015). CRISPR-Cas9-mediated single-gene and gene family disruption in Trypanosoma cruzi. *MBio* **6,** e02097-02014.

Santrich, C., Moore, L., Sherwin, T., Bastin, P., Brokaw, C., Gull, K. and LeBowitz, J.H. (1997). A motility function for the paraflagellar rod of Leishmania parasites revealed by PFR-2 gene knockouts. *Mol Biochem Parasitol* **90,** 95-109.

This article is protected by copyright. All rights reserved.

Accepted Article

Shen, B., Brown, K.M., Lee, T.D. and Sibley, L.D. (2014). Efficient gene disruption in diverse strains of Toxoplasma gondii using CRISPR/CAS9. *MBio* **5,** e01114-01114.

Sterkers, Y., Lachaud, L., Crobu, L., Bastien, P. and Pagès, M. (2011). FISH analysis reveals aneuploidy and continual generation of chromosomal mosaicism in Leishmania major. *Cell Microbiol* **13,** 274-283.

Wagner, J.C., Platt, R.J., Goldfless, S.J., Zhang, F. and Niles, J.C. (2014). Efficient CRISPR-Cas9-mediated genome editing in Plasmodium falciparum. *Nat Methods* **11,** 915-918.

Accepted Article

This article is protected by copyright. All rights reserved.

Accepted Article

Figure legends

Fig 1. CRISPR-Cas9-strategy to mediate deletion of LmjF_LEM3171 PFR2 locus

A. pTCAS9 bears the endonuclease Cas9 gene together with Nuclear Localization Signal (NLS) motifs and FLAG tags. B. Correct Cas9 expression was controlled by western blot using an anti-FLAG antibody. C. pLS7 is the donor plasmid bearing a puromycin resistance cassette flanked by homology regions (HR, green boxes), the sgRNA expression cassette under a U6snRNA promoter (purple box) and a *Sca*I site for linearization. D. Map of the LmjF_3171 U6 snRNA locus used for the construction of the sgRNA expression cassette.

Fig. 2. Edition of the PFR2 loci using the CRISPR-Cas9 editing system in *Leishmania major*

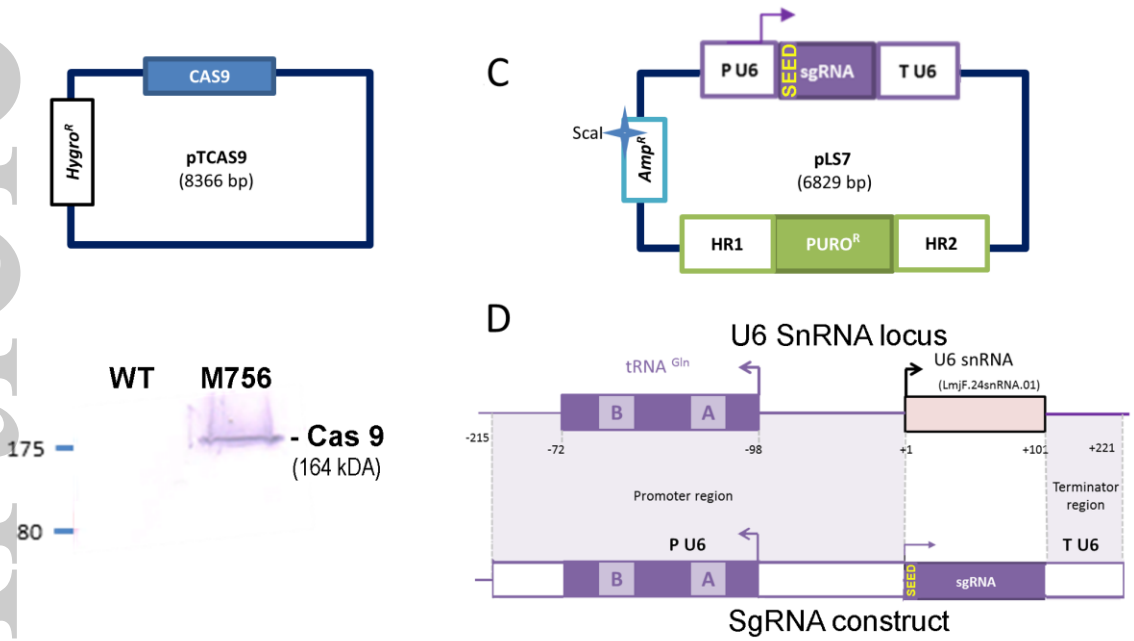
A. Schematic representation of the repaired locus after the double DNA strand break. pLS7 plasmid linearized by ScaI were positioned where the endonuclease produced the DSB. The seed targeted the 3'UTR next to HR1 (lighting). Red half-arrows (LS89 to LS91): primers used to check allele replacement in the PPR2 locus. Edition of the PFR2 locus was thoroughly verified using PCR (B), FISH (C), western blot (D), an Immunofluorescence assay (E) and whole genome sequencing (F). B. PFR2 genes were PCR-amplified in the wild-type (LmjF_3171) and M790 cell lines and found absent in clones M790_B4 and _D2. PCR testing of the 5' and 3' integration was positive in the clones and absent in the two parental lines. C. Using FISH, the PFR2 probe failed to detect chromosome 16 in the clones whereas the Lmj16.1030 probe detected it in all four cell lines. D. No PFR2 expression was found by western blot in the two clones. E. Immunofluorescence using a specific Ab showed a mixture of PFR2-positive cells and edited PFR2-negative cells in M790; Arrows: PFR-2 negative flagellum, bar = 5 µm. F. Coverage plot of the targeted region on chromosome 16 at positions 591295 through 599835. Mapping tracks are from B4, D2, and parental lines, from top to bottom, respectively. The three red bars indicate, from left to right, HR1, the target region, and HR2. The blue bars indicate gene annotations from GeneDB.

Fig. 3. Genome-wide scan of likely Cas9 targets given the seed sequence

Putative targets are ranked along the x-axis from most to least likely-to-cleave (according to Hsu et al.'s scoring algorithm). Y-axis 1 (blue) indicates the Hsu et al.'s pairwise score; Y-axis 2 (red) indicates the number of mismatches. The intended target in this figure is the top-left most point, at rank-1, as expected from any decently designed sgRNA. In this plot, the distance to the next best target and the drop-off rate is indicative of quality; here, the sgRNA sequence demonstrates a high quality sgRNA design.

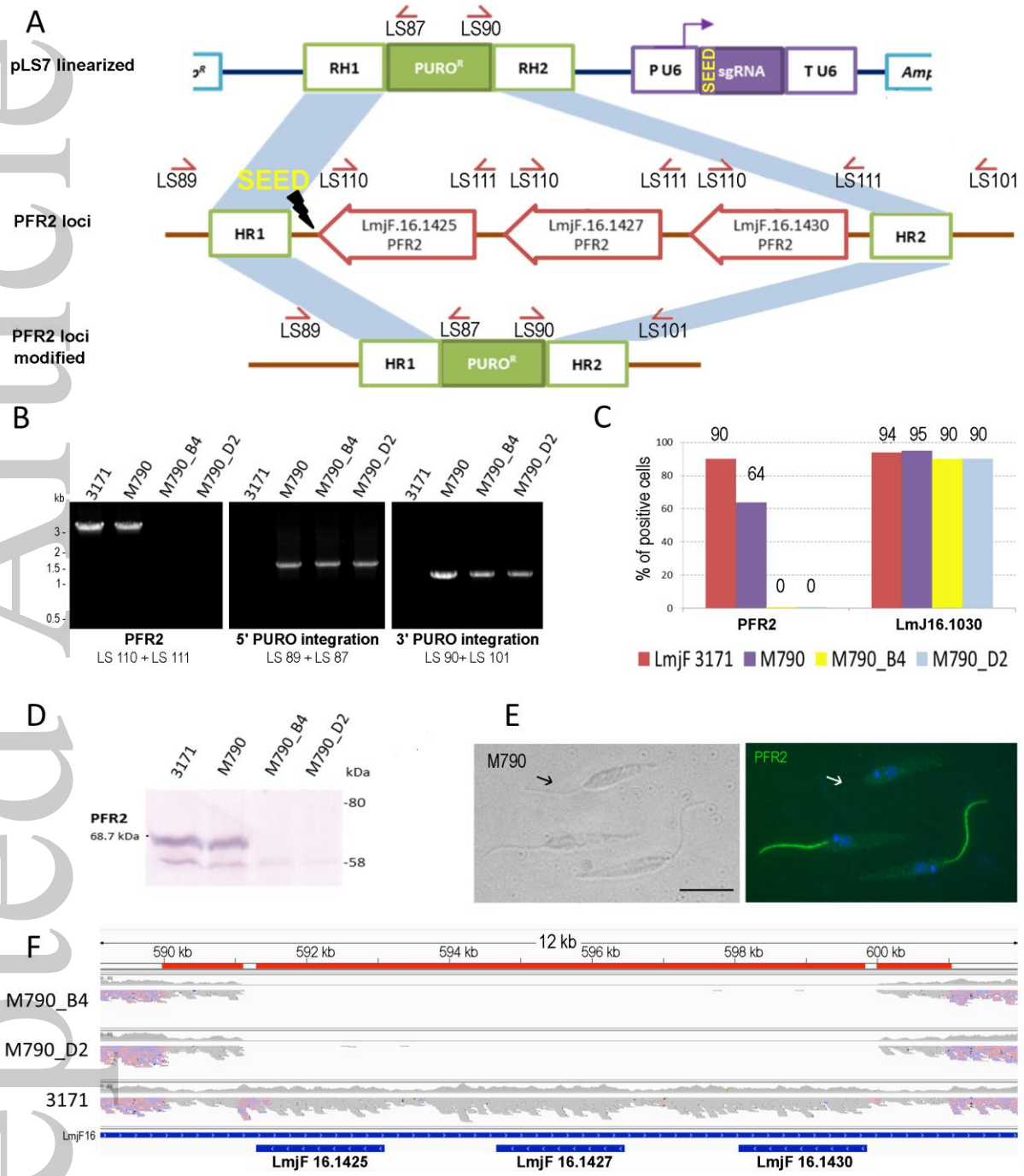
This article is protected by copyright. All rights reserved.

Accepted Article



CMI_12456_F1

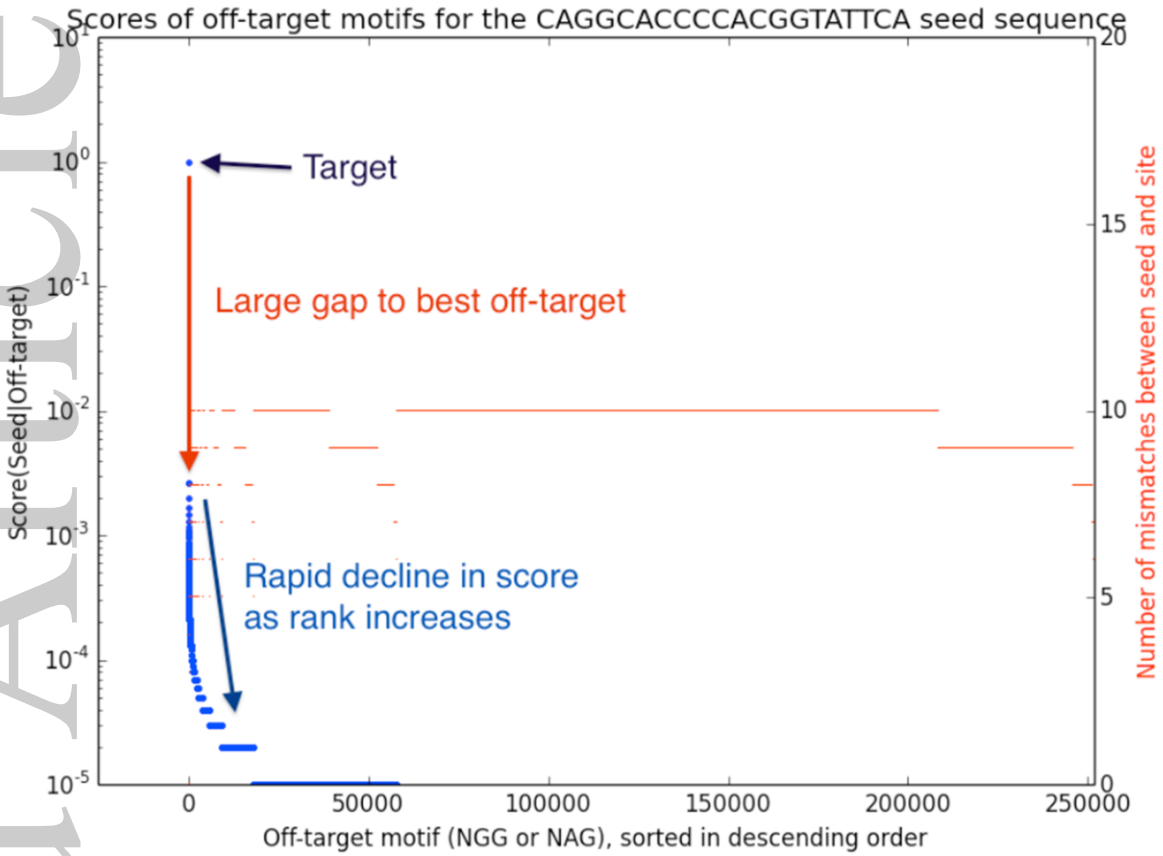
This article is protected by copyright. All rights reserved.



CMI_12456_F2

This article is protected by copyright. All rights reserved.

Accepted Article



CMI_12456_F3

This article is protected by copyright. All rights reserved.